Digit-T
DIGITAL MANUFACTURING TRAINING

DIGITAL MANUFACTURING FOR SMEs

An Introduction

Jack C Chaplin, Claudia Pagano, Santi Fort *Editors*

University of Nottingham UK | CHINA | MALAYSIA
STIIMA Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato Consiglio Nazionale delle Ricerche
eurecat Centre Tecnològic de Catalunya
AFIL Lombardia
Co-funded by the Erasmus+ Programme of the European Union

Jack C Chaplin · Claudia Pagano
Santi Fort *Editors*

Digital Manufacturing for SMEs

An Introduction

Digit-T
DIGITAL MANUFACTURING TRAINING

Editors

Jack C Chaplin
Institute for Advanced Manufacturing
University of Nottingham
Nottingham
UK

Claudia Pagano
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing
National Research Council
Milan
Italy

Santi Fort
Eurecat
Barcelona
Spain

Co-funded by the Erasmus+ Programme of the European Union

ISBN: 978-0-85358-339-4

Preface

The manufacturing industry is currently witnessing a transformation as it increasingly moves towards Digital Manufacturing - often known as Industry 4.0, Smart Manufacturing and Factory of the Future. It offers opportunities for companies to develop new products and ways of working while reducing costs. However, many people and organisations, in particular SMEs, struggle to access clear and useful information about Digital Manufacturing.

This book aims to provide SMEs and other users with an introduction to digital manufacturing. It has been compiled by experts who are members of the consortium of the ERASMUS+ project Digit-T: Digital Manufacturing Training System for SMEs (2017-1-UK01-KA202-036807). This includes University of Nottingham, Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA-CNR), EURECAT and AFIL.

This book collates contributions from within the field of digital manufacturing, and complements a free online training course developed by the Digit-T consortium which can be found at https://training.digit-t.eu/.

This book is divided into 3 parts.

Part I: Management in Industry 4.0 aims to provide a basic understanding of the concepts, trends and key technologies that characterise Industry 4.0. In addition it considers the roadmapping process which can support a company in prioritising and planning its own digital transformation. Finally, consideration is given to the impact of Industry 4.0 in human resource management, discussing the evolution and "buy-in" required in the work force, both leaders and employees, which is essential for a company to successfully move to a "Smart Factory".

Part II: Manufacturing Systems introduces the concepts of manufacturing analysis and decision making and breaks them down into formal processes which can be followed. It starts by introducing conventional decision making – methods for analysing manufacturing systems and networks to calculate key performance indicators or to identify areas of concern. The limitations of these methods are discussed before moving to discuss modern methods for manufacturing systems analysis, using offline modelling and simulation and state-of-the-art integrated digital twins and decision support systems.

Part III: Intelligent Robotics starts by explaining robotics generally, the different types of robots which exist, and providing a simple introduction to robot terminology, their different structures, their components and applications. The impact of Industry 4.0 in robotics is then considered and includes discussions on Cyber Physical Systems, collaborative robots which support human-robot collaboration, microrobotics and mobile robots. Finally, Artificial Intelligence and ethical issues associated with autonomous systems is considered.

This book is designed to be of interest to managers, engineers, researchers, students and lay people either operating or having interest in the manufacturing sector and who wish to gain a general understanding of the area of digital manufacturing and Industry 4.0. We hope we have achieved this goal.

Nottingham, UK — Jack C Chaplin
Milan, Italy — Claudia Pagano
Barcelona, Spain — Santi Fort

Contents

Acronyms

1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
5D	Seiri (Sort), Seiton (Set in Order), Seiso (Shine), Seiketsu (Standardise) and Shitsuke (Self-Discipline)
5G	Fifth generation cellular protocols
6LowPAN	Internet Protocol version 6 over Low-Power Wireless Personal Area Networks
AC	Alternating Current
ADKAR	Awareness, Desire, Knowledge, Action, Refinement
AGV	Automated Guided Vehicle
AHP	Analytical Hierarchy Process
AI	Artificial Intelligence
AIDC	Automatic Identification and Data Capture
AMM	Assemble-Measure-Move
ANN	Artificial Neural Network
ANSI	American National Standard Institute
API	Application Programming Interface
AR	Augmented Reality
AS/EN	Aerospace Standard / European Standard
ASME	American Society of Mechanical Engineers
B2C	Business to Consumer
BI	Business Intelligence
CAD	Computer Aided Design
CAE	Computer Aided Engineering
CAM	Computer Aided Manufacturing
CAN	Controller Area Network
CAPP	Computer Aided Process Planning
CAT	Computer Aided Tolerancing
CD-DSS	Communications-Driven Decision Support System
CDO	Chief Digital Officer
CEN	European Committee for Standardization

CENELEC	European Committee for Electrotechnical Standardization
CEO	Chief Executive Officer
CFRP	Carbon Fiber Reinforced Plastic
CIA	Confidentiality, Integrity and Availability
CL	Centreline of a control chart
CMM	Coordinate Measurement Machine
CMR	Crawler Mobile Robot
CNC	Computer Numerical Control
Cobot	Collaborative Robot
CPDMS	Capacitive PolyDiMethylSiloxane
CPS	Cyber Physical System
CRCPS	Collaborative Robotics Cyber-Physical Systems
CRM	Customer Relationship Management
CSV	Comma Separated Values
DAM	Decision Analysis Models
DBMS	Database Management System
DD-DSS	Data-Driven Decision Support System
DDS	Data Distribution Service
DES	Discrete Event Simulation
DfA	Design for Assembly
DfM	Design for Manufacture
DIS	Draft International Standard
DKIW	Data, Information, Knowledge, Wisdom
DMAIC	Define, Measure, Analyse, Improve, Control
DMS	Dedicated Manufacturing System
DOF	Degree(s) of Freedom
DSS	Decision Support System
EAS	Evolvable Assembly System
EASA	European Aviation Safety Agency
EBITDA	Earnings before Interest, Taxes, Depreciation, and Amortisation
EMC	Electromagnetic Compatibility
EN	Issued by CEN
ERP	Enterprise Resource Planning
ESB	Enterprise Service Bus
EtherCAT	Ethernet for Control Automation Technology
ETL	Extract, Transform, Load
Euro NCAP	European New Car Assessment Programme
FDM	Fused Deposition Modelling
FIFO	First In, First Out
FMS	Flexible Manufacturing System
GA	Genetic Algorithm
Gbps	Gigabits per second
GDPR	General Data Protection Regulation
GFSi	Global Food Safety Initiative

GHz	Gigahertz
GMAW	Gas Metal Arc Welding
GTAW	Gas Tungsten Arc Welding
HANA	HAsso's New Architecture
HMI	Human-Machine Interface
HRC	Human-Robot Cooperation
Hz	Hertz (frequency)
I4.0	Industry 4.0
ICT	Information and Communications Technology
ICY	Interchangeability
ID	Identifier
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IFR	International Federation of Robotics
IIC	Industrial Internet Consortium
IIoT	Industrial Internet of Things
IMU	Inertial measurement Unit
IoT	Internet of Things
IP	Internet Protocol
IP67	Ingress Protection code 67
Ipv6	Internet Protocol version 6
IR	InfraRed
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
JTC	Joint Technical Committee
kbps	Kilobits per second
KD-DSS	Knowledge-Driven Decision Support System
KNIME	Konstanz Information Miner
KPI	Key Performance Indicator
LAN	Local Area Network
LCL	Lower Control Limit of a control chart
LED	Light-Emitting Diode
LIFO	Last In, First Out
LMR	Legged Mobile Robot
LMS	Learning Management System
LWR	Light Weight Robot
M2M	Machine to Machine
MAA	Measurement / Metrology-Assisted Assembly
MADA	Measurement-Assisted Determinate Assembly
MAG	Metal Active Gas welding
MAST	Manufacturing Agent Simulation Tool
MCS	Monte Carlo Simulation
MD-DSS	Model-Driven Decision Support System
MEMS	Micro-Electrical-Mechanical Systems

MES	Manufacturing Execution System
MESA	Manufacturing Enterprise Solutions Association
MHS	Material Handling System
MHz	Megahertz
MIG	Metal Inert Gas welding
ML	Machine Learning
MQTT	MQ Telemetry Transport
MTBF	MTBF: Mean Time Between Failure
NASA	National Aeronautics and Space Administration of the United States of America
NoSQL	Not Only Structured Query Language
NP	Nondeterministic Polynomial Time
NPI	New Product Introduction
OCR	Optical Character Recognition
OEE	Overall Equipment Effectiveness
OMG	Object Management Group
OPC UA	Open Platform Communications – Unified Architecture
OT	Operational Technologies
PAM	Pneumatic Air Muscles
PCA	Principal Component Analysis
PCDA	Plan, Do, Check, Act
PD	Proportional Derivative
PDA	Personal Digital Assistant
PESTLE	Political, Economic, Sociological, Technological, Legal and Environmental
PI	Proportional Integral
PID	Proportional Integral Derivative
PL	Performance Level
PLC	Programmable Logic Controller
PLM	Product Lifecycle Management
PN	Petri Nets
POS	Point of Sale
Profibus	Process Field Bus
PROFINET	Process Field Net
PSO	Particle Swarm Optimisation
PZT	Lead zirconate titanate
QN	Queueing Network
R&D	Research and Development
RCC	Remote Center of Compliance
RFID	Radio Frequency Identification.
RIA	Robotic Industries Association (US)
RL-DSS	Reciprocal Learning-Based Decision Support System
RMS	Reconfigurable Manufacturing System
ROCE	Return on Capital Employed Ratio

RoHS	Restriction of Hazardous Substances
ROI	Return on Investment
ROS	Robot Operating System
S3	Simple Storage Service
SA	Simulated Annealing
SaaS	Software as a Service
SAP	A software corporation specialising in enterprise support software
SC	Subcommittee
SCADA	Supervisory Control And Data Acquisition
SCARA	Selective Compliance Assembly Robot Arm
SCM	Supply Chain Management
SD	System Dynamics
SDK	Software Development Kit
SEM	Search Engine Marketing
SEO	Search Engine Optimization
SMA	Shape-Memory Alloy
SMAC	Social, Mobile, Analytics and Cloud
SME	Small and Medium size Enterprise
SPC	Statistical Process Control
SQC	Statistical Quality Control
SQL	Structured Query Language
SRP/CSs	Safety related Parts of Control Systems
SWOT	Strengths, Weaknesses, Opportunities and Threats
TC	Technical Committee
TCP	Tool Centre Point
TIG	Tungsten Inert Gas welding
TQM	Total Quality Management
TR	Technical Report
TS	Technical Specification
UCL	Upper Control Limit
USB	Universal Serial Bus
USD	United States Dollar
UX/UI	User Experience/User Interface
VR	Virtual Reality
WIP	Work in Progress
WMR	Wheeled mobile robot
WPPF	Whole-Part Predictive Fettling
WSN	Wireless Sensor Network
XML	eXtensible Mark-up Language

Part I
Management in Industry 4.0

Part I Overview

Companies are navigating an era of profound change. Compared to earlier industrial revolutions, the so called **fourth revolution or Industry 4.0 (I4.0)** encompasses not just one but a combination of disruptive technologies (e.g. artificial intelligence, additive manufacturing, cloud computing and cybersecurity) moving at breakneck speed and providing unparalleled transformations across multiple industries.

The advancements in technology have been tremendous over the past decade, and it only continues to improve at an extremely rapid pace. The opportunities that have opened up for the enterprises capable of embracing emergent technologies are immense: novel value chains, new profit streams, new business models, streamlining operations… Nevertheless, not every company has a consolidated understanding, specialised team and overall vision covering all the components of this modern industrial revolution.

The next three chapters of this book focusses on understanding the basic concepts of the Industry 4.0 phenomenon, its strategic implementation and management in the business field. They emphasized that technology transformation requires strategic leaders who can engage and examine technology through both a business and technology lens.

Chapter 1 establishes the baseline of Industry 4.0 (I4.0), presenting a broad definition of (I4.0) and its derived key concepts and technologies. Chapter 2 reflects on the I4.0 strategic implementation, explaining how it is much more than just a sole and simple technology integration. I4.0 mandates a corporate design and management strategy to be set in place. In this context, technology roadmapping is presented as a tool for Industry 4.0 transformation to facilitate the planning and implementation process. Finally, Chapter 3 highlights how people are at the centre of I4.0, discussing the impact of Industry 4.0 on the employment landscape, Human Resources Management and changes on the way people work, learn and businesses produce value spans.

Chapter 1 Fundamental Concepts of Industry 4.0

Mireia Dilmé i Martínez de Huete

1.1 Introduction

Businesses have always been evolving and innovating, new technologies have always come with challenges and opportunities, and ecosystems have always been changing. So what makes the current era of technological change so different?

It's in the degree of interconnectedness, the speed of acceleration of change and the very nature of the change itself. Disruption isn't new, but the speed, complexity and global nature of the disruption is at a scale never seen before.

Businesses in today's competitive environment are increasingly being transformed by technology. Such technology transformation enables companies to grasp opportunities to increase revenues, improve efficiency and flexibility, and deliver more value to end-users.

Technology transformation requires a strategic leader who can engage and examine technology through both a business and technology lens. This book aims to empower users with the basic knowledge of the Industry 4.0 technology domain, its management and strategic implementation.

The present chapter aims to provide a basic understanding to the concepts, trends and key technologies that characterize Industry 4.0. Most importantly, it aims to reflect upon the reasons why enterprises should bet on digital transformation and the change of paradigm in Industry 4.0. This chapter serves as an introduction to the potential that Industry 4.0 unveils. Thus, it can be used as a general framework to then go into further details in understanding the relationship amongst each technology blocks presented in the subsequent book chapters.

M. Dilmé i Martínez de Huete (🖂)
Eurecat, Barcelona, Spain
e-mail: mireia.dilme@Eurecat.org

J.C. Chaplin et al. (eds), *Digital Manufacturing for SMEs*
DOI: 10.17639/9vtj-m308

1.2 Introduction to Industry 4.0

If you are a small-to-medium enterprise (SME), it is quite likely that you have already heard the term *Industry 4.0* (I4.0). You may not, however, be persuaded that it is vital to your business. Many SMEs erroneously believe that I4.0 only relates to initiatives pursued by multinational corporations with sizeable budgets like General Electric, Bosch or Boeing. Regardless of the organisation size, the enterprises maintaining this conservative mind-set are actually missing out on the opportunities to increase competitiveness and to innovate. For SMEs in particular, **Industry 4.0 should not merely be perceived as an end goal, but a direction which every enterprise can move toward in some degree.**

What exactly is Industry 4.0? What changes is it bringing? Why does it matter for small and mid-sized industrial enterprises?

1.2.1 Welcome to the 4th Industrial Revolution (Industry 4.0)

Did you know that…?

- 90% of the data in the world today has been created in the last two years alone (IBM research).
- 30% of companies started monetizing their data assets in 2017 (Microsoft research).
- The average lifespan of an S&P 500 company has decreased by 50 years in the last century, from 67 years in the 1920s to just 15 years today (Yale research).
- 86% of CEOs consider digital as their first priority (Microsoft research).
- 76% of millennials believe that innovations their most valuable trait (Deloitte research).

Industrial SMEs need to stay ahead of the ever-changing market requests: hyper-personalised products, shorter product and service life-cycles, global interconnection, the high speed of technological change, and the constant drive for delivering quality at reduced cost.

The rules of the game are changing. How enterprises cope with the transformational challenges is not only by refining their corporate strategies but also creating unparalleled opportunities. What makes the current era of technological change stand out?

- *Interconnectedness:* Now more than ever equipment, personnel and industrial processes are interconnected. Data is gathered and analysed at a global scale, constantly optimizing and facilitating the decision making processes. The physical world is becoming an information system itself. Companies like FedEx, have already stated that *"information is more valuable than any transported good"*.

- *Pace of Change:* Things move at an extremely fast pace in a race where it is hard to keep up. Previously, enterprises had time to track trends and wait for proof of the success of new applications in various settings before adopting them in-house. Now, however, new capabilities are rolled out faster every year, disruptive technological innovations pop up constantly and users are embracing novel technologies at an unrelenting speed.

  For example, as illustrated in Figure 1.2-1, it took about 75 years for the telephone to connect 50 million people. Today a simple iPhone app can reach that milestone in a matter of days.

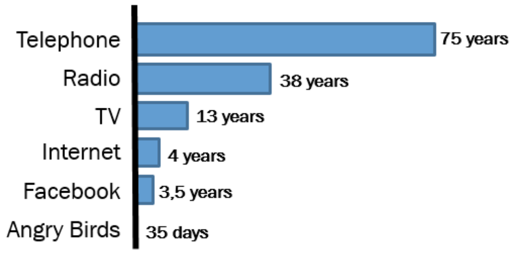


Figure 1.2-1 Time to reach 50 million users [1].

- *Nature of the latest technological change:* Adaptability is key to survival in the age of digital Darwinism, an era in which market drivers and demands are constantly changing. The long-term winners will not be the ones who simply try to make it to the next level but those who constantly adapt. Over the next few years, technologies that have not yet been understood completely by people (e.g. quantum computing) may tremendously influence industrial systems. Enterprises need to prepare their teams, infrastructure and capabilities to successfully embrace the potential of those technologies. This also includes learning to make data-driven and fast-paced decisions and progress within highly uncertain environments.

  As illustrated in Figure 1.2-2, Unicorns, a moniker applied to private start-ups valued at > $1 billion, are no longer so mythical in nature. Technological and digitalized enterprises are leading the unicorn club due to their capacity to operate in uncertain environments and their predisposition to adopt and bet on highly disruptive technology.

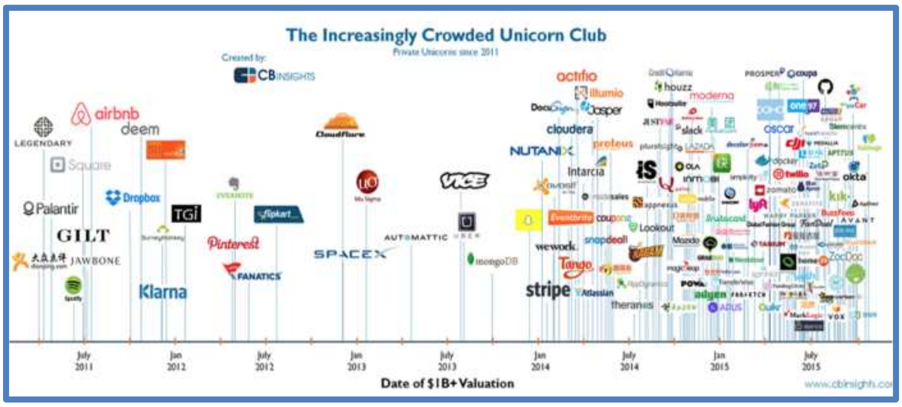


Figure 1.2-2 Technological and digitalised enterprises are leading the unicorn club [2].

These three factors, when combined, result in an unstoppable and self-reinforcing force, boosting and speeding each other up: massive usage of technology accelerates the development and introduction of more technology.

These novel market rules open tremendous opportunities but they can only be met by radical advances in current manufacturing technology. Within this context, Industry 4.0 is a game changer: a key enabler for enterprises to stay ahead in terms of innovation. To achieve this, Industry 4.0 is based on the integration of the company's value chain (suppliers, partners and customers), business and manufacturing processes as well as the adoption of ICT (both hardware and software) to current industrial production systems.

The concept of Industry 4.0 signifies the so-called fourth industrial revolution. Industry 4.0 is derived from the ongoing transformation in the industrial sector preceded by three other revolutions. The first revolution, around 1784, refers to the mechanization of work: incorporating the water/steam engine into mechanical manufacturing facilities. This consisted of the introduction of steam into the tasks that were previously performed by hand.

The second industrial revolution, around 1870, followed the introduction of electrically-powered mass production based on the division of labour. The introduction of electricity into various manufacturing processes made the use of assembly lines possible.

The third industrial revolution, around 1970, was based on the introduction of Programmable Logic Controller (PLCs) – electronics and IT – to achieve further automation of manufacturing. The third industrial revolution was already a huge leap ahead where, in the advent of automation, electronics and computers ruled in the industrial scene. During this era, robots and programmable machinery were increasingly used to perform industrial tasks.

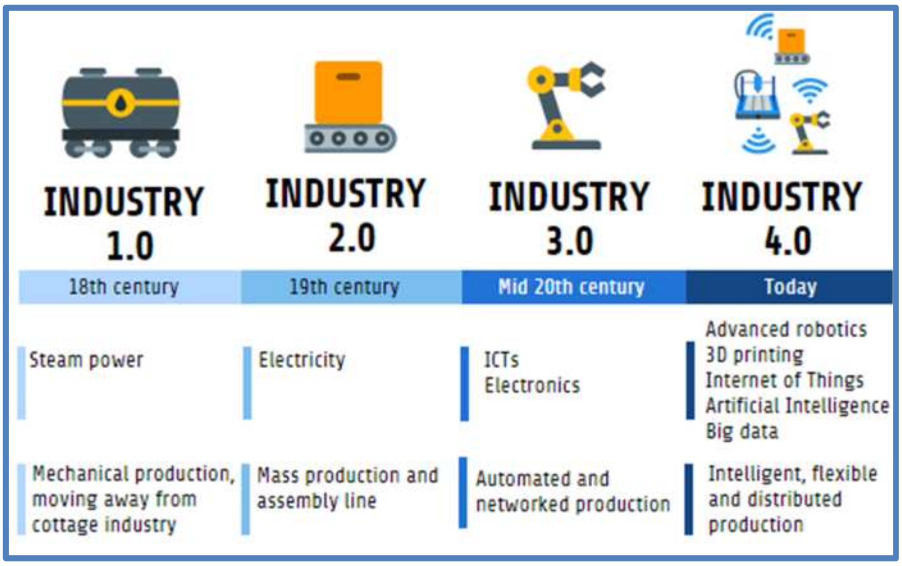


Figure 1.2-3 Evolution from Industry 1.0 to Industry 4.0 [3].

Today, Industry 4.0 is a change of paradigm: not only one but a myriad of cyber-physical technologies are combined to digitally transform industrial activities. Cyber Physical Systems (CPS) are comprised of storage systems, data processing capabilities, smart machines and manufacturing facilities capable of autonomously exchanging information, driving actions and controlling each other independently. An enterprise leveraging Industry 4.0 is not merely investing in production line automation, but it is transforming its production lines through IIoT (Industrial Internet of Things), using cloud technology and applying advanced software and data analytics. Through IIoT, a large number of synchronized sensors provide real time data to an enterprise's computer servers (local or cloud). All this data provides highly valuable information to decision-making processes and it is the key baseline to nourish predictive models which help enterprises to anticipate irregularities in its systems, operations and processes and, hence, actions can be properly taken before errors or major breakdowns occur. This data analysis (if in huge amounts known as Big Data) is the key to maintaining and improving the supply chain, industrial processes and product lifecycle management. Therefore, the result of Industry 4.0 is to create a highly flexible, intelligent, distributed production and service network. The end goal: to pave the way to reach the concept of a Smart Factory that is characterized by adaptability, flexibility and efficiency whilst improving the value delivered to targeted customers.

To ensure that the key difference between the third industrial revolution and Industry 4.0 is fully understood, let's consider the example of a CNC machining centre. If the machine is part of the third industrial revolution era, the tool change can be done automatically but an operator should manually keep track, observe and correct the actions regarding, for example, the spindle speed. However, if the machine is updated to the Industry 4.0 era, the tool changes are also done automatically but, at the same time, the spindle speeds and many other crucial process parameters are automatically recorded by the sensors integrated in the machine. Due to large data processing capabilities, proper settings are automatically calculated by the machine on its own and the process is automatically optimized [4].

1.2.2 Change of Paradigm in Industry 4.0

The first official use of *Industrie 4.0* was coined in Germany around 2011, as a strategic initiative introduced by the German government under the goals to:

1. Identify various trends that were taking place.
2. Encourage projects for the digitalization and introduction of high level technology in manufacturing.

In the following years (mainly after 2014), companies and governments outside of Germany began to step in. The most important move came when the European Commission established a priority: setting a target for the industrial sector to represent the 20% of the European economy up to 2020 thus increasing productivity, competitiveness and overall enterprises added value. To achieve this target, government initiatives, dissemination efforts, specialized financing policies and tools were established.

Industry 4.0 represents a qualitative leap in organization management, control of the entire value chain and monitoring of the entire product lifecycle. In fact, it is a paradigm shift for industries, requiring novel capacities and opening new windows of opportunity – some of which are described in Table 1.2-1.

Change of paradigm in Industry 4.0	
Traditional industry	**Industry 4.0**
Mass production	Hyper-personalisation based on customers' demands
Large factories to manufacture large volumes of one specific product	Intelligent factories with flexible production lines to produce at competitive costs
Rigid production planning based on stock forecasts	Dynamic production based on market demand
Revenues derived from product sales	Revenues derived from product as a service
Cost minimization	Maximization of ROCE: profitability / capital used.
Labour rigidity	Flexibility in the organization of work.

Table 1.2-1 Maturity model for the adoption of Industry 4.0 [5].

The shift in Industry 4.0 is based on the following principles:

- *Ensuring Interoperability:* The communication capacity of all elements of the factory. There is a need to create common standards that facilitate data flows amongst the cyber-physical systems, robots, corporate information systems, intelligent products and people, as well as third-party systems.
- *Decentralization:* Emphasizing greater autonomy and putting intelligence at the lowest practical level. For instance, implementing cyber-physical elements with the capacity to make decisions autonomously in order to reduce production time and costs. Coordination must be ensured, but a rigid, top-down organization is seen as undesirable.
- *Real Time Analytics:* Massive data collection and analysis (Big Data) in real time that allows the monitoring, control and optimization of processes, facilitating any decision derived from the process immediately.
- *Virtualization:* The ability to generate a virtual copy of the factory through the data collected; in other words, to digitize physical elements. Virtual models of the plant and modelling of industrial processes enable simulation models to perform experiments and better identify and compare alternatives that improve the current production systems.
- *Orientation to Service:* The ability to transfer greater value directly to the customer. This value signifies a better product, novel service or even improved business models.
- *Modularity and Flexibility:* Flexibility and elasticity to constantly adapt to the needs of the industry.

As a result of this paradigm shift, what does an enterprise in the Industry 4.0/Digital era look like?

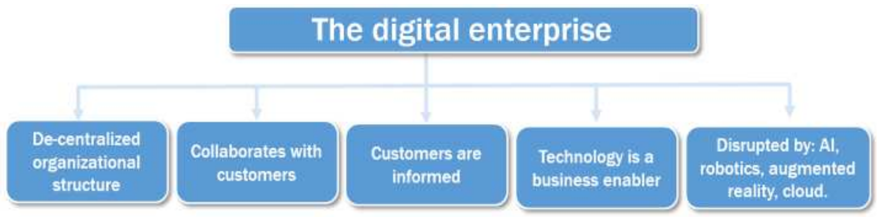


Figure 1.2-4 An enterprise in the Industry 4.0 era. Image rights: Eurecat, adapted from [6].

If successfully implemented, Industry 4.0 principles can impact performance across a myriad of enterprise functions. Several studies, such as McKinsey “Industry 4.0: How to navigate digitization of the manufacturing sector” have quantified this performance gain. The assessment by McKinsey is given in Figure 1.2-5.

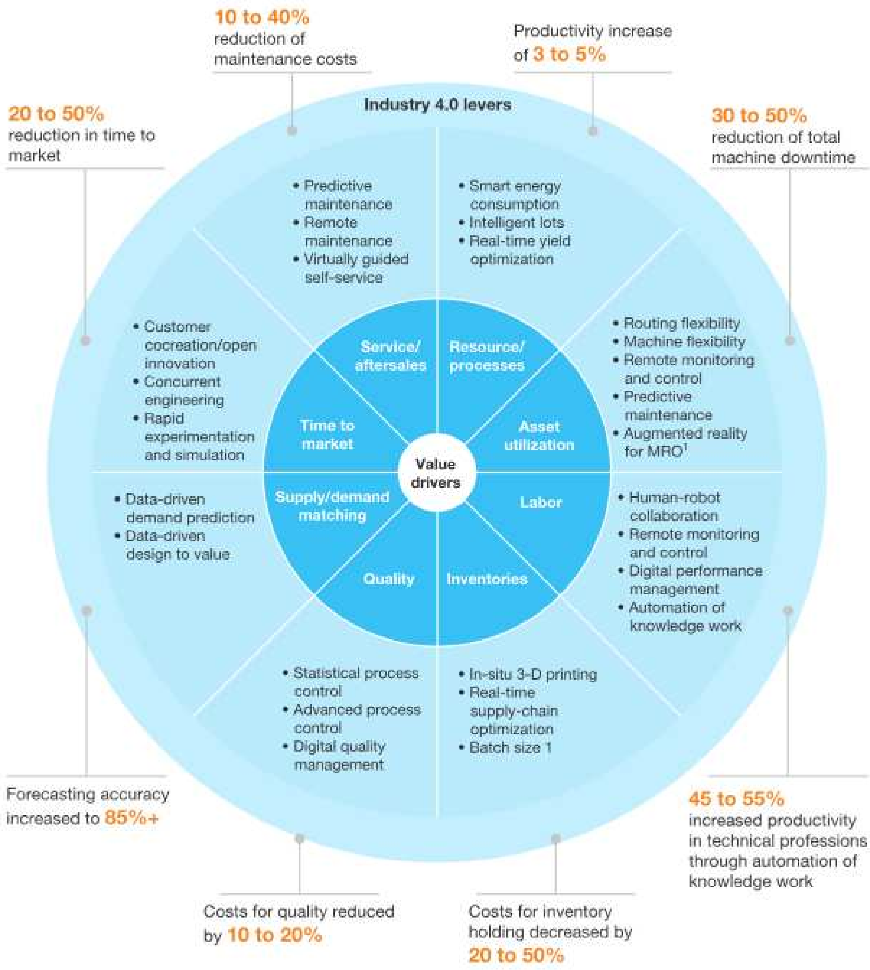


Figure 1.2-5 Performance benefits of digitalization in Industry 4.0 [7].

1.2.3 Industry 4.0, Building the Digital Enterprise

Paperless processes, robotic applications, Internet of Things, digital marketing, digital habits of customers, increasing mobile access. In today's business world everybody is talking about digital technologies and digital transformation. But how does it relate to Industry 4.0?

Some literature defines Industry 4.0 as "*profound transformation of business models by enabling the fusion of virtual and real words and the application of digitization, automation, and robotics in manufacturing*" (Gotz and Jankowska

2017). In short, when planning to transform into a Smart Factory, or be Industry 4.0 ready, a key step is to embrace digitalization. Industry 4.0 requires end-to-end digitisation of all physical assets and integration into digital ecosystems with value chain partners. The ability to generate, analyse and communicate data seamlessly through digitalised processes is key to underpin the Industry 4.0 gains.

In this context, digitisation, digitalization and digital transformation are three terms that tend to be erroneously used as synonyms but have distinctive and important meanings. Let's introduce key definitions to better understand the progressive path towards digital transformation:

- *Digitisation* is the conversion from analogue to digital. Analogue information is encoded and become bits (i.e. digitization of data). Converting handwritten, typewritten or "paper-based" text into digital form is the most simplistic example of digitization. For instance, for a service technician undertaking a field visit to a customer, digitization would imply that the technician is able to easily access all the customer's files, repair reports and product manuals in a digital format wherever he/she is and before, during and/or after field visit. Standalone digitisation doesn't necessary bring monetary benefits (savings or earnings). Nevertheless, it is a must in order to advance towards the digital transformation path.
- *Digitalization* refers to using digital technology in specific operations and the impact it has, generally in terms of cost saving (e.g. digitalization of a process reduces the amount of low-added value human time and effort). In other words, whilst digitisation was merely bringing the information to the digital realm, digitalization is the process of making digitised information work for you. Going back to the service technician example, centralized information about product history and customers (previous issues, replacement history, online manuals, customer contact, etc.) can assist technicians to achieve a first-time-fix contributing towards a smoother and more effective service. Therefore, the technician avoids consulting beforehand an immense amount of obsolete papers to get a deep understanding of the potential solutions and customer pains before the on-site visit.
- *Digital transformation* is a digital-first mentality that encompasses all aspects of business, no matter its nature (if it concerns a digital business or not). Digital transformation leads to the creation of entirely new markets, customers and businesses (capabilities, processes, revenue and operating models). Digital transformation is not something that organisations can implement as individual projects. It rather implies a transversal impact across the organisation: devising new business models, streamlining operations, entering novel markets and disruptively changing how operations are done. As a result, digital transformation can create novel profit streams as well as enable huge savings in the most valuable corporate resources: money and time.

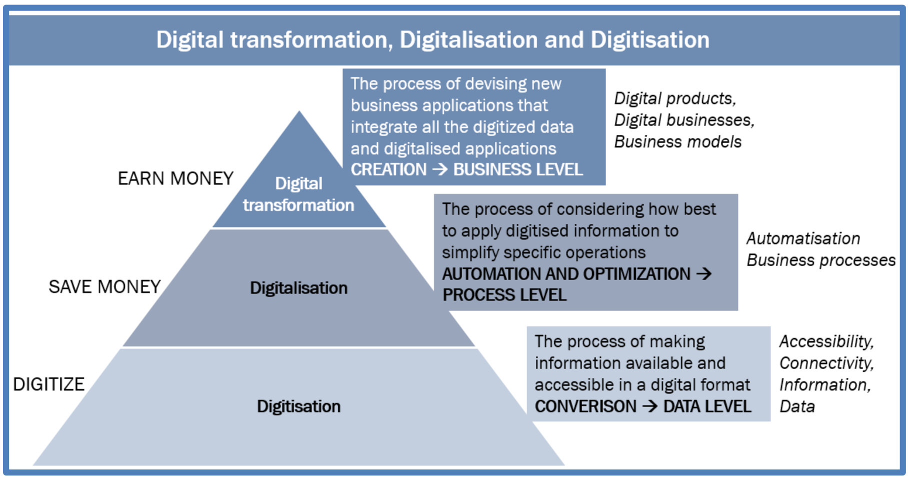


Figure 1.2-6 Digital transformation, digitalization and digitisation. Image rights: Eurecat, adapted from [8].

To summarise, we digitize information, we digitalize processes and roles that make up the operations of a business, and we digitally transform the organisation and its strategy.

1.3 Key Enabling Technologies for Industry 4.0

There is no single list of technologies related to Industry 4.0. In recent years, many consultancies and other organizations have published schemes representing the main technologies, with each scheme deriving from slightly different perspectives.

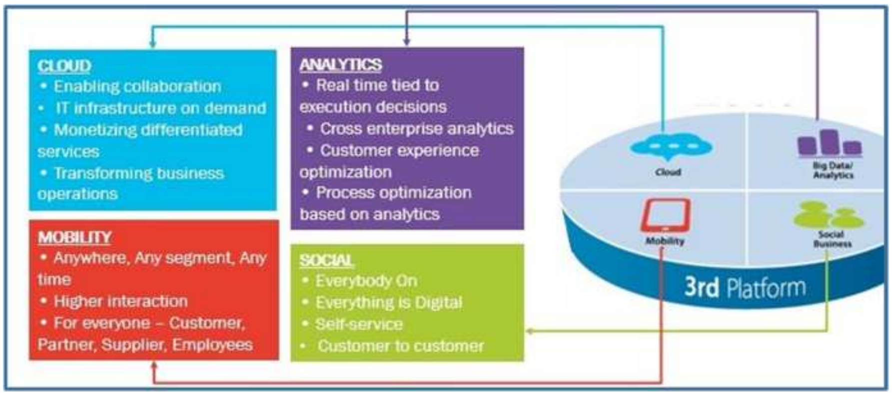


Figure 1.3-1 SMAC forces. Image rights: Eurecat, adapted from [9].

From a digital transformation perspective, in 2012, Gartner introduced the 'nexus of forces' **SMAC (Social, Mobile, Analytics and Cloud)** as the emerging technologies contributing to digital business transformation. In the past few decades the driving forces, behind business agility, were mainly **systems and IT abilities**. However, the current main driver is **information: how information is obtained, managed and used.**

DIGIT-T introduces some of the key technology trends most commonly referred to be leading the road to smart factories, cyber physical systems and end-to-end value chains with Industrial Internet of Things (IoT) and decentralized intelligence in manufacturing, production, logistics and the industry. The technologies defined as the 9 key drivers or building blocks in Industry 4.0 comprise:

1. Autonomous Robots
2. Simulation
3. System Integration
4. Internet of Things
5. Cybersecurity
6. Cloud Computing
7. Additive Manufacturing
8. Augmented Reality
9. Big Data

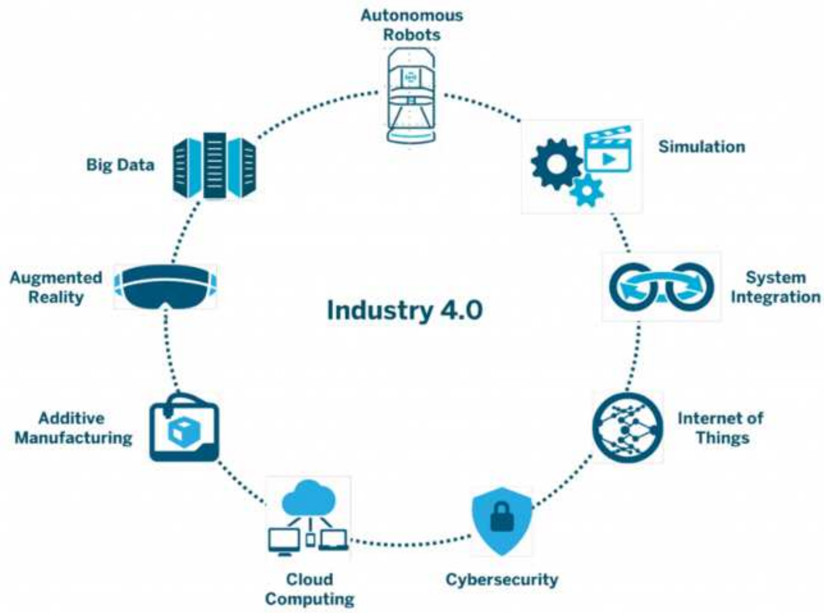


Figure 1.3-2 Key Enabling Technologies in I4.0. Image rights: Eurecat, adapted from [10].

These technologies will now be briefly described.

It is important to highlight that Industry 4.0 is not focused on a specific technology but rather on how to use and combine these technologies to achieve the organizations' planned objectives.

In the following chapters some of these key technologies will be further analysed further to provide a deeper understanding on how and why they are transforming industrial production.

1.3.1 The Internet of Things

The *Internet of Things* (IoT) is the extension of internet connectivity into physical devices and everyday objects. By incorporating embedded electronics, internet connectivity and other forms of hardware (such as sensors) within physical devices, they are capable of communicating and interacting with other physical devices over the internet and can be remotely monitored and controlled.

IoT has many diverse applications in different sectors, such as smart homes, healthy aging, medical and healthcare, transport, etc. The term Industrial Internet of things (IIoT) is often encountered in the manufacturing sector, referring to the industrial subset of the IoT.

Through the IIoT it is possible to connect any element of an industrial plant, transmitting and/or receiving information thus permitting real time monitoring and control and subsequent data analytics. The elements that can be connected include elements inside the plant such as machinery, personnel, tools and raw materials, together with external elements such as vehicles, manufactured products and even customers.

There are several basic building blocks involved in the development of an IoT application:

1. *The Connected Devices:* The physical devices we want to control and manage.
2. *The Gateway:* The element that connects the device to the Internet.
3. *The Internet:* The infrastructure that allows objects and other elements such as computers, servers and data centres to communicate with each other.
4. *The Cloud:* The set of servers and data centres which contain the platform where the information is stored and processed.
5. *The App or Software:* IoT applications usually have an App which allows the users to interact with the platform and visualize the results, control the devices, etc.

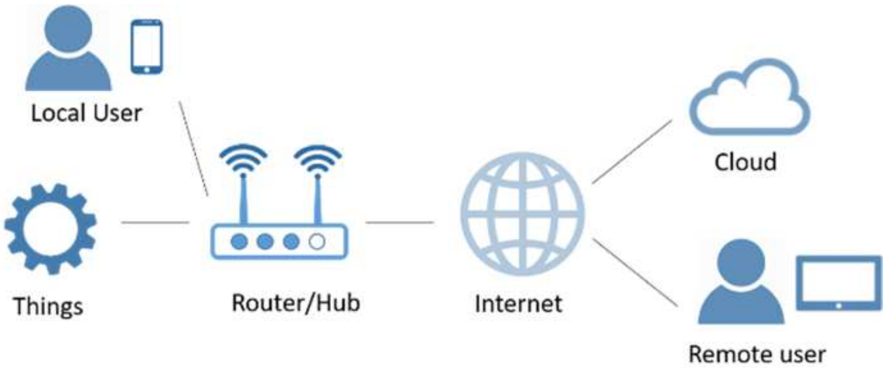


Figure 1.3-3 Elements of IoT. Image rights: Eurecat, adapted from [11].

1.3.2 Analytics and Big Data

Data analytics is the science of analysing raw data in order to make conclusions about that information, capable of permitting businesses to optimize their performance. As illustrated in Figure 1.3-4, data analytics is divided into four different types:

1. *Descriptive Analytics:* Describes what has happened over a given period of time.
2. *Diagnostic Analytics:* Focuses more on why something happened.
3. *Predictive Analytics:* Concentrates on what is likely to happen in the short term.
4. *Prescriptive Analytics:* Suggests a course of action.

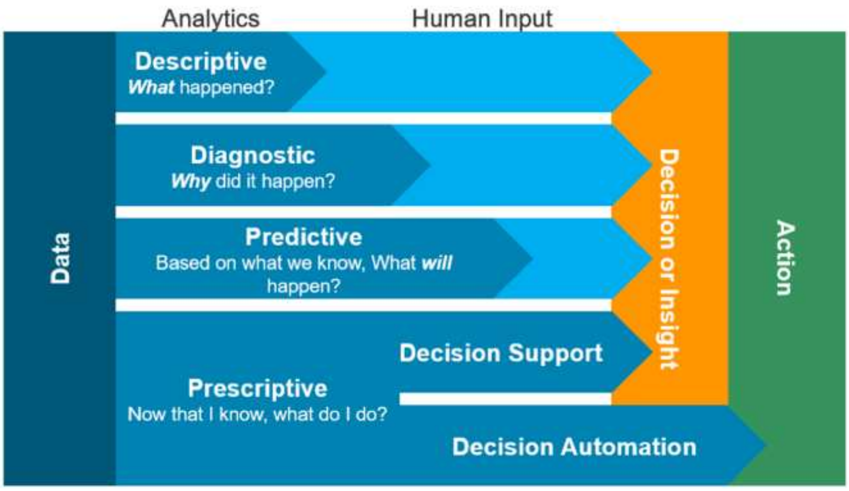


Figure 1.3-4 Types of Data Analytics. Image rights: Eurecat, adapted from [12].

The first type, descriptive analytics, requires more human intervention, given that the person in question must understand the problem, take decisions and finally act. At the other extreme, prescriptive analytics, human intervention is minimal, because the system is capable of suggesting a course of action or even taking a decision by itself.

The term *Big Data* refers to an enormous amount of diverse information which is beyond the capacity of conventional data base systems to manage and analyse in a specific time period.

Big Data is a set of technologies, algorithms and systems designed to collate large quantities of distinct data from which valuable information is extracted by employing advanced high speed analytical systems in real time. In an industrial scenario, the sources of information are multiple and diverse: sensors, equipment and installations, HMI interfaces, applications and information systems, operators, web, social networks, emails, cameras, etc.

Big Data is generally defined by the “5 V’s” which refer to:

1. *Volume:* The size of the data generated.
2. *Velocity:* The speed at which the data is generated, collected and analysed.
3. *Variety:* The different types of data collected.
4. *Veracity:* Trustworthiness of the data in terms of accuracy.
5. *Value:* Just having big data is no use unless we have value.

Within this context of Big Data, the application of Artificial Intelligence and particularly machine learning techniques, using mathematical algorithms, allows the development of automatic learning systems with the aim of creating expert recommendation systems.

Machine learning is the scientific study of algorithms and statistical models that computer systems use in order to perform a specific task effectively without using explicit instructions, relying on patterns and inference instead.

Both Big Data and machine learning are key tools and of great importance with diverse uses in industry. They offer many advantages including the support and automation of decision making, the intelligent planning of work within a factory, the auto-configuration of machines according to work orders, the optimization of quality control, and predictive and prescriptive maintenance.

1.3.3 Cloud Computing

Cloud computing implies a paradigm shift in relation to the traditional model which has always been based on progressively acquiring and installing new hardware and which is, as a result, extremely limited due to its cost (e.g. purchase of equipment, software licenses, maintenance, etc.).

Cloud computing allows the use of computing services over a network, usually the Internet, in such a way that the company only pays for the resources it uses, making it technically and economically viable to obtain access to large computing resources.

The model offers important advantages for a company, because it permits it to access only those resources which are required in an agile and cost-effective manner but with the additional advantage of having the capacity of adjusting the scale and/or increasing the resources as is necessary at any given moment.

The cloud-based model is a key element so as to be able to obtain other technologies such as Big Data, machine learning techniques or simulation, and is therefore indispensable for any industry that wishes to adopt the I4.0.

The cloud is also extremely useful to facilitate data sharing across sites and company boundaries. The performance of cloud technologies will improve, achieving reaction times of milliseconds. As a result, machine data and functionality will increasingly be deployed on the cloud, enabling more data-driven services for production systems.

1.3.4 Cybersecurity

Cybersecurity is an indispensable element without which the adoption of I4.0 cannot be successfully addressed. With the increased connectivity that results from Industry 4.0 it is necessary to protect critical industrial systems and production lines from cyber threats.

The three main pillars of information security are Confidentiality, Integrity and Availability, also known as the CIA triad:

1. *Confidentiality:* Only individuals with the legitimate authorisation to access the required information should be permitted to do so. The goal of confidentiality is to prevent sensitive data from being accessed by the wrong people.
2. *Integrity:* This principle seeks to ensure the accuracy, trustworthiness and validity of information throughout its life-cycle.
3. *Availability:* Availability refers to information being accessible to authorised personnel as and when it is needed.

In this sense, it is necessary for companies to adopt security models which respect the sets of relevant existing standards, in particular IEC 62443, a series of standards including technical reports to secure Industrial Automation and Control Systems [13]. On the other hand, it is also important that companies adopt the concept of *security by design* whereby security is taken into account from the early stage of design and conceptualization of new products, processes, systems and services. Therefore, by incorporating security measures and criteria from the beginning, it is possible to minimize and, to a great extent, avoid risks and impact against possible future attacks or accidents as the whole system grows and evolves.

1.3.5 Horizontal and Vertical System Integration

Vertical integration consists of integrating production systems to other areas and departments of a company (e.g. management, sales, finance, human resources, production, etc.). Meanwhile, *horizontal integration* consists of integrating the entire value chain of the product life cycle, thus including an interaction between suppliers, partners and customers.

1.3.6 Augmented Reality

The manner in which one interacts with computers and machines is poised to change in the coming decades. The term *Human-Machine Interface* (HMI) describes the methods in which people interact with computers.

According to Gartner, *Augmented Reality* is the real-time use of information in the form of text, graphics, audio and other virtual enhancements integrated with real-world objects. Creating real-time mixed realities that combine the real world with virtual elements offers extraordinary applications in the industrial environment and facilitates the work and productivity of workers by providing them with the ability to interact and access information of interest, *in situ* and in a systematic way, pertaining to any real element.

Augmented Reality applications are numerous. They include step-by-step instructions on how to assemble a product, field personnel drawing help from experts at remote locations, training, quality control, performance and productivity control, inventory, etc. Augmented Reality technology even permits novice employees to identify problems and perform repairs by following step-by-step instructions.

Another type of HMI is *Virtual Reality*. Augmented Reality alters one's ongoing perception of a real-world environment, whereas Virtual Reality completely replaces the user's real-world environment with a simulated one. Virtual Reality is principally employed for training (e.g. in the management of risk situations) and for industrial prototyping.

1.3.7 Simulation

A *simulation* is a model or representative example of the operation of a process, system or object over time. The simulation of products with virtual prototypes permits the optimization of the design phase of new products with a minimization of development costs and a reduction in the length of the marketing period. 3D product modelling techniques also allow the implementation of high-precision quality controls (e.g. metrology) of manufactured products.

The virtual reproduction of a factory (which can include machines, products and humans), whereby the performance of the plant in question is modelled, permits the evaluation in rationale cost and time spans of the suitability of different configuration alternatives in the plant and an analysis of its current response capacity when faced with different predicted demand scenarios.

1.3.8 Additive Manufacturing

Additive manufacturing is based on the creation layer by layer of an object through the use of different materials (e.g. plastic, resin, metal, etc.) with which it is possible to reproduce any 3D model as a real object.

This implies a paradigm shift because it allows:

- The redefinition of manufacturing processes since it allows production without moulds or tools.
- The minimizing or elimination of assembly pieces, thus reducing the amount of material employed to obtain much lighter objects and components.
- Flexibility and rapid adaptation to continuous changes in demand.
- Hyper-personalization of products and the viability of the production of reduced lots.
- The decentralisation capacity which permits production close to targeted customers thus reducing the cost of the associated logistics.

1.3.9 Autonomous Robots

Autonomous *robots* are the application of robotic systems capable of performing tasks with self-sufficiency, without explicit human control. According to the International Federation of Robotics (IFR) the demand for more productivity, the need to work with stricter standards in both industrial processes and the resulting products, the tendency towards mass customization, miniaturization requirements, and the evolution towards shorter product life-cycles has fuelled the use of robotic applications in recent years.

One of the most promising segments in autonomous robots are *collaborative robots* (often referred to as *cobots*). These are industrial robots designed specifically to work alongside humans in a shared workspace and to perform tasks in collaboration with them. These robots are designed with a variety of technical features that ensure they do not cause harm when a worker comes into direct contact. These features include lightweight materials, rounded contours and sensors at the robot base and joints that measure and control force and speed, and ensure these do not exceed defined thresholds if contact occurs.

Collaborative robotics enable manufacturers to improve productivity by using robots to complement human skills, relieving the workers of many non-ergonomic and tedious tasks and can be used to automate parts of a production line with very few changes in the rest of the process.

According to IFR, the market for collaborative robots is still in its infancy. Preliminary results show that, despite the claims of the mass media, less than 4% of the 381,000 industrial robots globally installed in 2017 were cobots. But this percentage is expected to grow in the near future when industry discovers its potential benefits.

1.4 Embracing Technology: Practical Case Studies

As previously highlighted, a company does not need to be a large enterprise or a multinational to benefit from the advantages of technology adoption as long as the technology adopted provide useful solutions to meet needs and challenges. In fact, technology allows the creation and delivery of value not only to customers but also to employees or even to stakeholders in general.

In the previous section Key Technology Enablers were identified. This section considers these in practise using some basic examples which illustrate the impact that technology adoption can have on enterprises' business models and value creation. These examples derive from diverse sectors and applications. On purpose, mostly outside the manufacturing field which could serve as inspiration if the principals are transferred to manufacturing.

1.4.1 Making Use of Open Data – Zaragoza Taxi Initiative

The Zaragoza Taxi Initiative is an example on how to deliver value to both customers and employees.

By using the opportunities opened up by data, some taxi companies are able to improve the performance of their company operations. For instance, in the city of Zaragoza, Spain, a taxi company is using the open data available on the municipality web site to improve its services. Every day the company downloads all public data related to scheduled activities (such as conferences, concerts, sport games, etc.) in the city. The taxi company gathers and processes that information in order to summarize activities, venues, timetable, etc.

The company has also developed an app for its taxi drivers so that they can be sent to the venues at the right time (right after the end of the activity) to pick up potential customers. By doing so the company is able to build value for their drivers to improve their chances of attracting customers and avoid unproductive driving around the city. In parallel, it also brings value for the customers who neither have to call nor wait for the taxis to get to their venues.

1.4.2 Optimizing Inventory Control – Casa Viva

This example shows how technology is able to build and bring value for employees whilst improving company performance is by simplifying internal operations.

Casa Viva is a Spanish home decor retailer with 36 stores in Spain and Andorra. Previously, employees managed inventory control with a purpose specific PDA.

Recently the company has introduced smartphones as the device to check and control all its products. Employees use a regular smartphone (such as the one they have in their personal life) with a built-in app to keep track of the availability and location of different products by merely taking pictures of their labels with embedded codes.

Product data introduction, updating and searching has become much simpler. Besides, in this particular case digital transformation has been characterised as a smooth implementation process since employees do not need to get used to a new technology. Instead, they are using smartphones the same way they use them in their normal lives.

1.4.3 Ground Robot Vehicles for Agriculture – Mas Llunes/ Grape Project

An important issue in business is environmental impact and ways to reduce the amount of chemical emissions.

In addition to robots being used in the manufacturing industry, they are also being introduced into other sectors such as agriculture. Apart from their utility to perform heavy tasks in the field, precision agriculture practices can reduce significantly the environmental impact of farming due to the over application of chemicals.

Mas Llunes, a vineyard company, has introduced unmanned Ground Vehicles (in the context of the EU H2020 GRAPE project [14]) to apply pesticides and fungicides with high precision saving huge quantities of those chemicals and preventing the plantation to be exposed unnecessarily to them. The robot is able to distribute up to 500 pheromone dispensers and allocate them on the vine branches by using an articulated arm. The purpose of dispensers is to spray the pheromones with precision to control plagues.

Advanced sensing capabilities also allow for monitoring at the plant level: the robot can monitor the health status of the vineyard, tracking the colours of the leaves, the dryness and helping the owners take decisions about the plantation and treatments.

1.4.4 Risk Management – IDP / BIM4Safety project

An important issue for companies is risk management. Risk in terms of people safety and in terms of assets security.

For industrial activities in general, and in the construction industry in particular, there is a high risk of injuries and fatalities due to the hazardous placement of people and machinery. There is also concern about people location in case of a major problem when the construction takes up large amounts of space (such as in civil infrastructure construction).

Companies such as IDP, an engineering company, have introduced a combination of BIM (Building Information Modelling) with Internet of Things to improve safety and risk management. BIM are software programmes able to model and manage physical and functional characteristics of places (mainly buildings). They are designed to help architecture, engineering and construction professionals to efficiently plan, design and manage those buildings and their infrastructures.

By combining the functionalities of BIM with the benefits of the Internet of Things companies can integrate within the software, not only the static assets (walls, pipelines, etc.) but also movable assets such as machinery and also workers. Sensors and wearables are attached to people and physical assets to track their location. The real-time localization can be used when needed (e.g. for inventory, in case of danger, layout planning)

1.4.5 Drones for Sewer Inspection – FCC

Decreasing operations and maintenance costs are a key issue for companies navigating today’s markets.

Fomento de Construcciones y Contratas (FCC) is a Spanish company offering urban sanitation services to municipalities and civil infrastructure operators. A drone, which in fact is an unmanned micro aerial vehicle (MAV) equipped with navigation sensors, aims to reduce labour risks associated with the operations and maintenanceactivities as well as to cut maintenance costs, thanks to quicker and more precise inspections.

The drones are also able to reach areas which cannot be reached by terrestrial vehicles, mainly due to dust and rainwater, narrow tunnels, physical obstacles, gases, etc. Drones are ideal for areas difficult to access or dangerous for human beings to work in; these vehicles can work in tunnels up to 80 cm wide and high.

Such drones can inspect 300 metres in 10 minutes. Therefore, it enables teams to inspect almost 2.5 km per day with a drastic reduction on costs and inconvenience, and increasing productivity.

Productivity is also fostered since the amount of data that the drone is able to gather by recording the full inspection on video for later processing is much larger than if acquired by humans.

The drone operates in an autonomous flight mode but it can be seen from the exterior of the tunnels in order to make decisions in case some specific measures and observations are necessary. Since no GPS or external positioning signals are available, the drone has to calculate its position and speed on its own.

1.4.6 Improving Patient Monitoring – Skintemp

Improving the usability and comfort to users is a key issue when it comes to delivering value or utility.

SkinTemp is producing and selling a strip with a sensor, similar to a plaster. That sensor is able to measure the body temperature, glucose level, oxygen saturation in blood, heartbeat rate and even blood pressure. Besides, the user can read (and record) the status and evolution of all those indicators in an app available for smartphones.

Since the strip is a really tiny device it can be worn discretely and comfortably for a long time, it's very convenient for the population in general, but especially for children or elderly people who have more trouble to use a normal thermometer on repeated samples.

1.4.7 Finding Novel Distribution Channels – Homeplus by Tesco

This example considers on value and utility to consumers.

Homeplus is a South Korean supermarket chain owned by Tesco until 2015 when Tesco eventually sold the company to investment fund MBK Partners.

When Tesco and Homeplus landed in South Korea, their main concern was to get a position in the mind of the South Korean consumer and, therefore, how to differentiate themselves from their competitors. Tesco realized that Korea is a hard working society: Koreans stay in the workplace until late. They also realized that Koreans spend a lot of time commuting using public transport, mainly on trains or the underground. As a result, they also spend a lot of time waiting on platforms and don't have much time left for shopping in the supermarket. In fact, when they go shopping, they also experience overcrowded supermarkets and long queues at the counters.

Tesco and Homeplus had the brilliant idea to replicate the supermarkets on the underground platforms so that Koreans could maximize their waiting time in the public transportation system.

Tesco replicated its supermarket shelves by covering the walls with billboards which contained the same layout of products and same packaging designs. Pictures of products also included an embedded QR code so that people were able to buy by simply taking a picture of the code, adding it to their shopping cart and finishing the process with a simple click.

By the time the customers got home the product was delivered to their door. Koreans were able to make the most of their time by transforming their waiting time into buying time.

1.5 Conclusions

Industry 4.0 started in Germany in 2011 and progressively expanded to other European and international markets. The end goal behind any initiative fostering Industry 4.0 is to connect and transform physical systems with cyber technology in order to gain adaptability, flexibility and production efficiency. The current era of technological change differs from previous revolutions not only by the technology itself but also due to the degree of interconnectedness, speed of acceleration and uncertainty of that change. Industry 4.0 is the systematic digitalization of an organization's processes, combining digital, industrial technologies with business transformational processes. We finished the chapter by moving from theory to practice: covering some practical cases of enterprises digitalizing their processes under the I4.0 realm.

1.6 References

[1] Source: Citi Digital Strategy Team.
[2] Source: CB Insights
[3] Lasi, H., Fettke, P., Kemper, H. G., Feld, T. And Hoffmann, M. (2014) 'Industry 4.0.' Business And Information Systems Engineering, 6(4), Pp. 239–242
[4] VIDEO: WHAT IS THE FOURTH INDUSTRIAL REVOLUTION? – World Economic Forum. https://www.youtube.com/watch?time_continue=570&v=kpW9JcWxKq0
[5] Source: Roland Berger
[6] Based on: "Document Model maduresa per l'adopció de la Indústria 4.0 en l'empresa" (SmartCatalonia, created by Eurecat Technology Centre). http://smartcatalonia.gencat.cat/web/.content/02_Projectes/documents/SmartCAT_Model-maduresa-I4.0.pdf
[7] Source: "Industry 4.0: How to navigate digitization of the manufacturing sector." McKinsey Digital 2015. Mckinsey&Company
[8] Source: Own Creation Inspired By: Api Product Management.
[9] Source: What Industry 4.0 Means For Manufacturers. Aethon Graphic Inspired By Boston Consulting Group Discussion On Industry 4.0
[10] Source: What Industry 4.0 Means For Manufacturers. Aethon Graphic Inspired By Boston Consulting Group Discussion On Industry 4.0
[11] Source: Inspired By Http://Archive.Rtcmagazine.Com/Articles/View/106730
[12] Source: Adapted From Gartner Report July 2015
[13] IEC- International Electrotechnical Commission. Publicly available documentation at: https://webstore.iec.ch/searchform&q=IEC%2062443
[14] Grape project website: www.grape-project.eu

Chapter 2
Technology Strategy

Mireia Dilmé i Martínez de Huete

2.1 Introduction

Industry 4.0 is the systematic digitalization of an organization's processes (maintenance, production or management) in order to collect, store and analyze its data. By exploring new production approaches, Industry 4.0 provides companies with a different economic model driven by new technologies.

The fourth industrial revolution (Industry 4.0) and its derived digital transformation goes beyond intra-organization challenges: optimizing processes, departments and the business ecosystem of a hyper-connected era.

Technology doesn't provide value to a business on its own. Instead, as written in a recent MIT Sloan Management article "*technology's value comes from doing business differently because technology makes it possible*" [1].

The days are over in which enterprises implement digital elements and think of them as solely technologies. These days, enterprises should consider technology implementation as a strategy. Consider Customer Relationship Management (CRM) software as an example. It's easy to think of CRM as a software application or even as a database to merely keep track of a company's interaction with its customers, but in practice it's neither. CRM is actually a business methodology: it organizes, automates and synchronize business processes (sales, marketing, customer service and technical support). CRM brings a layer of intelligence that allows an enterprise to really understand its customers – their situation, their preferences – and then use this to plan and design business strategies. All in all, to leverage the potential of CRM and many other technologies a strategy, set of goals and roadmap of the combination and interaction of all technologies across an enterprise needs to be defined.

M. Dilmé i Martínez de Huete (✉)
Eurecat, Barcelona, Spain
e-mail: mireia.dilme@eurecat.org

J.C. Chaplin et al. (eds), *Digital Manufacturing for SMEs*
DOI: 10.17639/QVCX-9K17

Digital transformation requires innovative approaches regarding cultural, technological, operational and strategic change.

The present chapter focuses on the challenges and the key actions that a company must take into account in order to carry out a successful digital transformation journey towards an intelligent and connected industry.

This chapter clarifies that Industry 4.0 involves the digital transformation of the industry with the integration and digitalization of all the industrial processes that make up the value chain, characterized by its adaptability, flexibility and efficiency that allows to cover customer's needs in the current market.

We will discuss underlying reasons behind the creation of technology roadmaps, and the necessary steps to carry out a diagnosis followed by the generation, prioritization and planning of I4.0 and digital transformation opportunities. Finally, the key steps to go from definition to action once the roadmap is defined.

2.2 The Path towards Digital Transformation

The manufacturing industry is currently witnessing a transformation as it increasingly moves towards Digital Manufacturing – often known as Industry 4.0, smart manufacturing or factory of the future. Most enterprises are responding to this move to some degree, albeit often cautiously. Caught up in day-to-day business operations, many SMEs and large businesses are struggling to understand current and future technology needs, drive business priorities and define technological implementation plans. In most cases, they do not know how to visualize the Industry 4.0 paradigm (which encompasses the inherent digitalisation of production processes) to their specific situation.

Certainly, any digital transformation can be challenging: it needs to be resource efficient, leverage internal capabilities and have employee buy-in. Figure 2.2-1 shows the main restraints to digitalization, identified through a European SMEs survey run by Canon Research.

Figure 2.2-1 What is restraining business transformation [2]?

Nevertheless, SMEs should see I4.0 transformation as an opportunity, as they have a huge advantage in terms of agility – for them it is much easier and faster to change approaches and resources than it is for big businesses struggling with massive infrastructure, large teams and long service contracts with suppliers.

Even if an SME has already made the most important decision of all – to embrace digital transformation – the large amount of investment options available, technologies to rely on, and decisions to make can be overwhelming. How and where to start can be daunting. To guide corporate efforts, technology roadmaps can be a crucial tool to clarify the technological readiness of an organization and help prioritize the right areas to focus on when building a technology strategy. Furthermore, it can also be used to effectively identify, plan and define digital opportunities across areas of the business.

In fact, as the years go by, the business world is leaning more and more towards technology, making it almost impossible for a lot of firms to separate business strategy from technology strategy. Technology provides new ways to create and capture value, as well as revenue. The right combination of information, digital technology, business know-how and physical assets can constitute a competitive advantage in key business model pillars (improve customer relationships, improve products and services offered, create novel revenue streams, increase operations' efficiencies and so forth). Due to the drastic impact technology can have on a company's success it should be considered as **one of the transversal pillars that define** an organizations' unique Business Model as seen in Figure 2.2-2.

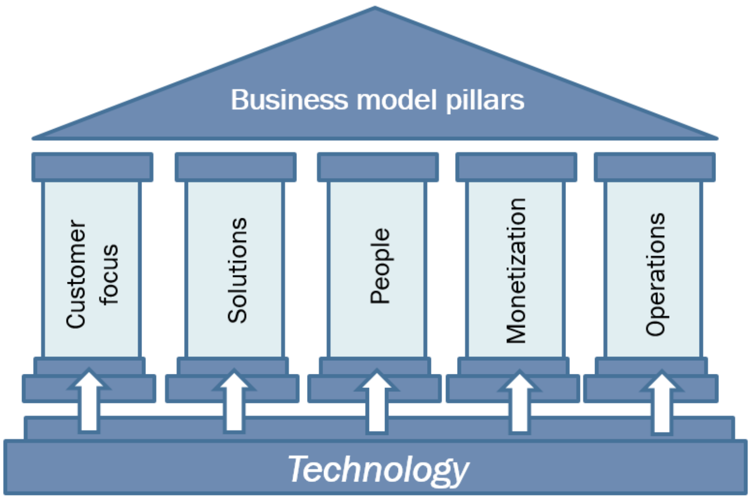


Figure 2.2-2 Technology as part of the business model pillars.

2.2.1 Strategic Roadmapping

Roadmapping is a strategic planning process that helps identify, align and communicate a business need *(Know Why)*, converted into a realistic action plan *(Know What)* and the required underpinning resources *(Know How)* during a specific time span *(Know When).* All of this is summarized in Figure 2.2-3.

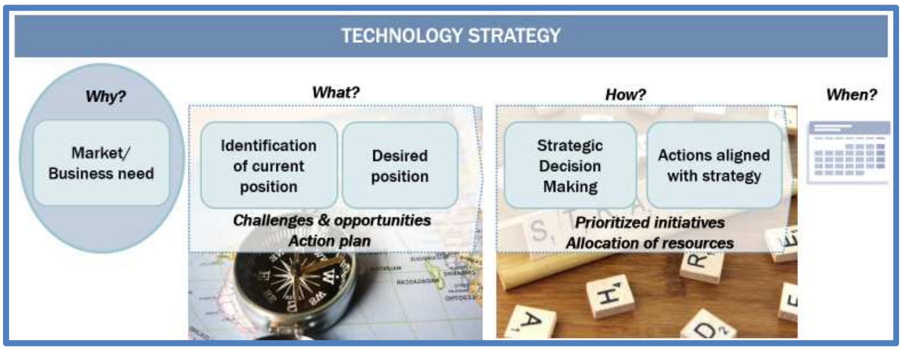


Figure 2.2-3 The four key questions of a technology strategy.

Strategic roadmapping and technology planning as a process was initially developed during the 1970s. Motorola became one of the first companies in the world to develop a formal technology mapping approach. They make use of the roadmap strategy in order to enhance and guide its product developments. On top of that, Motorola used the technology roadmap as a communication tool: to inform and align the needs of its target customers and its workforce with the disruptive novel products to be developed. Around the 1990s it expanded to the electronics and semiconductor industry and since then has exponentially expanded across many other industries.

There are many types of roadmaps that can be used in an organization and selected depending on the end-goal. Some examples include market and strategy roadmaps, knowledge asset planning, product roadmaps and ICT planning. Roadmaps can be presented in a broad range of formats, for example pictorial (flows), text, graphs, tables and bars.

The final roadmap(s) can take many forms, although generally the focus is a graphical representation that provides a top-level strategic view. The process of developing a roadmap is more important than the roadmap itself. The starting point is the future: where your company aims to be. From there you move backwards, defining the initiatives, investments, partnerships and resources that need to be established.

A simplified visual roadmap is depicted in Figure 2.2-4.

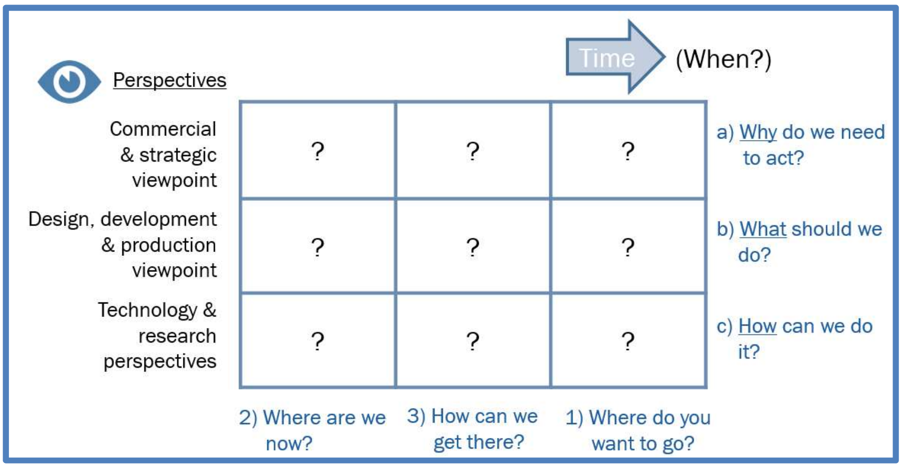


Figure 2.2-4 Simplistic roadmap framework [3].

From all the potential roadmaps, a *technological roadmap* (or digital transformation roadmap) allows you to represent the technology currently available to an organization over a period of time, as well as the best technological options that could be developed or acquired in a specific timespan. In many ways, a technology roadmap can be interpreted as a GPS system for a digital strategy road: it enables businesses to see where they are, where they want to go and how to get there. Therefore, it helps organizations plan which, when, and why certain technologies will be on-boarded while avoiding expensive mistakes and even planning for technologies becoming obsolete.

Some of the key benefits of technology roadmaps are:

- Facilitates the integration of technology into a business.
- Applies to processes, products, customer relationships and asset management.
- Facilitates consensus about needs and the technologies required to satisfy those needs.
- Provides a mechanism to forecast, plan and coordinate technology developments.
- Helps identify new business opportunities and exploit technology.

There are a lot of ways in which a company or industry can create, define and implement a technology roadmap. However, consolidated technology roadmaps share common features. Key considerations when building a technology roadmap are shown in Figure 2.2-5.

ROADMAP CREATION CHECKLIST

- **Identify and prioritize** the business and market needs.
- Harvest value as you go, **phase the process** and ensure **early delivery of benefits**. Therefore, include initiatives that generate short-term benefits already in the beginning of the plan (to encourage team motivation when pursuing the other initiatives included in the roadmap).
- Ensure **buy-in and support/endorsement from senior management**.
- **Allocate resources** (support, time and financial resources).
- Pay special focus on the **impact on people** (skills, process, culture and organisational change).
- Create the roadmap as a **team activity**, include diverse viewpoints and knowledge to enrich the plan. Process includes a lot of discussion and may need a facilitator.
- Keep it **simple and concise**.
- **Iterate and learn** from the experience. When developing the initiatives included in the roadmap use critical insights gained over time to modify and make the roadmap even more concise.
- Be ready to adjust the roadmap if necessary. The roadmap should be **flexible** and somehow adapted to the constantly changing environment
- Make the roadmap **accessible** to all parts of the organisation
- **Collaborate with relevant stakeholders**. Leveraging external resources such as consultants can be an effective way to access valuable knowledge and experience, helping you avoid pitfalls and learn from the success of others.

Figure 2.2-5 Roadmap creation checklist.

Moreover, suggested roadmaps should answer 4 key questions:

1. *WHY?*
   It is important to describe the current situation of the company focusing on the dimensions which the roadmap will impact.
   - Analyse company, define
   - Canvas, Pestle, SWOT, technology watch
2. *WHAT?*
   Define the vision and mission of the company and the strategic objectives in order to align the initiatives to implement on the roadmap.
   - Identification of current position and desired position
   - Technology maturity level

3. *HOW?*
   Describe the actions and initiatives to be carried out. These may be technology projects, new business models, organizational changes and modifications to the operational processes, changes to the customer and supplier relations or other activity. Also included the description of technical requirements, costs and execution time, expected results, KPIs, control measures, resourcing, possibility of obtaining financing... These activities will be steps on the critical path. It is also important to describe if these initiatives require a proof of concept or a prototype, and the implications of its scalability.
   - Define initiatives. Create sheets to describe each initiative, prioritization methodologies, define KPIs, create working groups, monitor target attainment

4. *WHEN?*
   Representation of the roadmap in the form of a diagram or schema with the projects or initiatives planned and sequenced over time. The roadmap should include intermediate target states or milestones on the process that the company will achieve ensuring the changing process is on course.
   - Calendarization

The roadmap creation process can help clients understand what threats and opportunities digital might pose to the business and to what degree. It can also help organizations get a better view of where their opportunities and threats are coming from, and what strategy and actions are needed within their organization to build competitive advantage in a quickly evolving digital business environment.

As in any strategy, we need to define the "*what?*". That is our starting point. What is our current situation and where do we want to go when it comes to technology? And then, the "*how?*". What are our strategic priorities and our processes (organizational structure, decision making, incentive systems, and norms and values)?

Although the technological roadmap is a governance document it should be regularly updated as the business evolves. It is a document that should be revised as the external and internal framework of the company changes with new opportunities and challenges that may change the strategy.

2.3 Technology Roadmap Approaches

As previously discussed, methodologies to create digital transformation/ technology roadmaps are very diverse.

Below are 4 summarized examples of different approaches which are taken to develop a digital transformation roadmap.

2.3.1 Partners in Performance

Partners in Performance, a global management consulting firm, proposes an approach that quickly develops a value-centric, prioritised and actionable digital roadmap enabling a company to:

- Identify what customers (and other users, like employees) really value.
- Understand the market and competition, including potential substitutes from outside the industry.
- Apply the right technologies, in the right context.
- Create organisational alignment and set up for execution success.

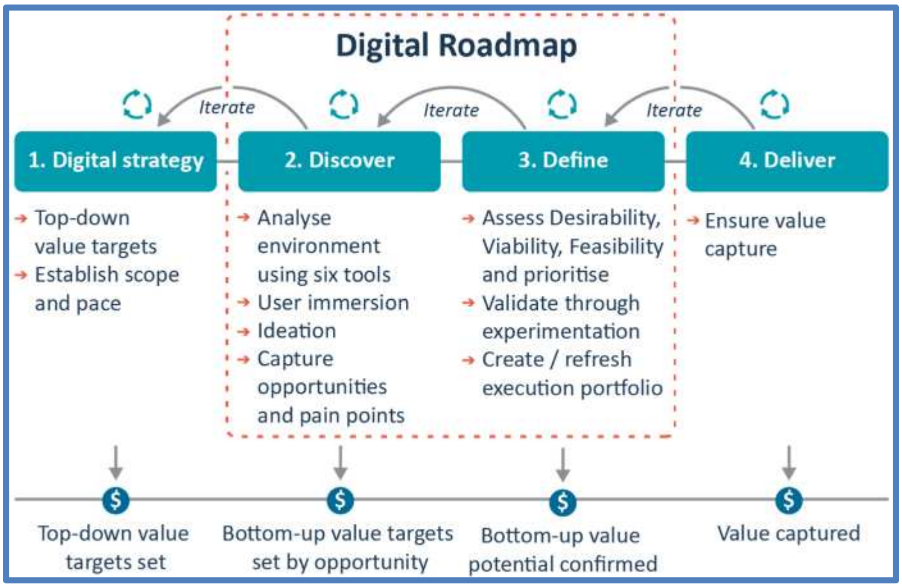


Figure 2.3-1 Partners in Performance proposed roadmap [4].

Further information regarding the Partners in Performance's roadmap methodology can be found at https://www.pip.global/es/services/digital-roadmap

2.3.2 Navvia

Navvia, a company specialised in business process modelling tools, proposes 5 best practices for establishing the digital transformation roadmap:

Figure 2.3-2 Navvia proposed roadmap [5].

1. Define what success looks like for your company.
2. Separate the change into manageable parts.
3. Harvest value as you go, instead of waiting until the end.
4. Focus on the impacts on people (culture, process, skills and organizational change).
5. Adjust course if necessary (the environment is continuously changing).

Further information regarding the Navvia's roadmap methodology can be found at https://navvia.com/digital-transformation-roadmap/

2.3.3 Mckinsey&Company

Mckinsey&Company, a global management consulting firm, proposes 10 guiding principles of a digital transformation divided into 3 principle stages:

1. *Defining value:* Place digital transformation at the core of your agenda. Make significant investments, and set clear, ambitious targets.
2. *Launch and acceleration:* Consider carefully which projects to start with and support them with the necessary resources. Prerequisites include a high-calibre launch team often led by a chief digital officer (CDO), consideration of organizational structure, and the nurturing of a digital culture.
3. *Scaling up:* When the company has a handful of initiatives up and running and starts to capture value, this is also the time to supercharge the transformation and do everything on a bigger scale. The thoughtful sequencing of subsequent initiatives is key to this. In addition, close attention will need to be paid to building more capabilities and eventually an entirely new operating model will be required.

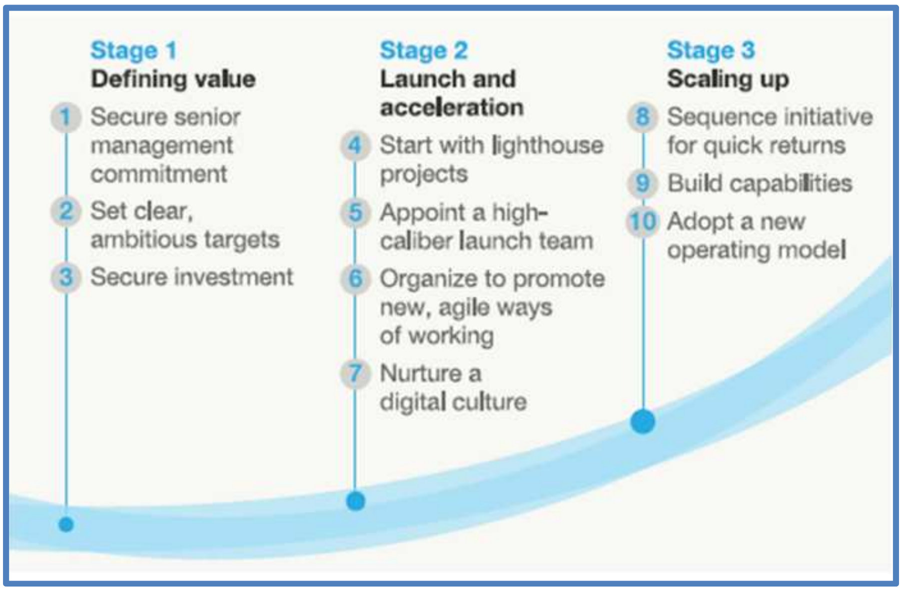


Figure 2.3-3 Ten guiding principles of a digital transformation [6].

Further information regarding the McKinsey's roadmap methodology can be found at https://www.mckinsey.com/industries/financial-services/our-insights/a-roadmap-for-a-digital-transformation

2.3.4 Eurecat

Eurecat, the Technology Centre of Catalonia, has an approach to improve the competitiveness of companies through the development of a transformation plan that allows them to incorporate digitization into their DNA. Defining a new strategy based on the incorporation of new digital technologies that allows the optimization of costs and generation of new business models. Figure 2.3-4 shows the 5 steps:

Figure 2.3-4 Key steps of a technology roadmap.

1. *Diagnosis:* External analysis of technologies and trends in the sector. Internal analysis through face-to-face interviews with the relevant team.
2. *Strategy:* Define the digital vision of the company. Set strategic objectives.

3. *Initiatives:* Generate opportunities and initiatives to boost digitalization for each of the dimensions: products, processes, technology and infrastructure, ecosystem, culture and organization, and information and data.
4. *Planning:* Establish a digital transformation plan minimizing risks and maximizing the impact for the company.
5. *Communication and implementation:* Communicate the company's digital transformation strategy. Sensitize and inspire about digital transformation. Implement the roadmap initiatives.

Further information regarding Eurecat's roadmap methodology can be found at http://smartcatalonia.gencat.cat/web/.content/02_Projectes/documents/SmartCAT_Model-maduresa-I4.0.pdf

2.4 Technology Roadmap Process

In this section further details are presented about the generation process of I4.0/digitalisation opportunities, their prioritization and planning to agree a roadmap that will define the company's strategy.

To illustrate the thinking and exemplify in further details the roadmap sequential methodology the aforementioned I4.0/digital transformation roadmap model developed by EURECAT is presented and described below.

2.4.1 Step 1: Diagnosis

A *diagnosis* is a key tool that allows a company to understand its current status as well as the areas where it should focus its efforts to undertake the digital transformation and progressive path towards I4.0 adoption. Diagnosis can take many forms and outcomes (reports, working groups, interviews, external audits, working sessions, etc). Amongst them, maturity models are a good example of a diagnostic.

Maturity models propose different levels of maturity for different dimensions within the operation of the company. Therefore, these models allow to establish the degree of implementation and development of Industry 4.0 solutions in different business areas, from human resources to IT architectures or production processes.

Several consulting firms and leading market companies have developed digital transformation maturity models. Examples include:

- *PWC Digital Transformation Maturity Model [7]:*
  - Industry 4.0 Capacity Maturity Model.
  - It is based on conclusions derived from a survey of 2,000 participants from leading companies, represented by nine industrial sectors and 26 countries.
  - Key areas that the model analyses include digital business models and customer access, product digitization, integration of value chains, data and

analytics, agile IT architecture, security, legal aspects and taxes, organization, workers and digital culture.

- *Bosch Maturity Model [8]:*
  - Maturity Model of the Internet of Things: How to Succeed in a Connected World.
  - It focuses on planning keys to embrace digital transformation.
  - Key areas that the model analyses include users, companies, things and partners.
- *Rockwell Automation: The Connected Enterprise Maturity Model [9]:*
  - Maturity model of the connected organisation.
  - It focuses on making IT networks more intelligent aiming to improve organisations' capacities and decrease costs. It also includes cultural change management perspectives.
- *PTC (Axeda): Connected Product Maturity Model [10]:*
  - Maturity model of a connected commercial product.
  - Represents the progression of IoT technologies and the progressive return that the company could expect through the expansion of its capabilities.
- *Switzerland Global Enterprise (S-GE): Industry 4.0 Maturity Model [11]:*
  - Industry 4.0 maturity model focused in key success areas.
  - Key areas that the model analyses include products and services, market and customer access, value chain, processes, IT architecture, legal aspects, security, risks and taxes, culture and organization.

Maturity models are represented in levels and areas of action, as shown in the example in Figure 2.4-1. A common feature of the models is that they range from an initial stage, in which a company has a traditional behaviour with very basic use of technologies, to the ideal situation of a fully digitised industrial company, which is able to make the most of I4.0 technologies both for the optimisation of its industrial processes and for the adoption and exploitation of new business models. In this way, a process of evolution is described in which each area must advance. The levels between describe a continuous progression where each subsequent level can be identified by specific criteria and characteristics. It is not necessary for a company to reach the highest level in all areas, each company should define the desired level of progress based on its business strategy.

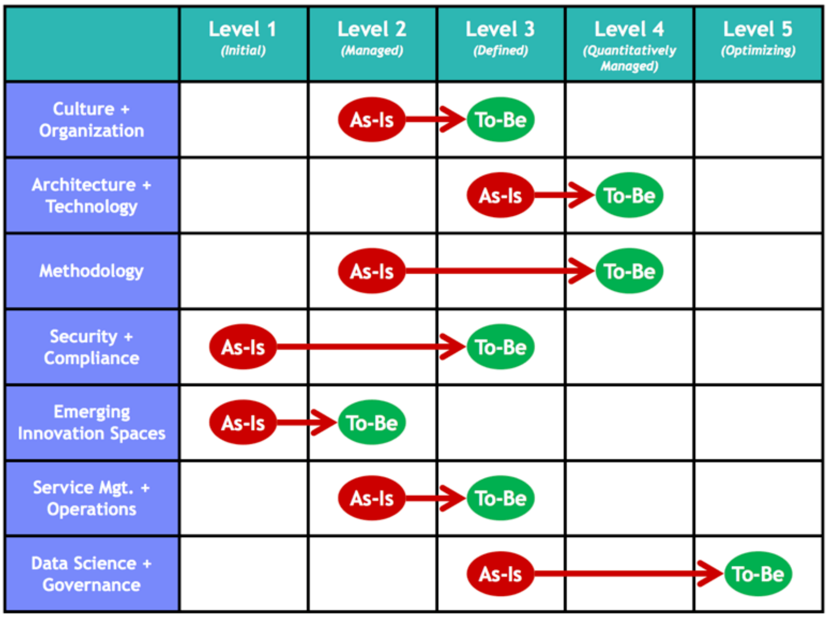


Figure 2.4-1 Simplified capability maturity roadmap (as-is, to-be) [12. Image Rights: IBM.

Following the maturity model created by Eurecat a suggested structure is illustrated in Figure 2.4-2

Maturity Level	Aware 1	Digital Novice 2	Competent 3	Expert 4	Digital leader 5
Products	No digital solutions	Roadmap for smart products	Products with digital solutions. New digital services	New business models	Full customization products Full Traceability
Productive Processes	No real time Reactive maintenance	Digitalisation roadmap	Pilots of integrated and digitalized production processes Control, planning and KPIs in real time. Traceability	Process optimization (artificial intelligence and cobots) Predictive Maintenance	Autonomous processes. Self-configurable Machines. Prescriptive Maintenance.
Technology and infrastructure	Non-integrated management systems	Production Management Systems (ERP, MES, PLM)	Integrated IoT platform with production and management systems	Recommenders and expert Systems	Simulation technologies and virtual models on plant.
Information And data	Disperse and uncomplete information	Data warehouse roadmap	Centralized knowledge database (Data Warehouse)	Machine learning analytics and techniques	Analysis of external company ecosystem data
Culture And organization	No digitalization strategy	Digital strategy and I4.0 Roadmap	Structure for innovation and digitalization management Coordination IT/OT	Training and more talent on data and I4.0 technologies	Continuous improvement plan
Ecosystem	Without use of digital channels Web 1.0	CRM. Web 2.0. Intranets and extranets.	Digital order management and quality control for providers.	System integration with third-parties.	Integral control and planning of the supply chain Autonomous and predictive management.

Figure 2.4-2 Maturity model on Industry 4.0 [13]. Image rights: Eurecat for Smart Catalonia

This maturity model identifies 5 levels or degrees of adoption of I4.0. These levels range from a minimum (Level 1 – Aware) with an absence of digitization in the organization to a complete integration (Level 5 – Digital Leader) and at different levels of digitization in all processes including decision making supported by data analysis.

The defined maturity levels are:

1. *Aware:* In general, the different functional departments of the company work in isolation and do not have access to integrated information of production and product. In spite of working with sensors, controllers, and monitoring and control systems that allow punctual automation the company does not generate knowledge based on empirical data on the company's productive processes, nor on the products and their associated services, such as maintenance, stock management or supplier management. This lack of knowledge limits the real capacity to take a leap forward and improve productivity, the quality offered, the response to demand, product innovation, or the establishment of new business models with new services, among others. Therefore, the company shows difficulties in satisfying the needs of the customer in the medium and long term with a competitive quality/price ratio with respect to the competition.

2. *Digital Novice:* The company is able to monitor and collect data associated with production processes and/or the performance of its products and/or services in real time. Having this information allows the company to obtain real knowledge (based on data) about its production processes, productivity, quality control and/or the performance of products and/or services. Thanks to this knowledge the company is able to define and implement a first set of indicators (KPIs) for better decision making.

3. *Competent:* The company is able to manage in an integrated way and with certain level of automation the different processes in the plant at the level of production, productivity, quality control and maintenance along with the rest of corporate management systems such as ERP, CRM or PLM. Information integration not only occurs at the factory level, but also at the product level. Thanks to this, the company enables and offers advanced solutions for automated management and control of the product for the customer.

4. *Expert:* The company implements improvement in automated processes. It is capable of optimizing processes, introducing certain intelligence through the generation of new knowledge thanks to the treatment of the information and data collected. This translates into greater efficiency and productivity, a higher quality product (e.g. a large reduction in the number of defective products per batch) and the satisfaction of the customer's needs. In short, it results in greater competitiveness.

5. *Digital Leader:* The company implements continuous improvement to continue advancing both the concept of the intelligent factory and intelligent product. The company is part of an ecosystem integrated with its partners and suppliers that gives it the option to compete at a much higher level, which otherwise would not be possible. The company is able to exploit the valuable information and knowledge gained as a result of its business through new business models.

This higher level (Digital Leader) results in what is known as the smart factory and represents the leap forward to a fully connected and flexible system which signifies the opportunity to drive greater value both within the four walls of the factory and across the supply network. This ideal level of digitization is described as a flexible system that can self-optimize performance across a broader network, self-adapt to and learn from new conditions in real or near-real time, and autonomously run entire production processes [14]. It is important to note that few companies are currently operating at this level.

At the same time, the I4.0 maturity model developed by Eurecat includes 6 areas of action or dimensions within the company in which it will be necessary to take measures to carry out the adoption of I4.0:

1. *Products, Services and New Business Models:* Creating products in a connected industry is different from traditional industry. Industry 4.0 provides the necessary tools to modify the product generated according to changes in demand, both in terms of volume and variability. There is a shift from product-centred production to customer-centred production. Added to this is the possibility of generating new services based on data and information and the use of technologies, which represent great added value for the customer.

2. *Productive Processes:* Activities performed on a recurring basis in business activities. In the case of industry, these include, among others, design, production, quality control, monitoring and stock control. Implementing the automation and digitalization of processes will allow a global vision of the entire value chain, maximizing efficiency and flexibility by producing more and better in less time.

3. *Technology and Infrastructure:* Transformation into Industry 4.0 requires a broad set of digital infrastructures. These infrastructures include:
   - *Software elements:* data processing systems; platforms for integrating the company's processes (customers, stocks, orders, energy expenditure, production time and efficiency), etc.
   - *Hardware elements:* sensors, PLCs, encoders, etc., which provide connectivity to the different machines and equipment in order to capture, store, monitor and analyse the information generated.

4. *Information and Data:* Data is the key to control, management and decision making in Industry 4.0. For this reason, companies must understand and manage data as one of their most important assets. Data must be actively and strategically managed throughout the value chain at all stages of its life cycle. Defining and implementing a plan for data collection, storage, analysis, valuation and sharing becomes basic and necessary to the success of the 4.0 industry implementation. At the same time, a balance must be found between the exchange and protection of data, guaranteeing their security at all times.

5. *Culture and Organisation:* Digital transformation requires a change of mentality throughout the company which in turn needs new organizational models in which the involvement of management is key. There is the need to establish new roles that allow the correct implementation of digitization initiatives fully aligned with strategic objectives. Moreover, the digitalisation demands specialization in IT at all levels of the organization. The recruitment of qualified personnel in this field, as well as the promotion of training programs qualification to existing personnel is key in a correct transformation in Industry 4.0.

6. *Ecosystem (Customers, Providers and Partners):* Companies are part of larger structures and value chains, which need the sharing of information and, therefore, the integration of data, processes and management systems with the aim of offering more value, more quality and more efficiency. This constant exchange of data between actors (suppliers, customers…) requires the correct protection. Therefore, it is necessary to establish alliances defining new cooperation frameworks in which all parties can benefit by overcoming possible initial barriers. In addition, it is necessary to take into account the central place that customers and their demands take in the ecosystem.

Thus, the maturity model provides a view across all capabilities, helping a company to prioritize focus areas and improve aligned with its desired outcome and digital ambitions.

To build the maturity model and identify in which level the enterprise is, both in-depth internal and external reflections are valuable.

In-depth internal analysis

An internal analysis is essential to understanding the company's current digital state. It is important to undertake the analysis with as many roles as possible to get the real situation of the company; interviewing only the management team could lead to a misguided perception of reality. Getting insights from the bottom up and the top down across all the dimensions of the company (the 6 areas in the maturity model previously described) will allow a company to identify where technology and organizational transformation is needed most and where investment is required.

A workshop can be a useful method to include representatives from different areas of the company and discuss together the situation of the company, even reaching an agreement. But the best way to identify the current situation, needs, deficiencies and areas for improvement is through in-depth interviews with those responsible for the different functional areas of the company – personal meetings with the different people involved in the day-to-day activities. Conducting one-on-one interviews will lead to knowing the limitations of each team, the team's information needs, the hours invested in spreadsheets and printed documents, the lack of communication with other units, the lack of IT software in order to be efficient and even solutions and projects in mind but never implemented.

External analysis: technology surveillance

Internal analysis must be complemented with knowledge of the reality of the environment, monitoring the innovations and technological developments taking place, focusing on their business area, and to be aware of the opportunities and threats presented to them.

Monitoring, environmental scanning and technology watch will make the company aware of the trends in the sector, the state of the art of the technology and the changes on the horizon that could have impact. This process favours the implementation of new projects, the success in the generation of new products and the success in the decision making on technologies of interest.

Figure 2.4-3 Benefits of technology surveillance.

A successful technology watch also charts out future strategies based on the insights gathered. To undertake the technology watch there are surveillance tools that help gather data from various sources and managing the large volume of information, categorizing the information into areas of interest to the company so that it can be analysed. Examples of simple surveillance mechanisms can be found in Figure 2.4-4

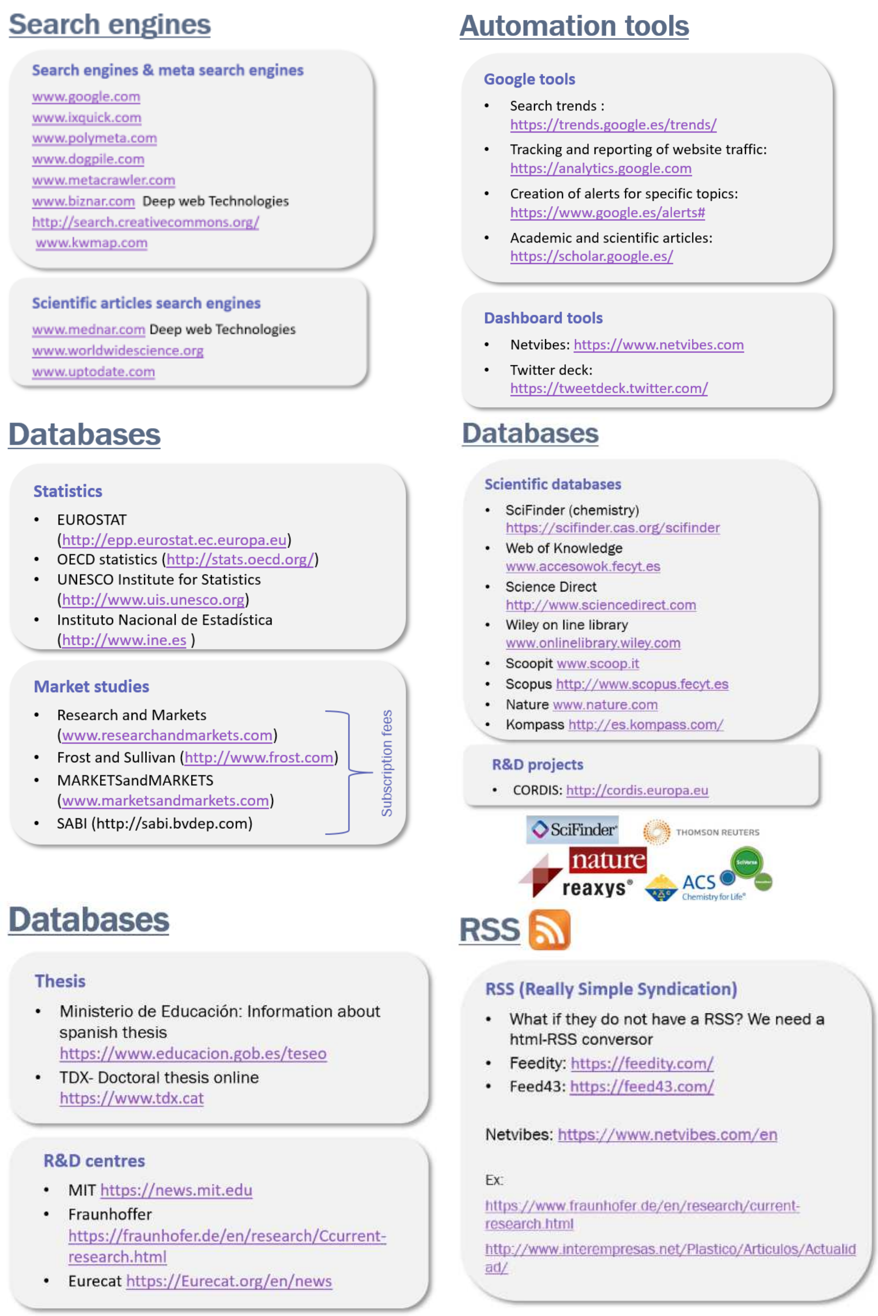


Figure 2.4-4 Simple surveillance mechanisms.

It is also important to highlight that patents can be a great source of valuable information. Around 95% of the patent applications that have been pursued since 1883 are available in the public domain†. To allocate specific resources in analysing patent databases can not only support enterprises analyse its competition and get information about state-of-the-art technologies but also ensure that no monetary resources are spent on already developed R&D. Some studies *(Enric Escorsa, 2015)* point out that each year 50bn€ are spent across the EU undertaking new R&D in developments that are already patented. In fact, 80% of technical information could be properly found in patent databases.

2.4.2 Step 2: Strategy

Complementary to knowing its current digital reality and, in turn, its current level for each dimension in the maturity model, the company must define its ambition and strategy for digital transformation: "*where do you want to go and what are you trying to do?*" When defining the future state to be reached, it is important to bear in mind that the objective is not to reach the maximum level of digitisation and that all levels must not necessarily progress to the same level. Each dimension must progress to the level that matches the digital ambitions and the strategy of the company itself. Tools and techniques including SWOT analysis [15], Porter's Five Forces [16] and Canvas [17], among others are useful to sort and visualize the results obtained in the diagnosis process.

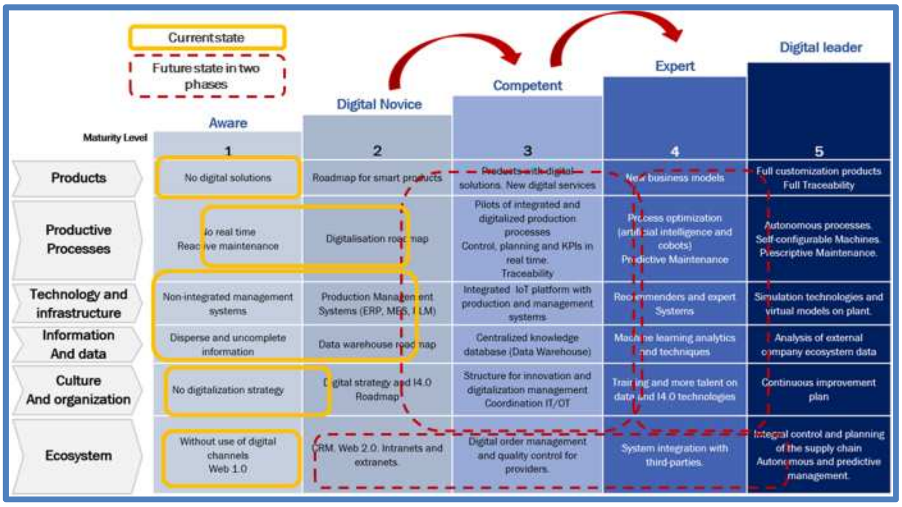


Figure 2.4-5 Maturity Model on Industry 4.0. Current state vs. envisioned state. Example of current and future state definition in the I4.0 adoption maturity model.

† James Conley et. al. "Study on Patents and the Public Domain (II)", 2013

Once the present state assessment and future vision definition are completed it becomes possible to identify the systemic gaps. These gaps represent both challenges that need to be addressed and opportunities for digitalisation and improvement.

The identification of opportunities to be developed and the challenges to be addressed will have an impact on different areas of the company in order to achieve the desired level of digitalization. In the Industry 4.0 maturity model presented in Figure 2.4-6 there are 6 dimensions defined when analysing the digitalisation level of an organisation:

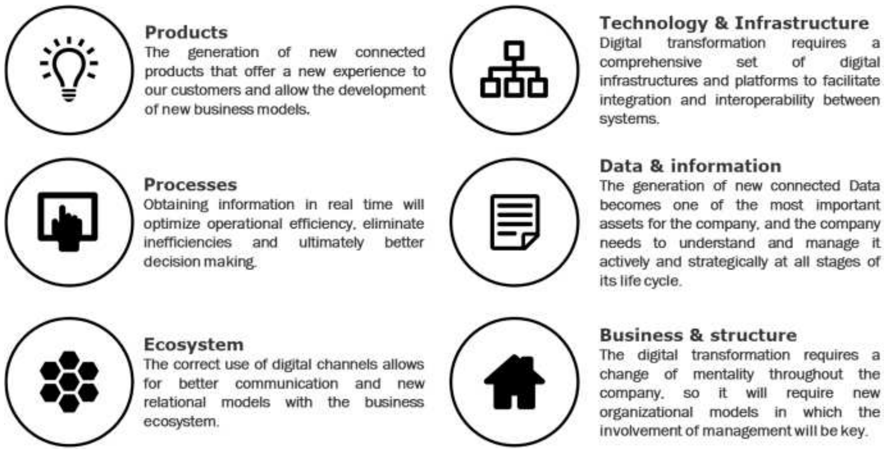


Figure 2.4-6 Six dimensions when analysing the digitalisation level of an organisation.

2.4.3 Step 3: Initiatives

After the challenges and digitalisation opportunities for improvement have been identified, a company can then create initiatives in order to reach the level of digitization desired. In generating and agreeing the initiatives (projects) to be implemented it is important to involve as many heads of departments and team leaders from within the company as possible, in the same way the diagnostics. Generally, this activity will involve several workshops where the result(s) of the diagnostic (maturity model assessment) is presented in a visual way with the current and future desired states the starting point for discussion. From this point the workshop will then work on the different opportunities identified, thus generating the initiatives (projects) the company will implement. Figure 2.4-7 presents some basic examples of digitalisation initiatives for each dimension of the Eurecat maturity model.

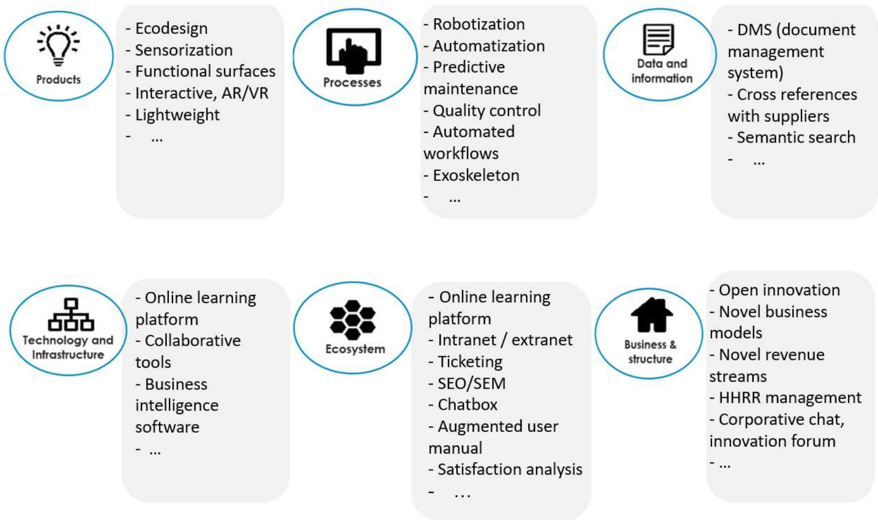


Figure 2.4-7 Examples of digital transformation/ I4.0 initiatives.

For each dimension, several lines of action can be defined. Sample initiatives for digitalisation of core smart factory production processes are illustrated in Figure 2.4-8.

Process	Sample digitalisation Lines of Action
Manufacturing operations	• **Additive manufacturing** to produce rapid prototypes or low-volume spare parts • **Advanced planning and scheduling** using real-time production and inventory data to minimize waste and cycle time • **Cognitive bots and autonomous robots** to effectively execute routine processes at minimal cost with high accuracy • **Digital twin** to digitize an operation and move beyond automation and integration to predictive analyses
Warehouse operations	• **Augmented reality** to assist personnel with pick-and-place tasks • **Autonomous robots** to execute warehouse operations
Inventory tracking	• **Sensors** to track real-time movements and locations of raw materials, work-in-progress and finished goods, and high-value tooling • **Analytics** to optimize inventory on hand and automatically signal for replenishment
Quality	• In-line quality testing using **optical-based analytics** • **Real-time equipment monitoring** to predict potential quality issues
Maintenance	• **Augmented reality** to assist maintenance personnel in maintaining and repairing equipment • **Sensors** on equipment to drive predictive and cognitive maintenance analytics
Environmental, health, and safety	• **Sensors** to geofence dangerous equipment from operating in close proximity to personnel • **Sensors** on personnel to monitor environmental conditions, lack of movement, or other potential threats

Figure 2.4-8 Digitalization initiatives for processes dimension [18].

POTENTIAL LINES OF ACTIONS
Ensure comprehensive and scalable data availability in a holistic view.
Improving information security.
Optimisation of production and quality control.
Automation of operations.
Optimisation of supply planning and internal logistics.
Improved product development management.
Improved maintenance management.
Marketing and sales empowerment.
Integration with the ecosystem.
New customer value services.
Support for HR management, training and change management.
Optimization efficiency financial administration.
Adaptation of the organization to the digital transformation.

Figure 2.4-9 Digitalisation opportunities.

Other possible opportunities are also proposed in Figure 2.4-9.

For this workshop it is important to involve experts on technologies in different fields mainly related to the nine technologies commonly defined as key drivers in Industry 4.0. In this way, the inclusion of external consultants, experts in digitalization or people who have participated in similar processes of digital transformation will facilitate the process of generating initiatives. In this way, with the combination of the company's internal knowledge and external expertise in technology and digital transformation, initiatives will be defined to solve the identified opportunities.

For each initiative it is important to define as much is possible its objectives, necessary steps for implementation, estimated costs as well as the expected benefits and risks. One possible way to define the initiatives is through a one-pager definition. This could be defined as a canvas template. Canvas is an easy approach for systematically understanding, designing and starting new projects. It is defined on one page and covers the necessary information for a project definition. It uses neutral language and it is easy to access and understand by all stakeholders of the project.

The most common information elements presented on a canvas for an initiative definition are:

- *Goal:* The main objective of the implementation of the initiative.
- *Description:* Explanation of the initiative. The challenge it intends to solve and how it will be carried out.
- *Benchmark:* Suppliers or solutions identified in order to implement the initiative
- *Planned Actions:* The main steps to carry out the implementation of the initiative.
- *Visual Representation:* Image, diagram or scheme to help understand the initiative.
- *Technology:* The technologies that are necessary to carry out the implementation of the initiative.
- *Estimated Cost:* An estimate of the cost of acquiring and/or implementing the initiative. As detailed as possible. It can also include information on the need for external funding and the time needed to implement it.
- *Benefits:* The main benefits to the company of implementing the initiative. It is important to be able to estimate a payback. This makes it possible to quantify the benefit and facilitates the subsequent prioritization of the initiatives.
- *Risk & Recommendations*: Impediments and problems that may occur during the implementation of the initiative. It also includes dependencies with other initiatives and recommendations for successful implementation.

As an example, Figure 2.4-10 presents the canvas for an initiative of the automatisation of the warehouse through automated guided vehicles.

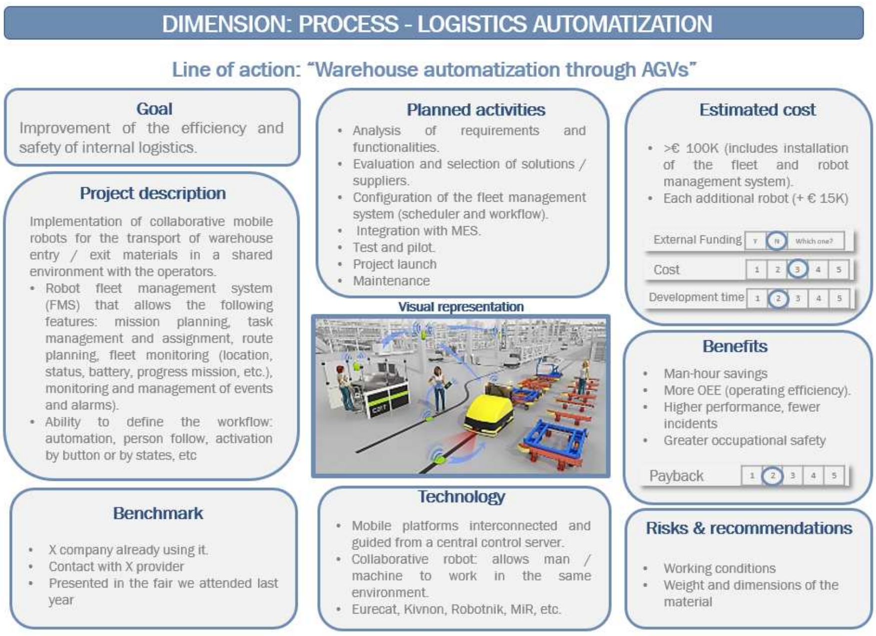


Figure 2.4-10 Canvas project example for Warehouse automatisation through AGVs. Image rights: Eurecat Technology Centre – Consultancy department

2.4.4 Step 4: Planning

Having the map of initiatives is not enough to move to action in terms of digital transformation. The set of initiatives generated must be prioritized, grouped if possible and ultimately sequenced in time to create the digitalisation roadmap for the company in the short, medium and long term. This can be a difficult process because each initiative often has its own business justification and funding. However, it is essential because the roadmap is the guideline and reference to keep the company moving towards its goal.

To prioritise correctly enterprises should select those projects that create maximum impact (e.g. ROI). Notwithstanding, they should also evaluate opportunity costs, dependencies among the initiatives and bundling activities that have synergistic opportunities.

When prioritizing initiatives some criteria that could be taken into account include:

- *Strategic Alignment:* How well the initiative aligns with the business strategy. As most projects will generally claim to align with the digitalisation strategy it is important to define the degree of alignment or even detect if this project would position the company among the best in the sector (top 1, top 3 or top 10) or, on the contrary, it is a commodity that everyone has.
- *Impact or Benefit:* Each initiative generated should have some impact on advancing the company in the process of digital transformation. What is important to evaluate is how far it will move the company forward. Estimating this impact can be related to increased profits, reduced costs or more qualitative elements such as loyalty or sales driver.
- *Cost Estimation:* Knowing the economic cost for the implementation of the initiative is essential to carrying out the prioritization. Economic resources are limited, so the more realistic the estimate can be the more refined can be the prioritization of projects. However, estimating the cost is not an easy task and a company will need to take into account the technologies involved, the workforce, etc.
- *Technology Maturity:* Often, the technology solutions that are required to implement an initiative have been available in the market for several years, with several providers able to supply the technology required. However, it is possible that the technology required is still a beta version in the process of being validated, or has not even been developed. Thus, it is important to identify the maturity of the technology required to implement an initiative – this translates into technological risk involved in carrying out that initiative.

An effective approach to carrying out the prioritisation of all initiatives is to attribute weight factors according to its importance and set levels for each of the evaluation criteria (strategic alignment, impact, cost estimation and technology maturity). And, through a workshop with the designated people, carry out a vote for each of the initiatives. A visual matrix can be established to facilitate prioritisation and group the initiatives into the following typologies, as shown in Figure 2.4-11.

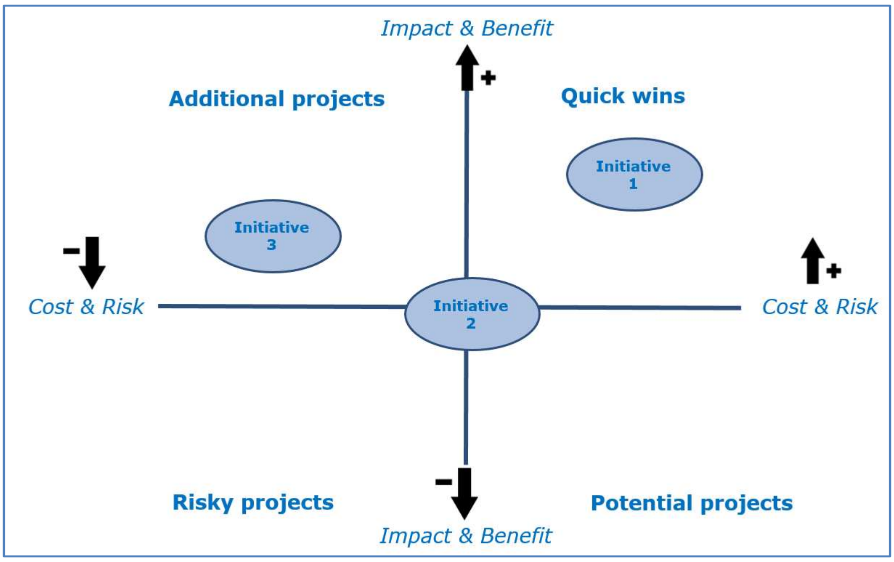


Figure 2.4-11 Example visual matrix to facilitate prioritisation and grouping of initiatives.

This matrix is divided into 4 quadrants:

- *Quick Wins:* Upper right quadrant. Those initiatives that have high impact with low cost and risk. Usually, these are the priority initiatives as they allow results to be achieved early on and gets staff/stakeholders motivated and establishes pace for the whole transformation process.
- *Potential Projects:* Lower right quadrant. Those projects that have a high value impact for the digitalisation of the company but have a higher cost and/or risk associated with them compared to the "quick wins". The degree of strategic alignment may determine an initiatives position in the implementation sequence.
- *Additional Projects:* Upper left quadrant. Those initiatives with a reduced impact in terms of the digitalisation process but have a low risk and cost which make them interesting for implementation in the medium/long term.
- *Risky Projects:* Lower right quadrant. Those projects that due to their low impact in digital transformation and their high cost and risk are discarded in the first instance so as not to waste resources in their implementation.

Ideally the initiatives with high impact and low costs and risks of implementation would be the ones with highest execution priority. However, sometimes is necessary to take a risk and position the company strategically by betting on a high-impact initiative, despite the high costs or risks involved.

The cost/benefit matrix will result in a first ranking of the initiatives having in the first instance the quick wins followed by the potential projects and ending by the additional projects. However, this arrangement requires an iteration taking into account dependencies between initiatives. That is to say, it is possible that some initiatives cannot be carried out if another initiative has not been implemented previously. For example, when creating an algorithm to predict demand, it is impossible to make the model work if shop data, market data, productivity information, sales values, etc. have not been previously collected. In the same way, if predictive maintenance is to be carried out, it will be impossible without having the machines first connected in a network.

Therefore, taking into account these dependencies, as well as other organizational and structural considerations or financing possibilities, the definitive prioritization of the initiatives to different time horizons will be carried out thus defining the strategy for the digital transformation.

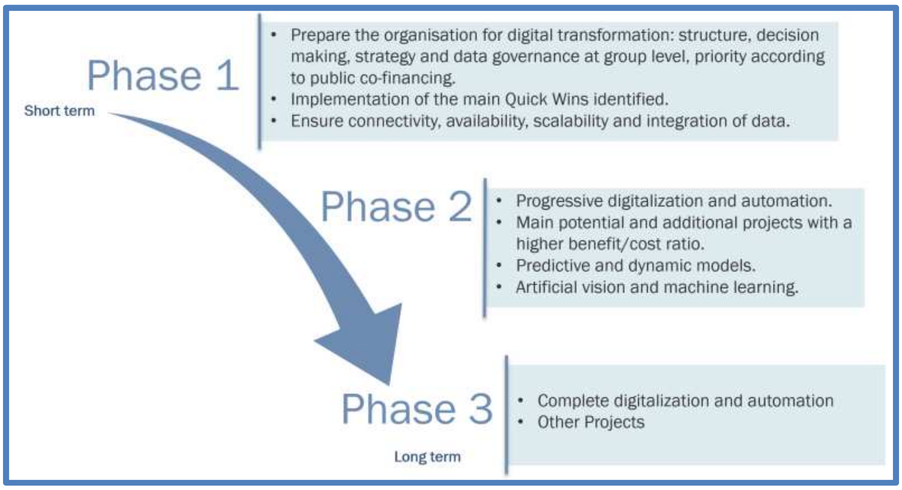


Figure 2.4-12 Roadmap execution time process. Image rights: Eurecat Technology Centre – Consultancy department

As introduced earlier in the chapter, a roadmap can be visually represented in many different ways. Initiatives can be grouped by functional area (production, sales, product development, quality assessment); by key challenges/goals (reducing costs, increasing sales, developing new products or services); or by key strategic themes (International presence, business expansion, higher production capacity, costs reduction). The columns then depict the stages of progression. They can be established by phase; by timing (months, years, quarters); by capability positioning (foundation, competitive, leading); by stage of the strategy (expand product range, market domination…). There is no right or wrong way to do this, the best way is the one that makes sense for the team. The options are endless, but key is to keep it clear, simple and focused so that all parts of the business can immediatly understand it. Each initiative will have a position in the sequence, being placed under a stage at which it will be implemented and grouped with other initiatives if this is the case. The initiatives may be linked sequentially or plotted in parallel with one or more initiatives. Simplified visual roadmap representations can be found in Figure 2.4-13.

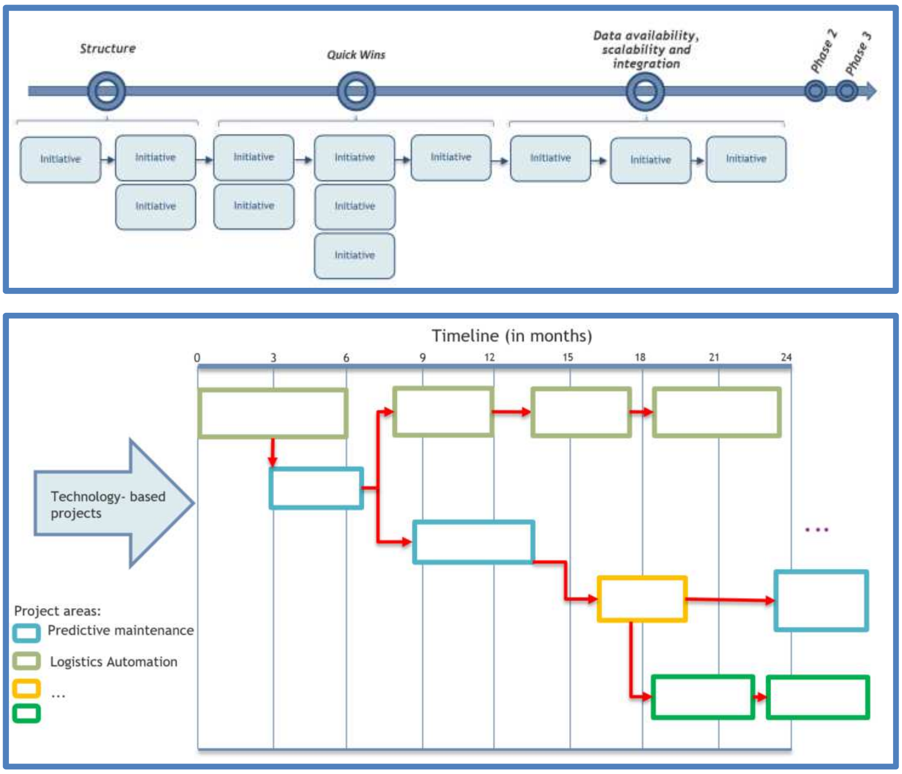


Figure 2.4-13 Roadmap representation examples. Image rights: Eurecat Technology Centre – Consultancy department

When plotting the initiatives it is important to take into account the resources available in terms of people, money and time and the pace that is being set for the company. The answers to some of the following questions will need to be taken into account:

- How many projects can be carried out at the same time?
- Even if I have the funding, do I have the staff to implement it?
- Is the company prepared for the change involved in certain initiatives?

Knowing the reality of the company, its willingness to change and the ability to implement projects will define the final roadmap.

Performing different iterations of the roadmap, with different scenarios to a greater or lesser time horizon can help to define the definitive roadmap.

2.4.5 Step 5: Communication and Implementation

Once the roadmap with the planned digitization strategy has been developed, the time comes for its communication and implementation.

For communication it is recommended that an organisation arrange events, workshops and collecting best practices to inspire your staff, share the common vision and train them on the new tools or technologies.

For the implementation, it is necessary to form a process management team, allocate financial resources (including the creation of a funding/financial plan) and identify appropriate people to implement projects associated with the strategy and make decisions on how to carry them out.

Once the necessary structure is in place to carry out the transformation process, the initiatives proposed in the roadmap must be implemented. Each initiative is a project in its own right and should be treated as such. Therefore, starting from the definition carried out in the roadmap development process, the proposed implementation steps should be followed.

In each case the company must decide how to carry out the implementation. You can acquire technology, develop your own, or collaborate with third parties for the development of it. The decision will have to be made whether to buy, make or collaborate. Although in many projects the decision may seem obvious, in most cases it is not. Each option has its pros and cons so the company must weigh in each case which is the strategy that can provide greater value.

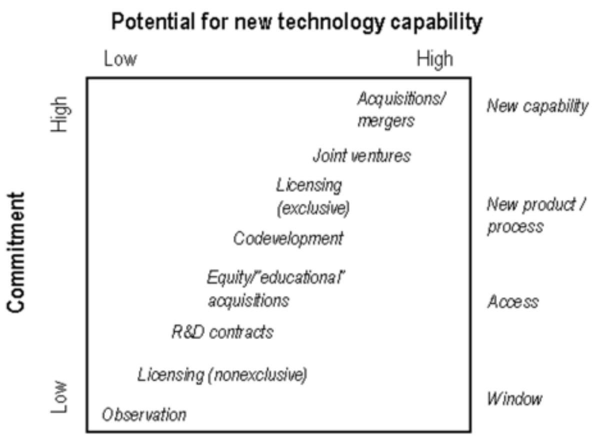


Figure 2.4-14 Matrix for acquisition of technology decision [19]. Source: Mechanism of Technology Sourcing, Leonard-Barton.

Below is a description of each potential strategy:

- *Observation:* Continuous monitoring of competitors' and market approaches on how processes, innovations, products and services are manufactured or offered.
- *Licensing (nonexclusive):* Grants to a third-party (licensee) the right to use the intellectual property, but means that the licensor and any number of other licensees to also exploit the same intellectual property.
- *R&D contracts*: Outsourcing R&D processes to reduce cost and enhance performance. Service agreements could fall within the definition of paid-for R&D.
- *Equity/Educational acquisitions:* Opportunity to investigate a technology in depth through investing, trusting and contributing to foster a venture.
- *Co-development:* To develop some projects, technologies or services jointly with a third-party enterprise. Knowledge "bleed through" from one company to another is inevitably envisaged.
- *Licensing (exclusive):* Gives the highest degree of exclusivity to the licensee. It indeed excludes everyone, even the licensor, from the use of the technology (clauses in the contract can be established to limit this). The licensee firm will enjoy a monopoly position over the use and commercialization of the invention.
- *Joint-Ventures:* Arrangement in which two or more businesses agree to pool their efforts and resources for the purpose of accomplishing a specific task, mainly to create a new business or project.
- *Acquisitions/mergers:* When one company takes over another and clearly establishes itself as the new owner, the purchase is called an acquisition. On the other hand, a merger is when two firms, generally about the same size, agree to go forward as a single new company.

Although each project is different, establishing alliances with technology providers or R&D entities such as technology centres and universities will surely enhance the process and provide support in the execution.

Additionally, when the roadmap is put into action and projects and technological or organisational solutions are implemented, it is essential to monitor progress and ensure that it remains in track. In order to monitor evolution, it is important to define metrics and indicators in order to detect whether the implementation process achieves the desired results and is aligned with the business objectives and strategy.

To create a transformation KPI, Gartner recommends asking these 5 key questions:

1. What is being measured? An example might be the percentage of customer interactions that are virtual/digital.
2. Where are we today?
3. What is our target goal?
4. What is our desired business outcome/benefit? For example, 50% better customer outcomes and 20% lower cost.

5. What is our balance point? (A “balance point” defines the reasons why a company shouldn’t over-digitalize. The law of diminishing returns applies also, and sometimes it makes no sense to have 100% as a goal. For example, a South American company might want to move all of its customers to mobile transactions, but in some countries 100% of consumers use a smartphone, while in other countries only 15% do).

Some examples of general milestones are proposed below:

- Digital proficiency
  - Reach of the organization in the market
  - I4.0 and digital maturity quotient of the employees including board and senior leaders
- Customer focus
  - Net promoter score
  - The rate of new customer acquisition
  - Number of customer touch points addressed to improve customer experience positively
  - Reduction in time to market new products to customers
  - Change in customer behaviour over time across channels
- Return on innovation
  - Percentage of revenue from new products/services introduced
  - Percentage of the profit from new ideas implemented
  - Number of innovative ideas that are effectively implemented
  - Number of new products or services launched in the market
  - Number of new business models adopted for different classes of customers

It should also be borne in mind that the roadmap must be a living document that has to adapt and grow with the company. It must be revised and analyzed frequently to update it in line with changes of the company’s own business requirements, external influence of competitors, customer demands or disruption of new technologies.

In the following chapter we will discuss that the integration of I4.0 and people’s perspectives. Discovering how to understand, reflect, lead, and apply the right tools to manage well people through the Industry 4.0 realm.

2.5 Building a Technology Roadmap: Practical Case Study

As already presented, technology roadmaps can be essential tools for assessing opportunities and defining how we accomplish our I4.0 objectives. It is a statement of intent and direction, coordinating the strategic options for a path to get there.

Technology roadmapping techniques can be used by organisations of all sizes, from SMEs to big corporations. The case study below outlines the range of challenges faced by a specific an SME (Casa Ametller) and how roadmapping has been used to visualise and develop strategies.

2.5.1 Interview with the Innovation Manager at Casa Ametller

How would you describe Casa Ametller?

Casa Ametller is a supermarket chain specialized in fresh products. The company, founded in 2001, opted for a vertical integration strategy: managing the entire value chain from the fields (currently owning more than 1,500 hectares), to the in-house production of food, vegetables and even prepared dishes such as tortillas, soups and creams amongst many others. Through the quality and proximity added value the segment of population that they have captivated the most are citizens interested in a healthy diet that are somehow aware of the ecological impact of the food industry.

The Casa Ametller group has more than 94 stores, a fixed workforce of 2,000 people and a turnover of more than 162 million euros (2017).

The company Ametller Origen Obradors (AOO) is integrated within the Prepared Food Division of the Ametller Group. It produces about 160 product references, highlighting the pre-cooked lines and the dairy line. AOO generated around 12 million euros in revenues, with a workforce of 90 workers.

What factors (external or internal) have been key to consider the need to create a technology roadmap?

The incorporation of technology into any productive processes must be guided, under a general plant strategy and following the objectives set by the company incorporating not only a short-sighted vision but keeping in mind:

1. The impact that a technology can have when creating new products and opening up to new markets,
2. The interaction with other machines, processes and know-how in the present and future.
3. Value the impact new technological projects have as an investment to improve quality, efficiency, productivity and serve more and better the final customer.

With an annual growth of more than 17%, the improvement in efficiency and control of processes is crucial for AOO to be able to respond to the increase in demand in the coming years. In this regards, technology plays a very important role.

Let's explain a specific case faced by Casa Ametller: tortilla flip and the subsequent packing used to be a laborious task that caused frequent wounds to the operators. The recent automation of this task, to be carried out by a machine with more than 30 pans, makes the process almost automatic. The staff have been relocated to process monitoring and controlling tasks – guaranteeing at all times the quality, agility and correct execution of the process. This example is transferable to the rest of the factory. Are there processes where technology can play a key role? Where could novel machinery improve the speed and control that will enable the proper continuity and growth of the company, the welfare of its workers and the quality improvement of the final products? Without a doubt, a technological roadmap is key to guide the identification, definition, prioritization and interrelation of these processes.

Did you find it useful to develop a technological roadmap?

Yes, because it has been possible to visualize the most urgent technological needs for our enterprise without becoming biased and focused only on extinguishing fires in the short term. Allocating time and resources to share points of view between the team and also with moderators and external technological experts has allowed us to have a more global and disruptive vision. Conceptualizing projects also in the medium and long term, prioritizing their place and visualizing the expected impacts and the way forward to achieve a correct implementation.

From the technological roadmap we have extracted the opportunities derived from the technological improvement proposed, identifying a clear and common way to make the current plant more competitive, to know what elements to modify to improve the quality and efficiency of the processes, to set the path towards to the greater standardization of the quality of the products and, in particular, to design small-scale pilots that serve us for the design of the new plant.

What advice would you give to a company that has not yet started the process?

In our case, we would highly recommend using external support that provides methodological expertise on how to create and implement the roadmap itself, as well as providing an overview of what is happening to the market and which technologies could be applicable and most appropriate to our processes.

What are some of the so-called Industry 4.0 technologies that have been included in your roadmap? Can you give an example of a project or specific technology application?

To be able to implement the roadmap several complementary technologies are being considered to be gradually incorporated in order to improve our production efficiency. Some of these technologies are technologies based on the cloud and big data. Another technology is artificial intelligence, specifically techniques based on automatic learning which will help us analyse the data we have in order to detect patterns and be able to design decision support systems that allow us to optimize certain activities of the production process, such as the daily planning of production orders in the plant. A multitude of variables will be taken into account, such as unscheduled stops due to incidents, or the needs of maintenance actions on the lines and equipment to minimize the number of these unplanned stops.

How do you expect to track the roadmap implementation success?

In order to ensure the correct implementation of the roadmap, technological change will be incorporated into the weekly meetings of the plant team. Resources will be allocated to establish a monitoring committee and execute the specific projects selected. In addition, external consultancy services will be contracted to develop and provide external support to certain projects. Methods to track and record the success of the roadmap specific KPIs are being established.

2.6 Conclusions

This chapter introduced the generation I4.0/digital transformation roadmap. Firstly, we discovered the importance of maturity models in orders to identify the gap between the actual digital state of the company and the desired future vision defined for each dimension of the company. Once the gap is identified, it is possible to generate I4.0 / digitalisation opportunities.

We learned how having an opportunities map it is not enough to implement the digital transformation to be successful. It requires a prioritization process where a cost/benefit matrix helps to get a first ranking dividing the initiatives into quick wins, potential projects, additional projects and risky projects. Setting priorities for different time horizons is essential to align the technology strategy with the business strategy and not overload the resources of the company. Finally, the way to represent the initiatives in the final roadmap was briefly discussed. In this case the representation can take the form that suits best the needs of each organization.

2.7 References

[1] Your Company doesn't need a digital strategy – George Westerman. MIT Sloan Management 2017.

[2] The Digital Transformation Journey: How to modernize your SME. Canon research 2016. Interviews to SME CEOs across the EU.

[3] Albright strategy (Roadmaps and Roadmapping Technology Futures Strategy).

[4] https://www.pip.global/es/services/digital-roadmap

[5] https://navvia.com/digital-transformation-roadmap/

[6] MacKinsey&Company

[7] R.Geissbauer, J.Vedso, I S.Schrauf, 2016, "2016 Global Industry 4.0 Survey", PricewaterhouseCoopers

[8] Febrer 2014, "Capitalizing on the Internet of Things", Bosch, White Paper Series, Part I

[9] Juliol 2014, "The Connected Enterprise Maturity Model", Rockwell Automation, Publication CIE-WP002-EN-P

[10] 2014, "Connected Product Maturity Model", Axeda, white paper

[11] Abril 2016, "INDUSTRY 4.0", Switzerland Global Enterprise, Whitepaper-Opportunities for the Swiss Export Industry

[12] IT maturity level example. IBM Cloud practices. https://www.ibm.com/garage/method/practices/think/it-maturity-model/

[13] Maturity model Industry 4.0. Smart Catalonia

[14] https://www2.deloitte.com/insights/us/en/focus/industry-4-0/smart-factory-connected-manufacturing.html

[15] "SWOT Analysis: Discover New Opportunities, Manage and Eliminate Threats". *www.mindtools.com*. 1006. Retrieved 24 February 2018.

[16] Michael E. Porter, "How Competitive Forces Shape Strategy," May 1979 (Vol. 57, No. 2), pp. 137-145.

[17] Business Model Generation, A. Osterwalder, Yves Pigneur, Alan Smith, and 470 practitioners from 45 countries, self published, 2010

[18] Deloitte University Press. Deloitte Analysis Industry 4.0. Smart Factory https://www2.deloitte.com/insights/us/en/focus/industry--0/smart-factory-connected-manufacturing.html

[19] Leonard-Barton D. 1995. Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation. Harvard Business School Press: Boston, Massachusetts.

Chapter 3
People at the Centre of I4.0

Mireia Dilmé i Martínez de Huete

3.1 Introduction

As presented in previous chapters, our world is rapidly changing economically and technologically. In order to stay relevant, every enterprise has to adapt to these changes. The Fourth Industrial Revolution represents one of the major drivers of change today. By investing in innovative technology that better connects people, machines and systems, manufacturers can become more efficient and agile than ever before.

These changes directly affect the ways of managing an industrial organization and the development of people working in this new digital and hyper-connected environment. The ability of an organization to adapt is a great competitive advantage.

I4.0 is about people as well as about machines and processes. Having the right workforce, with the necessary skills and training for implementing I4.0 is no less important than being financially ready and equipped with the required machines and upgraded systems.

Traditional manufacturing skills will need to evolve (and, even be replaced) with new capabilities such as programming, data and analytics, software development, etc. Moreover, whilst an organisation can buy technology. It can't buy employee engagement. Key success factors for organisational adaptation to Industry 4.0 such as teamwork management, cultural transformation and adaptability to change need to be progressively developed and properly managed.

M. Dilmé i Martínez de Huete (✉)
Eurecat, Barcelona, Spain
e-mail: mireia.dilme@eurecat.org

J.C. Chaplin et al. (eds), *Digital Manufacturing for SMEs*
DOI: doi.org/10.17639/Y73W-T460

The present chapter deals with the impact Industry 4.0 in the human resources management. Taking into account that people are central to the success of Industry 4.0 strategies.

3.2 Workforce Management and Teamwork Evolution in the Connected Industry

Technology is an agent for change. The advent of Industry 4.0 is giving manufacturers more flexible, faster and more efficient processes to produce higher-quality products at reduced costs. But one of the most critical disruptions in manufacturing is happening around the global workforce. I4.0 and digital transformation is less about technology and more about people. Companies must be aware that they need to take into account people ready to take on and develop the changes brought by the digital revolution. Progressively, the company will have to review all its operating and functioning values, its corporate policy, its organizational chart as well as the profile, competencies, responsibilities and functions of its employees.

The process of moving to a smart factory requires an organisation to adapt and change beyond the simple introduction of new technology. One of the key challenges that derives from the implementation of modern technology is the new skillsets that are needed from employees to code up new processes, run machines and fix new devices. Across the board, skills need to be developed. Thus, it is important to consider the people who are affected by the change at all stages in the transformation process. The evolution and buy-in of the workforce is integral to the interconnected and digital industry.

3.2.1 Focusing on People and Culture to Drive Transformation

The first phase of the transformation process concerns developing the culture and essential values for the organisation and workforce, as well as rethinking conventional operating models for the business. Aspects which should be considered include:

- The degree of commitment of what/whom?
- *Unlearning / Learning:* Having the capacity to unlearn knowledge, methodologies and old assumptions. Unlearning challenges assumptions in the conventional wisdom that may have become invalid and obsolete under the disruptive Industry 4.0 era.
- *Self-criticism capacity:* It is critical for leaders to recognize and accept fears, risks and errors, whilst learning from each experience and initiative undertaken.
- *Transparency of information:* Clear communication is critical during a digital transformation.

- *Ambition and new challenges:* Paving the way for new, often disruptive business models. Provide employees with personal growth opportunities, which drives engagement and retention. Therefore, be open to accepting those changes and have the ability to transform challenges into opportunities. In this transformation towards the digitalised, automated or robotised company, the role of people evolves towards an analytical and problems solving role, so that people are more autonomous and flexible in their jobs.

The management of an organisational change must be oriented to ensuring people enjoy their work, and are willing and able to learn new models and work systems.

To get the expected results, change must be occur at three levels, each of which are interconnected:

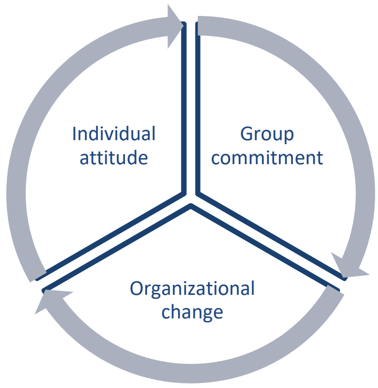


Figure 3.2-1 Keys to success and adapting change towards Industry 4.0.

- *At the individual level of attitudes and aptitudes:* This process is done through experimentation and learning. Adaptation is the result of a circular and constant process of exploration, understanding and modelling. Innovative initiatives are fundamental and mistake tolerance with failures with limited impact versus avoiding failure are indispensable.
- *At the group level with commitment and alignment:* When people interact in groups, it is essential to generate a long-term commitment to building a real team. Around a common project, members can share ideas and unify the analysis, thus spreading a true enthusiasm. Cohesion is the most critical factor in teamwork and it is a prerequisite to achieving total collaboration.
- *At the organizational level taking in account the cultural change:* Empowerment and delegation to collaborators to make decisions through established mechanisms is necessary to facilitate the emergence of self-leadership.

The necessary collaborative culture to drive an organization oriented to the connected industry affects the three levels and it implies a certain tolerance of error for people, aligning the objectives around a common project for the teams and a decentralized organizational structure.

One of the biggest mistakes that can be made in the digital transformation, and in any transformation, is to assume and expect that everyone's reaction will be positive and biased. If employees don't buy into the reason for change, don't trust leadership and don't share the organization's vision, there will be no successful change, regardless of how brilliant the strategy. Therefore, it is essential to apply change management strategies to prepare and support all the employees.

One of the most widely adopted change management frameworks is the Prosci ADKAR model which presents a step-by-step approach to help companies achieve success on a transformation process. The process, illustrated in Figure 3.2-2, focuses on the individual and the steps are as follows:

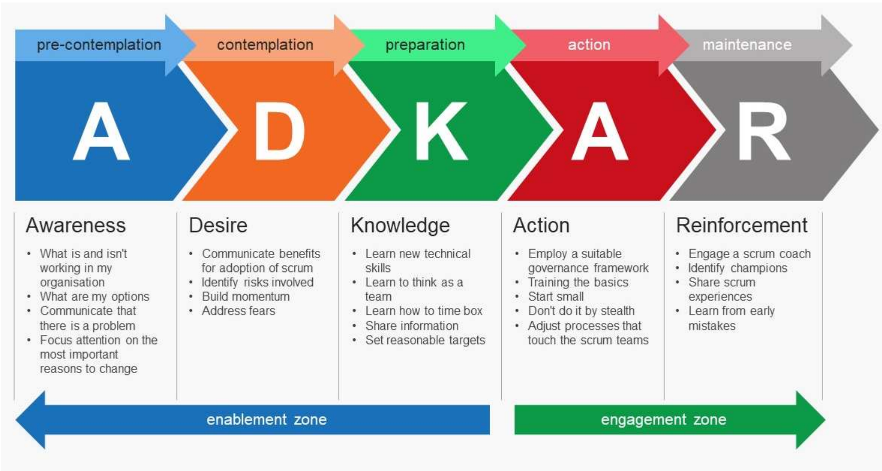


Figure 3.2-2 ADKAR methodology for change management [1].

- *Awareness of the need for change:* It represents workers' understanding of the rationale/drivers of that change, why it is happening and what are the effects of it. It helps answer "*what's in it for me*?"

  Actions that can be undertaken: Share vision and reflection of the future with all workforce. Demonstrate need for change, build trust and commitment.

- *Desire:* Relates to the personal choice and willingness to engage in a change and support it.

  Actions that can be undertaken: Clear communication of objectives and purposes. Resistance management.

- *Knowledge:* Represents the training and information (skills, tools, systems, responsibilities, etc) necessary to know how to be an active agent of that change.

  Actions that can be undertaken: workers' empowerment, developing talent. Streamline: flexibility, adaptation, evolution, speed.

- *Ability:* Is the capacity to turn knowledge into action. Having resources, time and support to develop those changes.

  Actions that can be undertaken: Creation of the work team (internal / external) as well as specialised change agents.

- *Reinforcement:* Is the critical and final milestone. While making a change is hard, sustaining it over time is even more difficult. It is a natural human tendency to revert back to what we know.

  Actions that can be undertaken: Constant monitoring and piloting. Feedback, dissemination and integration into culture.

3.2.2 The Development of Work Teams in Industry 4.0

Industry 4.0 seeks to move from a factory focused on its product offer to a service company focused on its customers, and for this a company has to be outward oriented. The traditional industry has to change its models, systems, ways to share information, data, make joint decisions, and adapt to the demand. Just as quickly as this, it also needs to accept that changes are continuous. Industry 4.0 has to innovate in new processes to increase operational efficiency and, at the same time, redesign the supply chain. Connectivity and collaboration with the client/customer allows the creation of new business models, such as buying for hours the use of a product, or renting systems management services.

The orientation that the organization should follow for adapting to the many changes of I.4.0 includes:

- *Ability to interact with the client (service):* Means managing the change of orientation of the traditional industry towards a collaborative industry. This implies an open attitude of the employees who are willing to accept [suggestions for] modifications and improvements to products and new product ideas coming from suppliers and/or clients. This vision implies a change of mentality, viewing the client as a partner in the definition of the proposal or the project. In this case, the product offer does not have to force the demand but it can adapt and meet the needs or requirements of the demand.
- *Adaptability and evolutionary flexibility:* Industry 4.0 has mechanisms and systems to have data available in real time (hyperconnected industry) which streamlines manufacturing processes, services and distribution points as well as the changes or modifications required by the client. Employees must adapt their mind-sets to these new paradigms, through the analysis of data, shared decision-

making, flexibility to move towards new processes or materials, and agile supply management systems.

- *Innovation in new processes:* The client orientation generates new solutions and new manufacturing processes. This implies that engineers have several opportunities to be able to give a more efficient service through economies of scale and, in addition, a much more personalised service using flexible manufacturing. On the manufacturing floor a few decades ago, fixed automation produced a limited assortment of part types manufactured in very large batches. At present, flexible automation technologies, robotics and 3D printing technologies enable manufacturers to quickly modify the layout the production plants orienting them to the needs of the client. Therefore, bringing the capability of making wider variety of part types/ products in shorter time frame and smaller batch sizes. These manufacturing premises encourage the creation of new processes closer to the customer, offering competitive advantages which were previously not possible.
- *Connection, collaboration:* Information technologies are used to generate data in real time. Individual work stations, computers, mobile devices and local area networks connect to the global internet, creating a collaborative environment that allows different partners to share the same information. This allows a continuous virtual connectivity between the team members, external partners and clients.

The evolution of the ways of working in the digital era are generating many changes in both the aptitudes and the attitudes of the manufacturing companies' workforce as well as team management practices, as illustrated in Table 3.2-1

Work evolution in the digital era	
PAST	**FUTURE**
Hierarchy	Flat structure
Fixed work schedule	Flexible schedule
Confidential information	Shared information
Direct the employees	Empower and inspire people to lead
Fixed technology	Cloud technology
E-mail as a method of communication	New digital communication methods
Corporate ladder	Opportunities based on competency
Fragmented company	Connected and interactive company
Office work	Work from anywhere

Table 3.2-1 Work evolution in the digital era.

One example of change in team management practices is the increasing emergence and support of high-performance teams. These teams refer to interdependent, stable, role-defined working groups with a strong structure, compelling direction, mutual trust, values and a supportive context that outperform all similar teams.

Principles of the actions required to create a shared vision of high-performance teams include:

- *Transparency:* Direct and clear communication. Encourage opinions with respect.
- *Participation:* The team’s vision is built between all team members. Promote that all members can have an important role, since teams have different workloads and relationships.
- *Pride of belonging:* Highlight the importance of the team's mission. Always celebrate successes.
- *Define the mission, the objectives and the strategy of the team:* To specify together where the team are going, how it will do this and what it wants to achieve. Set effective performance goals.
- *Maintain periodic communications:* Information exchange meetings. Facilitate positive feedback. Let everyone think and contribute with new ideas. Coach the team as a team, not as a group of individuals with individual skill sets.
- *Create a feeling of shared responsibility:* Socialize the destination, the successes, the efforts and the problems.

3.2.3 Adaptation of the Leadership for Industry 4.0

In the face of the changes provided by digitization, managers have to develop and adapt to the team’s needs.

With the analysis of the different theories of leadership, we can see that managers have to change beliefs of previous industrial eras. We understand that the difference between managers and leaders is that managers try, in a hierarchical way, to achieve their objectives through negative feelings (fear, intimidation, criticism, domination), while leaders try to motivate their collaborators through positive actions, such as inspiration, encouragement and training. A boss manages his employees, while a leader inspires them to innovate, to think creatively and to strive for perfection. Each team has a boss, but what the teams need in the connected industry are leaders that accompany them to reach excellence.

One of the keys is delegation. Delegation increases morale, productivity and trust. An excessive control of the tasks assigned to employees generates discomfort and dissatisfaction. Employees will constantly feel under scrutiny and will not feel confident towards the manager. To assign important responsibilities to teams and give them freedom to collaborate and complete their tasks, strengthens morale, innovation and teams’ satisfaction. Leaders must not only trust their employees, but also show them that trust.

Another key factor is to encourage collaboration among different knowledge units, The different disciplines and knowledge that we find in Industry 4.0 open up the possibility to combine technologies, (IOT, Artificial intelligence, predictive maintenance and machine learning etc) with which the specialists of each unit can collaborate to take advantage of technical and digital capabilities, how is the data processing and its subsequent applications to reach higher levels of quality and productivity, or generate disruptive solutions to its customers. Leaders should animate multidisciplinary teams to encourage cooperation, motivation and commitment, and guide to changes and innovation.

The team leader in the connected industry has a genuine interest for people in their technical and personal development, which allows a maximum performance of your team.

Evolution of competences from Boss to Leader for the team management in Industry 4.0		
Manager	**Leader**	**Leader in the digital era**
Knows everything	Open to new learning	Creativity and intralearning promoted (no barriers, no limits)
Speaks more than listens	Listens more than speaks	Constant digital communication
Gives solutions	Searches for shared solutions	Searches for global solutions including from the client
Critique	Gives encouragement	Gives encouragement and provide solutions
Highlights the weak points	Recognizes effort	Motivation based on particular needs
Directive	Coaches	Connected coach
Oriented to self	Oriented to the team	Part of the high performing team
Blames others	Assumes responsibility	Shared and assumed responsibility and commitment

Table 3.2-2 Evolution of competences of Boss to Leader for the direction of equipment in Industry 4.0.

3.3 New Professional Competences and Lifelong Learning Strategies, Paradigms for I4.0

The speed of change and the need to be permanently updated in today’s world requires a never-ending specialization and practical knowledge. As a result, there is a need for new forms of training and new models to access the training.

Lifelong learning is the search for continuous, voluntary and self-motivated knowledge, whether for personal or professional reasons. It is not only “training” and is not only “professional”, it is a predisposition to acquire continuously new lasting knowledge and not necessarily for only professional reasons but also includes personal development and cultural enrichment.

Rapid changes are taking place in access to information and in the capacity to generate new information or materials. These changes have challenged conventional training models or at least relegated them to a different function than they have had previously.

The new forms of learning and consumption of training must be, due to the need of the demand, ubiquitous (they should be able to be accessed anywhere), informal and personalized.

New models of intelligent learning are expected to emerge together with new paradigms for learning. Here for instance, machines, through the use of artificial intelligence and thanks to the observation of our activities, will decide which training materials can be of interest to increase our productivity and skills.

Traditionally, the learning process throughout a person's professional career has involved a linear progression from the apprentice or "novice" stage to the "master" stage. Throughout this long period, the person acquired knowledge and skills at an intermediate level before consolidating them at an "expert" or teacher stage.

At the present time these levels of knowledge are blurred and chaotic in nature. The process of learning and developing skills ceases to be linear and is now structured to be experiential or experimental.

These new forms of professional development are based on practice and experimentation, on realizing what is being done and learning from our achievements and mistakes. They necessarily involve phases of discovery, creation and application, and others of reformulation, trial and error until achieving incremental improvements and effective learning.

Figure 3.3-1 Traditional approach to professional career development. Image rights: Heather McGowan @HEATHERMCGOWAN

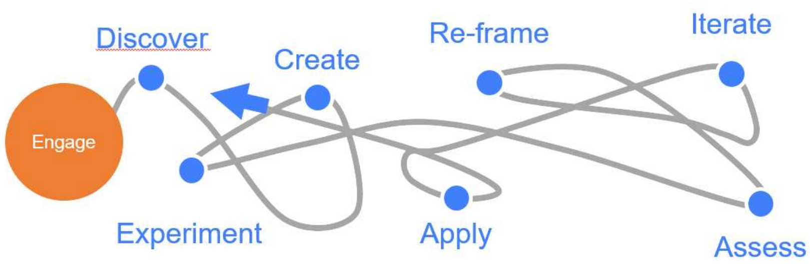


Figure 3.3-2 New approach to professional career development. Image rights: Heather McGowan @HEATHERMCGOWAN

3.3.1 How will I4.0 Impact the Job Market?

From a worker's perspective, the digital revolution will drastically change the labour market. As some studies point out, around 35% of current jobs are likely to be automated over the next 20 years. From 2015 to 2035, 8.3 million industrial jobs are expected to be lost, mainly due to a lack of productivity and competitiveness. McKinsey Global Institute estimated that intelligent automation technologies could save employers worldwide over $15 trillion in wages by 2030. Those massive labour savings raise a crucial point which needs to be considered: how will Industry 4.0 impact the job market?

In an extensive research paper released on 2019, McKinsey estimated that between 400 and 800 million current occupations could be displaced by 2030. The jobs most vulnerable to being replaced relate to data gathering and processing, physical labour, retail, manufacturing and repetitive tasks. When researchers analysed the bigger picture, however, it actually appears to be far brighter; it is estimated that 555 to 890 million new jobs will be created by 2030, with McKinsey noting "*this job growth could more than offset the jobs lost to automation.*"

The reason behind this growth is the growing demand of employees in non-repetitive, creative, mental tasks that can only be performed by humans and employment growth in sectors such as green technology, elderly care, and consumer goods and services.

Jobs expected be created relate to relocation of activities leveraging Industry 4.0 business models by:

1. Reinvestment in new industrial products and equipment.
2. Reinvestment in new services activities.

Companies will struggle without talented employees who are able to use existing digital technologies and adapt to evolving methods and new approaches. Companies will need people with digital design and technical skills, including UX/UI experts, human-centred design talent, data scientists, and data and technology engineers. Employees that are flexible, able (and willing) to learn and adapt, will be at a premium. To attract, develop, and retain the people they need, leaders need to adapt their organizations in multiple areas.

When embracing Industry 4.0 we should be constructive, asking how we can upgrade what we already have in terms of staff skills and their valuable capacities and reinventing ourselves. Engaging in continuous learning to develop new skills and lead the change. Upskilling and reskilling should be a priority to ensure that people stay at the forefront of the digital revolution and no one is left behind. Keeping in mind that people are the most important factor in digital, technological and knowledge developments.

3.3.2 The New Professional Competences and Profiles in Industry 4.0

Are the same competences and learning skills from the last century, important and relevant nowadays? Absolutely.

Previously, perseverance, procedure, order and deductive reasoning were competencies that could guarantee successful learning. These competencies continue to be important and valued skills. Today, however, today the competencies to guarantee success also include creativity, critical analysis, the resolution of complex problems and the ability to learn.

A report from the World Economic Forum 2018 lists which competences are professional development directors looking today compared with those pursued a few years ago:

	in 2020		in 2015
1.	Complex Problem Solving	1.	Complex Problem Solving
2.	Critical Thinking	2.	Coordinating with Others
3.	Creativity	3.	People Management
4.	People Management	4.	Critical Thinking
5.	Coordinating with Others	5.	Negotiation
6.	Emotional Intelligence	6.	Quality Control
7.	Judgment and Decision Making	7.	Service Orientation
8.	Service Orientation	8.	Judgment and Decision Making
9.	Negotiation	9.	Active Listening
10.	Cognitive Flexibility	10.	Creativity

Figure 3.3-3 Change in professional competences [2].

The implementation of Industry 4.0 requires a broad range of new professional roles which draw on skills from the STEM subject areas (Science, Technology, Engineering and Mathematics). Examples of areas where new skills are required include:

- *PLC/SCADA:* Professionals with knowledge in systems integration, automation, and supervision and control.
- *MES:* Professionals for the management of operations, integration with ERP, document management or management of manufacturing orders, among others.
- *M2M/IoT:* Engineers in charge of all technologies related to automation and sensors within the smart factory.
- *ESB:* Professionals in Big Data, Open Source software and Middleware to integrate all systems.
- *Operational Intelligence:* Experts in the analysis of real-time data for the improvement of the industrial plant and business operations.
- *Cloud:* Cloud experts to make the connected industry possible.

- *Machine Learning:* Experts to improve processes and perform preventive maintenance actions.
- *3D printing and additive manufacturing:* Experts for the development of processes, materials and design of parts.

Technical and social competences are considered equally important by companies who are transitioning to or have already transitioned to Industry 4.0. Therefore, the new technical profiles have and will have to be more and more trained in “soft skills” in order to achieve more senior roles within a company, and be able to join excellent companies. Examples of key competencies and soft-skills that will become more and more important for the change adaptation of digital industry workers include:

- *Open to Change and Flexibility:* This is the ability to adapt to different mentalities. It also extends to working in multiple disciplines, and extracting and integrating data and concepts from different fields.
- *Analysis and Solutions of Problems:* In the connected industry, this is the cognitive competence focused around the analysis of very detailed data (indicators, sensors, etc) that allows competence to be developed oriented to critical decision making.
- *Service Orientation:* In Industry 4.0 this is the ability to become a reliable "partner" of a client and the ability to share information with all project partners. Industry 4.0 cannot depend on the offer but must be oriented to meeting demand.
- *Computational, Data and Non-linear Thinking:* Big Data is a reality in our companies, so the ability to adapt to different approaches and thinking are skills that skilled workers must enhance. Some of the approaches that should be considered as fundamental thinking approaches to foster include:
  - *Design Thinking:* Non-linear, iterative processes that teams use to understand users, challenge assumptions, redefine problems and create innovative solutions to prototype and test.
  - *Lateral Thinking:* A new vision for solving problems using a creative approach. This involves ideas that may not be obtained using only traditional step-by-step logic.
  - *Systems Engineering:* Focussing on how to design, integrate, and manage complex systems over their life cycle.
- *Virtual Collaboration:* This is a key competence to share with virtual teams. It is necessary to promote knowledge of the functioning of technology platforms. It is the basis for sharing tasks and common work spaces. Technology is the facilitator, and the skills to take advantage of multiple tools effectively are key factors.
- *Independence:* This is the ability to work on your own, with little or no supervision; this is having the ability to monitor and evaluate your performance and be aware of your own strengths and weaknesses.

- *Achievement Orientation:* Give priority to actions that contribute to achieving the stated objectives and demonstrate an approach towards high levels of performance and quality. To be actively looking for continuous performance improvement.
- *Communication:* The ability to exchange information and ideas effectively through verbal and non-verbal communication. Adapt the communication style for different situations, people and media. You listen to understand, clarify understanding and take into account different points of view.
- *Teamwork:* This is about working with confidence within a group and assuming a specialist role to enable the team to achieve both personal and collective goals.
- *Critical Thinking:* This is asking the right questions to obtain quality information for a purpose. It is about interpreting the information in its context.
- *Ethical Behaviour:* This is about acting in accordance with agreed principles. It is to apply criteria of impartiality and transparency. It is about being generous, honest and acting with empathy as well as being oriented towards social objectives.

3.3.3 Training within Organisations

When, during a professional career, should we be trained through micro- (brief, concise, direct) or macro-(extensive) learning? Both have a place in lifelong learning and each one has its space and its objectives:

Figure 3.3-4 suggests that macrolearning is useful both at the start of a professional career and in the medium term, to provide support and updating of skills. Among these "periods" of formal learning, micro learning takes on special significance, as a permanent, fluid and agile update.

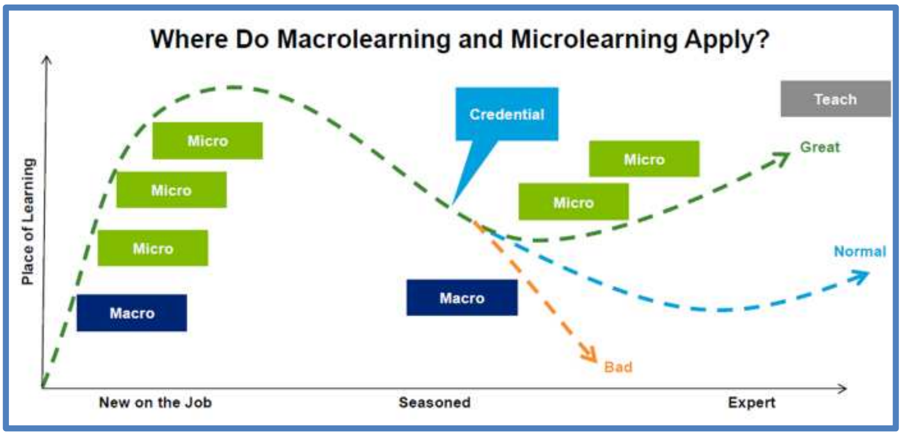


Figure 3.3-4 Understanding Macrolearning versus Microlearning [3].

As can be seen in Figure 3.3-5, there are countless sources of training that we can access almost instantaneously, free or almost free of charge, and through the internet.

Figure 3.3-5 Training resources [4].

Figure 3.3-6 Digital Learning Architecture [5].

Trying to put a bit of order in between so much training resources, the scheme in Figure 3.3-6 systematizes the type of materials that we can use to improve our learning.

Traditionally, the role of a training unit within a company has been to develop materials and training programs to ensure that the organization has the necessary talent at the right time. However, nowadays the previous concept has changed radically.

It is not efficient or even possible nowadays for the training unit department to respond to a permanent and almost erratic demand. Thus, it is no longer feasible to plan long-term training itineraries in a context in which everything is changing and evolving very quickly. The training unit should instead support three fundamental functions:

1. To promote a culture of continuous and self-directed learning in its organization, ensuring that all staff understand the importance of getting involved and committing to their continuous updating of skills.
2. To facilitate access to relevant (external) training resources that members of staff may use and teach techniques for its use.
3. To advise staff on the competences that they should develop, the resources they can access and lead these people through the difficult world of selection and search for quality resources, organized and systematized for a better understanding and use.

3.4 Bring innovation into your organisation: Practical Case Study

Encouraging workplace innovation not only helps SMEs stay on top of the market but also creates happier workplaces with higher levels of employee involvement and retention. As already mentioned, it is highly recommended that workers are directly involved in the innovation process. One way to achieve this is through creativity workshops. Teams from all areas work together with the aim of developing new ideas and solutions. Workers may feel inspired and mobilized towareds innovation if the workshops are well moderated and if they have the opportunity to solve key problems the enterprise may face. Of course, it is also crucial that the ideas don't end up in a drawer, but rather have a real effect.

In the case study below we present an example of how an enterprise used two creative workshop methodologies “role storming” and “Walt Disney model of creativity” to stimulate their team collaboration, problem-solving capacity and bring technical solutions to reality.

3.4.1 ABB Case Study – Collaborative Creativity Session to Optimize Line Operations

Foundation Year: 1934
Type: Large Company
Industry: Manufacturing of Electrical Components

Company Description

The ABB plant in Dalmine (Bergamo, Lombardy) can be considered the largest medium voltage switch factory in the world: it manufactures medium voltage devices and switchgear, with about 750 employees and an annual turnover of over 250 million dollars. The production process consists of the assembly and testing of configurable products, starting with ten thousand components purchased from suppliers.

Objective

Complete revamping of the materials and product handling systems along the assembly line. The scope was to improve the efficiency of the production process by optimizing the line operations.

Starting Situation

The initial assembly and testing process was completely manual handling, based on workbenches, rollers and forklifts with operators. Automatic handling systems were not available across the production line.

Proposed Solution

From a methodological perspective, the material handling system was revised according to the “role storming” and “Walt Disney model of creativity”, two approaches aimed at improving the problem-solving stimulating creativity.

The “role storming” and the “Walt Disney model of creativity” have been inspired by Walt Disney. In line with these approaches, people assume different thinking styles in order to propose ideas and suggestions: they act as an outsider to gain an analytical and external perspective. In the second thinking style, they act as a dreamer to propose radical ideas. In the third thinking style, they adopt a pragmatic point of view to select the best idea. Lastly, in the fourth thinking style they act in a critical ways to review and improve the idea.

From a technological perspective, the following solutions have been adopted:

1. Autonomous guided vehicles (AGV) equipped with artificial intelligence for an independent and efficient management of routes and material collection.
2. System of localization of the load units through the use of RFID technologies.
3. Central Internet of Things-based directing and coordination unit fulfilling the role of the Manufacturing Execution System (MES), which makes it possible to autonomously communicate multiple interconnected systems (Enterprise

Resource Planning, localization system, the AGVs, Programmable Logic Controllers on machinery, etc.).

More specifically, AGVs with magnetic guides were used for the handling of semi-finished products and the assembly of components. Moreover, they have been interconnected with the MES for the quality control of each single component and the management of production activities (visualization of the bill of materials, assembly drawings, production orders etc.) using tablets and barcode readers.

A SCADA (Supervisory Control and Data Acquisition) system has been adopted to monitors AGVs remotely and in real time, allowing preventive and predictive maintenance through the use of intelligent algorithms.

Further, to assist the operator in lifting and assembling materials, an intelligent electronic manipulator – equipped with a diagnostic app allowing the real time connection with the supplier's service / maintenance department and controlled by tablets and smartphones – has been introduced.

Barriers Encountered

During the phase of analysis of the technologies necessary for the implementation of the project, the company realized that the technologies on the market were not ready to satisfy the initial ambition. This caused difficulty in identifying the correct partner able to develop and implement the chosen solution in a short time.

Advantages Achieved

From an impact point of view, the case illustrates how the application of new automation systems from an Industry 4.0 perspective allows enhancement of the material handling flows, accelerate the production process and improve quality through real-time control, increasing the overall efficiency of the production process. Moreover, the case exemplifies how employee involvement and collaborative innovation methodologies supported the enterprise brainstorm on the final implemented solution.

3.5 Conclusions

The digitization of the industrial environment has caused vast changes in the way we work, think and interact with each other. Therefore, previous models (and techniques) for learning, training and managing teams have also been affected and even changed.

The evolution of the labour market in Industry 4.0 represents a paradigm change. Companies wishing to attract STEM, digital and technical talent need to evolve their processes for talent selection, training, staff management and retention. Technical and social competences are considered equally important by companies who are transitioning to or have already transitioned to Industry 4.0.

This leads to changes in jobs and workspaces, and requires teams to shape these changes. Therefore, in many places of work it will be more and more important to learn from and with others. A culture of cooperation is needed. The teams that are better prepared for digital transformation are those in which team members help each other to acquire professional knowledge. Teams that continue their education and that enjoy learning and trying new things. In the coming decades it will be necessary to consider the opinion of all – large and small teams, decision makers and team leaders, each individual employee.

3.6 References

[1] SOURCE: PROSCI. https://www.prosci.com/adkar/adkar-model
[2] SOURCE: WORLD ECONOMIC FORUM 2018
[3] The Disruption of Digital Learning. 10 Things We Have Learned. bersin.com Copyright © 2017 Deloitte Development LLC. All rights reserved. SOURCE: BERSIN BY DELOITTE, 2017
[4] SOURCE: WORLD ECONOMIC FORUM 2018. DEGREED. The Disruption of Digital Learning. 10 Things We Have Learned. bersin.com Copyright © 2017 Deloitte Development LLC. All rights reserved.
[5] SOURCE: BERSIN BY DELOITTE, 2017

Part II
Decision Making for Manufacturing Systems

Part II Overview

Manufacturing plants and lines are highly complex systems, with large numbers of interacting components required to not only produce products, but to do so in an efficient, cost-effective, and reliable manner. A manufacturing system extends beyond the machines or equipment used to create or assemble parts into products, but extends into the business processes, decisions made, people involved, and wider supply network. With the increasing adoption of digital manufacturing or “Industry 4.0” technologies, systems are getting exponentially more complex and difficult to understand. What is more, it is rare for a manufacturing system to be in a steady state. New products must be manufactured, tooling wears or fails, supply quality can vary. Understanding the current state of a manufacturing system is a significant challenge, predicting the future state even more so.

Understanding how all these elements interact and effect each other is key to effective decision making. Without understanding the system, it is not possible to make the correct decisions at the correct time to keep the manufacturing enterprise productive and profitable. Key to this understanding is the process of analysis – examining the system, understanding the current state, identifying what factors influence the state of the system, and using these to move the system towards an improved state.

The next three chapters of this book represent a progression through levels of sophistication of manufacturing systems analysis, from static mathematical methods focusing on the key performance indicators of the current state of the system, through to digital modelling and simulation to predict the outcomes of proposed changes, and into a tightly integrated real-time digital twin that mimics the current system and can update the model to improve the accuracy of simulations.

This is not to say, however, that this part details methods from bad to good. Traditional mathematical manufacturing analysis has been used for decades and provides effective and useful results. It is recommended that if you have not started formally analysing and modelling your manufacturing systems that you start with production analysis first, and then progress to modelling and beyond.

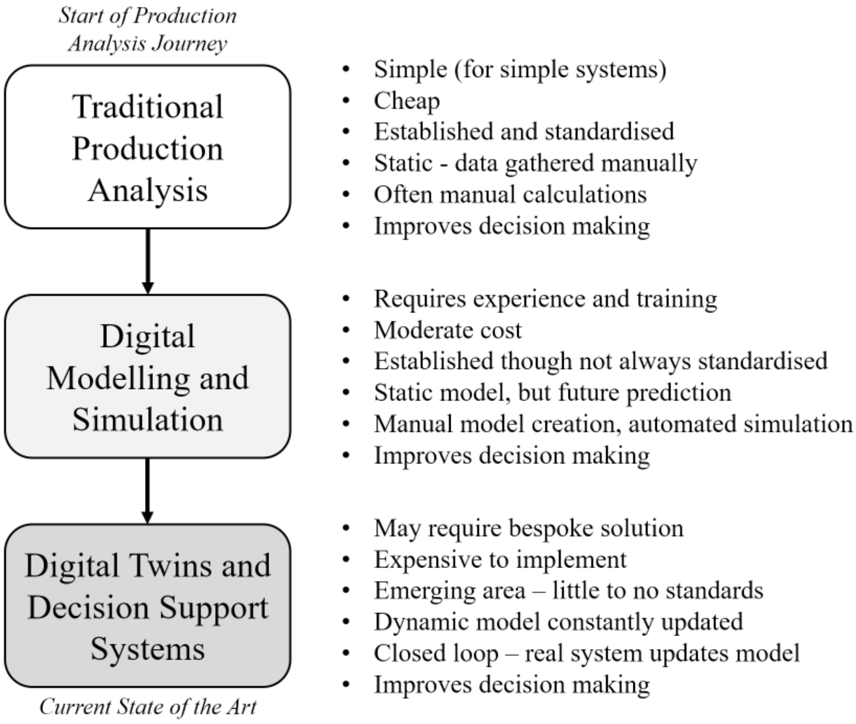


Figure II-1 Production systems analysis is not a choice of different technologies, but more of a journey. Each step gives richer data in a more automated fashion, but also typically costs more and is more complex. Every stage of the journey improves decision making, and there is no obligation to continue to the end. For many companies, digital modelling is sufficient, and a digital twin would not be cost effective.

These chapters introduce the concepts of analysis and decision making and breaks them down into formal processes which can be followed. The first chapter (Chapter 4) introduces conventional decision making – methods for analysing manufacturing systems and networks to calculate key performance indicators or to identify areas of concern. The limitations of these methods are discussed, and Chapters 5 and 6 discuss modern methods for manufacturing systems analysis, using offline modelling and simulation (Chapter 5) and state-of-the-art integrated digital twins and decision support systems (Chapter 6).

Chapter 4 Manufacturing Systems Analysis

Jack C Chaplin and Giovanna Martinez-Arellano

4.1 Introduction

Analysis (noun): Detailed examination of the elements or structure of something.

- Oxford University Press

Effective decision-making is critical in manufacturing enterprises, and selecting the correct course of action can be the difference between a successful and competitive enterprise, and falling behind the competition. Decision-making is the process of the selection of a course of action to best achieve your goals, given available options and the available information. Decisions are rarely simple, and the decision maker must select the best choice to maximise one or more criteria, and must often do so with incomplete information.

Within the manufacturing domain, effective decision-making is crucial to remaining competitive. For example, making large investment decisions carries significant risk and significant reward, and this is especially true for smaller companies. Effective decision making is best achieved with access to accurate and high-quality data which is then converted to and presented in more usable forms. Data can be converted into more usable forms to achieve maximum value and to inform the manufacturing enterprise. This is often described as the difference between data and information, with the former being the raw numbers or measurements, and the latter being actionable insight into the manufacturing process.

J.C. Chaplin (✉), G. Martinez-Arellano
Institute for Advanced Manufacturing, University of Nottingham, Nottingham, UK
e-mail: jack.chaplin@nottingham.ac.uk, e-mail: giovanna.martinez@nottingham.ac.uk

J.C. Chaplin et al (eds), *Digital Manufacturing for SMEs*
DOI: https://doi.org/10.17639/7CSC-VM73

A common representation of this is the Data, Information, Knowledge, Wisdom (DIKW) pyramid, shown in Figure 4.1-1. Each step in the pyramid can be analysed and compressed to the next stage up, slowly reducing the raw volume of information, and increasing the insight and value. Within this framework, the steps are defined as:

- *Data:* Signals or numbers representing physical phenomena from sensors, but without context or metadata.
- *Information:* Contextualised and inferred from data, information has meaning and context, and can be used to answer questions.
- *Knowledge:* Processed information that has been compared with previous experiences to enable the user to determine why things have happened.
- *Wisdom:* Sometimes omitted from this model or combined with knowledge, wisdom is using knowledge to intuitively understand a process and to determine the best future actions.

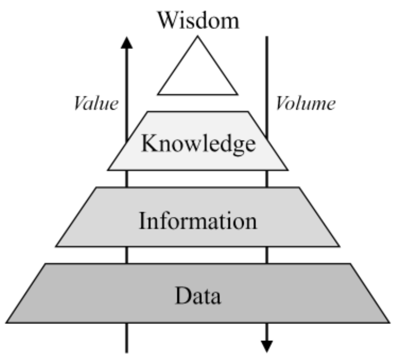


Figure 4.1-1 The DIKW Pyramid shows how data, information, knowledge, and wisdom are interrelated and how they compare in terms of storage volume required, and inherent value.

There are many sources of data in manufacturing enterprises, but one of the most important is data generated by the manufacturing lines and systems themselves. Manufacturing systems analysis, modelling and simulation deals with methods to better understand manufacturing processes and manufacturing lines. By understanding the current situation, and by testing potential changes, it is possible to ensure that the chances of success are maximised before making potentially costly changes (in terms of equipment or line downtime) to the manufacturing enterprise. What's more, these methods will enable tracking the performance of production systems over time, allowing the tracking of improvements as they are implemented or to spot issues before they become critical. The techniques here can be applied from the smallest manufacturing lines (which could just be a machining centre with a worker loading and unloading parts) to larger, more complex manufacturing lines.

However, simulation and modelling can often be overkill for many manufacturing systems. Standard mathematical evaluation can reveal insights into manufacturing systems that are not immediately obvious. Understanding utilisation, reliability, throughputs and capacities is an essential first step to maximising productivity by identifying areas for improvement. These areas will enable targeting interventions to the areas that will give the greatest return.

This book chapter discusses the pre-requisites for manufacturing analysis; understanding the type of system to be evaluated and understanding the question that needs to be answered. It then moves on to discuss two key methods of offline mathematical manufacturing analysis – conventional production analysis and queueing theory.

4.1.1 Manufacturing Systems Morphologies

Morphology (noun): The study of the forms of things.

- Oxford University Press

To understand how to analyse a manufacturing system, it is first important to understand the form and method of operation of the system, as this will influence the methods and formulae used. Knowing the terminology for your manufacturing system also simplifies searching for applicable resources and advice. A diagrammatic summary of these manufacturing systems morphologies can be found in Figures 4.1-2 to 4.1-5.

Dedicated Manufacturing Systems (DMS) use fixed automation to produce core products at high-volume with maximum cost effectiveness. When a single product is likely to be manufactured in large quantities without major alterations for the foreseeable future, a DMS is almost always the best choice, while simultaneously often being simpler to implement. Equipment is typically arranged in a linear manner – the stereotypical *production line* – and linked with a material handling system to move the parts along.

Where volumes are lower and multiple parts are to be manufactured, *batch manufacturing* is one of the most common manufacturing strategies, enabling mid-volume manufacturing by batching production together. The time required to change a production line between product types is significant, so batching ensures this changeover time happens as infrequently as possible.

Group technology is the strategy of gaining efficiency when many similar products are to be manufactured by grouping products into part families with similar features and manufacturing processes. For example, a company making bearings may make the same bearing in multiple different sizes, with multiple finishes and multiple different lubrication methods. Though this would constitute a very large number of potential part variants, the commonality between them allows them to be produced with the same machinery (albeit with different settings and tooling) and therefore comprise a part family. By comparison, though similar in function, a ball bearing and a cylindrical roller bearing require sufficiently different processes that they would not constitute a single part family.

Utilisation of group technology enables a production line to be set up and configured to produce any members of the part family quickly. This makes changing between products in the same family much quicker and simpler, reducing batch change over times and making smaller batch sizes more cost effective.

Cellular manufacturing groups machines into cells, where each cell is specialised in completing a step or closely related group of steps required to manufacture a product or family of products. Manufacturing time is reduced by bringing the machines into close proximity; the characteristic U shape as shown in Figure 4.1-3 is a common cell layout and enables a single operator who specialises in the cell's task to oversee all machines, and the order in which machines are used in the cell can be changed without significant disruption. Cellular manufacturing is a great choice for lower volumes of highly variable products. Multiples of the same machine in a cell might be grouped together into co-located stations to take advantage of parallelism, and machines which are almost always used together might also be grouped into stations.

Flexible Manufacturing Systems (FMSs) are highly automated manufacturing cells, able to automatically route parts between the constituent machines, enabling flexibility in terms of the parts and part families produced. Shown in Figure 4.1-4, one of the defining differences from a cellular system is the use of an automated material handling system to route products around the system. The stations in an FMS are themselves general and flexible, predominantly computer numerical control (CNC) machines, often with automatic tool changers. This allows for the stations to be utilised for a wider range of tasks. FMSs are ideal for lower volumes of highly variable products similar to cellular manufacturing, but the high levels of automation increases labour productivity and allows for unattended production.

Reconfigurable Manufacturing Systems (RMSs) utilise both a structural design and a digital control system that allows the constituent stations to easily be changed, so the function and capacity of the system can be rapidly altered. Stations can be plugged in and out of the system to change the cell's function. Additional functional modules can also be added or removed from machine and/or stations to change their functionality or increase capacity. The system can scale up or down capacity by adding or removing modules, machines, or stations. Individual stations are often less flexible than would be used in an FMS – system flexibility is given through system structure, not individual flexibility.

Traditional mathematical analysis techniques generally focus on dedicated manufacturing lines and cellular manufacturing morphologies, with variants for when batch manufacturing is implemented on either. FMS and RMS are comparatively recent developments, and their evolving and changing nature makes them less suited to these styles of analysis. Using digital tools and particularly automatically updating digital twins may be more appropriate here to help manage the changes, and digital twins will be discussed in Chapter 6.

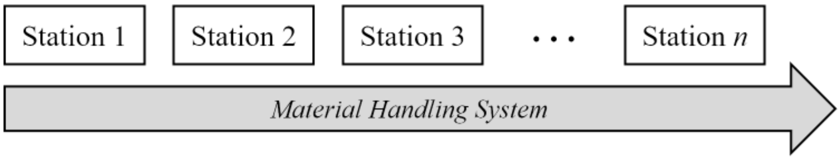


Figure 4.1-2 An example dedicated manufacturing system, showing the typical production line morphology. A single linear material handling system transfers parts or products between processes in sequence. Processes could be single pieces of equipment, manual process stations, or multiple similar pieces of equipment to form a processing station with parallel processing capability.

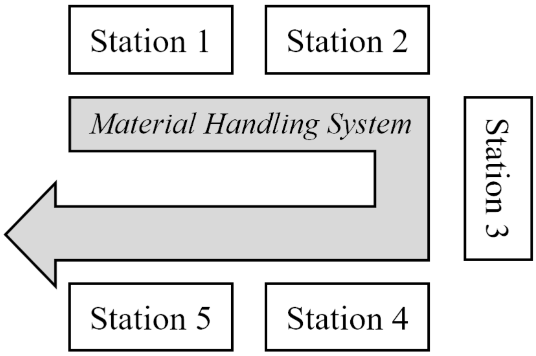


Figure 4.1-3 An example manufacturing cell layout. The arrow represents the most common flow of parts through the cell, but cellular manufacturing can alter the path to make a wider variety of products without significant disruption.

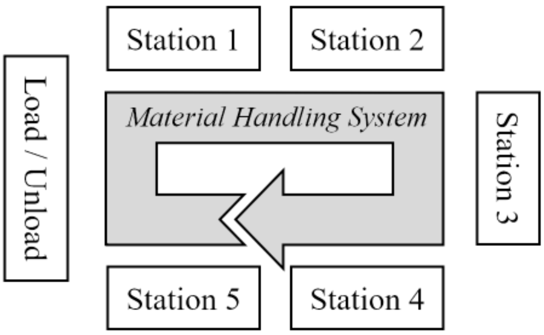


Figure 4.1-4 Example of a Flexible Manufacturing System layout. Parts can be routed between any stations in any order using the Material Handling System.

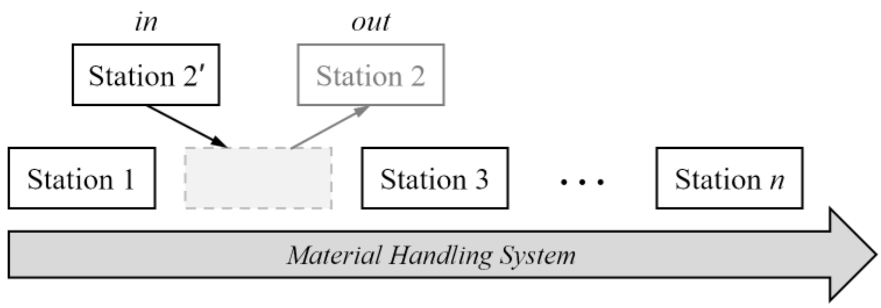


Figure 4.1-5 Reconfigurable Manufacturing Systems allow for rapid structural changes to the cell to change its functionality.

4.1.2 Decision Making

Before beginning analysis of a manufacturing system, it is important to understand what question you are actually trying to answer. Analysis usually serves to inform decision making by measuring key performance indicators (KPIs). By discovering values for KPIs you can better understand the strengths and weaknesses of the current system, and make decisions on how to improve it.

Decisions in manufacturing can mean a wide variety of different choices, but Hayes and Wheelwright [1] separate this into ten possible categories of decisions. These are detailed below together with non-exhaustive lists of what sort of choices these categories include:

- *Capacity:* How much capacity flexibility should be made available, what shift patterns should be utilised, and what strategies are available for subcontracting for temporary over or under-capacity situations.
- *Facilities:* The size, maximum capacity, physical location, and primary task assignment of physical manufacturing facilities.
- *Human Resources:* The policies around recruitment of new employees, the training and development of existing employees, and the culture and management style that the business adopts.
- *New Product Introduction:* How new products are selected and developed, as well as procedures for design (including design for manufacture), and how products are introduced and ramped up on the shop floor.
- *Organisation:* The structure of the manufacturing enterprise, as well as accountabilities, roles, and responsibilities.
- *Performance Measurement:* How the processes and people in a manufacturing enterprise are evaluated and monitored for productivity and other performance measures, as well as any recognition and reward schemes for employees.
- *Production Equipment:* The equipment and technologies chosen for manufacturing products, the physical layout of this equipment into cells or lines,

and the level of automation within these lines. This also covers maintenance approaches and policies, and how much in-house development of new or updated processes is possible.

- *Production Planning and Control:* How production is controlled (either via automated systems or by manual processes), how orders should be assigned and scheduled, and how materials are stored and moved around the manufacturing operation.
- *Quality:* The quality goals adopted by the business, as well as the quality assurance and quality control methods and policies used to reach these goals.
- *Vertical Integration:* High-level strategic decisions such as make versus buy, policies on supplier selection and continued relationships, reliance on single or multiple suppliers to spread risk.

To make effective decisions, it is important to understand exactly what the decision-making process is. The decision-making process [2] is a series of phases, and follows the flow shown in Figure 4.1-6:

1. *Intelligence – Problem Discovery:* First, it is necessary to identify a problem and to determine who the decision-makers and stakeholders in the decision process are. Once the problem has been recognized, the problem should be defined more formally, determining the requirements from the stakeholders to obtain a list of considerations and goals. It is important to take time with this phase, as the group of stakeholders may all think they share a common view of the problem, but there are often details that are assumed and may not be shared between all parties.
2. *Design – Solution Discovery:* Finding possible alternatives that could be implemented, and assessing their possible contributions to the problem. This could be as simple as a discussion amongst the stakeholders to brainstorm ideas, but for critical decision making it is recommended that a more rigorous process is followed. Though commonly overlooked, it is recommended that "doing nothing" is always one of the possible choices. A common recommendation is for the problem to be modelled, as this will enable potential solutions to be tested to enable the selection process in the next phase. This is the phase where the model is created. Models could be numerical models developed in a spreadsheet, full simulations of productions lines, or anything in between. The possible approaches for modelling here depend on the nature of the decision being made.
3. *Choice – Solution Selection:* The possible options developed are then evaluated for their contribution to the problem. Depending on the approach taken in the Design phase, this could either be a continuation of the discussion, gathering data to inform the decision, or running the possibilities on the model to see what impact they have on the defined goals. Even with a fully developed model, the choice of the solution is rarely simple. Many different criteria need to be evaluated, including the time and cost required to implement the solution, and risk and reward of doing so, the possible disruption while the solution is implemented, and the availability of skills to enact it. Remember to evaluate

doing nothing, as the disruption and cost of implementing a solution may be too high.

4. *Implementation – Solution Deployment and Testing:* Lastly, the solution should be deployed and tested. The solution needs to meet the goals defined in the first phase of this process, and as many choices are not instantly implemented or their effects instantly felt it is important to keep monitoring and testing. If the solution is failing to meet expectations or goals, the entire decision making process is an iterative one. It is always possible to return, rethink, and re-plan. It is very common for companies to fall into the trap of continuing with a bad decision when all evidence suggests that things will not improve (the "sunk-cost" fallacy). However, a properly executed decision making process will reduce the chances of this occurring.

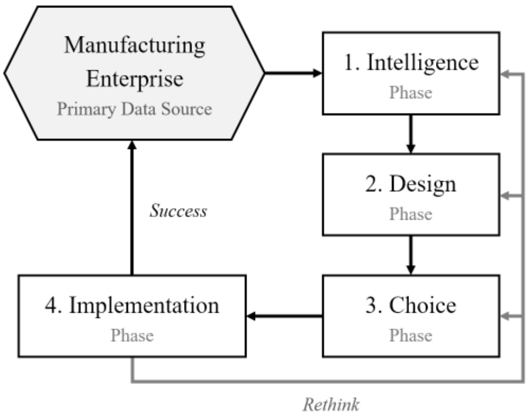


Figure 4.1-6 Phases of the decision-making process. It is never too late to go back and rethink a decision, especially as new information becomes available.

A key aspect of this process is the ability to monitor and check how successful a decision has been, and whether the changes to the manufacturing or business processes are yielding the expected and intended results. Key Performance Indicators (KPIs) are the tool with which the success of a decision can be measured.

4.1.3 Key Performance Indicators (KPIs)

To achieve an optimal manufacturing process, you must first define what optimal is. No manufacturing process can simultaneously maximise yield, productivity, uptime etc. while also minimising material wastage, energy usage, and downtime. Careful selection of KPIs is required to understand what is important and what you are trying to achieve. The indicators through which the performance of a manufacturing process is characterised have evolved and diversified. Traditional KPIs include:

- *Productivity:* The efficiency of production, or the ratio of output to input. Productivity is a KPI you want to maximise, and there are many ways to calculate it, though Overall Equipment Effectiveness (OEE) is a common and useful tool for manufacturing equipment.

$$OEE = AUYr_{os} \tag{4.1.1}$$

  Where:
  - *A: Availability.* The uptime of the equipment, which is reduced by maintenance, breakdowns etc.
  - *U: Utilisation.* What percentage of possible usable time is actually being utilised. Poor scheduling or a lack of available upstream parts (also called starvation) will reduce this.
  - *Y: Yield.* The percentage first-pass yield of the process, referring to the quality of produced components.
  - r_{os}: Operating Capability. The percentage of the maximum throughput the equipment works at.

  An optimal process will never break down or require maintenance, will be active 100% of the time, will produce consistently high quality parts, and will operate at its maximum processing speed, giving an OEE of 100%. Though an OEE of 100% is unrealistic, it remains a useful tool to identify areas of concern. The percentage of first-pass yield of a process is often overlooked, and is an important part of lean production.

- *Cost:* Reduction of processing costs is another common goal and can be an effective KPI. The cost of a process is typically obtained by combining overhead and labour rates with material costs, processing costs, energy costs, and the cost of waste.
- *Quality:* Maximising the quality and yield of processes is another very common goal. How this is measured will depend on the process and application, but improving quality reduces the need for rework or customer returns.

However, these are far from the only KPIs available to manufacturing enterprises. As part of the most recent metrics survey conducted by the Manufacturing Enterprise Solutions Association (MESA), 28 manufacturing metrics were identified as being the most utilised by discrete, process, and hybrid/batch manufacturers [3]. These are organised in categories based on what aspect of the business they represent and serve as a starting point for creating KPIs to represent them.

Overall equipment effectiveness may be a common metric for *productivity*, but it is far from the only one. Productivity and efficiency is key to a company's profitability and ability to compete, so productivity goals are extremely common. Some KPIs include:

- *Throughput:* How much product is produced at a machine, line, unit or plant over a given period. Formulas and methods for calculating this are discussed in section 4.2 Conventional Manufacturing Systems Analysis.
- *Capacity utilisation:* How much of the total manufacturing output capacity is actually being utilised over a given period. Again, see section 4.2 Conventional Manufacturing Systems Analysis.
- *Overall equipment effectiveness:* Shown in equation (4.1.1), this is a common measure of the overall effectiveness of a piece of equipment based on the availability, performance and quality.
- *Production attainment:* Percentage of time a target level of production is achieved, enabling an enterprise to meet the schedule agreed with its customers.

Improving *the quality of products* is another common aim for companies, with improved quality resulting in less waste and rework, and improved customer satisfaction. Quality, and improving it, is a vast topic in its own right, but some common KPIs are:

- *Yield:* Percentage of products correctly manufactured without a need for rework or scrap.
- *Customer rejects/returns:* How many times the customer rejects a product. Correct quality assurance policies should reduce this figure.
- *Supplier quality incoming:* Percentage of good quality materials coming from the suppliers is not as uncontrollable as you might believe. Working closely with suppliers can improve the quality of supply, but you may also have to consider alternative suppliers.

Improving *customer experience and responsiveness* is a common goal for companies but measuring it can be a tricky task. Three common KPIs for this are:

- *On-time delivery to commit:* The percentage of time a completed product is delivered to the customer on the schedule agreed.
- *Perfect order percentage:* The percentage of times customers have received a complete correct order on time.
- *Manufacturing lead time:* The time it takes to manufacture a product, from when the order is accepted to the finished product(s) being dispatched.
- *Time to make changeovers:* The time it takes to change a production line from producing one product to a different one, to meet customer demands in a changing market.

Regulatory compliance is conforming to relevant policies, laws and standards in areas such as health and safety, environmental protection, and data security. This is a clearly important area with non-compliance leading to fines and penalties, combined with the risks the regulations are designed to protect you from. Some KPIs in this area include:

- *Reportable Health and Safety incidents:* Measure of the number of reported health and safety incidents over a period of time, including both injuries and near misses that require action to stop them happening again.
- *Reportable environmental incidents:* Number of reported incidents over a period of time, including chemical spills, waste issues, air contaminants, etc.
- *Number of non-compliance events:* Number of times the plant was operating under non-compliant conditions over a period of time.

Profitability and reducing costs are a broad set of KPIs that can include aspects beyond the manufacturing process and into the wider business. The KPIs listed here are generally the KPIs which deal with the manufacturing processes rather than the business as a whole:

- *Total manufacturing cost per unit:* Typically represented excluding materials, this how much the production process alone costs to manufacture a single product.
- *Manufacturing costs as a percentage of revenue:* Related to the previous KPI, what is the ratio of manufacturing costs to the overall revenue of the enterprise?
- *Revenue generated per employee:* Typically a comparison between multiple manufacturing sites, what is the revenue divided by the number of employees?
- *Average unit contribution margin:* Profit made per manufactured product.
- *Return on assets:* Profit made divided by the value of the assets and deployed capital equipment required to generate that profit.
- *Energy cost per unit:* Energy costs incurred per produced unit or volume.
- *Cash-to-cash cycle time:* Time between the purchase of a product by a customer and the collection of payments from the sale of the product.
- *EBITDA:* Earnings before interest, taxes, depreciation and amortisation – a common metric for the profitability of a business.
- *Net operating profit:* One of the purest measures of cost effectiveness, what is the profitability of the enterprise?

Other KPIs worth considering which do not fit into broader categories include:

- *Work in Progress (WIP) Inventory:* Measurement of the efficient use of inventory materials. WIP represents unattained value, and is often a risk to the business if it cannot be rapidly converted into products.
- *Planned vs emergency maintenance:* Ratio of how often scheduled maintenance occurs versus the need for disruptive and unplanned maintenance.
- *Downtime vs operating time:* Asset availability and reliability.
- *Rate of new product introduction:* How quickly new products can be introduced to the market, including the product design, process planning, ramp up, and manufacture.
- *Engineering-change order cycle time:* How quickly modifications to existing products and process plans can be processed and implemented.

At their core, KPIs are a relatively simple concept – they are the measurable goals by which a production process or changes to that process can be monitored and evaluated. However, there are some considerations to make when implementing KPIs to maximise their usefulness.

4.1.4 Implementing Key Performance Indicators

Effectively measuring, analysing, and improving KPIs is not as simple as it may appear. While certain metrics work well for specific processes, it is often the case that there are multiple combinations of metric indicators needed to ensure that a larger business objective is being met. Implementation of KPIs follows a cyclic pattern with five stages:

1. *Definition:* The first phase of the life cycle is defining the KPIs to be used. Although there are thousands of KPIs already defined and used by manufacturers KPIs might sometimes need to be redefined as a more focused or aggregate KPI depending on the exact business objectives.
2. *Collection:* The second phase of the KPI life cycle consists of bringing together candidate KPIs for consideration. Of particular importance is excluding any obviously irrelevant KPIs, as it is common to end up with far too many.
3. *Set Composition:* The third step of the KPI life cycle is to choose from the KPI collection the specific set to implement. ISO 22400-1:2014 Part 1 [4] can help with this process, but the main focus is to ensure the chosen KPIs form a comprehensive set of measuring the business' objectives without being too onerous to implement and monitor.
4. *Implementation:* In this phase, the stakeholders define the process for assessment, examining KPI values and trends periodically and describing action plans for improving process control from KPI values.
5. *Assessment:* Stakeholders evaluate the relevance of the KPIs i.e. how well they align with the current performance objectives of the process, and how well they were implemented. If necessary, the implementation can be adjusted to improve the process.

The process is a cycle. KPIs should be periodically re-evaluated to ensure they're still meeting the needs and requirements of the business. As each KPI is chosen, the data to measure the KPI must be collected, and for that the data must be visible and transparent.

Visibility and transparency are key prerequisites for the optimisation of manufacturing processes. The more information available about a production process, the better performance can be measured through KPIs and better decisions can be made about how to react to events and issues.

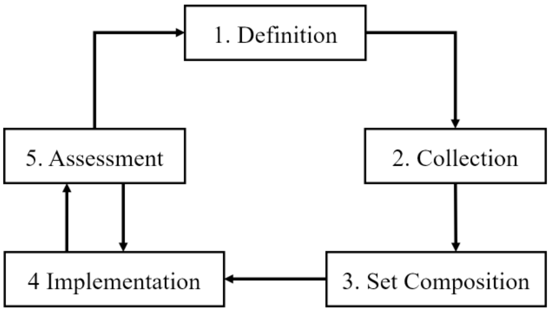


Figure 4.1-7 The KPI Life Cycle.

- *Visible Data:* Data which is easy to access with well-defined processes for doing so.
- *Transparent Data:* Data which is easy to understand and make decisions based upon it.

Despite the importance of process visibility and transparency, real-time reporting schemes with standardized KPIs are still missing in many enterprises. Even when KPI collecting is implemented these often require a manual input process through paper-based forms. Reported data is often based on quantities, e.g. produced units per shift, rather than more insightful metrics. Important data for process optimization such as setup times, change over times, processing times or downtimes are often missing and therefore cannot be reported. This is where modern technologies such as smart sensors and the Internet of Things can be beneficial and will be discussed more in Chapter 6.

4.2 Conventional Manufacturing Systems Analysis

Digital manufacturing techniques are changing the implementation of manufacturing systems and their analysis, with simulation and modelling now common practise across a wide range of sectors and company sizes. One key area in which these techniques can improve the performance of a system is through analysis of the system and identifying areas for improvement. However, the methods used to analyse systems with digital tools are built on a robust framework of more conventional production systems analysis, and understanding these will aid in understanding where and why more complex techniques should be used.

4.2.1 Production Analysis

Productivity (noun): The effectiveness of productive effort, especially in industry, as measured in terms of the rate of output per unit of input.

- Oxford University Press

Of all KPIs to monitor, the productivity of a manufacturing system is one of the most commonly desired, but also one of the most commonly misunderstood due its often-abstract nature. We all want to improve productivity, but what does that really mean?

Productivity is the ratio of output to input from a process. More formally:

$$Productivity = \frac{Units\ of\ Output}{Units\ of\ Input} \tag{4.2.1}$$

The units of output are the products the manufacturing process is creating, but the units of input can be one of several types, including:

- *Capital:* The produced output per unit of capital i.e. products manufactured per Euro/Pound/Dollar spent. This can include the non-recurring value of assets such as tools, plants, and equipment, as well as recurring costs such as maintenance or utility costs.
- *Labour:* The output per person, or more specifically per person-hour. Where manual processes are key, improving labour productivity can have significant effects. This could involve improved ergonomics are processing station, implementation of assistive technologies, or simple morale and motivation.
- *Material:* The output per unit of material, such as raw materials or parts used in the manufacturing process. Improving quality and reducing waste will improve material productivity.

Productivity is a top level goal, and identifying how to increase it requires identifying areas for improvement in a manufacturing system. Productivity is a key performance indicator – a measurement of how well the manufacturing process is performing, a desirable outcome. However, a lower level concept is that of a *metric*. A metric is a measured value which isn't necessarily desirable in isolation, but contributes to a KPI. Traditional manufacturing analysis typically measures metrics, which in turn identify areas for attention which – if improved – could affect the KPI which is the actual goal of the improvement process.

Measuring metrics will enable tracking the performance of production systems over time, allowing the tracking of improvements as they are implemented or to spot issues before they become critical. The techniques here can be applied from the smallest manufacturing lines (which could just be a machining centre with a worker loading and unloading parts) to larger, more complex manufacturing lines.

Calculating metrics does require some application of mathematical formulae, and also having data about the production operations. If you do not have access to this data, it is strongly recommended you start to collect it. Though it is possible to analyse your manufacturing systems via the methods in this chapter or with simulation and modelling tools using only estimated values, the results will be more accurate with real, measured data. Without access to real performance data, identifying ways to improve will be more guesswork than strategic, adding risk that investment in improvements (in terms of time or money) may be wasted. However, the results of calculations performed with informed estimates of performance figures may still offer broad insights and identification of large problems.

Note that these methods are generally concerned with discrete manufacturing systems, rather than continuous ones. There are alternative methods for analysing continuous processes such as chemical production or material processing.

4.2.2 Production Rate

The *production rate* is the number of work pieces a specific production process can produce per hour. Calculating this will enable understanding how long it will take to produce an order, the levels of utilisation of the system, and where bottlenecks might exist. The production capacity of the entire manufacturing line will be discussed in section 4.2.3. The production rate of a process can be calculated with a few short steps, starting with determining the cycle time and then calculating the production rate based on the batching strategy. The calculations should be based on real measured data to get the best results, but approximate measures could be used as a first estimation.

The first step to calculating production rate is to understand the *cycle times.* However, to calculate the cycle time, you must first identify what the work pieces that the process produces are.

- *Work Piece:* The discrete part or product being manufactured by the production system.

  These can be either complete products or parts depending on context. For example, a production line may manufacture entire watches, making each watch a work piece. Alternatively, the company's production system could manufacture only the watch casing as a supplier to a watch manufacturer, making those casings the work pieces. Work piece(s) is typically abbreviated as *pc* (pieces).

- *Cycle Time:* The time a work piece takes to have a single operation performed on it.

For example, the time it takes to mill the watch casing in a machining centre is a cycle time. The time it takes to polish the casing after milling will also have a cycle time.

This is expressed as a time, typically minutes (*m*). To calculate the cycle time, three pieces of information are required:

1. *Operation Time (T_o):* The time a work piece actually spends getting processed.
2. *Handling Time (T_h):* The time a work piece spends being loaded and unloaded from the production process.
3. *Tooling Time (T_t):* The *average* time necessary to set up the tools for the operation, including replacing worn out tools. A tool does not need to be replaced for every part, so the time that takes should be averaged out.

The cycle time is the sum of these three times.

$$Cycle\ Time = Operation\ Time + Handling\ Time + Tooling\ Time$$

$$T_c = T_o + T_h + T_t \tag{4.2.2}$$

For example, a watch casing that takes 5 minutes to mill, 1 minute to load into the machine, 30 seconds to unload from the machine, and requires 10 minutes to replace the mill head every 50 casings would have a cycle time of 402 seconds, or 6.7 minutes.

When you have calculated the cycle time for a process involved in creating the work piece, the production rate can be calculated.

- *Production Rate:* The number of work pieces produced by a production process per hour (pc/hour).

Given that the cycle time of a production process has been calculated, it might be assumed that the production rate is simply the number of cycles that fit into one hour. But this is an approximation that ignores the aspects of how a company batches its jobs together, which can have a significant impact on the production rate. There are four strategies for production that are considered here. These are:

- *Batch-Size-of-One:* Every product being manufactured is unique, and will require the manufacturing process to be setup specifically for each product every time. This is at the extreme end of customisation, and the cost and time of the setup will significantly increase the production cost of the product.
- *Sequential Batch Processing:* Grouping similar products together into batches, but each product still needs to be processed individually. This is the most common strategy for manufacturing, and helps reduce set-up costs.

- *Simultaneous Batch Processing:* A specialisation of sequential batch processing, simultaneous batch processing uses processes that can enable multiple work pieces to be processed simultaneously, such as heat treating.
- *Mass Production:* When a company manufactures parts in very high volumes, the costs of the setup times become so low they can effectively be ignored. This requires a product that almost never changes, and is in high demand.

Each method of production has a slightly different method of calculating the production rate, which are detailed in the following sections.

4.2.2.1 Batch-Size-of-One

At the extreme end of customisation, *batch-size-of-one* production rate is dominated by the change-over and set up times between the different products being manufactured. A company which specialises in job shop production may find itself dealing with low quantities of a product to manufacture, and one-offs are not outside the realm of possibility. The production time for a single item is the sum of two things:

1. *Set up Time (T_{su}):* The time it takes to set up the production process for the unique work piece. For example, loading the required CNC program, and adjusting the clamps to hold the work pieces in a milling machine.
2. *Cycle Time (T_c):* Equation 4.2.2, the time to process a single work piece.

$$Production\ Time\ per\ Work\ Piece = Setup\ Time + Cycle\ Time$$

$$T_p = T_{su} + T_c \tag{4.2.3}$$

The production rate (R_p) is then how many products can be made in one hour. These equations assume all times are expressed in minutes.

$$Production\ Rate = \frac{60}{Production\ Time\ per\ Work\ Piece}$$

$$R_p = \frac{60}{T_p} \tag{4.2.4}$$

For example, a small company makes bespoke watch casings. Each casing takes an average of 10 minutes to mill, but also takes 20 minutes to set up the milling machine for each unique casing. Using equation 4.2.3, *the production time is therefore 30 minutes per casing, and the production rate of the milling machine for bespoke casings is* 2 pc/hr *on average.*

4.2.2.2 Sequential Batch Processing

Batch processing is an extremely common approach to manufacturing, where a company makes a fixed quantity of identical products before switching over the production processes to a new product type. This minimises the impact of changeover time, while still allowing for changing products. Most batch processing is *sequential batch processing* – the products are batched but are still processed individually. To calculate the time required to process an entire batch on a production machine (T_b), three pieces of information are required:

1. *Set up Time (T_{su}):* The time it takes to set up the production process to process the products in the batch. For example, loading the required CNC program, and adjusting the clamps to hold the work pieces in a milling machine.
2. *Cycle Time (T_c):* Equation 4.2.2, the time to process a single work piece.
3. *Batch Quantity (Q):* The number of items in the batch, after which the process will be changed over for the next batch.

$$Batch\ Processing\ Time = Setup\ Time + (Cycle\ Time\ \times Batch\ Quantity)$$

$$T_b = T_{su} + (T_c Q) \tag{4.2.5}$$

The average time per work piece (T_p) is calculated by dividing the batch processing time (T_b) by the batch quantity (Q):

$$Production\ Time\ per\ Work\ Piece = \frac{Batch\ Processing\ Time}{Batch\ Quantity}$$

$$T_p = \frac{T_b}{Q} \tag{4.2.6}$$

The application of equation 4.2.4 will allow calculation of the production rate per hour.

It can be seen that the larger the batch, the smaller the impact of the setup time, as it is spread out over a larger number of products. Optimisation of the batching strategy is important to minimising the cost of each work piece.

For example, consider a company that is making a number of watch casings as a supplier for a company. They want to manufacture 100 casings. The cycle time is 10 minutes, and the setup time is 120 minutes. The company wants to split the order into two batches to allow them to produce other items between these batches to meet their other order deadlines. Each batch of 50 would therefore have a batch processing time of 620 minutes, and a production rate of 4.84 pc/hr.

Alternatively, if the company organised their scheduling to fit all 100 casings into one batch, the production rate would be 5.35 pc/hr, an improvement of 10.5%

over the two-batch strategy. It's up to the company to determine how best to batch its products while still delivering orders on time.

4.2.2.3 Simultaneous Batch Processing

Not all batch processing needs to occur sequentially like milling. Some batching can occur *simultaneously*, such as heat treating or electroplating components. Calculation of the production rate for a process that enables simultaneous processing is hence performed differently to sequential processing. The same method is used as for sequential batch processing, but replaces equation 4.2.5 with the following equation:

$$Batch\ Processing\ Time = Setup\ Time + Cycle\ Time$$

$$T_b = T_{su} + T_c \tag{4.2.7}$$

As can be seen, the processing time is no longer dependant on the size of the batch, assuming the entire batch can be processed in a single cycle. The batch quantity is still required for equation 4.2.6 however.

Consider the company making batches of watch casings. They electroplate the casings in silver, in batches of up to 50. The electroplating has a cycle time of 60 minutes. The setup time is 20 minutes. The batch processing time would be 80 minutes. And if the process is run at the maximum batch size of 50, the production time per work unit would be just 1.6 minutes, giving a production rate of 37.5 pc/hr. The company may not be able to mill casings fast enough to have a high degree of utilisation for their electroplating process unless they purchase multiple milling machines.

4.2.2.4 Mass Production

Mass production is a situation where a company effectively never stops producing a single product, as demand for the product is sufficiently large that a dedicated production process is financially viable. In this circumstance, the impact of the setup time is negligible. The mass-production rate (R_{mp}) is then simply the number of work pieces that a process can produce per hour:

$$Mass - Production\ Rate = \frac{60}{Cycle\ Time}$$

$$R_{mp} = \frac{60}{T_c} \tag{4.2.8}$$

A major watch manufacturer produces cases for their most popular product continuously, including a milling process. Though the milling process did take time *to set up originally, the time this took divided by the tens of thousands of parts produced since then is a tiny fraction of a second and can be ignored.*

Note that though the setup time can be ignored, the handling times and tooling times in equation 4.2.2 cannot be ignored and will affect the cycle time of the process.

4.2.3 Production Capacity

Whereas section 4.2.2 was concerned with the expected production rate of individual processes and pieces of equipment per hour, this section looks at the overall production capacity that the equipment enables the company to achieve. This represents the maximum number of work pieces that can be manufactured in a time period, such as pieces per day, week, or year.

Understanding the maximum possible production rate you can achieve with a production line or other sequential set of processes is important for several reasons. It ensures jobs are not over allocated to a facility, as this would result in missed deadlines and delays. It also helps understand the utilisation of production processes, and identify underutilised areas where more value could be made.

4.2.3.1 Production Operating Hours

In section 4.2.2 the number of work pieces which can be produced by individual pieces of equipment per hour was calculated. The next step to understanding the production capacity is to understand how many hours a day the production process is operating. Some companies work a single shift on week days. Others can approach 24 hours a day, 7 days a week. Understanding this is the first step to calculating the production capacity.

The production operating hours per year (assuming all shifts are the same length) is calculated by:

$$Hours\ of\ Production\ = Number\ of\ Shifts\ \times Shift\ Length\ \times Days\ per\ Week\ \times Weeks\ per\ Year \tag{4.2.9}$$

For example, a company which operates a single 8-hour shift on weekdays, and operates 50 weeks a year has 2000 hours of production per year. A different company which operates two 8-hour shifts 7 days a week, 50 weeks a year would have 5600 hours of production per year.

To approximate the hours of production per day or per week, divide the production per year by 365 or 52 respectively. The figures may need adjusting if you’re calculating for a period that includes a Christmas break, for example.

4.2.3.2 Simple Production Capacity

In many cases a company has a quantity of machines, and they produce parts at a roughly similar rate. For example, a company has five milling machines which produce watch casings, with each machine producing at a similar production rate. In this case, calculation of the production capacity of the facility can be calculated with the following information:

1. *Number of Machines* (n): The number of similar machines in the company that produce parts at approximately the same rate.
2. *Hours of Production* (H_{pc}): The number of hours over which to calculate the production capacity, calculated with equation 4.2.9. Hours per week or per month can be used here to calculate the production capacity for periods shorter than a year.
3. *Production Rate* (R_p): The number of work pieces each machine produces per hour. This is calculated using the methods in section 4.2.2, specifically equation 4.2.4.

$$Production\ Capacity = Number\ of\ Machines \times Hours\ of\ Production \times Production\ Rate$$

$$PC = nH_{pc}R_p \tag{4.2.10}$$

For example, the company with five similar milling machines operates one, eight-hour shift five days a week, 50 weeks a year. The milling machines have a production rate of 4.84 pc/hr. This company therefore operates 2000 hours a year, and has a maximum yearly production capacity of 48,400 pc/year.

4.2.3.3 Advanced Production Capacity

For situations where different pieces of equipment operate at different production rates (PR), a modification to equation 4.2.10 is used. Instead, the production rate of each individual machine needs to be considered separately:

- *Production Rate of Machine i* (R_{pi})*:* for a set of *n* machines, the production rate of a specific one. This is calculated using the methods in section 4.2.2, specifically equation 4.2.4.

$$\begin{aligned} Production\ Capacity \\ &= Hours\ of\ Production \\ &\times (PR\ of\ Machine\ 1 + PR\ of\ Machine\ 2\ etc) \end{aligned}$$

$$PC = H_{pc} \sum_{i=1}^{n} R_{pi} \tag{4.2.11}$$

Consider a company with three milling machines. They all produce watch casings, but operate at different speeds due to being different models from different manufacturers. The machines are numbered, their production rates are calculated individually as per section 4.2.1, and the results are put in a table below:

Machine Number	Machine Name	Production Rate
1	Faithful Workhorse	4 pc/hr
2	Cheap and Cheerful	3 pc/hr
3	State of the Art	6 pc/hr

Table 4.2-1 The milling machines available to the watch casing manufacturer.

The company works 2000 hours per year. Their weekly production capacity is hence:

$$Weekly\ Production\ Capacity = \frac{Yearly\ Production\ Hours}{Weeks\ in\ a\ Year} \times \left(R_{p1} + R_{p2} + R_{p3}\right)$$

$$Weekly\ PC = \frac{2000}{52} \times (4 + 3 + 6)$$

$$Weekly\ PC = 38.46 \times 13$$

$$Weekly\ PC = 500\ pc/week$$

4.2.4 Capacity Insights

Gathering data and processing the calculations for individual production stations is simply the first step in analysing manufacturing processes. It’s important to look at the results and understand what it’s telling you. This will enable a manufacturing engineer to make better decisions about their business, improving productivity and profitability. Of particular concern is matching production capacity to the capacity required. Failure to meet the required production capacity will cause order backlogs. Having unused production capacity represents unutilised equipment which could otherwise be producing value for the company.

A company cannot make more pieces per time period than its calculated production capacity. If the order book requires rates higher than your production capacity, the company needs to increase the production capacity or risk delays to delivery times. Similarly, making fewer pieces than the production capacity implies the company could be generating more value. If equipment is sitting idle, it is not generating as much revenue as it could be.

There are many ways to adjust production capacity up or down as required, some which can be short term considerations, and some which are longer term. It’s important to consider the time scales involved. In this section, we describe some options to increase capacity or to mitigate overcapacity in the short, medium, and long term.

4.2.4.1 Increasing Capacity

When an enterprise needs to produce more work pieces than is has the capacity to deliver, there are a number of options available. Which option is chosen will depend on how long the enterprise expects to be over capacity for.

- *Increase hours worked per shift* [Short Term]: For short-term capacity issues, one of the simplest ways is to increase production capacity is to ask existing workers to work overtime on their existing working days. This will increase labour costs, especially if the company needs to offer improved hourly pay to incentivise working overtime, but this is quick and simple to implement.
- *Repurpose existing equipment* [Short Term]: Where an enterprise has multiple production lines and makes multiple products, changing some equipment from one process to another is a way to boost production capacity. The availability of tooling to do this, or the time required to reconfigure and/or reprogram the equipment will determine the speed of this approach, but it is often short compared to other approaches. This can also include moving workers from one line to another.
- *Backlog orders* [Short Term]: Depending on the nature of the orders and the enterprise's relationship with the customer, deliberately delaying orders during short periods of overcapacity may be less financially damaging than adding more workers or equipment. The effect on the company's reputation must be considered.
- *Subcontract out work* [Short Term]: If the period of overcapacity is expected to be short, and delaying the delivery of products is not an option, subcontracting some of the work to other companies may be a solution. This can help fix bottlenecks in the manufacturing process and improve overall production capacity, but beware of increased costs and overheads related to organising the subcontract.
- *Increase the number of shifts per day/week* [Medium Term]: If overcapacity issues are likely to continue beyond the short term, establishing an additional shift to maximise machine processing times may be an option.
- *Increase production rate of bottleneck processes* [Medium Term]: The limiting process of a manufacturing line is the bottleneck. Improving the production rate of that process will improve the production rate of the whole production line. This could involve retraining operators, optimising CNC programs, improved tooling, or other technical improvements.
- *Purchase additional equipment* [Long Term]: If there is a real opportunity to increase revenues by increasing the production capacity, acquisition of more equipment (or more manual workers) may be the ideal choice. Be aware of the potentially long lead times on equipment, and that it will be difficult to get optimised output from new equipment until employees are experienced in its use.

Focus equipment acquisition on bottleneck processes, as these will enable higher production capacities.
- *Redesign manufacturing process* [Long Term]: If a product line has been produced for a long period of time without changing the way it is manufactured, possible efficiencies could be made by changing the manufacturing process and taking advantage of new experience and equipment. Also consider whether some degree of redesigning the product could improve production rates.

4.2.5 Mitigating Unused Capacity

Where far more production capacity exists than is being utilised, cost savings could be made by reducing the production capacity of a production facility or line. Alternatively, strategies to make use of idle capacity could be adopted, creating value from your assets.

- *Repurpose existing equipment* [Short Term]: Where an enterprise has multiple production lines and makes multiple products, if demand for a product is low, consider re-using equipment for other products where demand is higher. This can also include moving workers from one line to another.
- *Stockpile inventory* [Short Term]: If the under-capacity is temporary, and the enterprise makes products that they know will continue to sell in the future, is it possible to use unused capacity to stockpile inventory which will help smooth out overcapacity periods later. This is effectively gambling against the demand for the products, as the stockpiled inventory has no value until sold.
- *Reduce the number of shifts per week* [Short Term]: One of the simplest (and least popular) ways to address unused labour capacity is to reduce the workforce. This could mean implementing redundancy for workers, but an alternative is to reduce the number of shifts. Workers may accept a move from a five day week to a four day week if it means surviving a short-term period with a low number of orders without laying anyone off permanently.
- *Take on extra work* [Medium Term]: If production capacity is unused, consider offering it as subcontracted capacity to other companies. The ability to do this relies on the relationship with other manufacturers and the nature of the spare capacity, but taking on additional work can create value from otherwise idle equipment.
- *Sell equipment* [Long Term]: If equipment is unlikely to be made use of in the medium to long term, it may be worthwhile to claim back some of the value of the equipment by selling it, which can often be a significant return. The ability to sell equipment, and the price it sells for depends on the demand for that equipment. Carefully consider expected future requirements and the value of the equipment before committing to sell.

4.2.5.1 Bottlenecks

Many manufacturing plants do not manufacture single-stage products such as the watch casing discussed in the previous sections, where the casings only have one process applied to them. Instead, multiple operations are executed in sequence to produce the part, forming a production line. A production line may not be a physical entity on the shop floor, but a process followed using multiple distributed pieces of hardware instead, but the approach is the same.

However, calculating the production capacity of a manufacturing line is simplified by a single problem: in almost all manufacturing lines there is a bottleneck. The bottleneck is the process that limits the production capacity, by producing slower than any other process. By calculating the production rates of every process, you can identify the bottleneck as the machine (or set of machines) with the lowest summed production rate. The calculation of the production capacity of the entire manufacturing line is then simply the production capacity of the bottleneck process.

For example, the watch manufacturer mills watch casings with 2 milling machines at a rate of 5 pc/hr per machine, batch electroplates them at a rate of 20 pc/hr, and polishes the casings at a rate of 6 pc/hr. The two milling machines can together produce 10 cases per hour, the electroplating process can electroplate 20 per hour, but the bottleneck is the polishing process. You can increase the production rates of the milling and the electroplating, but you'll just end up with bigger piles of unpolished cases.

Informally, bottlenecks are often easily identified by looking at what equipment is constantly in use, and which equipment typically has large queues of work waiting in front of them. However, formally understanding queues and buffers in a manufacturing system can give insight into how best to arrange your manufacturing processes and at what speed to run them. The next section discusses queueing analysis, and it is highly applicable to manufacturing system optimisation. Manufacturing systems are connected series of systems, either as connected manufacturing processes to form a production line, or as connected suppliers in a supply chain to produce a large or complex product. Understanding how products move and wait in these systems will aid in analysis where optimisations can be made to improve the flow of materials in the system.

4.3 Queueing Analysis

4.3.1 Introduction to Queueing Theory

Queue (noun, British): a line or sequence of people or vehicles awaiting their turn to be attended to or to proceed.

- Oxford University Press

Buffer (noun): a means or device used as a cushion against the shock of fluctuations in business or financial activity.

- Merriam-Webster.com. 2011

Manufacturing typically involves multiple steps as parts flow between machining stations and assembly processes on the shop floor. Unless you're using a pulse flow production line where everything moves between stations in sync it is inevitable that parts will arrive at stations before the station is ready, or stations will be ready before parts are available. Using and understanding queues and the flow of parts will aid in understanding and optimising the production line.

The manufacturing process involves multiple operational steps, converting raw materials into finished products. In order to make the process efficient (e.g. maximising the production rate of a line) and cost effective, analytical tools such as *queueing theory* have been used extensively. They play an important role in the performance analysis, design, planning and control of manufacturing processes, as parts and products in queues are not generating value. Queuing theory is a branch of mathematics that studies and models the act of waiting in lines, and originated in the analysis of telecommunication exchanges handling phone calls. Queuing theory has been successfully applied in modelling production lines to study performance. Queueing Theory requires simplifying the processes in your manufacturing system and modelling them so you can apply formulas to calculate performance measures. By understanding what affects the performance of queues, you can start to record that information, and begin to gather insights into how changes to your production system could affect queueing performance, and therefore overall system performance.

In order to understand how this theory works, first some basic components and characteristics of a manufacturing production line and its queues are first defined. Manufacturing systems have different product-flow layouts that are classified depending on the method of part transfer and the number of part types produced by the system. Part transfer can be carried out in three different ways:

1. *Synchronous* (also known as transfer lines): Where parts are transported simultaneously between workstations.

2. *Asynchronous* (also known as production lines): Where each part moves independently from other parts.
3. *Continuous:* Where parts move continuously at a constant speed.

This product-flow is a waiting line or *queue*, where a sequence of objects (in this case manufactured parts) are waiting to be processed. Owing to the origins of queueing theory lying in human queueing analysis, these objects are generally referred to as *customers*, and the process being carried out on them is referred as a *service*. The term *buffer* is often frequently used in the manufacturing context. A queue includes both a buffer (the 'waiting area') and the service that feeds from the buffer, which in a manufacturing context will mean the production process or station. In practise, the terms are often used interchangeably. Figure 4.3-1 shows the general product-flow layout of mass manufacturing systems.

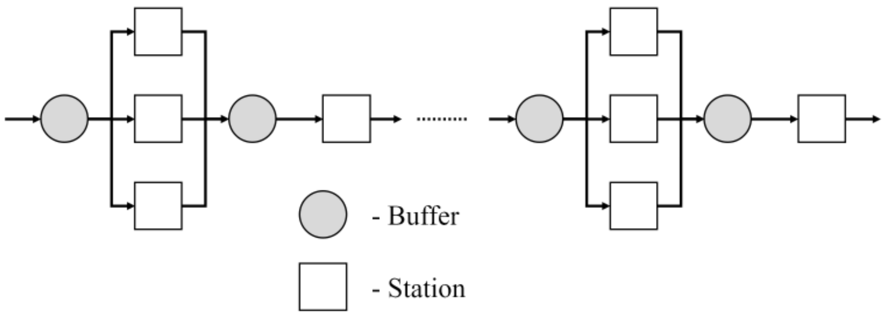


Figure 4.3-1 Product flow layout of mass production systems [5]. It can be seen that when there are parallel stations, those stations can take parts from a single shared buffer.

Production lines are used to produce parts which have a high-volume turnover, and they are characterised by a product-flow layout, low product flexibility (the line is restricted to producing a small variety of part types) and asynchronous part transfer. *Blocking* and *starving* of parts are the main causes of inefficiency in production lines.

- *Blocking:* A part is waiting to be processed, but cannot as the required machine is being used by another part.
- *Starving:* A machine is idle, as it has no input parts to start processing.

These phenomena are mainly caused by variable processing times, and by disruptions to the line caused by the unreliability of stations. To increase the efficiency of the lines, queues are placed between stations.

4.3.2 Queueing Analysis

The impact of blocking and starving on the productivity of a production line with a high production volume is significant. Understanding the queueing in your system will also help identify bottlenecks and inefficiencies as targets for specific processes to be improved. In many cases, simply being aware these problems exist can go a long way to intuitively solving issues. However, more detailed analysis can enable the identification of subtler problems and implementing broader optimisation approaches. To do this, you must model your manufacturing system.

4.3.2.1 Modelling the Problem

Model (noun): a simplified description, especially a mathematical one, of a system or process, to assist calculations and predictions.

- Oxford University Press

In order to understand the behaviour of a production system, it can be analysed as a stochastic process (a process with random elements to it). The main interest concerns the distribution of the number of jobs in the system at an arbitrary point in time. From this distribution, it is possible to define how the number of jobs in the system fluctuates, which will allow the computation of important performance characteristics such as the mean number of jobs in the system. The first step for modelling a production line is to characterise the system. The system has different features:

- First, the *arrival* process of the parts is characterised, defining how the parts arrive in the system.
- Then the *service* duration is characterised, which is the time the part is at the station where some operations are performed.
- It is also necessary to specify the *number of stations*, in which several parts could be processed in parallel.
- Parts wait in the *buffer* if all stations are busy, so the total number of parts in the system is estimated by including both the ones being served as well as the ones that are waiting.
- Finally, the *scheduling policy* needs to be specified. This determines in what order parts in the buffer are released for processing.

Common scheduling policies (also called service disciplines) include:

- *First in first out (FIFO):* First come, first served; the type of queue you are intuitively familiar with. Like a queue of customers in a shop, in manufacturing the part which has been waiting the longest is served next.

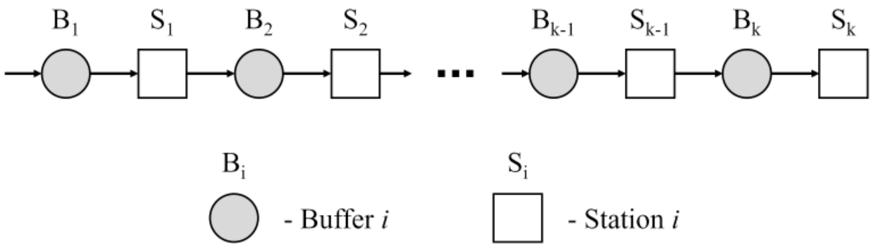


Figure 4.3-2 The basic queueing network model of a production line [5].

- *Last in first out (LIFO):* Imagine a stack of trays in a canteen. The mostly recently replaced tray is placed on the top of the pile, and is also the next tray that will be taken. Sometimes used in manufacturing where the parts are buffered in a stack.
- *Priority:* Where not all parts are the same, parts can be allocated a priority. Products that need to be shipped earlier could be assigned a higher priority, and the next part to be processed is the part with the highest priority assigned.
- *Shortest processing time:* The part that will take the least time to process is used next. This is sometimes used when buffers are reaching their limit and space needs to be cleared.

Figure 4.3-2 shows a common basic model of a production line. The production line consists of *k* stations arranged in series. Each station (S_i) has a buffer (B_i) preceding it. The buffer before the first station may be finite or infinite, all inter-station buffers are finite. Parts enter the system at station 1 and pass through all stations in order. At each station, an operation is performed on each of the parts by a single machine. The parts leave the final station (S_k) in finished form.

The common underlying assumptions when modelling a production line as a queueing network are:

- The line is operated at steady-state conditions (conditions always remain constant through all the production line).
- All random variables are independent.
- All the transport times between stations are zero.
- All failures are single-machine failures and they are operation-dependant (they can only fail while they are operating).
- No parts are scrapped.
- Only a single part type is modelled.
- All buffers utilise the FIFO policy.
- There are enough repair personnel.

4.3.2.2 Part Arrivals

To model the production line, the continuous arrival of the parts at a given workstation needs to be defined. In simple models, it is assumed that the successive inter-arrival times, U_1, U_2, U_3 ... U_i between parts are mutually independent and that they follow the same probability distribution.

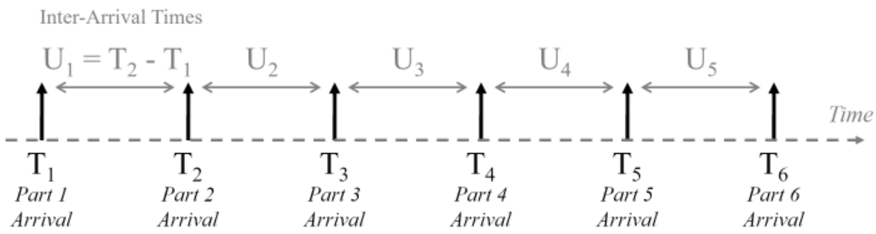


Figure 4.3-3 Part arrivals. T_i is the arrival time for Part *i*. U_i is the time between the arrival of two parts $U_i = T_{i+1}$-T_i

The arrival process of a queuing system is often modelled as a Poisson process. In this process, the inter-arrival times U_1, U_2, U_3… of each part are independent, and arrive according to an exponential distribution with mean λ. If you are familiar with normal distributions or "bell curves", a Poisson process is similar but is discrete (i.e. you can't have fractions of objects arriving). It is described with a single variable – the average number of events per time unit, in this case the number of part arrivals.

If variance is important, the probability that the inter-arrival time U_i is greater than a given value u is equal to $exp(-\lambda u)$. In this case, when the inter-arrival times are independent and identically distributed according to an exponential distribution with parameter λ, the arrival process is said to be a Poisson process of rate λ, where the arrival rate λ is the average number of arrivals per time unit.

For example, λ is typically expressed as the average number of arrivals per time unit. Where 12 parts arrive in a queue per hour, λ = 12/h. Ensure when calculating queue behaviour that all time quantities are expressed for the same time unit e.g. per hour.

4.3.2.3 Service Duration

In order to characterise the system, it is also necessary to properly define the service duration, which refers to the time it takes for the station to perform an operation on the part. The *time of service* is the elapsed time between beginning of service and departure, independently of any waiting time in the queue. For a production process, the time of service is equivalent to the cycle time. In general, it is assumed that:

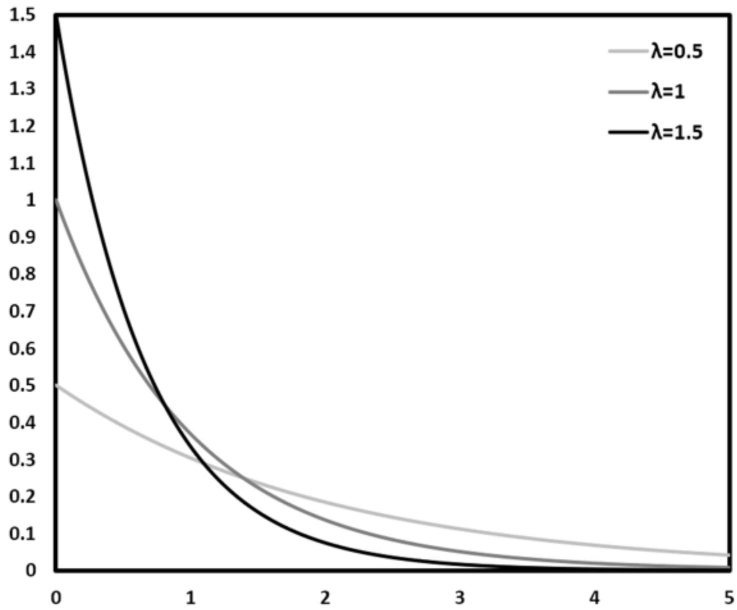


Figure 4.3-4 Three examples of exponential distributions with varying means. In all cases, it can be seen that shorter arrival times are much more likely than longer ones, but occasionally long arrival times will occur – a part has arrived late to a station. These rarer events are typically the ones that cause problems with queueing systems.

- Service times are independent and identically distributed.
- Service times are characterised by their probability distribution.

Much like part arrival times, the exponential distribution is a common model of service duration. It is expressed as the number of processed parts per time unit. In this case it is assumed that if S_i denotes the service duration for a part *i*, then the probability that S_i is greater than a given value *s* equals $exp(-\mu s)$, where μ is the mean of the exponential distribution and *s* is a duration.

A few metrics can be calculated with this information. The *service rate* is equal to μ and is the average number of parts processed (i.e. served) per time unit if the machine (the server) is always busy e.g. parts per minute.

For example, if a machine can process a part every 4 minutes, μ = 15/h (i.e. 15 processed parts per hour). Remember when calculating queue behaviour that all time quantities are expressed for the same time unit.

The *offered load* (sometimes called traffic intensity) is another important metric, representing the expected amount of traffic at a station. The offered load ρ is defined as:

$$\rho = \frac{\lambda}{\mu} \tag{4.3.1}$$

where λ is the average number of arrivals per time unit and μ is the average number of parts a station is able to handle if it is always busy. Offered load is the average proportion of time a server will be occupied, and is an important value for calculating many useful metrics. Offered load must be less than one. If it's greater than one, the station can't process parts fast enough, and the queue will continue to grow forever.

With our example, $\lambda = 12/h$ and $\mu = 15/h$ so the offered load is 0.8.

4.3.2.4 The M/M/1 Queue

Different classes of queues are defined using Kendall's Notation, which defines the features of the queue necessary to perform queueing analysis. In Kendall's Notation, a queue is described with 5 different parameters *A/B/c/K/Z* where the value for each parameter describes the following:

- *A* is the inter-arrival time distribution. M is for Markovian (i.e. exponential as discussed previously), D is for Deterministic (constant), and G is for General Distribution (i.e. an unknown distribution). Other values exist for less common distributions.
- *B* is the service time distribution, and can generally have the same values as the inter-arrival distribution.
- *c* is the number of servers that take parts from the queue.
- *K* is the system's capacity i.e. the maximum length of the queue, plus the number of servers. For this reason it is sometimes written as *K+c*. If the value is omitted, the queue is infinite.
- *Z* is the service discipline e.g. FIFO, LIFO, Priority. Where this is left blank, the discipline is assumed to be FIFO.

The simplest type of queue and the one discussed in this section is the M/M/1 queue, which is a queue with Markovian inter-arrival and service durations, a single server to process parts, and a queue with no maximum length. An M/M/1 queue could be more fully written as M/M/1/∞/FIFO.

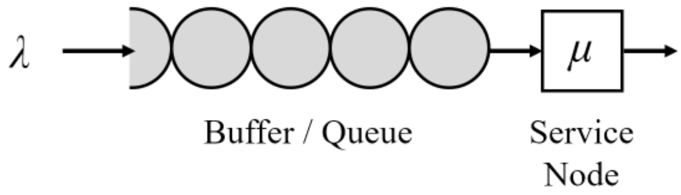


Figure 4.3-5 An example M/M/1 queue.

Although in reality a production system would never be as simple as this model, it contains most of the essential characteristics of a production system and shows the basic ideas and methods of Queuing Theory. By simplifying the model, it is possible to create "good enough" conclusions without the process to arrive at the conclusions being so complex as to no longer be useful.

The analysis of the system consists of studying the evolution of the value $N(t)$, which refers to the number of parts in the system N at time t. "Part in the system" is the number of parts in the queue, plus the parts actively being processed. The value of $N(t)$ can change in two different ways during a time period, representing a transition from $N(t)$ to $N(t+\Delta t)$. If there are n parts in the system, then the following changes can happen:

- If an arrival occurs, the state of the system increases from n to $n+1$. The rate of increase is represented by λ, the arrival rate.
- If a process is completed, the state of the system decreases from n to $n-1$. The rate of decrease is represented by μ, the service rate.

Predicting the value of $N(t)$ allows you to predict the average queue length of the modelled queue, as well as the proportion of the time the queue is over a specific length. Although an M/M/1 queue has an "infinite" queue to simplify the calculations, the real system being modelled will have a limit you don't want to exceed. The transition from one state to another state can take place as shown in Figure 4.3-6.

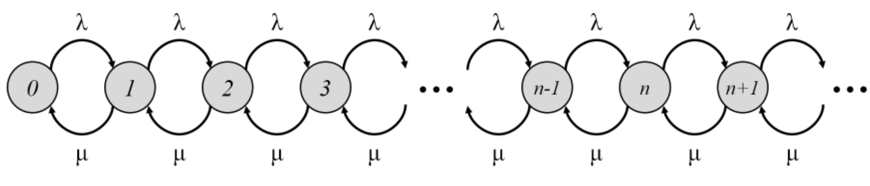


Figure 4.3-6 States of an M/M/1 queue.

4.3.3 Event Probabilities and Performance Measures

Having modelled the queue, you can start to calculate the probabilities of events occurring, such as the average number of parts in the queue, and the average time a part spends in the queue. This will enable identifying issues with your existing queueing system, where buffers may be too large or too small, or where a buffer is at risk of exceeding its capacity. The assumptions made about the system (i.e. Poisson arrivals, exponential processing times, FIFO) make it possible to describe the state of the system at an arbitrary point in time by simply specifying the number of parts in the system. Without these assumptions, the state description would be very complicated and would have to contain not only the number of parts in the system, but also, for example, the residual processing time of the part in service.

At any given point, if the system is in the state *n*, which is the number of parts in the queueing system (including in the service node(s) itself)*:*

- The state of the system moves from $n\text{-}1$ to *n* at the rate of $(P_{n-1})(\lambda)$
- The state of the system moves from *n* to $n\text{-}1$ at the rate of $(P_n)(\mu)$

Where P_{n-1} and P_n are the probabilities of being in states *n-1* and *n* respectively. Assuming the system is in a steady-state:

$$(P_n) = (\rho)(P_{n-1}) \tag{4.3.2}$$

Where ρ is the offered load discussed in section 4.3.2.3. Although queues are dynamic and changing systems, by computing the probabilities of being in each state a number of performance measure can be calculated to give insight into the queue and its expected behaviour.

4.3.3.1 Number of Parts in System Probabilities

To calculate the performance measures of the M/M/1 model, first the probability of having *n* parts in the system is defined. If we know the offered load as calculated in section 4.3.2.3, the odds of the system being empty and idle is:

$$P_0 = 1 - \rho \tag{4.3.3}$$

From equation 4.3.2, the odds of having one part in the system is:

$$P_1 = \rho P_0 \tag{4.3.4}$$

That is to say, the odds of having one part in the system is equal to the odds of having zero parts multiplied by the offered load. Then, similarly the odds of having two parts in the system is the odds of having one part multiplied by the offered load:

$$P_2 = \rho P_1 \tag{4.3.5}$$

And by extension the odds of having n parts in the system is calculated as:

$$P_n = \rho^n P_0 \tag{4.3.6}$$

As an M/M/1 queue can only process one part at a time, the queue length is *n-1* (to a minimum of zero). Though an M/M/1 queue is assumed to be infinite, in reality there is likely to be a practical limit above which significant inconvenience or extra cost will occur. The odds of the system having *n* or more parts in it is equal to:

$$P_{n\ or\ more} = \rho^n \tag{4.3.7}$$

The result will be the proportion of the time the system spends with a queue equal to or in excess of length *n-1* (as one part will be in the station being processed).

4.3.3.2 Performance Measures

With the aid of the formulas from the previous sub-section, the following performance measures can be calculated:

- Average number of parts in the queue (L_q).
- Average number of parts in the system i.e. the queue and the station (L_s).
- Average time a part spends in the queue (W_q).
- Average time a part spends in the system i.e. queueing time plus processing time (W_s).

The average number of parts in the entire system (i.e. both the queue and the processing station) L_s is calculated with the offered load ρ, as the offered load represents the amount of 'traffic' in the system:

$$L_s = \frac{\rho}{1-\rho} \tag{4.3.8}$$

The average number of parts in the queue L_q is calculated by multiplying the average number of parts in the system L_s by the offered load ρ:

$$L_q = \rho L_s \tag{4.3.9}$$

Which is equivalent to:

$$L_q = \frac{\rho^2}{1-\rho} \tag{4.3.10}$$

An important law in Queueing Theory is Little's Law, and it states that the number of parts in the system L_s is equal to the arrival rate λ multiplied by the time the part spends in the system W_s:

$$L_s = \lambda W_s \tag{4.3.11}$$

L_s and λ, this can be rearranged as:

$$W_s = \frac{L_s}{\lambda} \tag{4.3.12}$$

As L_s is calculated with the offered load ρ, and ρ depends on the arrival rate λ and service rate μ, this can be simplified even further to be:

$$W_s = \frac{1}{\mu - \lambda} \tag{4.3.13}$$

Note that the time unit of W_s is the same as the time unit of λ and μ. *If* $\lambda = 12$/h and $\mu = 15$/h then $W_s = 0.333$ hours, or 20 minutes.

Lastly, the time a part spends waiting in the queue W_q is the total time in the system, minus the processing time. The average processing time is equal to $1/\mu$ and therefore the average waiting time is simply:

$$W_q = W_s - \frac{1}{\mu} \tag{4.3.14}$$

$$W_q = \frac{1}{\mu - \lambda} - \frac{1}{\mu} \tag{4.3.15}$$

4.3.3.3 Queue Performance Example

Queueing Theory is an extremely effective tool for understanding networks of connected processes, but does require some practise and thought to understand. This section shows a worked example of a queueing theory applied to an example process, which should help you understand how to apply the formulas.

For a machining station on the watch casing manufacturing line, the average rate of arrival of parts is 10 per hour. On average, the station can process parts with a rate of one part every five minutes. Assume the arrival of parts follows a Poisson distribution and the processing of parts at the station follows an exponential distribution. Find the average number of parts waiting in the queue and the average number of parts in the system. Find the average waiting time in the queue and the overall time of a part in the system. Find the odds of the queue exceeding ten parts in length, as this is the longest the queue can be without blocking the previous process.

Poisson and exponential distributions may sound complex, but if your inter-process and service times follow a bell curve (so the average time is common, and much higher or lower times are uncommon) and you can calculate an average time, then they probably fall into this category. It is the variability in these two values that makes queueing theory important. If part arrival times and service times were constant, queues would be entirely predictable. However, processing is rarely entirely predictable, and even a short run of short part arrival times combined with long service times could overwhelm a buffering system.

Poisson arrivals, exponential service and a single station means this example follows an M/M/1 model and we can use the formulas we have learned. Before we can calculate the performance measures, we need to calculate the arrival and service rates, and the offered load.

- The part arrival rate is λ = 10/h

The part arrival rate is how often parts arrive at the processing station, on average. It is unlikely the parts will arrive exactly every 6 minutes, but instead some parts will arrive slightly faster, some will arrive slightly slower, and more rarely there may be more significant variation. But the average is ten parts per hour. This is a number you will have to measure from your actual production line.

- The part service rate is μ = 1 in 5 minutes = 12/h

The part service rate, is how fast the processing station can process parts, and is also called the cycle time. Like the part arrival rates, here we give an average but each specific processing time can fluctuate. Again, this is something you will have to measure from the processing station itself, remembering to include factors such as the loading and unloading of parts, and any tooling time (see section 4.2.2 for some considerations here). If your processing station is able to process parts in a constant time i.e. exactly five minutes for each part, the formulas will be slightly different. See the next section for information on M/D/1 queues.

- The offered load ρ is therefore:

$$\rho = {}^{\lambda}/_{\mu} = {}^{10}/_{12} = \mathbf{0.833}$$

The offered load is the part arrival rate divided per the service rate, and is a measure of how 'busy' the processing station is. It is an important value for the remaining calculations. Note that the offered load must be lower than 1, otherwise the processing station will be unable to keep up with demand.

- The average number of parts in the system L_s is:

$$L_s = {}^{\rho}/_{1-\rho} = {}^{0.833}/_{1-0.8333} = \mathbf{5}$$

The "queueing system" referred to here includes the processing station and the buffer before it. A single part in the system implies the part is being processed and *the buffer is empty. 5 parts here implies that on average there is one part being processed and 4 parts in the buffer.*

- The average number of parts waiting in the queue L_q is:

$$L_q = {\rho^2}/{1-\rho} = {0.8333^2}/{1-0.8333} = {0.694}/{0.167} = \mathbf{4.156}$$

It may seem counter intuitive that the average number of parts waiting is not just the average number of parts in the system minus one, but this is due to parts arriving at variable times during the part processing operation (called "residual time"), combined with the queue sometimes being empty.

- The average time a part spends in the system W_s is:

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{12-10} = \mathbf{0.5\ \mathit{hours}}\ (30\ minutes)$$

The total time a part spends in the queue is the combination of the waiting time, and the processing time. This will of course vary considerably as the length of the queue fluctuates.

- The average waiting time of a part in the queue W_q is:

$$W_q = \frac{1}{\mu - \lambda} - \frac{1}{\mu} = \frac{1}{12-10} - \frac{1}{12} = \mathbf{0.416\ \mathit{hours}}\ (25\ minutes)$$

As stated previously, the time a part spends in the system is the waiting time plus the processing time. So the time the part spends waiting is just the total system time minus the average processing time.

- The probability of the queue equalling or exceeding 10 parts is:

$$P_{11\ or\ more} = \rho^{11} = 0.8333^{11} = \mathbf{0.135}$$

The M/M/1 queue is assumed to have an infinite length buffer as this significantly simplifies the mathematics behind it. This is obviously not a realistic assumption for real systems however, so a common application of queueing theory is to calculate how often the maximum queue length is exceeded. Remember that P_n is the number of parts in the entire system (including being processed) not just the queue. Hence the probability of the queue exceeding 10 is the probability of the number of parts in the system exceeding 11.

The consequences of a queue exceeding its maximum limit can vary significantly, but a common issue would be that the previous production station in the production *line is "blocked" i.e. cannot output its parts as there is nowhere for them to go. This can negatively affect productivity, so understanding if this is going to be a regular occurrence or highly infrequent is important. Conversely, an oversized buffer can be expensive or take up a lot of space, particularly for larger parts and products. If the calculations show the system will never require such a large buffer, the buffer can be reduced in size or never purchased in the first place.*

4.3.4 Queueing Theory Conclusions

As Kendall's notation implies, M/M/1 queues are a single type in a huge range of possibilities, but are often discussed as they are both simple to calculate metrics for, and encompass a large number of real-world queueing systems with only a few simplifying assumptions. However, they do not cover all queueing types. For example, where the service duration is a known constant (for example, an automated CNC milling operation with a fixed program) you would have an M/D/1 queue. A buffer that feeds into three similar processing stations could be an M/M/3 queue.

For each of these types of queues formulas exist to calculate the performance measures. For example, the average number of parts in the system for an M/M/1 queue is given by equation 4.3.8:

$$L_s = \frac{\rho}{1-\rho} \tag{4.3.8}$$

By comparison, the average number of waiting parts in an M/M/c queue (i.e. an M/M queue with *c* servers) is given by:

$$L_s = \frac{\rho}{1-\rho} C\left(c, {}^{\lambda}/_{\mu}\right) + c\rho \tag{4.3.16}$$

where $C\left(c, {}^{\lambda}/_{\mu}\right)$ is equal to:

$$C\left(c, {}^{\lambda}/_{\mu}\right) = \frac{1}{1=(1-\rho)\left(\frac{c!}{(c\rho)^c}\right)\sum_{k=0}^{c-1}\frac{(c\rho)^k}{k!}} \tag{4.3.17}$$

which is quite clearly a much more complex situation!

This is where computer-based simulation and modelling packages become important – by hiding and solving the complexity for the users, and enabling the calculation of far more complex situations with a lower risk of error and less time investment on behalf of the user.

4.4 Conclusions

Manufacturing analysis is a combination of several aspects – understanding how to formally approach the decision making process, what the question being asked really is in terms of KPIs, gathering data to inform the analysis process, and finally performing calculation and modelling to find an answer to the question. Understanding the process is as important as the formulas and models themselves, as once the process is understood the relevant formulas can be looked up and implemented.

However, as shown at the end of the Queueing Theory section, seemingly simple calculations for manufacturing analysis can rapidly become more complex as the situation being analysed expands beyond small examples. Similarly for manufacturing capacity and production rate analysis, the simpler the situation, the simpler the mathematics required. However, real manufacturing systems are rarely so simple, and the calculations required can rapidly become complex, time consuming, and prone to error.

Though it is important to understand the basics of mathematical basis of manufacturing systems analysis, there exists a substantial range of computer tools to aid in analysis and decision making, removing or hiding much of the complexity and allowing the user to be more accurate, productive, and more able to respond to change. The decision-making process, understand and correctly applying KPIs, and understanding what a tool is suited for and its limitations are all as important for computer-based tools as they are for manual calculations.

The next chapter (Chapter 5) will focus on modelling and simulation computer-based digital tools which are offline i.e. not connected in real-time to a manufacturing system. Chapter 6 will discuss the developing area of digital twins, which are modelling tools connected directly to physical systems and update in real time. It will also discuss decision support systems, which are tools to directly aid in the formal decision-making process.

4.5 References

[1] R. H. Hayes and S. C. Wheelwright, Restoring our competitive edge: competing through manufacturing (Vol. 8), New York: Wiley, 1984.

[2] H. A. Simon, The new science of management decision, Harper & Brothers, 1960.

[3] M. Davidson, "28 Manufacturing Metrics that Actually Matter (The Ones We Rely On)," LNS Research, 9th October 2013. [Online]. Available: https://blog.lnsresearch.com/blog/bid/188295/28-manufacturing-metrics-that-actually-matter-the-ones-we-rely-on.

[4] International Organization for Standardization (ISO), *ISO 22400-1:2014 Automation systems and integration — Key performance indicators (KPIs) for manufacturing operations management — Part 1: Overview, concepts and terminology*, 2014.

[5] H. T. Papadopolous, C. Heavey and J. Browne, Queueing theory in manufacturing systems analysis and design., Springer Science & Business Media, 1992.

Chapter 5
Digital Modelling and Simulation of Manufacturing Systems

Jack C Chaplin and Giovanna Martinez-Arellano

5.1 Introduction

Digital (adjective): involving or relating to the use of computer technology.

Simulation (noun): the production of a computer model of something, especially for the purpose of study.

- Oxford University Press

In the previous chapter, we introduced concepts of analysis and modelling manufacturing systems, focusing on productivity and capacity analysis, followed by queue modelling to improve throughout in a larger system. Both these techniques require the use of mathematics to evaluate data gathered from the real system. However, manual mathematical processing can rapidly get extremely complex for anything other than the simplest examples of manufacturing systems. Fortunately, computer-based digital modelling and simulation will allow you to glean the insights you require, while simplifying the process significantly.

Modelling tools and software help engineers be more productive in the process of conceiving, designing, modelling, evaluating and planning the implementation

J.C. Chaplin (✉), G. Martinez-Arellano
Institute for Advanced Manufacturing, University of Nottingham, Nottingham, UK
e-mail: jack.chaplin@nottingham.ac.uk, giovanna.martinez@nottingham.ac.uk

J.C. Chaplin et al (eds), *Digital Manufacturing for SMEs*
DOI: https://doi.org/10.17639/nfxp-ve75

of manufacturing systems. Modelling and simulating the manufacturing processes and the wider production line enables past experience to be utilised to predict the performance, and to detect issues and improve productivity in a virtual model before the expensive process of deploying a manufacturing process to the shop floor.

Modelling and simulation tools allow manufacturing companies to test changes and improvements to their manufacturing systems before physical implementation. In many cases decisions regarding manufacturing system changes are often based on experience and intuition rather than on quantitative prediction. Though this can often work, it relies on the availability of human expertise and is prone to error.

Creating modelling and simulation tools is a highly specialised field, requiring expertise in probability, statistics, and scientific computing. Fortunately, the wide range of commercial solutions means enterprises can take advantage of these tools off-the-shelf. Despite this, it is important to understand the underlying techniques used by simulations and models, as it will influence which one is best applied to which area. Often, multiple simulation packages may be required to give you coverage of both the more detailed areas of your system such as machine performance, as well as the larger more strategic areas like stock levels.

Taking a step backwards however, we will first discuss data capture in section 5.2 Data Capture and Analysis. As discussed in section 4.1, the purpose of analysis is to transform data into knowledge. Modelling and simulation are simply additional tools for analysis, so also exist to transform data into knowledge. However, your ability to accurately model your manufacturing systems relies on access to accurate data. Without accurate data, the insights generated by the models and simulationswill also not be accurate.

Section 5.2 Data Capture and Analysis discusses the entire data processing cycle from data collection, through analysis and to decision making. Following this, section 5.3 Modelling and Simulation Approaches discusses the range of computer-aided tools available for modelling and simulating manufacturing systems, the underlying priciples and assumptions, and example applications of their use.

5.2 Data Capture and Analysis

5.2.1 The Data Processing Cycle

Data (noun): facts and statistics collected together for reference or analysis.

- Oxford University Press

Analysis of data is often assumed to be taking collected data and running it through a mathematical formula or model to find an insight or result that helps you understand the subject of the analysis better. And in the first instance this is true. However, for more complex data analysis when systems are larger, changing, and producing much higher volumes of data, there are additional steps that need to be considered to ensure the analysis is accurate and useful.

The Data Processing Cycle is a series of steps to extract information from data in a structured and organised manner. It represents a series of steps from the collection of data to the disposal of it. However, the information gathered in the cycle may inform what new data is required, turning it from a process into a cycle. There are different life cycle models available in the literature, depending on the application. However, as shown in Figure 5.2-1 the general steps are:

1. *Data Collection:* Gathering the data from sensors or manufacturing processes and centralising it for processing.
2. *Data Pre-processing:* Initial cleaning and refinement of data to solve issues that might impede the analysis process.
3. *Data Analysis:* Use of a variety of techniques and methods to extract patterns and information from data and to find hidden information.
4. *Data Visualisation:* Presentation of data and information to decision makers to enable effective decision making.
5. *Data Interpretation:* Correctly understanding the information presented and making effective decisions.
6. *Data Storage/Disposal:* Either storing the data or information for future reference, or determining it to no longer be useful and disposing of it safely.

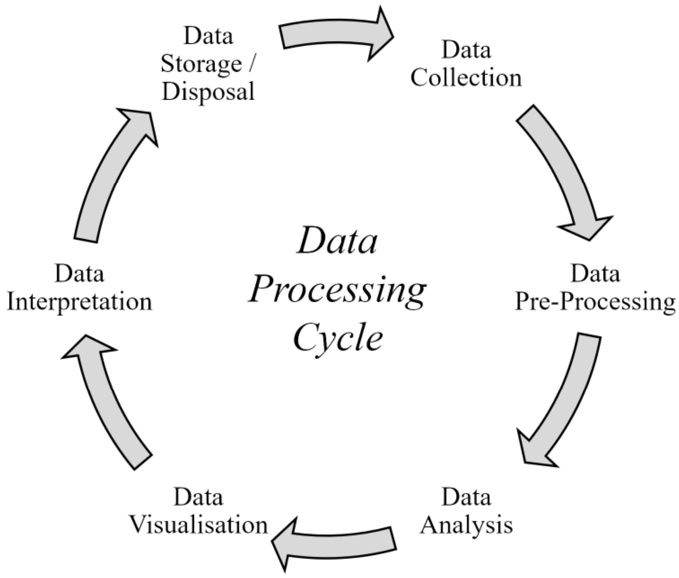


Figure 5.2-1 The data processing cycle.

We will discuss the stages in the data processing cycle in the next six sections.

5.2.2 Data Collection

Data collection is the first step in the data processing cycle and is one of the most critical to the success of the processing. The quality of the data that is acquired will impact on the quality of the analysis and decision-making. This has given rise to the common phrase “garbage in, garbage out” – a high quality data processing system will give bad outputs if given bad inputs.

Data gathering methods are often loosely controlled, resulting in out-of-range values (e.g., Volume = 100), impossible data combinations (e.g. Product Finish = Unpainted Steel, Paint Colour = Green), missing values, etc. Analysing data that has not been carefully screened for such problems can produce misleading results. Typically, the data collected falls in one of the following types:

- *Structured data:* Data that has been organised into a formatted repository, typically a relational database. Usually organized as a table in rows and columns, and its elements mapped to specific fields, structured data is how databases store data. Fields have tight restrictions on the formats they accept such as numbers within a certain range, a choice from a list of options, or a simple Boolean (yes/no) value. Structured data is generally the easiest to work with, as the meaning of fields and numbers are well defined and organised.
- *Semi-structured data:* Data that does not reside in a structured table but does have some organisational properties. An email for example, has structured elements such as the recipient, sender, timestamp etc., but also unstructured elements such as the message body or attachment content. XML is another example of a semi-structured format with tags marking up the content, but the content of the tags being unstructured.
- *Unstructured data:* Data that is not organised in any pre-defined manner. Though the file type may have some structured elements the vast majority of the data does not conform to any standardised structure. Audio files, video files, most text formats, and this book chapter are all examples of unstructured data. Unstructured data is harder to work with, but 80% of all data gathered is unstructured.
- *Metadata:* Different to the above classifications, metadata means “data about data”. Any of the three above forms of data may include metadata. For example, a Word document contains metadata about who has edited the document, when it was last edited, the file size, and much more. This metadata can be used to aid in understanding what the data relates to and how best to process it.

The source of the data will influence the level of structure to the data. There are three types of data source generally considered in manufacturing systems:

- *Primary source:* Primary sources are sensors, transforming physical phenomena into electrical signals either binary, digital or analogue. Temperature, humidity, pressure, force, motion, acceleration, light and object presence are examples of things that can be detected and measured with sensors. Sensors may come built into manufacturing equipment, retrofitted into existing manufacturing equipment or may be stand-alone. In most cases, a manufacturing database will not receive data directly from sensors; the data will be first received and processed by an industrial controller, industrial PC, or other data acquisition device. These devices are *secondary* data sources.

- *Secondary source:* Devices receiving data from primary sources are secondary data sources. These can include devices such as programmable logic controllers (PLC), industrial robot controllers, CNC machine controllers, embedded computers, human-machine interface devices, and smart sensors with built in memory and processing capability. The advantage of taking data from secondary sources rather than primary, is that secondary sources have computational capability and storage memory. This allows converting the sensor data from raw electrical signals to more easily interpreted standards. Signals can also be compressed and pre-processed, requiring less storage space in the data warehouse. The data from multiple sensors connected to the same secondary source can also be combined and compared for additional insights that a single sensor may not be able to offer.
- *External source:* Not all data related to a product and its manufacture and service life may originate inside a single company. Process data from suppliers producing constituent parts can be centralised as the complete product comes together, aiding in understanding how the quality of supply can affect the quality outcomes of production. Modern use of external data sources remains heavily reliant on paper-based solutions, with parts and materials arriving with documentation. Highly fragmented and variable manufacturing data systems throughout the supply chain make automating acquisition of data from external sources a challenge.

Data gathered from primary sources such as sensors will likely be unstructured or semi-structured, whereas the computing power of secondary sources such as PLCs or embedded computers allows for a more structured approach to collected data. It is often easier to structure data closer to where it is gathered than to send it all to a single server for structuring and adding to a database.

The methods of collecting of data will depend on the manufacturing process you are looking to analyse, and whether the data is coming from primary or secondary sources. Data collection does not have to be automatic or in real time, periodic data collection and reporting remains useful, even when done manually. Automatic real-time data collection does have advantages, and these will be discussed in Chapter 6.

5.2.3 Data Pre-Processing

Noise (noun): irregular fluctuations that accompany a transmitted electrical signal but are not part of it and tend to obscure it.

- Oxford University Press

Conversion of data to information is more difficult when data is noisy or unreliable. It is also hindered when there is irrelevant and redundant information present. For this reason data pre-processing needs to be carried out to ensure the quality of the data and efficiency of the analysis process. Data pre-processing includes many sub-steps depending on the type of data, and these steps are discussed over the next seven subsections.

5.2.3.1 Data Cleaning (Missing Data)

Data may be incomplete, which could mean missing some values that should have been recorded, or missing certain attributes of interest during some measurement periods. This could be due to data not always being available (e.g. equipment malfunction), equipment not being set up correctly for some periods of measurement, or simply because the attribute which is now considered important was not thought to be worth measuring at the time.

For example a primary source sensor records vibration in a CNC machine. These readings are collected by a secondary source PLC, which adds the timestamp of the recording, and the spindle speed the CNC machine was using. Each measurement then has three attributes – vibration, timestamp, and spindle speed. This data is then used for analysis of the relationship between spindle speed and vibration.

A soldered connection was loose in the sensor, and for some periods of high vibration, the sensor recorded no data. As a result, the occasional random measurement in the data set has no value for vibration.

The CNC was serviced, and the PLC was incorrectly restored to an earlier configuration where the spindle speed was not being included in the measured data. As a result, a solid hour of measurements have no spindle speed attribute before the error was spotted and could be fixed.

Many analysis methods require homogeneous data – i.e. data where all values to be analysed have the same format and attributes. If attributes have been omitted due to failure or error, several methods can be used:

- *Ignore the entire measurement*: Best used only as a last resort when multiple attributes were missed or there’s no way of reconstructing the data.

- *Fill in missing attribute manually:* Only usable if a human worker can quickly and easily predict what the missing value would have been, such as a timestamp where measurements are taken periodically.
- *Use a global constant:* Define a “default” value for any missing attributes to automatically fill them in. Only usable for some types of data that has a single value most of the time.
- *Use an averaged value:* Use an average calculation to fill in the value, either the average of adjacent measurements, average over the whole data set, or some other calculation appropriate to the data used.

5.2.3.2 Data Cleaning (Noise)

Data may be noisy due to collection instruments being faulty or there being errors in data collection. However, the vast majority of noisy signals are unavoidable, and are a consistent reality in even the best electronic equipment. Noise is an unwanted random error or variance in a measured value, and hides the true value.

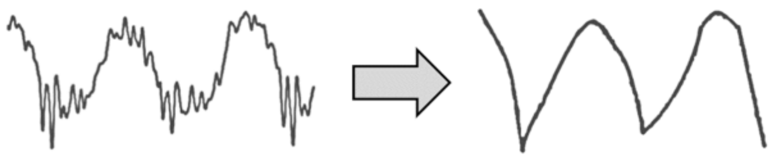


Figure 5.2-2 An example of a noise sensor measurement signal (left) and the same data after removing noise via smoothing (right).

Almost all electromagnetic devices and circuits will be subject to some degree of noise, and the study and characterisation of noise is a huge domain by itself. Noise can originate from the power supply of sensors, incorrectly grounded circuits, or the fundamental properties of electrons to name just a few. Several techniques to smooth the error can be applied such as:

- *Data binning:* These methods involve replacing values with intervals (“bins”) based on the neighbouring values.
- *Clustering:* These techniques are used to detect unusual values or ‘outliers’ and remove them.
- *Regression:* Smoothing the data by fitting it to an automatically discovered function.
- *Filtering:* A variety of functions to remove certain frequencies or other components from the signal. For example, UK mains power has a 50Hz AC signal, and this often causes noise in sensor measurements. This component of the signal can be removed with a filter.

5.2.3.3 Data Cleaning (Inconsistency)

There may be inconsistencies in the data recorded for some processes, particularly when the data is being recorded manually. For example when manually recording information different operators may have inconsistent naming conventions for error types. Even automated data collection may be inconsistent when set up with inconsistent file names or data formats. Data may also be mistakenly duplicated, either due to manual error or a networking issue.

These issues should generally be corrected at the source, by defining and enforcing naming conventions and formats. Some automated tools do exist such as knowledge engineering tools, and they may be used to detect the violation of known data constraints based on previous knowledge and automatically highlight the inconsistencies.

5.2.3.4 Data Integration

Data may come from multiple sources including databases, spreadsheets, XML files etc. and you may not always be able to define the format and data representation standards used if they are set by the equipment you're using or come from external sources. Combining these can be difficult if data is labelled or represented differently. Redundancy is another issue. An attribute can be redundant if it can be derived from another one.

Though the most common solution to this problem is manual integration, using a data warehouse and the Extract, Transform, Load (ETL) method can simplify this process, transforming all the data into a single homogeneous format.

A data warehouse draws together many different data sources and provides a single interface to them all, allowing queries to look inside the many data sources without extra effort by the user. Databases are an example of a data source in a data warehouse. Data warehouses need to unify and integrate multiple data sources to enable them to be searched and compared. This is achieved with the ETL process.

A database is typically updated in real-time with the most recent values for data to enable decisions to be rapidly made. By comparison, a data warehouse will keep all historical values of data. A data warehouse is hence slower to access and analyse, but allows for a deeper insight into data than a database.

A *data lake* is another term occasionally used. Whereas a data warehouse uses ETL to harmonise all data into compatible formats, a data lake keeps data in its original form and only transforms it into other formats on demand. This process is sometimes called Extract, Load, Transform or ELT.

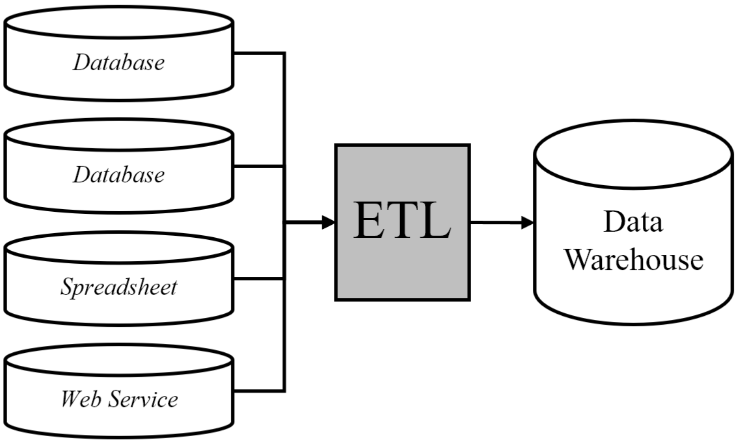


Figure 5.2-3 Data warehouses combine current and historical information from a wide variety of data sources using a method called ETL. ETL is an often complex process to regularly extract data, transform it into a common format, and load it into the warehouse. Though complex, this will save time in the long run when large amounts of data are stored.

5.2.3.5 Data Transformation

Whereas data integration converts an entire set of data from one format to another for use with different programs without changing the actual values of the data, data transformation changes the values in the data itself into a form more suitable for analysis. The transformations required will depend on the type of analysis, but can include:

- *Normalisation:* Changing the scale of values of the data to ensure compatibility. This often means using a specified range such as -1.0 to 1.0 or 0 to 1.0.
- *Aggregation:* Refers to creating summary or aggregation operations to attributes, for example calculating the monthly or annual amounts of daily sales rather than working with daily values.
- *Generalisation of the data:* Raw or 'low level' data is replaced by higher level concepts. For example, categorical attributes like street can be generalised to higher level concepts like city or county. Similarly, numerical values like age can be mapped to categories such as young, middle-aged and senior to move the focus from specific ages to categories.

For example, a company wants to compare the quality of parts manufactured with different machine centres on the shop floor, to see which machines might need maintenance or adjustment. There are a variety of machining centres of different *sizes and ages to be analysed. The workers have noticed that the newest machining centre produces more parts that fail quality checks that any other and are concerned about their investment.*

Critically, the only data the workers are looking at is the absolute number of parts which fail quality control. The new machine produces more parts overall than the other machining centres. If the data for each machining centre can be normalised into a percentage of failed parts rather than the absolute number, it becomes clear that the new machine has a lower percentage of rejected parts than any other, even if the absolute number is higher.

Note that a process also called normalisation is an important aspect of database design, but this process is distinct from the data normalisation discussed here.

5.2.3.6 Data Reduction

Reducing the volume or dimensions (i.e. the number of attributes) of the data. This technique is useful when data analysis on the complete data set is unfeasible or impractical. Data reduction techniques need to ensure the integrity of the original data is not compromised, and produces a reduced set that can still ensure quality knowledge extraction. Some strategies include:

- *Dimension reduction:* Redundant attributes may be removed, either manually or after an initial period of analysis.
- *Data compression:* Encoding mechanisms are used to reduce the size of the data set. Size can be a major problem in some applications, making it difficult or impossible to process given available computing power.
- *Numerosity reduction:* Data is replaced or estimated by alternative, smaller data representations, such using regression models or clustering techniques.
- *Discretisation and concept hierarchy generation:* Raw data values for attributes are replaced by ranges or higher conceptual levels. For example, replacing numerical values with concepts such as low, medium and high. This is a similar concept to generalisation for data transformation.

As an example, consider monitoring power quality disturbances. In order to effectively identify problem causes one needs to sample in the range of MHz [1]. *Compared to other applications where data is gathered at frequencies less than 25Hz, this represents 40,000 times more data. In this case techniques such as Wavelet Transform and Principal Component Analysis (PCA) can be applied to achieve a high reduction ratio, in some cases above 90%, depending on the technique used* [2].

5.2.3.7 Data Pre-Processing Comments

Once the data has been pre-processed, the outcome of this process needs to be saved. Often a simple comma separated value (CSV) format is more than sufficient, and can be read by all spreadsheet and analysis applications. If dealing with a complicated output format, databases, XML files, or JSON files (amongst many other formats) can be used. These formats are established standards and almost all technologies that are available today can understand them easily.

As can be seen by the size of the data pre-processing section, it's a complex area and often overlooked by those less familiar with data analysis. However, as the table below shows, data pre-processing takes up the majority of professional data scientist's time, showing how important but also difficult the task is.

Task Name	Percentage of Time
Data Collection	19%
Data Pre-Processing	60%
Data Analysis	9%
Refining Analysis Algorithms	4%
Building Training Sets	3%
Other	5%

Table 5.2-1 The proportion of time data scientists spend on tasks [3]. The majority by far is data pre-processing. The same study that collected this data also showed that data pre-processing is the least enjoyable aspect of being a data scientist.

5.2.4 Data Analysis

After ensuring the data is cleaned and in the appropriate format and structure, it is ready for *analysis*. Modern data analysis blends traditional statistical data analysis with new and emerging computational methods. Analysis uses a huge number of techniques, and many are specific to the type of data being processed. Some of the largest categories are discussed below, starting from highly numerical analysis and moving to more textual analysis:

- *Statistical analysis:* The process of generating statistics from stored data and analysing the results to deduce or infer meaning about the underlying dataset or the reality it attempts to describe. Some of these statistics include Bayesian analysis, conditional probability, data classification, linear regression, resampling, shrinkage, tree-based analysis, to name just a few [4]. Though advanced data analysis methods such as machine learning can be very effective, for many domains well executed classical statistical analysis can be as or more effective. Statistical analysis may not be the most modern or trendy method of data analysis for numerical data, but techniques are well understood and refined, and results are often extremely good for all but the most difficult of circumstances.
- *Quantitative analysis:* Techniques that try to understand behaviour by using mathematical and statistical modelling and measurement, to create a model of

the process being measured. This model allows analysts to examine and test the past, current, and anticipated future events.

- *Qualitative analysis:* This uses subjective judgment based on unquantifiable information, such as management expertise. While quantitative analysis uses exact inputs, qualitative analysis deals with intangible, inexact concerns that belong to the social and experiential realm rather than the mathematical one. Quantitative and qualitative analysis are often used together in order to examine a company's operations and evaluate potential investments.
- *Semantic analysis:* This is the use of ontologies to analyse content in text-based sources. Ontologies are data models of the formal naming and definition of categories, properties and relation between concepts, allowing automated systems to understand the approximate meaning of text. It uses text analytics to measure the relatedness of different ontological concepts. It can be used for natural language processing i.e. automatically understanding human speech. A common example of use is understanding if a tweet or review uses positive or negative language about products.

Two additional techniques have emerged in recent years, and are used alone or in combination with the above techniques~~:~~

- *Data mining:* This is the process of discovering patterns in large data sets (also called "big data"). These patterns can then be seen as a summary of the input data and may be used in further analysis and predictions.
- *Machine learning:* This exploits patterns found in historical data to identify risks and opportunities. Machine Learning refers to the use of algorithms to learn automatically from the data, without using explicit instructions or relying on models.

5.2.4.1 Data Mining

Knowledge is a very valuable asset in manufacturing, as it enables a business to differentiate itself from competitors and to compete efficiently and effectively to the best of its ability. The advancements in information technology (IT), data acquisition systems, and storage technology as well as the developments in machine learning (ML) tools have led towards new ways of knowledge discovery in manufacturing processes. Data from almost all the processes of a manufacturing business such as product and process design, material planning and control, assembly, scheduling, maintenance, recycling, etc., are recorded. These data stores offer enormous potential as sources of new knowledge. However, data must be analysed and converted into actionable knowledge to be useful. In addition, the volume of collected data is becoming an issue with insights buried in large volumes of other data.

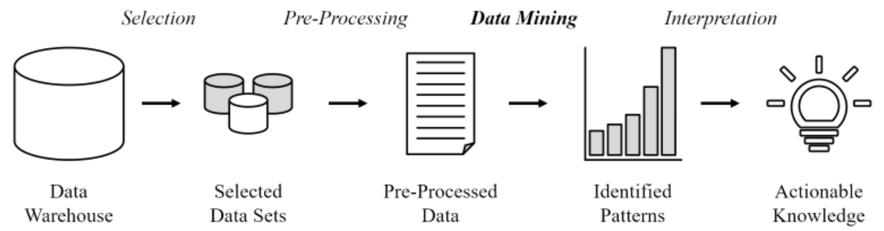


Figure 5.2-4 The data mining process. Sources of data are selected from the data warehouse, some pre-processing performed, and then patterns identified in the data. These patterns are then converted into actionable knowledge.

Data mining is an area of computational intelligence focusing on providing new systems, techniques, and theories for the discovery of hidden knowledge in large volumes of data. It is a blend of concepts and algorithms from statistics, artificial intelligence and data management. The term *big data* is also commonly used to refer to this area of data analysis.

The use of data mining in manufacturing began in the 1990s and is gaining traction in different aspects of the manufacturing process such as predictive maintenance, fault detection, design, production, quality assurance, scheduling and decision support systems. A typical data mining process is shown in Figure 5.2-4.

Data mining methodologies were introduced to provide a more holistic view to the knowledge discovery process, beyond the application of statistical or machine learning algorithms. Data mining looks at a single set of data to discover previously unknown patterns with no prior information, whereas machine learning uses past experience to find examples of known patterns in new data sets. The two domains cross over frequently, so the distinction is often blurry.

5.2.4.2 Machine Learning

Machine learning is a broad topic, but generally refers to algorithms that learn over time to give improved results. Machine learning as a term and as a concept dates back to 1959, but it is recent improvements in computational power that has brought it to the mainstream. Machine learning algorithms use a *training set* of data which has been labelled and uses that to automatically generated a predictive model. This model is then used on a new set of data and assigns a label automatically.

Machine learning is now in common use with applications including congestion prediction in satellite navigation apps, email spam filtering, fraud detection in banking transactions, and facial recognition systems. Anywhere in the manufacturing domain where there is a requirement to assign a class or type to a piece of data is a potential application of machine learning.

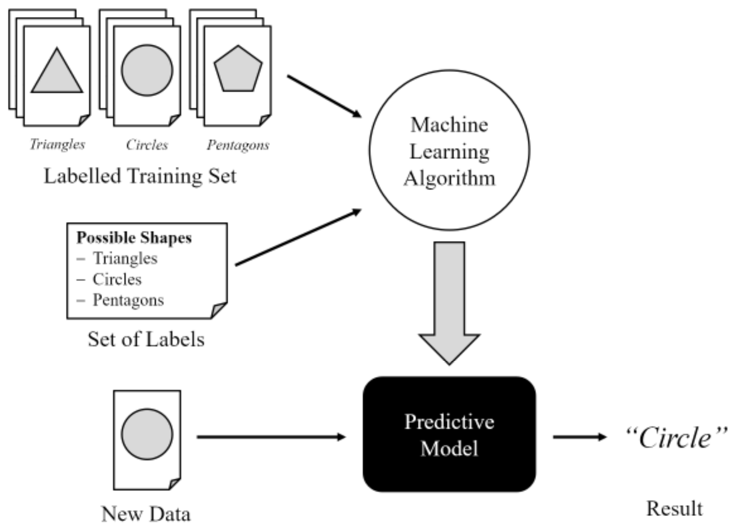


Figure 5.2-5 An overview of the machine learning process. A training set of labelled images of shapes (and a set of all possible labels) is used by the algorithm to generate a predictive model. Future unlabelled images of shapes will then automatically be classified, such as the pictured circle.

For example, a company has been using a vision system to identify defects in the paint finish of their products. The system can identify something out of the ordinary but cannot tell what type of fault has been detected without a worker to analyse the image. The worker is given a picture of the defect and assigns it a label. The label is the type of imperfections, and could be chip, abrasion, dent, scratch, bubbling, or more. Knowing the type is important as it helps identify the cause of the problem and stop it happening again. The company wishes to automate the labelling process as it is a time-consuming process and their workers could spend their time better elsewhere.

Here, the set of manually labelled images of imperfections becomes the training set for the machine learning algorithm. The algorithm creates a predictive model from this training set. Any future imperfections which are identified will be run through the algorithm, and the algorithm will automatically classify the imperfection type.

Machine learning is not without its drawbacks, however. Machine learning requires a large training set and access to a labelled set of data is often the largest hurdle in implementing a machine learning solution. The quality of the results is also highly dependent on the quality of the training set. Consider the example in Figure 5.2-5 – what if the new data were a picture of an octagon? The predictive model would likely classify it as a circle as it has no examples of octagons to use, and a circle would be the best fit from the available labels. Algorithms can state a level of certainty in their classification however, to flag issues like this. A final problem is that the predictive model is a "black box" – it was automatically generated and how a specific model works is extremely complex and almost always unknowable to a human. This means a model which has problems becomes very difficult to troubleshoot.

Both data mining and machine learning are made possible with extremely complex algorithms, but knowing exactly how they work is not necessary to start implementing them, as a great many software packages exist that handle the complexity. The main choice would be between free, open source packages such as Google's TensorFlow, Facebook's PyTorch, Keras, and Weka; or using proprietary integrated solutions such as those available in Amazon Web Services, Microsoft Azure, or Google Cloud.

5.2.4.3 Data Analytics

Also of note is the field of *Data Analytics*. Whereas analysis focuses primarily with data from the past and understand why events occur, analytics focuses on what is likely to occur in the future and how to respond to it. Data analytics is also sometimes called *Business Analytics* when applied to more general aspects of business performance. Analytics is typically defined as having three types:

- *Descriptive analytics:* What happened? This is effectively analysis, and focuses on what happened in the past.
- *Predictive analytics:* What will happen? Using predictive modelling such as machine learning and some statistical techniques to predict what is likely to occur in the future based on previous events.
- *Prescriptive analytics:* How can we make things happen? By understanding what events or parameters contributed to past events, automatically recommend changes to parameters to influence future events.

Many software packages or programming languages exist for data analysis and analytics, with free open source examples including KNIME (Konstanz Information Miner) for data analytics, R which is a dedicated data analysis programming language, and SciPy which is a data analysis library for Python. Many commercial software packages also exist, often described as *Business Intelligence* software.

5.2.5 Data Visualisation

Data analysis is often carried out in parallel with *data visualisation*, and indeed many data analysis and analytics tools will include data visualization capabilities. Their purpose is to show in a graphical and easily understood way, the trends and the structure of the information that has been discovered through the analysis process. Broad relationships and patterns can be brought out, as can emerging trends. Visualisation also helps to quickly narrow the search for information of interest. It can be static, in the form of charts and diagrams in reports, or live, such as manufacturing dashboards reporting live data from the shop floor.

Common examples of visualisation software include Microsoft Power BI, Dundas BI, and Tableau, but many companies also develop customised dashboards with visualisation libraries such as R Studio's Shiny or Grafana. Even spreadsheet software's reporting and graphing features can achieve these goals.

Data visualisation is an example of a data-driven decision support system. Decision support systems will be discussed more in Chapter 6.

5.2.6 Data Interpretation

Though automated tools exist to assist in making decisions the ultimate interpretation and decision making responsibility still typically lies with humans. The data processing cycle exists to facilitate the decision making process, but the decision itself should be in line with established company procedures.

It is important when a decision is being made to understand the type of problem that is actually being solved, as this will influence how much and what type of data is required to come to a good decision.

- *Unstructured problems:* These possess multiple solutions and there is no algorithm or formula that can lead to the optimal solution due to uncertainties in the problem, and there are few parameters that can be directly affected to solve the problem. Unstructured problems are often rare or novel, and do not have established responses. One example would be a decision around employee reward schemes where the results of the decision are more about employee satisfaction and company perception rather than the exact value of the reward. Unstructured problems are typically solved with informal intuitive decisions based on experience.
- *Structured problems:* By contrast, structured problems have very clear parameters and goals, and the results of a decision are possible to model and predict. Often there is a small number of criteria to be maximised, and algorithms exist to model the decision. Structured problems are often routine and well understood. An example would be the response to a common maintenance requirement for a piece of equipment or machine tool. The problem is common and well understood, and experience shows what the best solution is. The costs of replacement parts and downtime are all understood.
- *Semi-structured problems:* These are problems where only part of the problem can be optimally solved, requiring a combination of a standard solution and individual judgement. They are the grey area between structured and unstructured. An example would be creating a maintenance plan for a new piece of equipment. You know the cost of downtime and replacement parts, but you have no intuitive understanding of how frequently the machine will need maintenance performed until it has been in use for a considerable time. Failure modes are not fully understood.

The more structured a problem is, the easier it is to apply data and find an optimal solution. So far, the methods we have been discussing are best applied to structured and semi-structured problems. Unstructured problems are more challenging to take a data-driven approach too, but that does not mean it is impossible. Many unstructured problems can be reduced to semi-structured one upon closer analysis, and tools such as decision analysis models (DAMs) can start to decompose the problem into smaller, manageable components. DAMs are statistical tools and methods such as analytical hierarchy processes (AHP), decision tree analysis, multi-criteria decision analysis, and probabilistic forecasting to support decision making where multiple criteria need to be considered and there is no single optimal answer.

There are two common problems when it comes to making decisions assisted with analysed data:

- Decision makers make bad decisions because they do not have access to high quality data, and either are misinformed by bad data or instead use intuition or gut-feel.
- Data scientists attempt to make decisions based on data, when they do not have a detailed understanding of the real world in which the data originates.

The simple solution is to ensure data scientists and those with domain knowledge work together on the decision making process. Performing all the steps in the data analysis cycle will provide a solid, reliable set of analysed and visualised data with which to base decisions on, but this must also be combined with real-world experience.

5.2.7 Data Storage and Disposal

For the collected, processed, and analysed data this stage of the cycle represents a fork. Data can be stored and fed back into the data collection stage of the cycle for future use and analysis as new data is acquired. As the amount of data that can be generated for the production of a single manufactured item can be in the order of gigabytes per day, there is a need for a database that is capable of storing this data. Though the absolute cost of data storage continues to decrease (particularly when cloud-based solutions are considered), the amount of data that is being generated continues to increase, and the rate at which data can be sent and accessed is not increasing at the same speed.

Alternatively, data can be considered to have fulfilled its purpose, and be disposed of. Data can take up a large volume of storage space, and permanent storage is a costly endeavour. Though storage costs are decreasing, the volumes of data being gathered by companies are also increasing. An alternative is to dispose of the raw collected data, while keeping the analysed data. Analysed data typically requires a smaller amount of storage space, so you can keep the insights even if you dispose of the source of the insight.

When disposing of data, consider how well the data is erased. Simple file deletion may not remove data from the storage medium. This may not be a huge concern for gathered process data, but data which is commercially or personally sensitive should be securely deleted or it could be recovered by a malicious individual such as a corporate rival or disgruntled former employee. Consider also your responsibilities under the General Data Protection Regulation (GDPR) in the EU or equivalent local regulations – personal data must be gathered only when necessary, used for a clear and specified purpose, and disposed of securely when no longer required.

5.3 Modelling and Simulation Approaches

5.3.1 Introduction to Modelling and Simulation

A model is a mathematical representation (and often a simplification) of a system built using mathematics. A simulation is using the model to predict the behaviour of the system under specified parameters, by inputting values to the model and recording the results. Modelling and simulation tools allows manufacturing companies to test changes and improvements to their manufacturing systems before physical implementation. In many cases decisions regarding manufacturing system changes are often based on experience and intuition rather than on quantitative prediction. Though this can often work, it relies on the availability of human expertise and is prone to error.

The creation of a modelling or simulation package requires extremely specialised knowledge and expertise. Though these tools will hide many of these details of their implementation from the user, understanding the underlying model type will aid in understanding which package is best applicable to your needs. Modelling tools tend to specialise in specific areas, and often a combination of packages will be required to get good coverage of the problems you want to solve.

For manufacturing, these specialised areas can be broadly grouped into three domains, and it is often the case that you will need a separate package or solution for each domain at minimum:

- The *product domain* is related to the product that is to be manufactured, and can be classified into structure-oriented, geometry-oriented, feature-oriented, or knowledge-oriented models.
- The *process domain* represents the relations between events and activities in a manufacturing system, describing how a manufacturing system as a whole should act, without going into the specific details of each manufacturing machine and resource.
- The *resource domain* represents the operative instructions for a given manufacturing resource, such as a machining centre or robot. These instructions are typically specific to the resource being modelled, such as the G-code of a CNC machine.

The following section gives an overview to the most common modelling approaches used by softwarte packages, and details their applicability to help you evaluate a package for your needs.

5.3.2 Types of Modelling Approaches

5.3.2.1 Discrete Event Simulation

Discrete Event Simulation is a simulation methodology that has been widely adopted within industries to test manufacturing system changes virtually before implementing them physically. It allows for a high-level analysis of the system's performance by statistically and probabilistically reproducing the interactions of its components and resources.

The entire manufacturing facility can be modelled as a sequence of operations being performed on passive entities (e.g. components), as they pass through the processing sequences. Although the components are passive, they have attributes that affect the way they are handled and some of these attributes change as the component advances through the process. The simulation allows you to test changes to your production line and answers questions such as:

- How long will the cycle time be for this new product?
- What is the utilisation of equipment, do I need to invest in more of some resources or optimise the utilisation of others?
- Are my buffers likely to fill up and queues form in the process?
- What is the predicted error rate of this process, and is the process cost effective?

You may notice many of the questions above are similar to those answered in Chapter 4, and that's not a coincidence. DES is a statistical simulation rather than a physical one. It allows for the inclusion of existing data from your resources (e.g. what is the cycle time of a process, how often does a machine require maintenance, and how often does it fail) to predict future performance under different circumstances. It is often simple to implement within a company, provided you have historical data to base predictions on.

5.3.2.2 Agent-Based Modelling

Agent-based modelling is a tool that allows the modelling and simulation of complex systems by breaking the problem into simpler units and modelling those using intelligent agents. In this way, the modeller only needs to understand the behaviour of simple components to create the model, whereas techniques such as system dynamics require a better understanding of the system as a whole.

An *intelligent agent* (typically simply called an *agent*) is a piece of independent software is capable of exhibiting autonomous and intelligent behaviour [5]. An agent is an individual problem solver with some capability of sensing and acting upon is environment. They can be used both for the control of systems (where the environment is the physical system), or for modelling systems (where the environment is a model). We focus here on the application to modelling systems.

There are many types of agents, but they have some basic properties:

- An agent can observe its environment with sensors.
- An agent makes a decision based on the observations.
- An agent initiates and executes actions using actuators to change the environment.

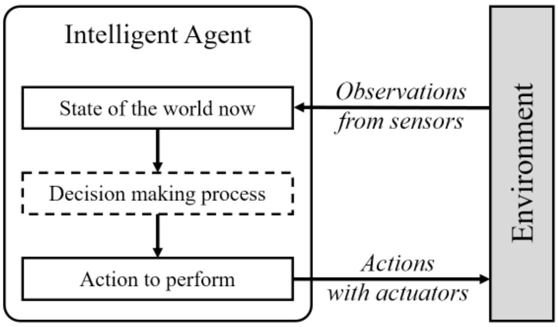


Figure 5.3-1 The basic function of an agent. The decision making process can vary from simple condition-action rules, to highly complex artificially intelligent learning processes.

Agent-based modelling breaks down complex problems by focussing on simpler components, with each agent representing the behaviour of a resource or worker in the system. The complexity then emerges from the interactions of these components, each represented by an agent. For modelling and simulation the environment, sensors, and actuators would all be virtual. However, agents are increasingly used as a control method for real physical systems, so the basic definition omits “virtual”.

In manufacturing, this tool has been typically used for modelling production planning and resource allocation, production scheduling and control, monitoring and diagnosis, production in networks and assembly and life-cycle management. Due to their distributed nature, agent-based systems provide modularity, robustness and autonomy, which offer an alternative way to design and control systems when compared to conventional approaches. An example of existing environment for the simulation of multi-agent systems in the manufacturing domain is MAST (Manufacturing Agent Simulation Tool) [6], developed by Rockwell Automation. This tool is particularly focused on dynamic product routing.

5.3.2.3 System Dynamics

System Dynamics is a technique for constructing models that focus on interdependencies, feedback effects, time dependencies, and causality in the object that is being represented. It allows for more complex modelling than DES, and is applied beyond manufacturing in domains such as finance, population growth, agriculture, and ecological behaviour. Feedback loops are one of the key considerations in system dynamics, and cause some of the least predicable behaviour in complex systems such as production lines and facilities, and can cause serious issues if not understood [7].

The system orientation of system dynamics makes it ideally suited for the analysis of future production dynamics, and can include operational and organisational aspects in addition to shop-floor elements. It allows studying how information flow affects the behaviour of productive systems, an aspect often not covered in other simulations, which only consider the flow of physical objects. A typical system dynamics model is built using the following elements:

- *Levels* (shown as rectangles) represent the quantity of some element of the system at a point in time, such as stock levels.
- *Rates* or *flow* variables (shown as valves) represent the rate of change levels in the system in an interval of time, such as utilisation of coolant.
- *Converters* (shown as circles) are additional equations or calculations that effect the system.
- *Connectors* (shown as arrows) are the information links in the system connecting other components together. The arrows are often given different weights to show the difference between information (thin arrows) and changing physical quantities (thick arrows).
- *System boundaries* (shown as clouds) are the edges of the system being modelled.

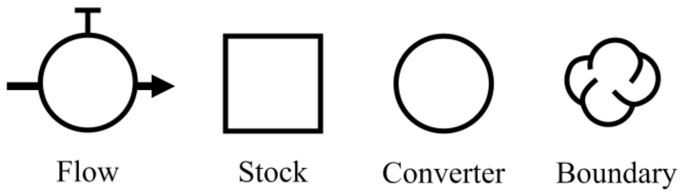


Figure 5.3-2 Systems Dynamics diagram components. Image rights: Author, adapted from [8].

Depending on the problem being modelled, system dynamics can be implemented in two phases:

1. *Qualitative system dynamics*: This phase involves the creation of cause and effect diagrams or system maps.
2. *Quantitative system dynamics*: This phase involves deriving the shape of relationships between all variables within the diagrams, the calibration of parameters, and the construction of simulation equations and experiments.

One of the most common diagrams in system dynamics is the levels and rates diagram (also known as the stock and flow diagram). Figure 5.3-3 shows an example.

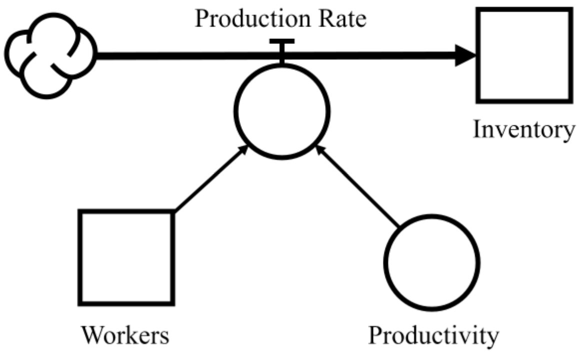


Figure 5.3-3 Example of a levels and rates diagram, which models a production line's production rate as being dependant on the quantity of workers available and the productivity of the system, resulting in finished products being added to the inventory. Image rights: Author, adapted from [8].

Once the system is defined with system dynamics, a detailed evaluation is used to find the ideal solution. Large amounts of data are required at this stage, making this evaluation a challenge for some manufacturing companies. In general, system dynamics is also a more complex simulation to create than DES, but offers considerably more possibilities for understanding highly complex systems.

5.3.2.4 Petri Nets

Petri nets are a graphical and mathematical modelling tool applicable to several systems. As a graphical tool, Petri nets have been used as a visual communication aid similar to flow charts or network diagrams. The classical Petri net is a directed graph with two node types called *places* and *transitions*. Places are represented by circles and transitions by rectangles. A petri net has states, represented by *tokens*. Places may contain zero or more tokens, which are represented by black dots. A transition is activated only if all the places that are inputs to the transitions have tokens, and a transition will consume the input tokens and move the tokens to the output places.

Figure 5.3-4 shows a Petri net that models a machine which processes jobs and has two states: free and busy. There are four places, *Storage* (1), *Available* (4), *Processing* (2) and *Complete* (3), and there are two transitions, process start and process finish. There are five tokens, 4 in place *storage* and 1 in place *available*. The tokens in *storage* represent parts to be processed, and the token in *available* represents the availability of the lathe for processing. The system proceeds in discrete time steps.

The transition representing the process can only activate if there is both a token in *storage* (representing the part to be processed) and a token in *available* (indicating that the process is not busy). The transition will consume these tokens, and place a token in *processing*. The process cannot start again as – although there are parts available in storage – the lathe is not available.

The transition representing the process finishing will consume an available token in *processing* and place a token in *complete* (representing a finished job) and in *available* (representing the lathe is available again). The Petri net is then ready to start a new job by taking a token from *storage*, and will repeat until all parts are processed.

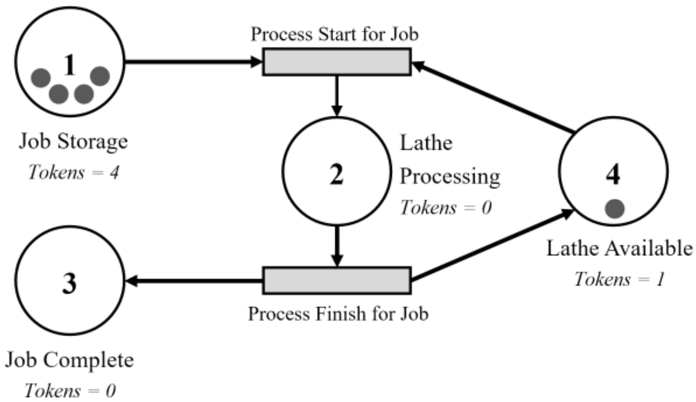


Figure 5.3-4 Petri net modelling a machine with two states: processing and available.

Petri nets are a powerful but difficult and verbose modelling method. They are often the underlying model for more user friendly modelling software, and many modelling tools will handle the lower-level modelling for you.

5.3.2.5 Monte Carlo Simulation

Monte Carlo simulation is a computerised analytical process that is used to evaluate and measure the risk associated with any given venture or project. It allows users to understand a given situation but also the impact of other possible scenarios.

Rather than using absolute values for a simulation (for example, an average run time for a milling process) it utilises a probability distribution of values.

An example of this probability distribution is the Normal distribution. During the simulation, samples are drawn randomly from the input ranges and the results recalculated over and over. The result is a range, or distribution, of possible outcome values and their associated probabilities of occurrence. Compared to deterministic methods, MCS offer the following advantages:

- Provides *probabilistic estimations of potential results*, rather than single values. For example, a conventional simulation might inform you of an estimated process time that is acceptable to you. By comparison, an MCS will inform you the odds of the process time falling outside of acceptable bands.
- Provides information of *which inputs contributed the most to a particular result*. Deviations in one process may have little to no impact on overall system performance, but a different process may have a huge effect with even small deviations.
- Allows the user to *understand the behaviour of inputs* related to a particular resulting scenario. What potential combination of variables are likely to cause your system to have problems? How can you protect yourself against this?

An example of the use of MCS in manufacturing companies is for minimising the number of disruptions caused by the supply chain. By using this type of simulation, variables such as sales demands, material costs, investments outlays, delivery and processing times, inventory uncertainty and disaster risk can be modelled, and both independent and interrelated variables considered. When the MCS runs, the user can consider thousands of scenarios and determine several outcomes such as which factors are most likely to impact supply chain; weather, material costs, overhead expenses, fluctuating prices, to name just a few.

5.3.2.6 Virtual Simulation

Many modelling methods now have three-dimensional rendering capacity, allowing the user to see the results of a simulation in a realistic and intuitively understandable way. Advanced in virtual reality and augmented reality technologies can even present the user the system in an immersive, interactive fashion.

However in some cases, creating a 3D model to get an intuitive idea of a design or system is sufficient to achieve the aims of the modelling process. Examples include computer-aided design (CAD) models of products to test for usability (though CAD offers a lot more also), and the creation of models of assembly stations or working areas for workers to get a feel for ergonomics and check for issues before the workstation is built.

5.3.2.7 Simulation Gaming

A model is a mathematical abstraction of a system, and the usual approach is to use the model to simulate the result of specified parameters, or to optimise the parameters for desired outcomes. Sometimes however, the model is used to test how humans will react to the situation. This is called simulation gaming and allows for testing human responses to unexpected or stressful situations such as incident management, or for testing the usability of interfaces and workstations.

5.3.2.8 Intelligent Simulation

Modelling and simulations can be extremely complex. These systems often have a lot of parameters and to explore all possibilities can require unfeasibly long run times, providing answers too late to be useful. *Intelligent simulation* is based on Artificial Intelligence (AI) techniques and typically supports other types of simulation techniques**.** For example, machine learning and intelligent sampling can be used for the optimisation and calibration tasks of an agent-based system, creating meta-models that can deliver dramatic speed increases for large-scale models. For digital twin and what-if analysis systems, AI components can be directly embedded in the simulation model to allow testing and forecasting.

Another application of Intelligent Simulation is the use of deep learning for the replacement of rule-based systems. PricewaterhouseCoopers, for example, currently supports a large car manufacturer to introduce autonomous vehicles for the public. Part of this work involves using deep reinforcement learning to determine optimal decision rules that allows the vehicles to maximise efficiency also satisfying customer trip demand. Similar approaches can be used to automated manufacturing lines.

5.3.2.9 Distributed Simulation

Where a model is so complex running a simulation would take an extremely long time, a distributed simulation may be used. Many underlying modelling methods can be used, the challenge here is to make model decomposable into smaller units which can then be distributed to multiple computing resources to execute the simulation in parallel.

Another use case for distributed simulation is where multiple entities (such as companies) wish to create a model and run a simulation together, but also wish to create their models using different modelling packages, or wish to keep elements of their model secret. Many of the standards in distributed simulation such as IEEE 1278 [9] have their origins in military wargaming, which required the distribution of simulation with elements kept secret.

5.3.3 Applications of Modelling

Modelling has a wide range of applications on manufacturing. Perhaps the most commonly used is CAD – Computer Aided Design. CAD allows for a model of a proposed design to be built, and virtually tested before prototypes are made. It allows for dimensional analysis and validation, as well as engineering analysis such as:

- *Mass properties analysis*, such as product volume, surface areas, weight, and centre of gravity.
- *Interference checking* in multi-component designs, checking components will not physically clash.
- *Tolerance analysis* to automatically determine what tolerances are required for correct product operation.
- *Finite element analysis* provides approximate solutions to tests such as stress-strain, heat transfer, or fluid flow which would otherwise require physical prototypes.
- *Kinematic and dynamic analysis* helps test the motion of multiple linked components and analyses their motion properties.

As well as this, quality of life services such as automated design reviews, automated drafting, and version control are typically integrated in CAD programs.

Within the context of manufacturing systems analysis, modelling has a large number of applications, with some listed below along with the underlying types of modelling typically used for these applications [10].

- *Assembly line balancing:* The design of assembly lines, and the process of balancing them. Balancing is ensuring that a line has sufficient resources (e.g. workers or equipment) to meet the required production rate, without excess spare capacity.

  Modelling Methods: Discrete Event Simulation.

- *Capacity planning:* Modelling external unpredictable elements of a business environment to ensure the business has sufficient capacity to deal with the fluctuations. For manufacturing this will include fluctuations in supply and demand, and ensuring the company has sufficient storage, buffers, and production capacity to cope, or identifying areas that could be changed or refined to meet the required capacity.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Monte Carlo Simulation, Petri-net Simulation.

- *Cellular manufacturing:* Optimising the layout of a cellular manufacturing set-up for increased production capacity and improved operator ergonomics. Also checking new scheduling schemes or process plans for potential operational issues.

  Modelling Methods: Virtual Simulation

- *Transportation management.* Supply chain modelling to evaluate the effectiveness of vehicle routing, truck loading, distribution centre utilisation, and incident management approaches.

  Modelling Methods: Discrete Event Simulation, Agent-based Simulation, Petri-Net Simulation, Other (Traffic simulation, a specialised set of modelling methods for transportation)

- *Forecasting.* Predicting future patterns and analysis of trends. In a manufacturing context this will generally mean market forecasting to predict the demand for a company's products under changing global situations.

  Modelling Methods: System Dynamics

- *Inventory management.* Evaluation of the cost and benefits of holding inventory, either to save space or insurance against supply disruption. Can also include elements of inventory utilisation policy (i.e. when to replenish stock, which stock to use first) and understanding at what level of stock should a replenishment order be placed.

  Modelling Methods: Discrete Event Simulation, Monte Carlo Simulation.

- *Just-in-time.* Design of just-in-time (JIT) manufacturing systems, which are systems where the receipt of parts from suppliers aligns extremely closely with when the parts are due to be used (i.e. the parts arrive "just in time"), resulting in low levels of inventory required and quicker cycle times. This requires a very carefully balanced production line, as even small disruptions to supply or process can cause significant system-wide delays.

  Modelling Methods: Discrete Event Simulation, Intelligent Simulation.

- *Process engineering - manufacturing.* Design of manufacturing processes including elements of process design, ramp-up optimisation, and performance measurement and optimisation. Can also be used to plan entire new facilities or to evaluate the potential impact of new equipment acquisitions.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Agent-based Simulation, Monte Carlo Simulation, Petri-net Simulation, Virtual Simulation, Intelligent Simulation.

- *Process engineering – Service.* Design of service processes such as logistics and distribution, waste handling, retail and the service industry, and financial services. Service processes share many similar concerns with manufacturing process, including scheduling, capacity, bottleneck analysis, and performance measurement and optimisation.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Distributed Simulation.

- *Production planning and inventory control.* A specialisation of process engineering – manufacturing, production planning and inventory control specifically focuses on batch size optimisation, bottleneck analysis, forecasting and scheduling.

  Modelling Methods: Discrete Event Simulation, Agent-based Simulation, Distributed Simulation.

- *Purchasing.* Optimisation of purchasing strategies for parts and supplies, including when to replenish stock, and understanding minimum purchase sizes and bulk discounts to keep sufficient stock with the minimum price and without excess inventory.

  Modelling Methods: Discrete Event Simulation.

- *Resource allocation.* Assigning equipment, workers, and time/overtime to tasks to improve process flows and productivity under varying circumstances. Can also include assignment of raw materials, tooling, and storage to processes.

  Modelling Methods: Discrete Event Simulation, Agent-based Simulation, Monte Carlo Simulation, Distributed Simulation, Intelligent Simulation.

- *Scheduling.* Modelling a production process to test potential job sequencing to optimise throughput and ensure orders are delivered on time. This is a broad subject area, including elements of throughput analysis and production capacity, resource allocation, workforce planning, and trade-off analysis between conflicting objectives such as delivery reliability and cost to manufacture.

  Modelling Methods: Discrete Event Simulation, Agent-based Simulation, Monte Carlo Simulation, Petri-net Simulation, Intelligent Simulation.

- *Strategy*. Modelling a business or system to simulate the effect of high-level policy change and proposed business strategies.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Agent-based Simulation, Monte Carlo Simulation, Simulation Gaming.

- *Supply chain management.* Modelling complex and interconnected supply chain systems to understand critical links, how inventory should be distributed, scheduling deliveries, and evaluating instability and robustness in a supply network.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Agent-based Simulation, Petri-net Simulation, Distribution Simulation, Simulation Gaming.

- *Workforce planning.* Modelling the impact of different shift patterns and staffing levels, the impact of training and cross-training (training employees in multiple jobs or processes so they can be more easily reassigned), and the impact of increased workforce versus investment in new equipment.

Modelling Methods: Discrete Event Simulation

- *Maintenance management.* Modelling proposed maintenance schemes and schedules against the predicted failures of equipment, to evaluate the trade-off between time and cost of the maintenance policy versus the likelihood and impact of a failure.

  Modelling Methods: Discrete Event Simulation, Monte Carlo Simulation, Virtual Simulation.

- *Knowledge management.* Modelling new product introduction and design, the learning curves of new processes or policies, as well as understanding the impact of organisation-level training.

  Modelling Methods: Discrete Event Simulation, System Dynamics.

- *Project management.* Modelling potential scenarios around project delivery and understanding which project management approach is most appropriate and understanding the propagated impact of delays or disruption.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Monte Carlo Simulation, Petri-net Simulation, Intelligent Simulation.

- *Organisational design.* Modelling changes to organisational structure and behaviour and analysing the effects on business outcomes.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Agent-based Simulation, Simulation Gaming.

- *Management training and education.* Modelling the outcomes of training and educating an organisation's management.

  Modelling Methods: Discrete Event Simulation, System Dynamics, Simulation Gaming, Distributed Simulation, Virtual Simulation.

- *Financial management.* Estimating the costs, risks, and outcomes of financial changes or capital acquisitions.

  Modelling Methods: Discrete Event Simulation, Monte Carlo Simulation.

- *Quality management.* A wide-ranging domain that models the outcomes of different quality paradigms on measurable outcomes in a manufacturing enterprise. Quality paradigms include continuous improvement, six sigma, total quality management, lean, and many more.

  Modelling Methods: Discrete Event Simulation, System Dynamics

5.3.4 Evaluating Modelling Tools

It can be seen that there exists a great many domains in which modelling and simulation can be applied in a manufacturing enterprise. Many software packages will cover multiple domains in a single package, or may be targeted to a specific sector. The range of modelling tools is constantly changing, and though some examples are listed below, it is important to evaluate a package against your own needs and to understand how well it fits within your business. Some criteria for evaluation include [8]:

- Handling heterogeneous information related to the design of products, processes and resources, and their operation during the manufacturing process. If a model cannot interpret the variety of information provided to it, the model cannot function.
- Integration of knowledge and information from different tools and techniques working at different levels of detail. Models and simulations work at different levels of accuracy and detail in accordance with their needs and requirements. Combining information from multiple sources and other models is required to make meaningful conclusions.
- Maintenance of the virtual representation of the manufacturing system so that it can be constantly synchronised with the physical counterpart. A model is only useful if it accurately models the physical world, so it is critical a tool makes it easy to update and change the model as the physical equivalent changes.
- Enable shop-floor engineers and technicians to make useful conclusions without the need for modelling and simulation specialists, to ensure the information is available to those who need it.
- Decrease of investment and operating costs, saving the company money if they take advantage of the model instead of adopting a "try and see" approach to process planning and design.

Modelling and simulations are highly effective tools for predicting the outcome of a manufacturing process, production line, or supply chain. However, most techniques and packages are best suited to specific areas, and have their own strengths and weaknesses. The best simulations and models are comprised of multiple smaller simulations and models sharing information to provide a more accurate level of prediction. To enable this approach, effective platforms support interoperability between digital factory tools. Not all platforms do however, and one should be cautious adopting or developing a tool that does not support the following [8]:

- A *common and standard data model* for representing entities related to production systems, resources, processes and products. Use of standards makes sharing data, or translating the data between different formats considerably simpler.
- A *shared data storage* accessible by different digital factory tools to retrieve and contribute with data. Standard data models are important, but the data needs to be stored somewhere that is easily accessible. Cloud data storage is increasingly common for this, allowing simple data access from anywhere.

- A *middleware* able to access shared data and correctly interpret/convert these according to the data model employed. It is less common for simulation models to communicate directly. Instead, middleware software acts as an intermediary.

Manufacturing simulation software must strike a balance between three key attributes:

- *Effectiveness*: How accurate are the results of the simulation, and how well does the software achieve the goals of the user?
- *Efficiency*: How long does it take the user to set up the simulation?
- *Ease of Use*: How easy is it to use the software? How steep is the learning curve?

An ideal simulation software will have excellent outcomes in all three aspects, but in practise different software packages will prioritise different attributes. In addition, simulation software usually specialises in different areas, such as line balancing, scheduling, supply chain modelling, inventory modelling, and many more. Choosing a package is about understanding what you are trying to model, and what attributes you want to prioritise. Other features you may want to consider when selecting a simulation package include:

- *Confidence:* Are there examples of successful implementations of this software package?
- *Cost:* How expensive is the package? Does support cost extra? Is it a one-time cost or a subscription?
- *Model runtime:* How long does the software take to produce an answer?
- *Reporting:* How are the results of the simulation reported, and is it a form that is useful to your business?
- *Support:* Is there a help desk you can ask for help? Is training available?

5.3.5 Example Manufacturing Modelling and Simulation Tools

The following is not an exhaustive list of modelling and simulation tools but serves as an example of the range of tools available. Most packages will cover several related domains described in section 5.3.3 so a single piece of software can solve multiple challenges in a common area. It's important to note that the field is constantly changing, and before committing time and money to a solution an enterprise should research currently available tools, and evaluate them with the questions and criteria described in section 5.3.4.

- *ARENA* is a simulation software that provides a fast, easy and intuitive way of building of a manufacturing process. It is based on drag-and-drop elements and structures with 2D and 3D visualisations. It also provides a dashboard for manufacturing optimisation, identify process bottlenecks, improve logistics and evaluate potential process changes.

- *AutoMod* is a 3D simulator able to model, analyse and emulate large and complex manufacturing, distribution and material handling systems by providing a simulation language. It has been largely employed in automation sector related to industries in automotive, airport, postal lines, warehousing and material handling.
- *Dassault Systèmes DELMIA* is a complete digital manufacturing and simulation software package, enabling users to analyse simulated and live production performance and document results for decision making. It features a collaborative 3D digital factory environment for process flow simulation and analysis, accuracy and profitability. DELMIA offers a flexible, object-based, discrete event simulation environment combined with visualisation and a colossal range of features and options.
- *Frepple* (Free Production Planning) is an open source supply chain planning software. It focuses on production planning constrained by machine/operator capacities, material availability and lead times. In addition, it provides inventory planning and demand forecasting.
- *Predator* offers stand alone or fully integrated manufacturing solutions for lean manufacturing and automation. Included in the suite are a variety of modelling and simulation tools primarily focussed on CNC machining. Predator Virtual CNC is an example of a resource domain simulation allowing for the verification and optimisation of CNC G-code offline before beginning the physical process.
- *FlexSim 3D* is a simulation software designed for modelling processes, including manufacturing, packaging, warehousing, material handling, etc. It imports relevant processing objects produced with CAD-based physical layouts. FlexSim allows end users to test all options in order to find the best combinations of operational characteristics to optimise performance and reduce costs. In addition, it provides support through accurate 3D animation and statistical reporting to both run "what-if" scenarios and make informed decisions.
- *Simio* provides a system composed by intelligent objects representing physical components such as forklifts and conveyors. This manufacturing solution encompass a set of industries like discrete manufacturing, automotive, consumer packaged goods, metals and plastics. The application areas are design of green field production plants, process improvement using Six Sigma and Lean Manufacturing, production planning and scheduling.
- *Tecnomatix* is Siemens' suite of manufacturing simulation technologies. It closely ties in with their automation products, allowing the results of simulations to be applied directly to (for example) PLC controllers, a process called "virtual commissioning".
- *Lanner Witness* is a process simulation package utilising both discrete and continuous event modelling, making it applicable to a wide range of modelling tasks. It also features an extensive library of 2D and 3D models of manufacturing resources, enabling both simulation and visualisation of proposed manufacturing lines.

5.4 Conclusions

Modelling tools are extremely valuable in the manufacturing domain. Manufacturing systems are extremely complex and extremely valuable, so understanding how they work, what the effects of proposed changes will be, and mitigating against risk and uncertainty are in term valuable pursuits. Use of these tools enables more complex calculations and simulations to be undertaken than performing the analysis by hand with mathematical formulae.

However, a manufacturing system is a dynamic, evolving, changing thing, and any model of a system must be accurate to give useful results. Even with modelling tools, creating an accurate useful model is time consuming and complex, and there is a risk that the final model may no longer accurately reflect the manufacturing system. There is also the issue of obtaining the data required to make the model accurate in the first place.

A *Digital Twin* is a live, automatically updating model of a manufacturing system which can perform simulations. Rather than being created offline by a worker, it is linked directly to the physical twin, and uses live data gathered from the real-world system to keep the model accurate and updated. Creating a digital twin is more involved that an offline model, but the cost of keeping it updated is significantly reduced. The next chapter will discuss digital twins, and how they can be implemented.

5.5 References

[1] S. Santoso, E. J. Powers and W. M. Grady, "Power quality disturbance data compression using wavelet transform methods," *IEEE Transactions on Power Delivery*, vol. 12, no. 3, pp. 1250-1257, 1997.

[2] M. K. Jeong, J. C. Lu, X. Huo, B. Vidakovic and D. Chen, "Wavelet-based data reduction techniques for process fault detection.," *Technometrics*, vol. 48, no. 1, pp. 26-40, 2006.

[3] G. Press, "Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says," Forbes, 23 March 2016. [Online]. Available: https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/.

[4] NIST/SEMATECH, "Engineering Statistics Handbook," 30 October 2013. [Online]. Available: https://www.itl.nist.gov/div898/handbook/. [Accessed 2 11 2020].

[5] L. Monostori, J. Váncza and S. R. Kumara, "Agent-based systems for manufacturing," *CIRP annals*, vol. 55, no. 2, pp. 697-720, 2006.

[6] P. Vrba and V. Marcík, "Simulation in agent-based control systems: MAST case study.," *IFAC Proceedings*, vol. 38, no. 1, pp. 145-152, 2005.

[7] A. Oyarbide, T. S. Baines, J. M. Kay and J. Ladbrook, "Manufacturing systems modelling using system dynamics: forming a dedicated modelling tool.," *Journal of Advanced Manufacturing Systems*, vol. 2, no. 1, pp. 71-87, 2003.

[8] T. Tolio, M. Sacco, W. Terkaj and M. Urgo, "Virtual factory: An integrated framework for manufacturing systems design and analysis.," *Procedia CIRP*, vol. 7, pp. 25-30, 2013.

[9] Institute of Electrical and Electronics Engineers Standards Association (IEEE SA), *IEEE 1278.1-2012 - IEEE Standard for Distributed Interactive Simulation--Application Protocols*, 2012.

[10] M. Jahangirian, T. Eldabi, A. Naseer, L. K. Stergioulas and T. Young, "Simulation in manufacturing and business: A review," *European Journal of Operational Research*, vol. 203, no. 1, pp. 1-13, 2010.

Chapter 6
Digital Twins and Intelligent Decision Making

Jack C Chaplin, Giovanna Martinez-Arellano and Andrea Mazzoleni

6.1 Introduction

Twin (noun): something containing or consisting of two matching or corresponding parts.

- Oxford University Press

Chapter 5 discussed ways of simulating and modelling manufacturing systems an offline manner. Offline in this context means the simulation is running disconnected from the real system and is reliant on the user to update parameters and data to maintain the accuracy of the model. In contrast, this chapter details online simulations, where the model is connected directly to the physical system and is automatically updated as the system changes. This approach is common called a *digital twin*. In addition, this chapter will discuss decision support systems, which are software packages intended to enhance the decision making process discussed in Chapter 4, and to make complex problems solvable.

A digital twin is a simulated replica of a complex system. Unlike more

J.C. Chaplin (✉), G. Martinez-Arellano
Institute for Advanced Manufacturing, University of Nottingham, Nottingham, UK
e-mail: jack.chaplin@nottingham.ac.uk, giovanna.martinez@nottingham.ac.uk

A. Mazzoleni
AFIL – Associazione Fabbrica Intelligente, Milan, Italy
e-mail: andrea.mazzoleni@afil.it

J.C. Chaplin et al. (eds), *Digital Manufacturing for SMEs*
DOI: https://doi.org/10.17639/91WZ-CP62

conventional simulation, the digital twin is connected in real-time to the physical equivalent and collects generated data. This allows the digital twin to improve its accuracy based on the real system, and for the digital twin to analyse the system or perform tests which would be too costly or time consuming to run on the real thing. Digital twins are used to model complex systems and have their origins in the modelling of air and spacecraft.

The first definition of a digital twin (then just a nameless concept) was coined by Michael Grieves in 2002 [1] as a concept for product-lifecycle management. After many different names including Mirrored-Spaces Model, Information Mirroring Model, and Virtual Twin, the now established name “digital twin” was subsequently introduced by John Vickers in a NASA report published in 2010 [2] and has become the standard terminology. A review of literature published between 2012 and 2016 by Negri and colleagues [3] as part of the MAYA project found 16 different definitions for digital twins across four research fields: aeronautics and space, robotics, manufacturing, and informatics. Some of the proposed definitions can be found in Table 6.1-1.

Year	Definition
2010	An integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin. The digital twin is ultra-realistic and may consider one or more important and interdependent vehicle systems [4].
2012	Ultra-realistic, cradle-to-grave computer model of an aircraft structure that is used to assess the aircraft’s ability to meet mission requirements [5].
2015	Very realistic models of the process current state and its behaviour in interaction with the environment in the real world [6].
2016	The simulation of the physical object itself to predict future states of the system [7].
2018	Digital twins are realistic digital representations of physical things [8].

Table 6.1-1 Definitions of digital twins in chronological order.

Grieves described the digital twin model at its most basic with the diagram seen in Figure 6.1-1. A digital twin always corresponds to a *physical twin* – the actual physical in stance of the product that is being simulated with the digital twin.

The physical twin and digital twin are connected. This is what distinguishes a digital twin from a more conventional simulation. Data is collected from the physical twin in real-time with sensors and used to improve and optimise the digital twin. The digital twin can run analysis methods, and multiple possible scenarios can be tested digitally. The processed information gained from these can be fed back to the physical twin to optimise the real-world performance.

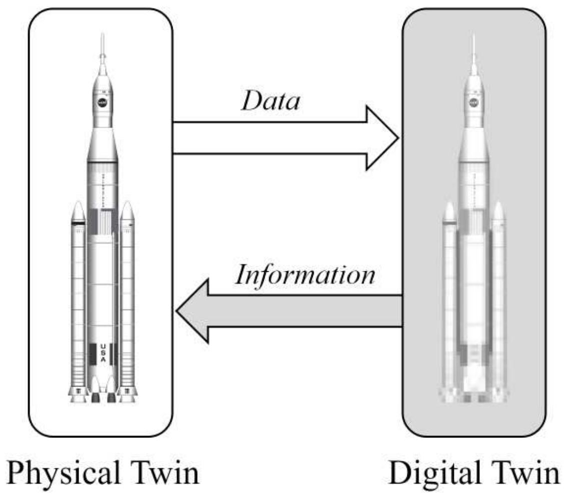


Figure 6.1-1 The digital twin is the virtual counterpart to the physical twin, and is used to better understand complex systems such as manufacturing lines or rockets. Data collected with sensors in the physical twin inform the digital twin and improve accuracy, and insight gained by analysing the digital twin can be used to control the physical twin. Image rights: Author, adapted from [1].

Simple systems can be simulated to a sufficient level of detail that the added complexity of a real-time connect may not be required. Digital twins are typically reserved for complex systems where creating an adequate simulation offline is not possible, and the simulation must be improved over time.

"Airplanes, rockets, manufacturing floor equipment, and even automobiles have or will have [digital twin instances]. Paper clips will not".

- Michael Grieves

The key differentiating factor between a conventional model and a digital twin is that the physical system is feeding real-time data into the digital twin to update the model. Though significant structural changes to the physical system may require manual intervention, data such as manufacturing throughput, buffer sizes, up/down times etc. can be automatically collected and used to keep the digital twin updated. Any simulations that run on the digital twin's model use the latest, most accurate information.

Before digital twins are discussed further, it is important to understand how it is possible to connect a digital twin simulation to the physical twin – the online aspect which differentiates the approach from the offline approaches in Chapter 5. This is achieved using sensors to monitor the physical twin, and report back its state, and section 6.2 will introduce sensors and the modern development of smart sensors.

6.2 Sensors

6.2.1 Introduction to Sensors

The manufacturing industry must achieve sustainable growth and increases in productivity to remain competitive on the global stage. Increasingly, access and exploitation of manufacturing data is contributing to these aims, enabling quicker, more effective decision-making.

One of the key technologies for data exploitation is the *Internet of Things* (IoT), which embeds sensors and communication equipment in manufacturing machineries and lines, each collecting and transmitting data to the manufacturing enterprise's network. The application of these techniques to manufacturing is also sometimes called the Industrial Internet of Things (IIoT).

IoT is a recent development and is a broad term for a set of technologies, systems, and design principles associated with using Internet-connected things that monitor and manipulate the physical environment. IoT connects sensors and actuators, to information and communication technologies (ICT) systems via wired or wireless networks. The most important technology at the device or hardware level in this network infrastructure is the *sensor* technology, as it is the most basic means for collecting and controlling data in real time.

A sensor is a device that observes and measures a physical property of a natural phenomenon or man-made process, and converts that measurement into a signal. That signal can then be reported to a worker, used to trigger an actuator, or collected for analysis. Sensors have played a role in manufacturing since their invention. They provide a means for gathering information on manufacturing operations and processes as they are performed. Typically, this means some property of the process (temperature, speed, location etc.) is converted into an electrical signal and collected by the process controller, often a programmable logic controller (PLC). The controller may use the sensor reading to modify the process, or the signal may just be logged for later inspection.

Process data are records of the processing performed to create products, including details such as the program performed (for a CNC machine), the user-set process parameters, and recorded data from sensors such as vibration, temperature, or cutting force (depending on the process). Correlating this with the individual product ID allows for analysing this data relative to the quality outcomes or service life of the specific product. Most traditional sensors convert their measured property into electricity, and hence require a wire or cable to connect with the external instruments that record this value. The cable can be copper wire, twisted pair or fibre optic.

The data from sensors must somehow be sent to a computer for interpretation or storage. Often this will be a local PLC or data capture device where the sensor data can immediately be acted upon. It may be a remote server where data can be logged in a database or simple spreadsheet. *Fieldbus* is the term used for industrial computer networks. Many standards exist including Ethernet, Industrial Ethernet, Controller Area Network (CAN), Process Field Bus (Profibus) and a wide variety of vendor specific technologies. Use of these will depend on equipment compatibility and the requirements for process control.

Wireless sensors offer flexibility of installation, resulting in improved process monitoring and control while simultaneously offering reduced installation and maintenance costs. Industrial applications offer a broad scope for growth in wireless sensor use, but this growth cannot be achieved without overcoming some of the key challenges facing the market:

- *Evolving standards*: New wireless communication technologies are still under research and development. Some of these standards might not be compatible with others, limiting the interoperability of the network.
- *Network size*: The application or use case scenario will determine the size of the network. Some wireless approaches are better for smaller or larger scales of deployment, and if the application demands a change in location, number of nodes or adjustments to topology the approach may have to be changed.
- *Open wireless frequency bands*: With the propagation of wireless technologies such as mobile phones and Wi-Fi, there is a lack of open frequency bands. Currently, most wireless sensor network devices operate in unlicensed bands such as 915 MHz and 2.4 GHz, and reliable communication can be affected by interference from other devices operating in the same frequency band.
- *Industrial safety*: Making a wireless sensor system fail-safe depends heavily on the type of application in which the wireless sensor is used.

On top of these networks, many higher level *machine to machine communication protocols* exist for exchanging process data (rather than control data) in industrial automation settings. These include Open Platform Communications – Unified Architecture (OPC UA), MTConnect, Data Distribution Services (DDS), and MQ Telemetry Transport (MQTT). The standards will significantly simplify the acquisition of process data by handling many networking aspects and providing a common standard for data.

6.2.2 Types of Sensors

Sensors convert physical phenomena into signals, and are an example of a primary data source (see section 5.2.2). Almost any physical phenomena can be measured with a sensor, but the most common types used are listed below.

- *Temperature*: This sensor gives temperature measurements as an electrical signal (e.g. voltage) proportional to the temperature measurement. There are various electrical temperature sensors such as the thermistor, thermocouple, resistance thermometer and the silicon band gap temperature sensor, and each have different properties which make them better or worse for certain situations.

- *Force, pressure:* There are devices that convert variations in applied force or pressure into an electrical signal. There are two principles that have become dominant in force measurement; strain gauge-based sensors and piezoelectric force sensors. Strain gauge sensors contain an electrically conductive foil which deforms as force is applied. This deformation changes the electrical resistance of the spring, and this property is converted to an electrical signal. Piezoelectric sensors contain two crystal disks with an electrode foil mounted in between. When applying force, this generates electrical charge that can be amplified and used as a signal. Piezoelectric sensors are first choice for fast measurements of small forces while strain gauge-based sensors are superior when larger forces are involved [9].
- *Level*: These sensors detect the level of liquids and other fluids and fluidised solids such as powders, and are common in industrial process control. Examples include hydrostatic sensors (which measure water pressure to deduce water level), and optical level sensors which use the attenuation of light to measure fluid depth.
- *Acceleration, vibration*: Motion can also be detected with sensors. Accelerometers measure acceleration in a single direction, and are often combined into units with two (bi-axial) or three (tri-axial) accelerometers at right angles to detect the direction of acceleration. Accelerometers can also be used to detect vibration. Accelerometers are often Micro-electro-mechanical Systems (MEMS) that converts the motion of a small mass into an electrical signal with piezoelectric crystals. These crystals generate small electrical signals when subjected to mechanical stress.
- *Orientation*: Gyroscopic sensors detect rotation around a single axis, and like accelerometers are often combined into bi-axial and tri-axial sensors. Tri-axial accelerometers can in fact detect the orientation of a stationary object as they detect Earth's gravity. However, accelerometers cannot do this while the object is moving. Tri-axial accelerometers and tri-axial gyroscopes can be combined into Inertial Measurement Units (IMUs). IMUs are used to determine the orientation of mobile phones, adjustment of suspension in cars, and control of aircraft.
- *Proximity*: Proximity sensors detect if an object is physically close to the sensor, and effectively detect the presence or absence of objects. They can be implemented with a variety of technologies including optical, inductive, magnetic and capacitive methods. They are widely used in industrial automation like conveyor lines for counting and jam detection, and in machine tools for safety interlock and sequencing.
- *Position*: Position sensors detect the mechanical position of objects. They often use similar technologies to proximity sensors, but can determine if an object is present, but also how far the object is from the sensor. They can measure absolute distance, angular rotation, and tilt angles depending on the technology used and application. These are commonly used to control the motion of robots and actuators, rotation of valves, and angles of actuators.

- Other sensors include *humidity*, *gas*, *biosensor*, *photoelectric*, *flow*, and many others, far more than can be described in this book.

A discussion of sensors with additional detail on their applications in robotics can be found in Chapter 8.

6.2.3 Smart Sensors

At their core, sensors are often very simple mechanical or electronic devices to convert physical stimuli into electrical signals. However, the progress and miniaturisation of technology, combined with the move towards the Internet of Things means many sensors (and a particularly high percentage of wireless sensors) have more features than just signal transducing.

Smart sensors are microprocessor driven and include features such as communication capability and on-board diagnostics that provide information to a monitoring system and/or operator to increase operational efficiency and reduce maintenance costs. They can perform some computation locally, reducing the amount of information that must be transmitted. This is particularly important for wireless sensors which have less bandwidth available than wired sensors. Common characteristics of smart sensors include:

- *Signal conditioning* that preserves integrity and ensures isolation in harsh industrial environments, smoothing out noise and amplifying weak signals.
- Using *local computing power* to process and interpret data locally; make decisions based on the physical parameters measured, adjust parameters autonomously, and be selective about what data is transmitted.
- *Built-in diagnostics* for simplified troubleshooting and maintenance.
- Complying with a *variety of communication standards* such as Wi-Fi, Bluetooth and Zigbee, rather than being constrained to a single technology.

Sensors have been utilised in manufacturing for decades, but smart sensors offer new and diverse benefits that can potentially lead to greater profitability and productivity. Principally, smart sensors offer richer data – the context and relevance of the data is recorded in addition to the signal itself, and smart sensors can be selective about the data which is sent and where it is sent. Applications of smart sensors to manufacturing include:

- *Aggregation and collection of big data*: Big data and data mining uses extremely large quantities of data to “mine” for insights. Smart sensors facilitate this in three ways;
  - Firstly, smart sensors are often simpler to implement than conventional sensors, as they include all the necessary equipment and wireless communication protocols to connect to existing manufacturing execution systems in a single package without running lots of cables.

  - Secondly, the ability for smart sensors to pre-process data allows them to aggregate larger quantities of data and transmit them effectively. Big data and data mining typically relies on very large quantities of data to achieve the best insights, and simple sensors may not be able to gather and share the required data volumes.
  - Thirdly, smart sensors can often communicate with one another, allowing data from sensors to be correlated to each other, facilitating the analysis of data from multiple sources.
- *Quality control:* Quality is critical to manufacturing competitiveness, and quality control must be an integral element of the manufacturing process. Identifying problems early reduces or removes the cost of scrapped parts of expensive re-work. Keeping processes in control requires monitoring, and this is traditionally done with conventional sensors to feed into control charts and statistical quality control methods. Smart sensors however are able to pre-process data to detect anomalies as they occur, and provide real-time alerts and flags when processes are deviating from nominal. The additional features of smart sensors allow them to send richer information than just a signal relative to the physical parameter, and to make decisions.
- *Improving and automating logistics and asset management:* Battery powered and communicating wirelessly, smart sensors can track the location of assets, vehicles, inventory, or people. The data can be used by manufacturers to track logistics and the supply chain, monitor the movement and utilisation of assets, and find lost parts and tools.
- *Regulatory compliance:* Many manufacturing sectors are highly regulated, with stringent rules on data collection and data storage to ensure compliance. Compiling the necessary reports from sensor data, logs, and records across multiple systems can be extremely time consuming and labour intensive. A well-managed data management system combined with smart sensors can dramatically simplify this process, with sensors able to collect environmental data such as temperature and humidity, as well as equipment utilisation data such as energy consumption, hours of operation, and maintenance information.

6.2.4 Product Tracking

The fundamental function of sensors is to detect physical phenomena and convert it to electronic signals, but sensors may not always be as simple as converting (for example) a temperature to a voltage. A carefully set up optical sensor can become are barcode reader, allowing for richer and more complex information about the world to be collected. *Product tracking* is the use of sensors to determine what specific item is at a location, allowing for unique products to be tracked and more granular information to be collected. This information could then be fed into a digital twin, giving it live data on the products being produced.

Product tracking requires three key features to be implemented:

1. A way of uniquely identifying the product (or part, or asset) being tracked.
2. A way of acquiring data about the product (such as measurement data or process data) that can be correlated to its identifier.
3. A way of storing the acquired data for future reference.

There is no inherent requirement for any of these three features to be automated. The identifier could be a number written on the part with a marker pen, the acquisition of data could be process parameters noted on a sheet of paper, and the data could be stored in a plastic folder for later reference. This may seem an exaggerated example, but this approach remains extremely common in modern manufacturing companies both large and small. Handwritten job cards are still the most common standard used, and can function extremely effectively. However, every instance of manual data entry raises the chances of a mistake being made. Manual data entry also represents a skilled worker performing a low skilled job, and their time could often be better spent.

Automated product tracking technologies can fit into each of these three categories. They enable the identification, acquisition, and storage of data with less worker effort or no effort at all. They also significantly reduce the probability of error, and make data recall and analysis easier by enforcing common data standards. Automated product identification, also referred to as *Automatic Identification and Data Capture* (AIDC) are technologies that allow data to be entered into a computer system with little to no human involvement. The most common example of AIDC is the use of barcodes in retail stores. The use of these codes shows the three critical stages of AIDC:

1. *Data encoding:* Human-readable characters or numbers are rarely the most efficient way for AIDC technologies to represent data. Instead, data (such as the product number) is encoded in some way, such as the width of the bars in a barcode.
2. *Data reader:* A device able to reliably read the encoded data, and convert into an appropriate data format for transmission. The barcode scanner is an example of a data reader.
3. *Data decoder:* The data decoder converts the signal from the data reader back into the original characters that were encoded. For example, this would give the barcode number representing the product. The Point of Sale (POS) software can then look up the number in a database and retrieve the name and price of the item.

AIDC uses many different technology types, with different advantages and disadvantages. These are detailed in the table below [10]:

Technique	Performance	Advantages	Disadvantages
Human Manual Entry	Sensor Type: *Manual* Entry Time: *Slow* Error Rate: *High* Cost: *Low*	Low initial cost. Simple to set up. Highly adaptable.	High ongoing cost (cost of a worker's time). Slow entry speeds. Prone to error or omission.
Biometric	Sensor Type: *Optical (typically)* Entry Time: *Medium* Error Rate: *Low* Cost: *Medium*	Intuitive. Rapid identification of people. Appropriate for security applications.	Niche applicability. Often unpopular. Initial setup requires physical presence of people.
OCR (Optical Character Recognition)	Sensor Type: *Optical* Entry Time: *Medium* Error Rate: *Medium* Cost: *Medium*	Data remains human readable.	Characters must be printed. Low data density. Error rate condition (e.g. lighting) dependant.
Machine Vision	Sensor Type: *Optical* Entry Time: *Fast* Error Rate: *Application Dependent* Cost: *High*	Versatile application. Equipment can be reused for different applications. High speed. Can read other forms of data encoding such as barcodes.	Success highly dependent on application and quality of implementation.
1D Barcode	Sensor Type: *Optical* Entry Time: *Medium* Error Rate: *Low* Cost: *Low*	Cheap to implement. Versatile. Easy to print and affix barcodes.	Lower data density than 2D barcodes.
2D Barcode	Sensor Type: *Optical* Entry Time: *Medium* Error Rate: *Low* Cost: *High*	High data densities. Reliable and versatile. Barcodes take up more space than 1D, but still easy to use.	Equipment cost higher than 1D barcodes, so consider if 2D is necessary.
RFID (Radio Frequency Identification Tags)	Sensor Type: *Electro-magnetic* Entry Time: *Fast* Error Rate: *Low* Cost: *Medium*	Functional without line of sight. Versatile: Read/write capable tags and battery powered tags exist. High data density (with more expensive tags).	More expensive per-use than optical methods. Tags with high data capacity or read/write capability are expensive. Quality and read range can degrade in metallic environments.
Smart Cards (subtype of RFID)	Sensor Type: *Electro-magnetic* Entry Time: *Fast* Error Rate: *Low* Cost: *Medium*	Identifies people without biometrics. Can be implemented without the person being present. Quick and simple to use.	Can get lost / stolen / abused in ways biometrics cannot.
Magnetic Stripe	Sensor Type: *Electro-magnetic* Entry Time: *Medium* Error Rate: *Low* Cost: *Medium*	High density of data. Read / write capability.	Physical contact required to read data. Damaged by electromagnetic fields.

Table 6.2-1 A summary of common AIDC technologies and their properties.

It is important to note that there is no “best” technology – these technologies are all in use because they offer unique advantages and disadvantages. These technologies may also be implemented in isolation or in combination for different results. When considering which AIDC to use, the encoding, reader, and decoder should all be considered. For example, what is the available area for the encoding? Does the encoding need to be human readable as well as machine readable? How much data must be encoded? In what conditions must the encoding be read – well illuminated, or irregular and dark? Can the encoding be seen, or is it hidden on the product?

It’s also important to note that though new and emerging technologies such as Machine Vision may offer significant new capabilities, tried and tested technologies such as 1D barcode remain in such broad use due to their efficient, reliable, and cost effective operation.

6.2.5 Sensors Conclusions

Sensors are of course important in manufacturing for the control of processes, and many are often built into equipment and integrated into the control systems. The positions of actuators, presence or absence of parts, rotational speeds of spindles or temperatures of chemical processes are all examples of process features that must be monitored for the process to be successful, and often you may not know the sensors are there at all.

However, a critical part of creating an effective and useful model or digital twin is access to accurate data about the system you are trying to model. Sensors integrated into processes can provide useful data for a model, but often the key missing data is external to processes. The movement of parts around a system, the use of tooling or materials, inventory management, and any processes that must collaborate can require additional sensors to be implemented to capture the performance and parameters.

Sensors can either be read directly (primary sources) or fed into a secondary control system such as a PLC or embedded computer (secondary source) depending on application. Data from these can be logged in a database, spreadsheet, bespoke monitoring software, even manually recorded, and then the data analysed with the data processing cycle discussed in section 5.2. The difference between the conventional modelling approaches from section 5.3 and digital twins, is that conventional modelling takes a snapshot of a system and builds a model that remains static, whereas a digital twin is connected to the data logging, and can constantly update the model.

6.3 Digital Twins

6.3.1 Categories of Digital Twins

"Digital Twin" is an increasingly common buzzword in manufacturing and is often misused to refer to offline modelling and simulation approaches. As an emerging area, there are few formal standards for digital twins. The ISO/IEC JTC1/AG 11 Digital Twin [11] working group and ISO/TC 184/SC Industrial Data [12] technical committee are working on standards such as ISO/DIS 23247-1 [13] Digital Twin Framework for Manufacturing, but they are currently not completed. The German Plattform Industrie 4.0 network also promoted their asset shell standards for digital twins, but this serves as a starting point rather than a complete reference architecture. Due to the lack of standardization, the concept of the digital twin is still diverse, and often understanding is driven by providers of digital twin software tools. However, some authoritative groups have proposed different classifications of digital twins. When evaluating software that claims to be a digital twin, it is worth considering where it falls under these classes.

According to Kritzinger *et al.* [14] from Fraunhofer Austria Research, the level of data integration between the physical and digital counterpart enables the classification of the approach:

1. A *Digital Model* uses no automatic data transfer between the real/physical asset and the virtual model, only manual data transfer. Most would not consider this a real digital twin as the manual data exchange prevents the digital object having access to real-time data.
2. A *Digital Shadow* includes a one-way automatic flow of data from the physical asset to the digital representation.
3. A *Digital Twin* has data automatically flowing in both directions between the physical asset and the virtual model. In this instance, the digital twin is able to control the physical object based on the decisions of digital twin.

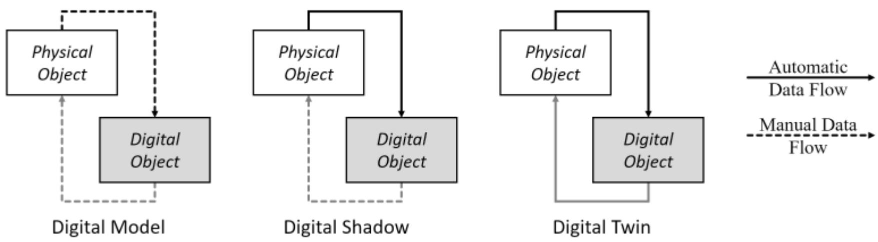


Figure 6.3-1 The classes of digital twins according to Kritzinger. The more automated the flow of information, the closer the result is to a "true" digital twin [14].

The Advanced Manufacturing Research Centre (AMRC) at the University of Sheffield in the UK has proposed classifying digital twins based on the capabilities and added value given by the digital twin [15]. The core functionality is presenting data in context, and the type of twin then depends on whether the optional added value features includes data analytics, control over the physical asset, and/or predictions via simulations. The three classes are:

- *Type 1 Supervisory Digital Twin*: A passive monitoring twin, where data is received from the physical system and combined into a single model. This model can be used to identify warning limits or thresholds on variables that indicate issues. Approximately equal to a Digital Shadow in Kritzinger's classification.
- *Type 2 Interactive Digital Twin*: The digital twin is able to influence the physical twin by setting parameters to improve key performance indicators with simple control algorithms. Control may be complete or partial. Approximately equal to the full digital twin in Kritzinger's classification.
- *Type 3 Predictive Digital Twin*: Integrates simulation and analysis methods to predict performance based on process data from the physical twin and use these to optimize processing parameters, and then proactively make these adjustments as needed in the physical twin.

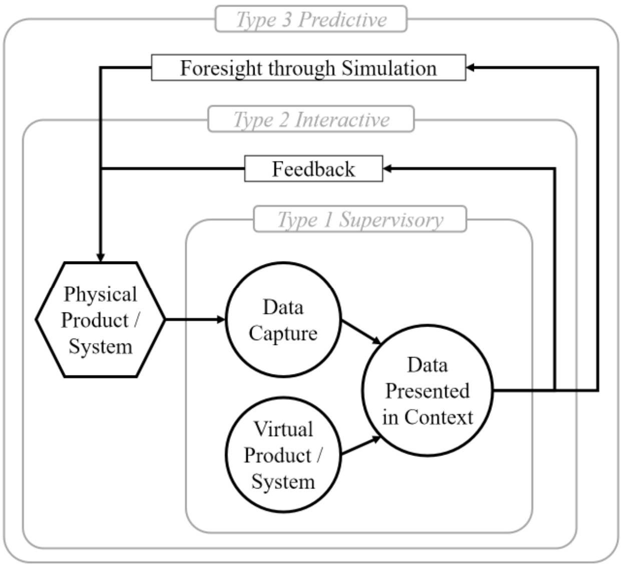


Figure 6.3-2 The types of digital twins are layered on top of one another, according to the AMRC classification [15].

6.3.2 Elements of Digital Twins

Due to the current lack of standards, the components and technical requirements of a digital twin vary with the type, classification, and supplier of the digital twin. However, some common elements are emerging.

- *Mandatory: Real-Time Connected Physical Twin.* Sometimes under considered, a digital twin requires a physical twin, typically a product or system. The physical twin can be a dumb and/or uncommunicative, or a smart product/system capable of machine-to-machine (M2M) communication or of communicating with humans. The digital twin will require sensors (either integrated into smart products / systems, or IoT solutions added afterwards), the communication standards to share the data with the digital twin, and the ability to do so in real-time. What qualifies as real-time is application dependant, but usually considered in the order of milliseconds.
- *Mandatory: Model.* The digital twin requires a virtual equivalent of the physical twin which gives context to the collected data, differentiating it from a database or data lake. Despite this difference, the digital twin still requires storage for collected data, and a database may be the underlying technical implementation of this. The model is one of the harder elements to define as it is highly application specific – a model of a factory may be implemented differently to a model of a jet engine for example. 3D models of the physical twin are a common consideration, but not mandatory.
- *Optional: Analytics and Simulation*. One of the most commonly envisaged use cases of digital twins is to use the twin to perform analysis and understand how better to control the physical twin. This is typically achieved through the use of simulation and analytics. It may seem strange that this is listed as optional – what is the point of a model if you never run simulations on it? However, as described by the AMRC classification [15], a digital twin can just be a way to present data in context without any analytics performed directly. Analytics and simulations remain a very common desirable feature of course but are not mandatory. Multiple different analytics and simulation techniques can be included in digital twins depending on the application, with the intention to identify problems or opportunities in advance based on the incoming data and state of the physical twin.
- *Optional: Control.* Giving a digital twin the ability to control the physical twin is not mandatory (see type 1 supervisory digital twins) but – like analytics and simulation – is a very common envisaged use. Sensors in the physical twin enable the data capture, but similarly actuators can enable the digital twin to change parameters or other elements of the physical twin's control system. This would typically be in conjunction with analytics or simulation tools to make the decisions to change parameters, which are then enacted in the physical twin.

6.3.3 Applications of Digital Twins

Digital twins can be used in a wide range of applications, including vehicle automation, power generation, traffic modelling and urban planning, healthcare, and more. Here however, we will focus on some applications specific to manufacturing [1].

- *Production line replication:* One of the most commonly envisioned applications of digital twins to manufacturing is the creation of a digital replica of a production line, or entire factory. Manufacturing simulations are not new but rely on a high degree of understanding of the process being simulated. The real-time connected nature of digital twins enables the live simulation to improve over time, to be monitored and optimised based on the real-time state of the system, and to respond to new and unexpected events. The digital twin serves two primary goals here. Firstly, the captured data for the production system is unified and kept in one place, simplifying monitoring and enabling historical snapshots of the system behaviour to be retrieved if a fault occurs. Secondly, the digital twin can run *front-running simulations*, using the faster-than-real time nature of simulations to “fast forward” and predict the state of the system in the future [16].
- *Product replication:* Even complex manufacturing systems may produce simple products that do not require a digital twin. However, when the product is as complex as the system that creates it, a product digital twin can offer some advantages. Similarly to production line replication, using a digital twin for a product aids in collecting all the data generated in its creation in one place, which is useful for highly regulated domains such as aerospace or pharmaceuticals. Digital twins can also be integrated with product lifecycle monitoring systems to ensure the expected processes are performed on the product, particularly when there are large sources of uncertainty in the manufacturing process or when configurations are changed as the product is being manufactured.
- *Preventative maintenance:* Preventative maintenance is a key technique for preventing costly breakdowns by scheduling maintenance during breaks in production to replace worn components before their expected failure. Preventative maintenance typically is regularly scheduled based on operator experience and the equipment manufacturer’s guidance. Predictive maintenance monitors equipment condition with sensors to improve the estimation of when maintenance must be performed. This both helps determine if equipment will fail earlier than anticipated, or if the maintenance can be delayed – saving money.

  A digital twin of production equipment aids in the implementation of predictive maintenance by enabling the collection and analysis of sensor data from the equipment and running analysis to predict the optimum maintenance window. Data from previous breakdowns can be used to help identify imminent new breakdowns, and real-time data from the equipment can be compared to past “ideal” data to look for deviations.

- *Robotic planning / cobotics*: Industrial robotic path programming is often developed with offline simulations to plot and test proposed programs, using software packages such as ABB RobotStudio, KUKA.Sim or Dassault Systemes DELMIA. These are excellent tools for highly predictable, repetitive tasks. Where digital twins can aid this process is where the robotic process is variable or unpredictable. This can occur when the product is large, flexible, or of an unknown quality. Unpredictability can also occur when humans are involved in the process, and the increasing popularity of collaborative robotics or cobotics where robots and humans work together on.

  Safety is critical to the successful implementation of cobotics and digital twins allow the simulation used for motion planning to be updated in real time to ensure the position and motion of the human worker is accounted for. They could also allow for virtual or augmented reality techniques to enhance the human worker's interaction and control of the cobot.

6.3.4 Examples of Digital Twin Software

Digital twins are an emerging technology, and a rapidly changing field. Without clearly defined standards for digital twins, what does and does not qualify as a digital twin is ambiguous. Moreover, many offered digital twin packages are actually multiple pieces of software which work together to implement the digital twin. These packages often use an existing modelling and simulation package (such as those mentioned in section 5.3.5) with add-ons to gather data in real time. Due to the early-stage nature of the technologies, it is recommended you research solutions and developments before committing to a specific product or package, and pay particular attention to the types and elements of digital twins so you can be sure what you're purchasing does what you expect. However, here are some examples of solutions or packages on the market.

- *Siemens Digital Enterprise:* Siemens' digital twin offering is a holistic combination of a range of their software products rather than a single dedicated piece of software, but these come under the Digital Enterprise portfolio. Depending on the application area, different tools will be used. For example, in a pharmaceutical setting STAR-CCM+ is used as the model, HEEDS for analysis, and SIMATIC SIPAT is used for monitoring quality in real time. As Siemens also provides a wide range of industrial automation and sensing equipment, integrating this approach into a real-time digital twin is simplified.
- *GE Digital Predix:* Predix is a cloud-based data platform that can gather data from IoT sources, contextualise it with a model, and use analytics algorithms to predict future events. It has multiple stated applications including component, asset, system, and process modelling.
- *Dassault Systemes 3DEXPERIENCE:* Another large portfolio of software, 3DEXPERIENCE covers product design, process planning, simulation and analytics, and product data management. Calling digital twins

"3DEXPERIENCE Twins", the package covers the range of elements required for a digital twin. Other companies have also used 3DEXPERIENCE as a basis for their own digital twin products, such as *Veristar's AIM^{3D}* which is specialised for large vessels and gas/oil platform modelling.

- *Microsoft Azure Digital Twins:* Currently in a public testing phase, Azure Digital Twins is a forthcoming cloud-based digital twin solution with a domain-neutral "spatial intelligence graph" approach to modelling, and integration with Azure's existing IoT data ingress, machine learning and AI capabilities. Microsoft have also partnered with *Ansys Twin Builder* to add additional capabilities around predictive maintenance.

6.3.5 Implementation Challenges

As an emerging area, there are several challenges for digital twins to overcome to become mainstream. These are summarized here, and are aspects to be mindful of before adopting any specific digital twin solution.

- *Standards and Interoperability:* As discussed in section 6.3.1, there is currently a lack of industrial standards for digital twins, and even when they are finalized the adoption will take time. As a result, interoperability between different vendors of digital twin solutions or components of a solution may be limited. Many current tools are based on existing suites of manufacturing software, and mixing and matching these may be costly or not possible.
- *Trust*: A key part of digital twin sales pitch is for a system to self-optimize, with analytics and simulations being run on live data and the findings being used to tune and optimize the physical twin. However this requires a level of trust on behalf of the operators – a few incorrect decisions or wrong deductions could cause digital twin adopters to sever the automated control loop and instead use the digital twin as advisory rather than fully integrated.
- *Data Quantities:* A key aspect of digital twins is to collect real-time data in a single place and be able to present it in context (see Figure 6.3-2). This approach is not dissimilar to data lakes and data warehouses but unlike these approaches digital twins require the centralized data in real-time. Depending on the volumes of data, the distance of the data sources, and the method of data transmission, sending raw data may simply not be possible. Pre-processing data before transmission may go some way to alleviating.
- *Cybersecurity:* Any company innovating in Industry 4.0 is unfortunately a target for malicious entities, from individuals, organized gangs, to nation states. Irdeto found that 79% of surveyed companies implementing IoT has suffered some sort of cyber-attack against their IoT systems in the previous year [17]. The consequences of a cyber-attack on a digital twin could be significant, ranging from disruption of the physical twin, ransomware attacks, or theft of intellectual property due to the richness of data stored in the digital twin. Typical industrial security methods such as hardware security and air-gapping are no longer

sufficient. Instead, software-based protection should be implemented, a well-defined business process for software/firmware updates should be enacted, and data encrypted where possible.

- *Implementation:* Digital twins require the involvement of a wide range of equipment and people within a business to take maximum advantage of the approach. A common buzzword for digital twins is "holistic" i.e. something that is more than the sum of its parts, but in this case the word is appropriate. Digital twins can draw data from across a business, and if the data isn't already organized or well understood the digital twin implementation process may be a painful one. Gartner [18] released a study conducted in 2017 among 202 companies, and described four best practices for implementing and maintaining digital twins:
  1. Investments in digital twin solutions could be driven by the product or process value chain, understanding why stakeholders need access to the data or control of the physical twin is key.
  2. Standardized documented procedures should be carried out during the creation to the digital twin to ensure this potentially highly complex system is well documented and understood, which will facilitate changes and upgrades.
  3. Use and access of data should be possible from multiple sources to allow interaction and evolution of the digital twin. This may require existing data silos to be broken down and standards implemented within the enterprise so that data can be accessed more widely.
  4. Proprietary software and non-standard formats should be avoided to ensure the company doesn't get locked into an approach or is unable to integrate new software components.

6.4 Decision Support Systems

6.4.1 Introduction

Throughout this book the concept of decision making has been discussed, and its importance emphasised. The role of analysis in manufacturing is to enable better decisions to be made, and therefore for the business to be more productive and more profitable. All the tools and methods we have discussed exist to inform decisions. However, another classification of tools exists that can aid with decision making – the Decision Support System (DSS).

A DSS is a software system to support decision making in an enterprise, including but not limited to manufacturing. They typically have a constrained area of interest (the *domain*) in which they aid decisions. For example, DSSs are increasingly common in the medical field to aid with diagnosis and treatment plans where the problems are far from structured. There are three classes of DSS, representing the type of support they give:

- *Passive*: Offers information and analysis to aide a human in the decision making process, but doesn't offer any direct suggestions or solutions.

- *Active*: Analyses available data to offer suggestions and solutions to the user.
- *Cooperative*: Offers suggestions and solutions to the user, but also takes feedback from the user to refine and improve the decisions and suggestions. These are rare outside of research however.

In addition, DSSs fall into four further classes based on the type of assistance they offer, and these are discussed in the following section.

6.4.2 Classes of Decision Support System

6.4.2.1 Communications-Driven

Perhaps the most common form of DSS, and one you may be using without realising that it is a decision support tool at all. *Communications-Driven DSS* (CD-DSS) facilitates decisions by allowing users to share information to collectively make a decision. They are often called Group Decision Support Systems for this reason. A CD-DSS does not generate or analyse data by itself, it makes it more available to users, including users who are distributed rather than co-located, and who are communicating asynchronously.

Examples include document sharing tools such as Google Docs or Microsoft SharePoint, and collaboration tools such as Slack or Microsoft Teams. Even teleconferencing solutions such as Skype or Zoom are sometimes described as CD-DSSs. These systems allow for users to make better decisions about unstructured and semi-structured problems by better pooling and sharing experience and knowledge, rather than the problem being tackled by a single individual. CD-DSSs are examples of passive DSSs.

6.4.2.2 Data-Driven

For semi-structured problems and even some structured ones, data may be available to inform the decision, but in a format which makes its use difficult or impossible for the decision maker to use effectively. Data-Driven DSSs (DD-DSS) take data (typically time-series data) and presents it to the user in a more informative manner, possibly after some initial analysis of the data. The data will be sourced from a company database or databases, and also sometimes include external data sources. Data is usually historical, but sometimes includes real-time data. The largest challenge with DD-DSS approaches is the integration of the data – how is machine-generated data captured and integrated with human-generated data.

A manufacturing dashboard is a common example of a DD-DSS in industry, displaying key data drawn from multiple sources to give a clear indication of the state of the shop floor, of the order book, or of productivity. Common examples include Microsoft Power BI, Dundas BI, and Tableau, but many companies also develop customised dashboards with visualisation libraries such as R Studio's Shiny or Grafana. Many approaches are directly integrated into the database such as SAP

HANA. Even a spreadsheet's reporting and graphing features can achieve these goals.

DD-DSSs are the second most common example of DSS, and the most common that people commonly consider as a DSS. They are often referred to as Business Intelligence packages, reflecting the increasing sophistication of the tools offered. They are usually passive, but increasingly have active decision support tools.

6.4.2.3 Model-Driven

Model-Driven DSSs (MD-DSS) use simulation models for decision support by offering predictions of the outcomes of changes to the existing circumstances. These can be numerical models build in spreadsheets, computer-aided design (CAD) models of products, all the way to multi-physics 3D simulations of manufacturing processes.

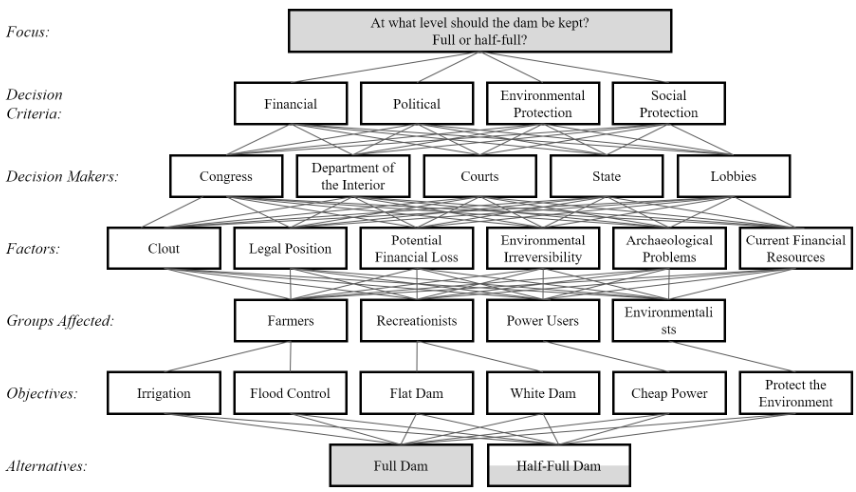


Figure 6.4-1 Analytical Hierarchy Processes are used to start deconstructing extremely complex and subjective problems to try to structure them into a model that allows alternatives to be tested.

Computer simulation offers the great advantage of studying and statistically analysing what-if scenarios, reducing the overall time and cost required for taking decisions. Monte Carlo simulation, discrete event simulation, agent and multi-agent simulation, system dynamics, and visual simulation are all in increasingly common use in manufacturing, and are discussed in section 5.3, The introduction of advanced simulation-based visualization of CAD designs with interaction and collaboration technologies such as augmented reality and virtual reality (AR/VR) is changing the product design process, enabling visual prototyping but also simpler collaboration between distributed teams, enabling CAD modelling to implement elements of

Communications-Driven DSS. Models and simulations are typically active DSS systems, giving the user the results of their proposed choices and in some cases running algorithms to show the optimal choice for a problem.

Also under the MD-DSS category are decision analysis models (DAM). DAMs are statistical tools and methods such as analytical hierarchy processes (AHP), decision tree analysis, multi-criteria decision analysis, and probabilistic forecasting to support decision making where multiple criteria needs to be considered and there is no single optimal answer.

6.4.3 Knowledge-Driven

Knowledge-Driven DSSs (KD-DSS) gather data, information, and knowledge from within the company and from external sources. They then use this database of past knowledge to make decisions with artificial intelligence (AI) techniques and recommends action. KD-DSS use AI techniques to combine large quantities of domain knowledge and past experiences to form new information. Expert Systems (which reached peak popularity in the 1980s and 90s) are a form of KD-DSS and used *if-then* rules and heuristics to solve problems. More modern KD-DSS use the latest developments in AI techniques such as neural networks, machine learning, fuzzy logic and genetic algorithms.

Reciprocal Learning-Based DSS (RL-DSS) are a sub-class of KD-DSS and can learn from common decisions and take over this decision making to reduce the human decision-making load. Routine decision tasks can be learned and programmed, and decision makers can update their knowledge and help create more intelligent decisions than previously possible for semi-structured problems.

In a manufacturing environment, capturing expert domain knowledge is a challenge. Instead, most modern manufacturing KD-DSS systems use data mining and machine learning to take large volumes of historical data and use this to determine the optimal course of action rather than relying on domain experts turning their knowledge into rules. Machine learning is proving extremely successful for classification problems such as defect classification, detecting problems with manufacturing processes so proactive and preventative maintenance can be performed, and creating models of manufacturing processes that were previously too complex for humans to fully understand

A famous example is IBM’s Watson. Watson is a natural language question-answer computer system and is a form of KD-DSS, using a huge databank of knowledge and rules to rapidly answer questions. It used this approach to win first place in the game show Jeopardy!, beating the previous champions by a significant margin. Watson uses a complex method answering questions, but still shares the same approach as any other KD-DSS. How to answer the question is the decision, and it uses databases of information to propose possible solutions, and then a second database to determine which solution is most likely to be successful. This process takes milliseconds.

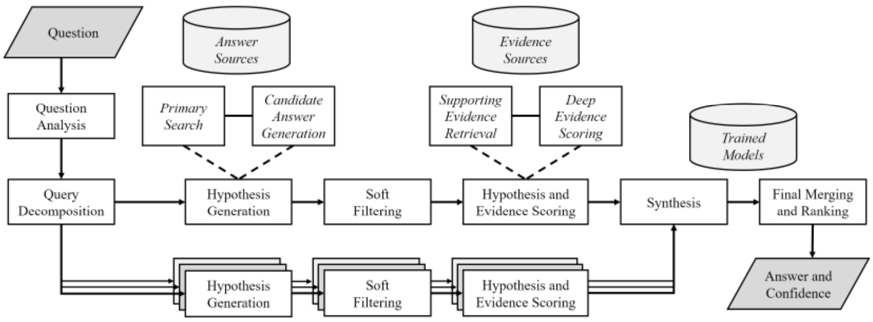


Figure 6.4-2 IBM's Watson used advanced natural language processing to analyse a question's meaning, and then draws upon a vast database of information to hypothesis answers and weigh them up with answers. Watson may be at the cutting edge of DSS systems, but the techniques used are increasingly common in commercial solutions.

6.5 Case Studies

6.5.1 Shop-Floor Monitoring – Rold

Rold are a medium-sized company in the domestic appliance components sector. The company has a high level of flexibility and responsibility due to being a family-run business, and have been able to combine this with the introduction of modern management approaches and novel technologies, constantly updating their standards in the search for the best performance to offer the market.

The company was experiencing the following issues in its production plant:

- Production machinery not connected to each other, thus not allowing, e.g., an efficient monitoring of the energy consumption as well as the analysis of the production data.
- Data not being available in real time and communicated in paper format, delaying the identification of problems and rectifying action.
- Widespread presence of subjective information and not objectified by the process, replying on operator expertise.
- Difficulty in having visibility of inter-plant processes, hindering identification of 'big picture' problems.
- Digital technologies not very common at the shop floor level.
- Need for operator empowerment in process control.

Many small to medium companies will look to external IT companies or solution vendors for assistance with digital twins or decision support systems. Due to Rold's size however, it possessed the in-house expertise to develop its own solution, which has since become a commercial product. Rold SmartFab is the result of Rold's research and development collaboration, which has put together internal company know-how and technologies to create a system for manufacturing SMEs that allows monitoring and analyzing information coming from the company's subsystem plants, and making them available on fixed, mobile and wearable devices. SmartFab is suited for companies in the mechatronic and manufacturing sectors, since it acquires the operating status of the machines and presents it in a user-friendly manner, in real-time and on smart devices anywhere. This allows for rapid identification of production issues, such as downtime and significant slowdowns. This digital manufacturing platform allows for:

- Real time monitoring of production lines.
- Real time data and alarms on touch screen devices and mobile devices, as well as wearable devices that are user-friendly (e.g. smartwatches).
- Using open standards and middleware, allowing modern machines to communicate with older ones.
- Cost reduction thorough a reduction of technical intervention and maintenance times.
- Empowerment of operators in the control of processes as they can now quickly access objective information.

A reduction of energy consumption and associated costs thanks to the possibility of measuring the energy usage of individual production resources in real time. Increasing the level of the operators’ soft skills in terms of digitalization, paving the way for the future. The company mainly needed to overcome technological barriers related to the necessity of connecting heterogeneous machines, including legacy systems without standard interfaces. This shows the importance of open standards in manufacturing digitalization. If you are trying to gather data from multiple machines old and new, from multiple vendors with multiple communication protocols, you may have issues unless you are able to use common standards.

6.5.2 Remote Sensing - Alascom

Alascom are a medium IT-technology services company founded in 2001 and based in Milan, Italy. Alascom has developed extensive telecommunications and system integration expertise. For several years, it has implemented a solution-providing strategy combined with the development of new innovative solutions to integrate the production domain with the Information and Communication Technology (ICT) domain, broadening the approach to cover the Internet of Things (IoT), Industry 4.0 and data analytics. Alascom is expert in statistical analysis and programming tools, mathematical models, AI machine learning, and in memoryDB.

Alascom were tasked with implementing monitoring systems for a range of biogas plants to provide a comprehensive view of the entire plant network as well as individual plants. Biogas plants are complex systems that obtain “clean” energy.

Due to their complexity, maintaining correct functioning is fundamental to guarantee constant revenues flows as well as to avoid extra costs due to breakdowns, malfunctions, and downtimes. Furthermore, improving the efficiency and effectiveness of the electrical energy production of the plants is another fundamental objective which may be possible through better monitoring. The geographic dispersion of the plants and the inability to keep specialized maintenance personnel onsite and to improve the production performances means any monitoring solution must be remote, centralized, and granular.

Monitoring provides a comprehensive view of the entire plant and the details of the individual monitored systems. The objective is to improve the efficiency of the production process through the use of software that allows maximization of the revenue/cost ratio, to have a greater understanding of the process itself and to better understand the impact of different operating modes to reduce the complexity of network management.

The software, designed to be adaptable to different production plants, is able to normalize data coming from multiple sites enabling comparison between different plants in scope and size; the greater quantity of data can be analyzed to benefit each plant. The software architecture takes data from newly installed smart meter sensors and performs data analysis. The steps are as follows:

1. Smart energy meters installed in the biogas plant are used to monitor the electrical consumption of auxiliary systems.
2. "Concentrators" gather and combine data from the smart meters at a specified frequency and return the values in a standards-compliant manner on the physical Ethernet channel and the Modbus TCP protocol.
3. IoT gateways acquire data from multiple sources, normalize the data, and send it to the database of a digital platform, where the data is then then processed and presented. This architectural element is able to communicate with different systems, in particular with the concentrators for the acquisition of electrical measurements of the biogas auxiliary systems, and with the process controllers for the acquisition of data available from the existing control system.
4. Finally, a digital platform composed of a cloud database and web-based user interface. The first architectural level acquires normalized data, that is then stored and made available to the web service interface. The front-end is developed with web services technologies, it allows the visualization of plant data with appropriate indicators, dashboards and tables defined with the support of the domain experts of the client company. The filters available in the web interface allow the user to enter specific requests and obtain analysed information - graphical or numerical - in either descriptive or statistical-predictive form. The system, through the use of appropriate algorithms that operate on available data, automatically suggests what actions need to be carried out and when, also taking into account any constraints imposed by the user.

The sensor integration and software solution was a success – all the biogas plants can be monitored remotely from a single location, reducing costs and improving the efficiency of the plants. The project was not without challenges however, and the integration of the client company's biogas plant experts, consulting smart sensor experts, and Alascom's own software engineers was the single largest challenge they faced.

6.5.3 Connected Transport Systems – Bellini

Bellini is a small company specialized in the development, production, marketing, and technical assistance of lubricants and fluids. Due to the complexities of chemical manufacturing and the weight of the materials and products, Bellini wanted to review the production process with the integration of software for remote monitoring, as well as integration of automated guided vehicles (AGVs) for automated materials and product transport.

Shop-floor operators currently use stand-alone programmable logic controllers (PLCs) to control the production process: the information is not shared with the enterprise resource planning (ERP) software. It is necessary to improve the user interface in the silos and mixers area so that the information necessary for carrying out the activities within the process are more complete and shared with the ERP system. Additionally, the handling of materials and products between the production and the logistics department takes place manually. It is necessary to implement a transportation system for automated handling from the production department to the shipping bay.

Given the specified needs, the company decided to purchase a software program to monitor the progress of activities. In particular, the intervention envisaged the introduction of a software for remote control of the PLCs, the revamping of the sensors, and the replacement of the current LCD panels on the equipment with intuitive and larger touch screens. Improved access to information, allowing it to be viewed remotely or simply in a clear and unified manner is an example of a data-driven decision support system. Being able to clearly see the state of a system at a glance makes quick and effective decision making simpler and less prone to error. Thanks to the remote access of the PLC data, it will be possible to identify equipment failures and other losses quicker. The remote access of the PLC data avoids manual data entry giving a reduction of worker time and fewer introduced errors.

Control of AGVs often requires a digital twin approach for highly dynamic job routing typical of smaller dynamic companies – the live location of the AGVs needs to be controlled to ensure material can be routed where and when it is required in a safe manner. The new automated material handling system will allow the movement of material between the two departments in an automated way, thanks to the use of an AGV connected to positioning sensors, both on board and in the plant. The AGVs are more energy efficient than the current manual handling machines (forklifts mainly) saving costs. Manual material handling time is reduced thanks to the use of the AGV.

The main challenge was the search for the appropriate technological solution. One was eventually found after 4 months of design study. It was also an extensive investment to both acquire and to plan the integration of the system. As with any significant investment in digital manufacturing technologies, it is critically important to have a solid business case developed that shows the value of the proposed solution to the business.

6.6 Conclusions

This chapter discussed the latest advances in production analysis, including the use of online modelling and simulation (digital twins) and support tools for decision making (decision support systems), as well as detailing some use cases of examples of how access to data and the use of simulations can support businesses.

As these fields are new and/or rapidly evolving, it is important to understand both their strengths and limitations. There is a lot of hype surrounding digital twins, and the lack of standards mean many products may be labelled as digital twins without necessarily having the functionality you might expect. Though the potential advantages of online systems are significant, there are risks of investing in technology which may not be fully supported in the future, and which may not deliver on investment.

Critically before making any investment in a new modelling package of digital twin system is to develop a business case for it. What questions will this new approach enable us to answer that we could not before? How will we use the new information to improve productivity (or any other key performance indicator)? What changes will we need to make to our working processes, and how long will training and acclimatisation take? Particularly for digital twins where established standards have not been settled and buzzwords are frequently used, it is important to understand exactly what is being offered and the potential costs and benefits thereof.

If there is one critical lesson in this book, it is that mathematical formulae, digital modelling software, decision support systems and digital twins are all tools to assist in the decision-making process. All these tools have their strengths and weaknesses, and areas of applicability; the chapters in this book are not a scale from ‘bad’ tools to ‘good’ tools, only from manual, to offline digital, and then online digital. Understanding the decision-making process, the measurable key performance indicators, and why an analysis tool is required, and which is most useful is a critical skill for effective decision making in a manufacturing context. Without understanding exactly what questions are being asked and why, supporting tool will not be able to offer their full capacity in aiding the enterprise in answering those questions.

6.7 References

[1] M. Grieves, "Virtually intelligent product systems: digital and physical twins," *Complex Systems Engineering: Theory and Practice,* pp. 175-200, 2019.

[2] B. Piascik, J. Vickers, D. Lowry, S. Scotti, J. Stewart and A. and Calomino, "Materials, structures, mechanical systems, and manufacturing roadmap," *NASA Technology Area 12-2,* 2012.

[3] E. Negri, L. Fumagalli and M. and Macchi, "A review of the roles of digital twin in cps-based production systems," *Procedia Manufacturing,* vol. 11, pp. 939-948, 2017.

[4] M. Shafto, M. Conroy, R. Doyle, E. Glaessgen, C. Kemp, J. LeMoigne and L. and Wang, "Modeling, simulation, information technology & processing roadmap.," National Aeronautics and Space Administration, 2012.

[5] B. Gockel, A. Tudor, M. Brandyberry, R. Penmetsa and E. and Tuegel, "Challenges with structural life forecasting using realistic mission profiles," *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference,* p. 1813, 2012.

[6] R. Rosen, G. Von Wichert, G. Lo and K. and Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," *IFAC-PapersOnLine,* vol. 48, no. 3, pp. 567-572, 2015.

[7] T. Gabor, L. Belzner, M. Kiermeier, M. Beck and A. and Neitz, "A simulation-based architecture for smart cyber-physical systems," *2016 IEEE international conference on autonomic computing (ICAC),* pp. 374-379, 2016.

[8] A. Bolton, L. Butler, I. Dabson, M. Enzer, M. Evans, T. Fenemore, H. F. and e. al, "Gemini Principles," Centre for Digital Built Britain, 2018.

[9] HBM, "Piezoelectric or Strain Gauge Based Force Transducers?," [Online]. Available: https://www.hbm.com/en/3719/piezoelectric-or-strain-gauge-based-force-transducers/. [Accessed 2020 11 23].

[10] M. Groover, Automation, production systems, and computer-integrated manufacturing, Pearson Education, 2016.

[11] International Electrotechnical Commision, “ISO/IEC JTC 1 Information Technology - AG 11 Digital Twin,” [Online]. Available: https://www.iec.ch/dyn/www/f?p=103:14:3438570860908:::: FSP_ORG_ID,FSP_LANG_ID:25023,25. [Accessed 23 11 2020].

[12] International Organization for Standardization, “ISO/TC 184/SC 4 - Industrial Data,” [Online]. Available: https://www.iso.org/committee/54158.html. [Accessed 23 11 2020].

[13] International Organization for Standardization, “ISO/DIS 23247-1: Automation systems and integration — Digital Twin framework for manufacturing — Part 1: Overview and general principles,” [Online]. Available: https://www.iso.org/standard/75066.html. [Accessed 23 11 2020].

[14] W. Kritzinger, M. Karner, G. Traar, J. Henjes and W. and Sihn, “Digital Twin in manufacturing: A categorical literature review and classification.,” *IFAC-PapersOnLine,* vol. 51, no. 11, pp. 1016-1022, 2018.

[15] High Value Manufacturing Catapult, Visualisation and VR Forum, “Feasibility of an Immersive Digital Twin,” 2018.

[16] M. Grieves and J. and Vickers, “Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems,” *Transdisciplinary perspectives on complex systems,* pp. 85-113, 2017.

[17] Irdeto, “Irdeto Global Connected Industries Cybersecurity Survey,” Amsterdam, 2019.

[18] R. Millman, “Gartner: Four best practices for managing digital twins,” 19 03 2018. [Online]. Available: https://internetofbusiness.com/half-of-businesses-with-iot-projects-planning-to-use-digital-twin/. [Accessed 23 11 2020].

Part III
Intelligent Robotics

Part III Overview

Traditionally, industrial robots have been developed to perform repetitive tasks in a structured environment, which means that the task is clearly and precisely known by the robots, no variables exist so that it is fully predictable and the consequences of the robot action are predictable too. However, the development and emergence of Industry 4.0 technologies, such as Internet of Things, Big Data, Cloud Computing and Artificial Intelligence, is enabling the emergence of a new generation of robotics. Intelligent and autonomous robots will achieve a significant breakthrough in future manufacturing. Smart factories will cause a wide range of changes to industrial production and productivity. The interconnectivity of intelligent machines and facility components leads to a flexible, reconfigurable and fully automatic production line. The incoming orders are processed by a machine, which defines the production process, orders the materials, which are handled by robots, as well the final product and the shipping. In such an interconnected environment, robots monitor their own health so that predictive and self maintenance is possible, reducing the downtime and increasing efficiency and productivity. The role of humans will be to supervise the correct operation.

Although there is huge opportunity, it is still challenging to achieve due to the implementation issues of Industry 4.0 technologies. Indeed, even though the level of the I4.0 enabling technologies is already quite high, their adoption is hindered by the complexity of the concept and the lack of detailed roadmaps, strategic guidance and implementation advice, together with the large investment which can be required.

In this context the final three chapters of this book present some insight into the potentiality of Industry 4.0 in robotics, providing basic and advanced knowledge on robotic systems and their applications in digital manufacturing. Chapter 7 considers robots generally, describing the different types of robots which currently exist, introducing common robot terminology, discussing robot applications and introducing innovative concepts on collaborative robots, human-robot collaboration and mobile robots. The subsequent chapters then consider the different elements which comprise a robot and discuss their development in the context of digital manufacturing. Chapter 8 considers one of the key elements of intelligent robotics, the hardware element, which has been developed in the past few decades and provides solid foundations for the smart factory. Basic and advanced principles of sensing and actuation in robots are discussed as the link between physical and cyber world. The chapter then discusses microrobotics, which introduces a core element of Industry 4.0: Cyber Physical Systems. Chapter 9 then considers the software, or cyber, elements which are essential to enabling the new generation of intelligent robots. It explains what a Cyber Physical System is and how it applies in the production environment, stressing the benefit and challenges of collaborative systems. This chapter also explains the different ways of programming a robot and the principals of robot control. Following this, Artificial Intelligence – an essential element to enabling intelligent robots – is discussed. The chapter, and this book, conclude with a brief overview on the ethical issues related to intelligence implemented in machines.

Chapter 7
Industrial Robots 4.0

Irene Fassi, Simone Pio Negri, Claudia Pagano, Lara Rebaioli and Marcello Valori

7.1 Introduction

The main goal of Industry 4.0 in Robotics is the development of smart industries with increased productivity of high-quality products meeting customer expectation. This requires flexibility, automation and interconnection. In smart factories the products will be handled by autonomous *mobile robots*, human and robot skills will be combined using *collaborative robots* (cobots), the huge amount of collected data will be used to predict maintenance and disruption and take decisions by robots with *artificial intelligence*.

The use of industrial robots in factories is becoming more widespread. They are typically used to perform tasks which are dangerous to humans, to enhance the throughput and quality of production, and reduce production costs. They work in a confined and structured space, are programmed to continuously perform a repetitive sequence of actions which hinders the flexibility and reconfigurability of production lines.

The new generation of robotics, integrating the technologies of Industry 4.0 such as Internet of Things, Big Data, Cloud Computing, Artificial Intelligence, creates the industry of the future able to automatically and efficiently reconfigure itself to fulfil individual customer requests. Machines and facility components are all interconnected so that the entire production process is fully automatic. The incoming order is processed by a machine which defines the production process, orders the materials, which are handled by robots, as well the final product and shipping. In such an interconnected system of the smart factory, robots monitor their own health so that predictive and self-maintenance is possible, reducing the

---

I. Fassi, S.P. Negri, C. Pagano (✉), L. Rebaioli, M. Valori
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, Milan, Italy
e-mail: claudia.pagano@stiima.cnr.it

J.C. Chaplin et al (eds), *Digital Manufacturing for SMEs*
DOI: https://doi.org/10.17639/vjvt-7681

downtime and increasing efficiency and productivity. The role of humans will be to supervise the robots and ensure correct functionality.

This chapter gives a short overview on industrial robots, providing the terminology and the basic concepts helpful for the comprehension of the next sections in the book. Then the main concepts of collaborative and mobile robots are introduced, before being discussed further in the next chapters.

7.2 Industrial Robot

Although several definitions of robot and robotic systems can be found in the literature, hereafter we adopt those provided in the ISO 8373:2012 [1] standard.

- A *robot* is an actuated mechanism programmable in two or more axes with a degree of autonomy, moving within its environment, to perform intended tasks. Autonomy in this context means the ability to perform intended tasks based on current state and sensing, without human intervention.
- A *service robot* is a robot that performs useful tasks for humans or equipment excluding industrial automation application. Robots that perform useful tasks for humans or equipment in industrial automation applications are referred to as *industrial robots*.
- A *personal service robot* or service robot for personal use is a service robot used for a non-commercial task, usually by the general public. Examples are domestic servant robots, such as robot vacuums, lawn mowers, personal robot assistants, automated wheelchairs, and personal mobility assist robots, such as wearable upper or lower limb exoskeletons.
- A *professional service robot* or a service robot for professional use is a service robot used for a commercial task, usually operated by a properly trained operator. Examples are cleaning robots for public places, delivery robots in offices or hospitals, fire-fighting robots, rehabilitation robots and surgery robots in hospitals. In this context, an operator is a person designated to start, monitor and stop the intended operation of a robot or a robot system.
- A *robot system* is a system comprising robot(s), end-effector(s) and any machinery, equipment, devices, or sensors supporting the robot performing its task.

In addition, a *collaborative industrial robot* is defined by the International Federation of Robotics [2] as an industrial robot that is designed in compliance with ISO 10218-1 “Robots and robotic devices – Safety requirements for industrial robots – Part 1: Robots” [3] and intended for collaborative use. A collaborative operation happens when a purposely designed robot works in direct cooperation with a human.

7.2.1 Robot Terminology

An industrial robot's mechanical apparatus is usually composed by sequences of articulations and rigid segments, respectively called *joints* and *links* (Figure 7.2-1); the resulting *kinematic chain* enables the motion of the *end-effector* according to the trajectories required for the specific task. As a result, the mobility level of the

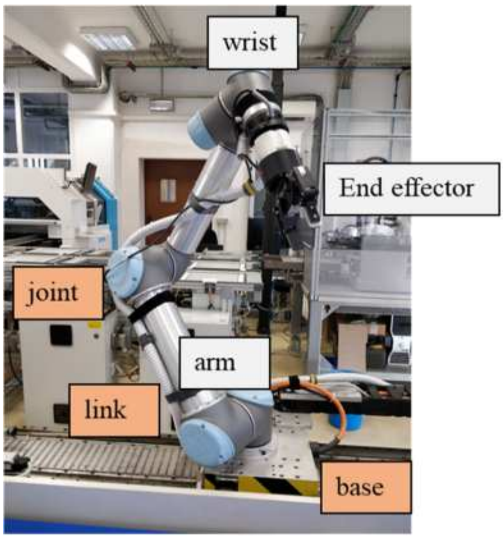


Figure 7.2-1 Robot main components.

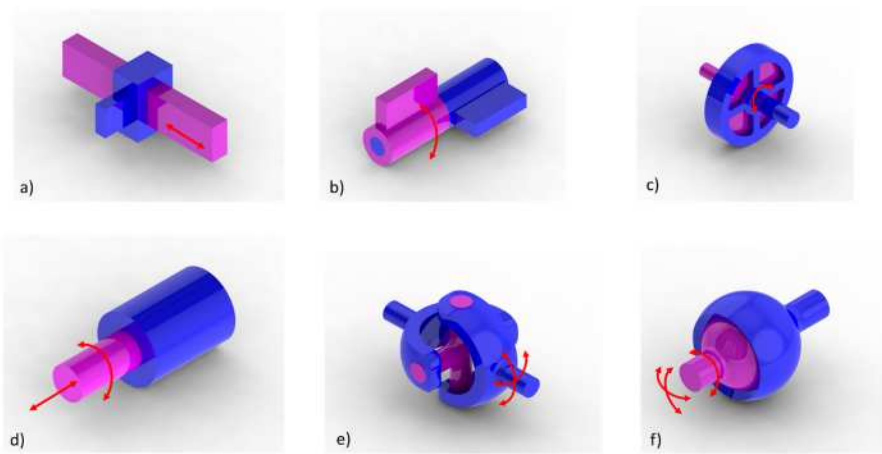


Figure 7.2-2 Typical joints used in robotics: a) prismatic (translation along one direction), b) and c) rotoidal (hinge, rotation about one axis), d) cylindrical (rotation and translation) e) cardanic (rotation about two orthogonal axes), f) spherical (rotation about 3 axes) [4]. Image rights: Authors.

end-effector directly depends on the number of joints (Figure 7.2-2) in the system. In order to represent this concept, the number of *degrees of freedom* (DOFs) of a robotic arm can be defined, on a general basis, as the number of joints making up arobotic arm's kinematic chain. This feature directly affects the possible motion of the end-effector: under appropriate design criteria, an equivalence can be established between the number of DOFs of the arm and the number of DOFs of the end-effector.

Industrial robots are usually classified by their kinematic chain (serial or parallel), the number of degrees of freedom, accuracy and repeatability, payload and workspace. A *serial* kinematic structure is normally realized with rigid links connected in series by means of 1 DOF actuated articulations (able to perform a rotation or translation), while a *parallel* kinematic structure is composed by several convergent kinematical chains, where not all the articulations are actuated, as shown in Figure 7.2-3.

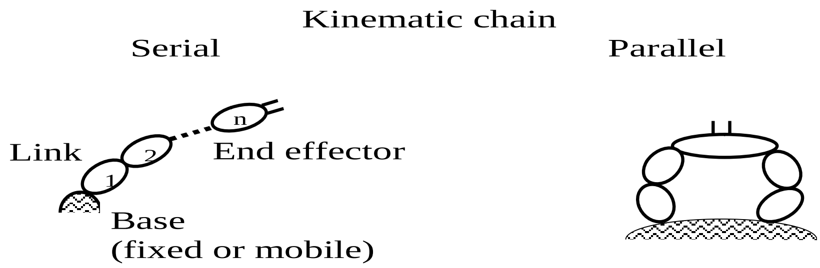


Figure 7.2-3 Example of serial and parallel kinematic chains [4].

The DOFs of the end-effector cannot exceed three translations (x,y,z in the Cartesian space) plus three rotations, which are the six DOFs which can characterize an object in a 3D space. As a result, a number of DOFs higher than six in the joint space has the effect of increasing the overall dexterity of the system, which is the ability to obtain a defined pose of the end-effector with different configurations.

Robots with more than 6 DOFs are defined as being *kinematically redundant*. This redundancy can be usefully applied to improve the dexterity of the end effector.

A robot is then characterized by its *accuracy,* which is defined as the deviation between the obtained and the expected (programmed) position, and its *repeatability*, which is a measure of the manipulator's ability to return to a previously reached position (from the same and from different directions).

The robot repeatability depends not only on the characteristics of its mechanical structure and dimensions but also on the transducers and controller. Typical repeatability values are in the order of few hundred microns, while precision robots have repeatability values up to 5 micrometres.

The *workspace* is generically defined as the 3D space the robot can access (Figure 7.2-4). The *space of movement* is the space achievable by every part of the robot. The *maximum space* is the space of movement and the portion of space achievable by the end effector. The *operating space* is the portion of maximum space reached in a specific task. The *safeguarded space* is the space delimited by physical barriers–to limit potential dangers. The *collaborative workspace* is the workspace within the safeguarded space where the robot and a human can perform tasks simultaneously.

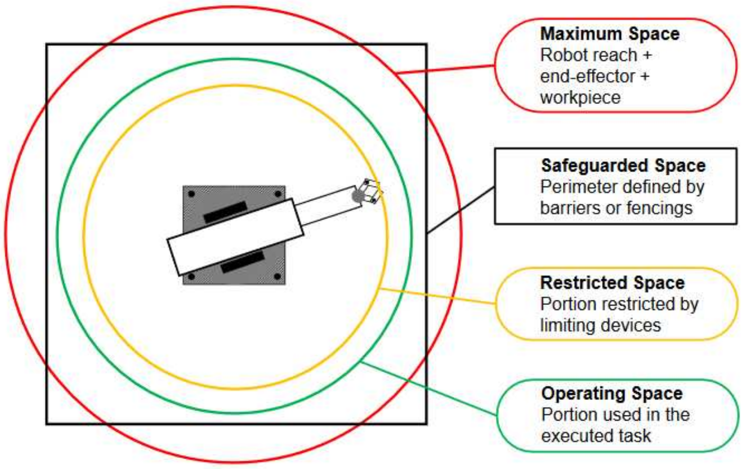


Figure 7.2-4 Working space zones according to the applicable industrial standard for traditional robotics [5]. Image rights: Authors

7.2.2 Mechanical Structures of Industrial Robots

The main categories of mechanical structures of industrial robots are:

- *Cartesian Robots*: Having three prismatic joints (Figure 7.2-2) whose axes are coincident with a Cartesian coordinate system.
- *SCARA (Selective Compliance Assembly Robot Arm) Robots:* Having two parallel rotary joints to provide compliance in a plane.
- *Articulated (or Anthropomorphic) Robots:* Having at least three rotary joints (Figure 7.2-2) placed in series with their interconnecting links.
- *Parallel Robots:* Whose arms have concurrent prismatic or rotary joints.

Each kinematic structure determines the corresponding shape of the workspace. The different industrial robots and their workspaces are illustrated in Figure 7.2-5.

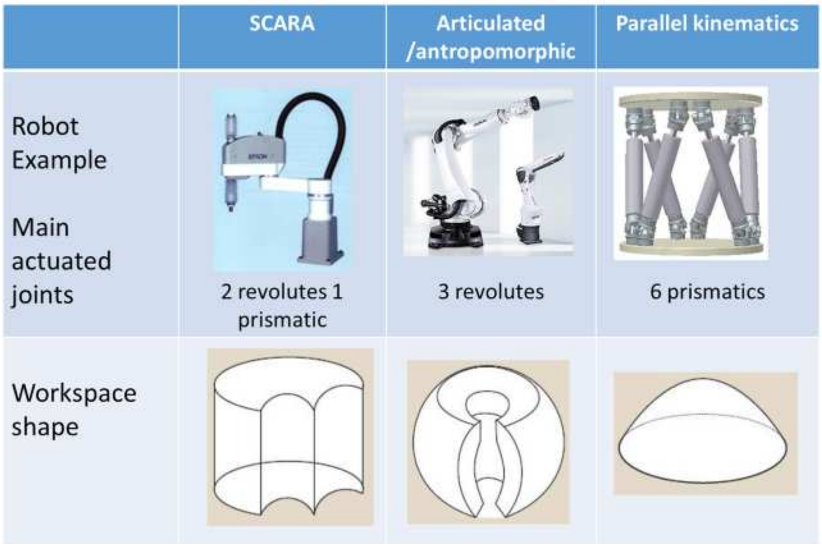


Figure 7.2-5 Main categories of mechanical structures of industrial robots.

The robots of the first three categories (Cartesian, SCARA and articulated) are serial manipulators, because they are realized by connecting several links in series with each one actuated by a rotoidal (revolute) or a prismatic joint. Each joint is actuated by a motor; the term *axis* usually denotes a joint and its actuator (linear or revolute motor). Parallel manipulators are characterized by several links connected in parallel to move the mobile base.

7.2.3 Robot Applications

The different structures of industrial robots are better suited to different tasks:

- *Articulated Robots:* These are used in almost all applications, but they are mostly associated with welding, dispensing and handling applications. Examples of different applications of articulated robots are shown in Figure 7.2-6.
- *Cartesian Robots:* These are mainly employed in plastic moulding, packaging, pick and place operations, and assembly applications. Examples of different applications of Cartesian robots are shown Figure 7.2-7.
- *SCARA Robots:* These are used in assembly and material handling, and also cleanroom applications. Examples of different applications of SCARA robots are shown in Figure 7.2-8.
- *Parallel Robots:* These are typically adopted for handling, packaging, pick and place operations. Examples of different applications of parallel robots are shown in Figure 7.2-8.

Handling for metal casting	Palletizing
 [photo courtesy of Global Casting Magazine]	 [photo courtesy of ABB]
Welding	**Painting**
 [photo courtesy of Roboteco S.p.A.]	 [photo courtesy of FANUC]

Figure 7.2-6 Examples of applications of articulated robots. Image rights: Global Casting Magazine, ABB, Roboteco S.p.A., FANUC.

Handling for plastic moulding	Dispensing
	
[photo courtesy of Dürrschmidt GmbH]	[photo courtesy of Steeldot]
Welding	**Pressing**
	
[photo courtesy of R&E Engineering]	[photo courtesy of Güdel Group AG]

Figure 7.2-7 Examples of applications of Cartesian robots. Image rights: Dürrshmidt GmbH, Steel dot, R&E Engineering, Güdel Group AG.

Assembly	Packaging
	
[photo courtesy of Dynamic Automation and Robotics]	[by Hirata Robotics GmbH - Bilder von Hirata Robotics, CC BY-SA 3.0 de]

Figure 7.2-8 Examples of applications of SCARA robots. Image rights: Dynamic Automation and Robotics, Hirata Robotics.

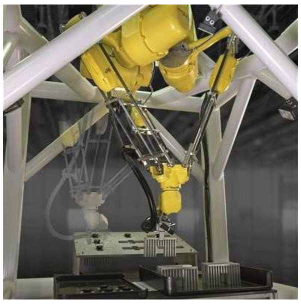
Pick and place operations	Assembly
 [photo by Humanrobo - Own work, CC BY-SA 3.0]	 [photo courtesy of FANUC America]

Figure 7.2-9 Examples of applications of parallel robots. Image rights: Humanrobo, FANUC America.

7.3 Collaborative Robot

The concept of human-centred factories is aimed at combining the new "social expectations on how factories are related to workers and neighbours as a working place and as a dynamic element of the environment" with the idea of a "production system that delivers cost-efficient and flexible manufacturing" [6].

Competences, awareness and individual preferences of the worker are promoted and working environment and equipment accordingly evolve in terms of ergonomics, user-friendliness and integration. On the robotic side, this results in the emergence of *collaborative robots*.

"A collaborative robot is a robot that CAN (capable) be applied for use in a collaborative operation", that is a *"state in which a purposely designed robot system and an operator work within a collaborative workspace"* [7]. By combining the potentialities of human (flexibility, critical thinking) and robots (precision, reliability), new automated production lines can be set up in a wide range of fields. *Human-Robot Collaboration* is driven by some innovative aspects of the smart factories, such as the centrality of human workers and the introduction of the CPPS as more "intelligent" production systems.

Advanced interfaces are enabled by the technological progress in actuators, sensors and control techniques. *Safety* and *ergonomics* are the main issues to consider in the definition of new human-robot interaction models. Analysing these aspects is useful to understand the flexibility introduced in workspace configuration by a "collaborative" approach.

The idea of *cobots* (collaborative robots) was first introduced in the late 1990s into the automotive industry. Nowadays, a wide variety of collaborative robots (Figure 7.3-1) are available on the market, with different features and capabilities, and at affordable prices.

YuMi IRB 14000 (ABB)	**Sawyer (HAHN Robotics)**
 Multi-purpose end-effector: vacuum and servo grippers, camera. Padded redundant arms. Collaborative assemblies, inspection and packaging. https://new.abb.com/	 Power and force limited compliant arms (elastic actuators, sensors). Force sensors in each joint. Cameras in the wrists and in the head. https://www.rethinkrobotics.com/
E-Series (Universal Robot)	**Aura (Comau)**
 Available in 3 sizes. Touch-screen tablet to set movement options. Programmable by demonstration. Low starting price. https://www.universal-robots.com/	 High payload and fully collaborative. Manual guidance functionality. Collision prevention. https://www.comau.com/
LBR iiwa (KUKA)	**duAro (Kawasaki)**
 Available in 2 sizes. Integrated sensors at each joint for position control and sensitivity. Collision detection control algorithms. Image rights: KUKA AG, https://www.kuka.com/	 Dual-arm SCARA (Articulated Arm in the horizontal plane and end-effector with vertical motion). Easy teaching by demonstration. Extreme repeatability. https://robotics.kawasaki.com/

Figure 7.3-1 Some commercial models of collaborative robots. Image rights: ABB, Rethink Robotics GmbH, University Robots, Comau, Kuka, Kawasaki.

Collaborative robot versatility is achieved thanks to several factors:

- The intrinsic safety facilitates the workcell design and reconfigurability.
- The lightweight robots are characterized by simpler transportation and installation.
- Fast programming methods accelerate task set up.

Collaborative robot flexibility makes them particularly attractive for SME environments, often poorly automatized, to carry out tasks such as fine manipulation and assembly, packaging, manufacturing, pick and place. Due to their easiness of installation and reprogramming, and intrinsic safety, they can be easily moved around the production lines and placed where needed when needed Furthermore, the price of collaborative robots can be conveniently affordable. On the other hand, collaborative robots are often characterized by low repeatability and payloads when compared with traditional industrial robots.

7.3.1 Robotic Devices for Enhancing Workers' Performance

Current developments in the industrial world are characterized by even more synergistic devices: with the innovation introduced by recent industrial trends, exoskeletons for industrial applications have seen a great development in recent years (Figure 7.3-2). They can be classified as follows [9]:

- *Tool Holding Exoskeletons*: These articulated arms are directly supported by a structure that moves with the legs of the wearer, transmitting the weight directly to the ground. In the arms' joints there are springs that make weights lighter to lift and hold, as they are not motorized.
- *Chairless Chairs*: These devices are wearable, compliant with the user's movements. If required, they provide suitable stool-like support.
- *Back Support*: The purpose is to provide a rigid support to the back during the lifting of heavy loads. They can be motorized or not.
- *Powered Gloves*: These gloves are used to enhance the operator grasp by the implementation of auxiliary "muscles".
- *Upper Body Exoskeletons*: Lightweight suites that support the wearer in performing tiresome tasks, sometimes in unnatural positions.
- *Full Body Powered Suits*: The original concept is based on heavy robotic devices. Nowadays, some simplified architectures are available, empowering different part of the body.

Figure 7.3-2 Exoskeletons for industrial use: 1. Fortis by Lockheed Martin Corporation (courtesy of Lockheed Martin), 2. Noonee "Chairless Chair" (courtesy of Noonee), 3. Powered Wear ATOUN MODEL Y (courtesy of ATOUN), 4. BIOSERVO Ironhand powered glove (courtesy of Bioservo Technologies AB), 5. ©EksoVest by Ekso Bionics upper body exoskeleton, 6. MAX© - Modular Agile eXoskeleton by suitX©.

7.3.2 Human-Robot Collaboration Safety

In traditional robotics, the operations of a robot and a human operator can be complementary but not collaborative. A robot working area is limited to a confined workspace with different *working space zones* defined around that workspace, as shown in Figure 7.2-4. On a general basis, a human operator can only access a robot workspace when it is not working.

In cases where human and robot are required to share a working space, the robot has to be in a "safe" mode before the operator accesses the robot working space.

In *collaborative robotic tasks*, the operations of a robot and a human operator are more integrated and they need to become collaborators, moving beyond simple coexistence. The first step towards this goal is the need to share the workspace: the optimal synergy is achieved by working hand in hand. Due to this, *new safety requirements* arise for the design of the production systems, affecting both robotic machines and workspaces.

The ISO TS 15066:2016 "Robots and Robotic Devices – Collaborative Robots" [8] provides a complete and comprehensive guideline for anyone involved in the risk assessments of collaborative robotic operations. The maximum power and speed of collaborative robots are described, as well as the design criteria for the realization of collaborative robots. Although this Technical Specification is not a mandatory standard, the design of collaborative operations is exhaustively detailed, even defining the impact limits for each human body region (Figure 7.3-4).

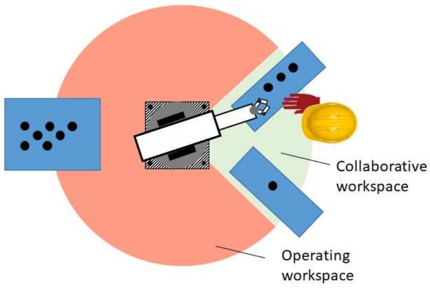


Figure 7.3-3 Workspace sharing in collaborative workcells. The operator can access only the collaborative workspace. Source: Author based on [8].

With reference to the risk assessment described in the standard UNI EN ISO 13849-1 "Safety of Machinery – Safety-related Parts of Control Systems", a simplification of the safety performance level is expected in the design of the safety functions of robot controllers (Figure 7.3-4).

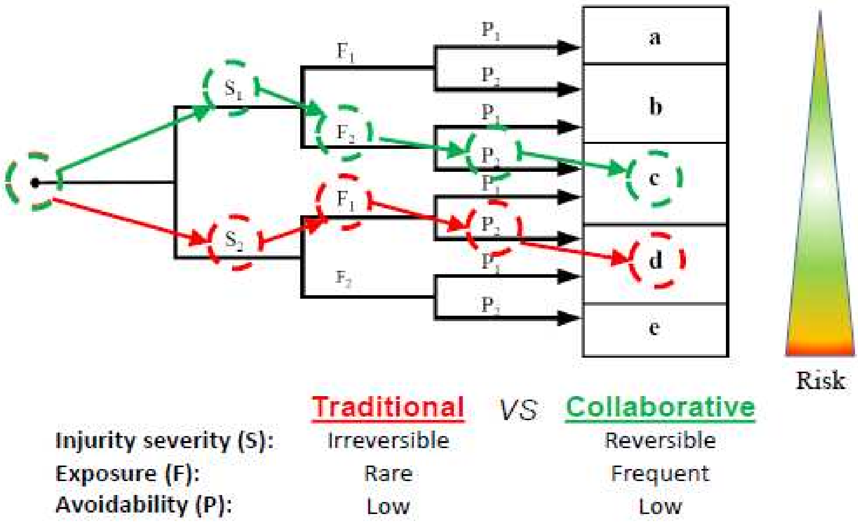


Figure 7.3-4 The Performance Level (PL) of safety functions in control systems is based on the probability of dangerous failure per hour. An example of a comparison between traditional (PL d) and collaborative (PL c) robots from the risk assessment perspective is reported, highlighting the assessment method as per the definition of the applicable standard. *Collaborative robots can be characterized by lower PL, as the intrinsic safety guarantees higher safety levels.*

According to the technical innovations on the robotic side, the role of the human operator changes considering the interaction with the robot and the robot programming. Building on existing standards, ISO TS 15066:2016 “Robots and Robotic Devices – Collaborative Robots” summarizes the types of collaborative robotic operations, defining requirements, safety measures and risk assessment, as illustrated in Figure 7.3-5.

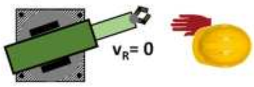
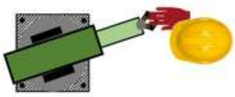
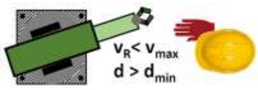
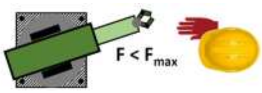
Method	Description	Risk reduction
Safety-rated monitored stop 	The workspace is shared, but simultaneous operation is not allowed.	When the operator is in the work space, the robot stops without interrupting drive power.
Hand guiding 	The end-effector is guided by the operator for collaboration or teaching/programming.	The robot stop and then moves guided in a *safety-rated monitored speed* mode.
Speed and separation monitoring 	The operator and robot can operate simultaneously; the robot is equipped with vision or proximity sensors.	The robot moves with safe dynamics and distance from the operator.
Power and force limiting 	The operator and robot can operate simultaneously; advanced force control strategies are used.	In a collision the robot can transfer a controlled amount of energy or immediately stop.

Figure 7.3-5 Types of collaborative operations.

7.4 Mobile Robots

Logistics in smart factories is usually performed by means of automated guided vehicles (AGVs) that move along predefined paths in the factory plant to transport heavy loads or materials. They are typically equipped with sensors for navigation and collision detection and path planning. The next generation of AGVs is represented by fully autonomous mobile robots, which can autonomously navigate along the production line and interact with machines and operators.

Ideally, it is possible to represent a mobile robot by four subsystems (Figure 7.4-1):

1. *Perception*: This is the ability to measure and perceive either the state of the external environment (by external sensors), or the internal state of the robot (by internal sensors) and to extract information useful to reach the goal.
2. *Information processing*: This is the computation that, using the information acquired by the sensors, determines the operations to be executed. For a mobile robot the most important processes are:
   - *Mapping:* The identification of a map of the environment in which the robot operates.
   - *Localization:* The determination of the robot state.
   - *Navigation:* Definition of the robot path to reach the goal, considering the obstacles in the environment.
   - *Kinematics:* Description of the robot movement, considering any constraints due to the robot's mechanical structure.
   - *Control:* Coordination of the whole process, comprising acquisition, computation and actuation to obtain the goal. The control process can include some input from the operator.
3. *Action*: This is the ability to translate the decisions coming from the control system into operations in the real world. This function is implemented by the actuators installed on the robot itself.
4. *Power Supply*: This system is responsible of energy generation and delivery to the whole system. It is usually supplied by chemical batteries, but other solutions are possible.

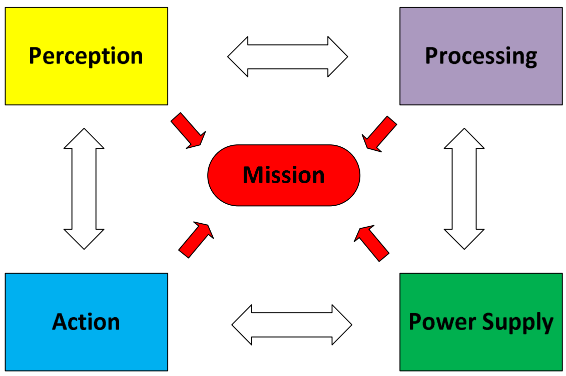


Figure 7.4-1 Conceptual diagram of a mobile robot.

The most ability of a mobile robot is locomotion. Different possibilities arise for how the locomotion [10] is implemented, many of them inspired by the biological world though some of the concepts implemented in nature are difficult to replicate from an engineering point of view.

Figure 7.4-2 divides the mobile robot set in two sub-sets: Wheeled Mobile Robots (WMR) and Not Wheeled Robots.

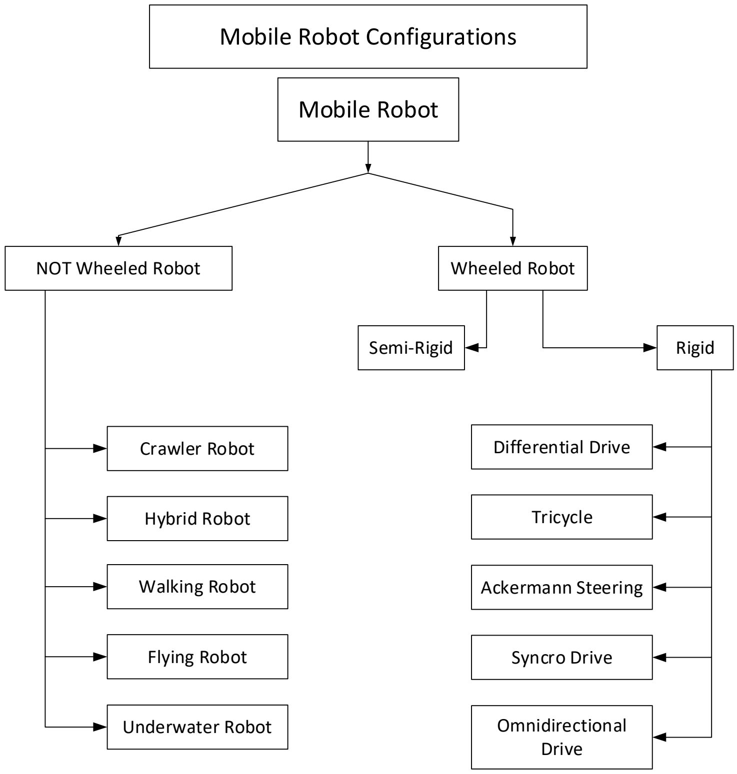


Figure 7.4-2 Classification of Mobile Robots.

7.4.1 Wheeled Mobile Robots (WMR)

Locomotion by wheels is the most used in mobile robotics. Robots equipped with wheels (also called “Rover”) achieve very good power to weight efficiencies. Moreover, their mechanical construction is simple.

WMR are commonly used on flat floors such as found in research laboratories, warehouses, factories, airports, exhibitions and other similar environments (indoor Rover). There are also outdoor applications on rough surfaces with obstacles and subsidence (outdoor Rover).

WRM stability on the ground is assured by using at least three wheels, though there are robots with only two wheels that can be stable under specified conditions.

A WMR is made of a rigid or semi-rigid frame, on which the wheels are installed according specific criteria to make it stable and to guarantee good traction. Each kind of wheel has its specific kinematics (i.e. movement) so the wheel choice influences the WMR whole kinematics.

There are four major wheel classes, as illustrated in Figure 7.4-3**:**

a b

c d

Figure 7.4-3 Wheel classification: (a) standard wheel, (b) Caster wheel, (c) omnidirectional wheel, and (d) spherical (ball) wheel.

- *Standard Wheel:* This can be fixed or steerable. It is the most common wheel used in robotics, and also the simplest from both a mechanical and kinematic point of view. The fixed standard wheel has a rotational axis, through its center and parallel to the ground. The movement of the robot is due to the rotation around this axis (1 degree of freedom). The steerable wheel has the same axis of the fixed wheel for the traction plus a steering axis orthogonal to the first. This steering axis goes through the instantaneous contact point between the wheel and the ground, and to the wheel center. The steerable wheel has a variable orientation with the frame (in total 2 d.o.f.).
- *Caster Wheel*: This is similar to the steerable wheel, the only difference is in the steering axis that does not intersect the horizontal axis, but it passes at a fixed distance. This distance must be considered in the wheel design, because it generates a static torque on the steering axis. In this way the Caster wheel can move in every direction, potentially without sliding. The steering process can be passive or active; in the latter case a steering actuator is needed.
- *Omni-directional Wheel* (also called *Swedish Wheel*): is a particular wheel that can go forward, backward and sideways. It is complex from a mechanical point

of view (so has a high cost), but enables potentially every movement and this is a big advantage in many situations.
- *Spherical Wheel:* This is made using a sphere, without a main rotational axis. The advantage is that it is conceptually simple; the drawbacks are that it has a big volume compared with the payload and the difficulty to actuate it to make a traction wheel.

Different kinds of wheels can be installed on the same robot to exploit the peculiarities of each one of them.

7.4.2 Crawler Mobile Robots (CMR)

WMR have some limitations due to the fact that the ground must be sufficiently flat so that no obstacles prevent the wheels' movement. This is a drawback that does not fit with an increasing number of applications, for example on sandy or rocky ground, big slopes or other planet exploration. For these reasons and in this context CMR are preferred.

Crawlers allow a correct movement of the mobile robot on rough ground, giving better traction and distributing better the pressure due to the weight. On the other hand, crawler mobile robots are slower than WMR and are more energy hungry. Their control is more complex because the point of contact between the crawler and the ground is not well defined and an estimate of their position is necessary to evaluate their movement.

7.4.3 Legged Mobile Robots (LMR)

Legged robots are usually bio-inspired robots. They are used in unstructured environments because they have a great mobility; one typical application is climbing stairs. However, legged robots have a slow speed and relatively high energy consumption. Their mechanical structure is very complex compared with WMR; a legged mobile robot has a high number of links and actuators, so its control is quite difficult.

The number of legs can vary: robots with four legs are called *quadrupeds* and exhibit a very stable and robust locomotion. Using four legs guarantees at least three contact points with the ground at all times; in this way the robot is always in the condition of static equilibrium. The advantage of quadrupeds is that dynamic effects do not affect their gait, so these effects are not considered in the control system, making it simpler with respect to other legged robots with a lower number of legs. To guarantee equilibrium the robot's center of mass has to be inside the projection of the robot on the ground. Moreover, quadrupeds enable dynamic gaits: under certain conditions the number of contact points can be reduced to two or even one. Although more complex, this kind of gait can reach high speed. *Exapods* are legged mobile robots with six legs. Usually the gaits implemented on exapods are only static.

A particular class of legged mobile robot is that of *bipeds*. These robots can implement either static or dynamic gaits[*]. Dynamic gaits are more interesting, because they are more efficient and guarantee higher speed. The advantages of bipeds are:

- The low number of legs imply a lower number of links and actuators.
- They can easily move and work in the same environments designed for humans beings (stairs, lifts, …).
- The interaction can be more friendly due to their humanoid aspect.

Indeed, research on humanoid robots is in progress; and many prototypes have been built, especially from USA (e. g. Boston Dynamics [11]) and Japan (Honda [12]).

7.4.4 Hybrid Mobile Robots

This class of robot is reserved for very special applications, for example planetary exploration or very rough ground, such as steps or surfaces with high slopes. The robots belonging to this class rely on a mixed "wheels-legs" traction system that tries to combine the advantage of both. One of the most famous robots in this class is the NASA All-Terrain Hex-Limbed Extra-Terrestrial Explorer (ATHLETE,) prototype [13].

The ATHLETE (Figure 7.4-4) vehicle concept is based on six 6 d.o.f. limbs, each with a 1 d.o.f. wheel. ATHLETE uses its wheels for efficient driving over stable, gently rolling terrain, but each limb can also be used as a general purpose leg. In the latter case, wheels can be locked and used as feet to walk out of excessively soft, obstacle laden, steep, or otherwise extreme terrain. ATHLETE is envisioned as a heavy-lift utility vehicle to support human exploration of the lunar surface, useful for unloading bulky cargo from stationary landers and transporting it across long distances.

[*] The word "static gait" is referred to a movement of the robot where the dynamic forces are negligible, so they are not considered into the dynamic model used to control the robot. When these forces are important the movement is classified as "dynamic gait" and these forces must be carefully considered for the motion control of the robot

Figure 7.4-4 The robot ATHLETE, climbing a hill with high slope. Image rights: reproduced courtesy NASA/JPL-Caltech.

Another way to integrate the advantages of legged and wheeled mobile robots are *Whegs* (Wheel + Leg). Wheel+Leg is a mechanism that includes good qualities of both wheels and legs. The result of that is a good passing ability in different terrain including stairs and steps. In the smooth terrain the wheel regime is used. When terrain changes to hardly passable the wheel-leg adjusts itself.

A famous example of robot based on *whegs* is RHex by Boston Dynamics Figure 7.4-5. RHex is an autonomous robot, based on a hexapod with compliant legs and one actuator per leg. It has shown good mobility over a wide range of terrains at speeds of 2.7 m/s, climbs stairs and slopes exceeding 45 degrees and swims [14][15].

Figure 7.4-5 Rhex robot images provided courtesy of Boston Dynamics, Inc. For more information see the video at this link: https://www.bostondynamics.com/rhex. Image rights: Boston Dynamics, Inc.

7.5 Conclusions

The role played by robots in modern manufacturing industry is already essential and forecasted to increase. However, the majority of current industrial robots lack sufficient intelligence and flexibility to accomplish complex tasks. The uncertainties and variables of unstructured environments require sensing and adaptive features not yet fully present in the robots. Industry 4.0 aims at providing the robots with innovative technological features, focusing on safety, flexibility, versatility and collaboration, enabling them to complete tasks intelligently. These innovative concepts will have a huge impact on both processes and business models. Flexibility and reconfigurability will allow the production of small lots, and even single unit, at low price due to the ability to rapidly configure machines and fulfil customer's requests. This will boost innovation since prototypes and several variants can be easily and quickly produced. Innovation will be further promoted by digital integration since modelling and simulation will reduce the time between design and production. Production quality will improve due to the extended sensorization which allows monitoring every product rather than just a few samples. It will not only be errors that can be immediately detected, but also machine malfunctioning, improving also the production efficiency. This new generation of robots is also suitable for smaller companies since the investment is largely repaid thanks to the flexibility and high number of usage options. Overall, thus, Industry 4.0 robotics will increase competitiveness and economic growth.

The next chapter, Chapter 8, will focus on sensors and actuators, providing basic and advanced principles, and introducing the link between physical and cyber world.

Chapter 9 will go deeper into the main concepts of Industry 4.0, considering the software, or cyber, elements, describing the cyber-physical system and some applications. Finally, an overview on Artificial Intelligence and the related ethical issues is presented.

7.6 References

[1] International Organization for Standardization. “ISO 8373: 2012 (en) Robots and robotic devices–Vocabulary”.

[2] IFR World Robotics 2019.

[3] International Organization for Standardization. “EN ISO 10218-1: 2011 Robots and robotic devices — Safety requirements for industrial robots — Part 1: Robots”.

[4] Legnani G. Fassi I. Robotica industriale. Modellazione, pianificazione, controllo, programmazione, componentistica, normativa e sicurezza, Citta Studi edizioni, 2019.

[5] EN ISO 10218-2:2011 - Robots and robotic devices - Safety requirements for industrial robots - Part 2: Robot systems and integration.

[6] J. Lee, S.D. Noh, H.J. Kim, and Y.S. Kang, “Implementation of cyber-physical production systems for quality prediction and operation control in metal casting”, *Sensors*, *18* (5), p.1428, 2018.

[7] https://ec.europa.eu/assets/jrc/ilv/edge/pcs/index-63.html.

[8] ISO/TS 15066:2016 – Robots and robotic devices – Collaborative robots (Technical specification).

[9] https://exoskeletonreport.com/2016/04/exoskeletons-for-industry-and-work/

[10] Roland Siegwart, Illah R. Nourbakhsh, Davide Scaramuzza – Introduction to Autonomous Mobile Robots – Second Edition, The MIT press, ISBN 978-0-262-01535-6, 2011.

[11] https://www.bostondynamics.com/ [Accessed 16 07 2019]

[12] https://it.wikipedia.org/wiki/ASIMO [Accessed 16 07 2019]

[13] https://www-robotics.jpl.nasa.gov/systems/system.cfm?System=11 [Accessed 16 07 2019]

[14] https://en.wikipedia.org/wiki/Rhex [Accessed 16 07 2019]

[15] https://www.bostondynamics.com/rhex [Accessed 16 07 2019]

Chapter 8
Robot Components

Gianmauro Fontana, Claudia Pagano and Marcello Valori

8.1 Introduction

The availability of digital technologies, in terms of both hardware (sensors, high computation capabilities, PLC, etc.) and software (big data analytics, artificial intelligence, etc.) tools, enables the development of advanced smart factories, based on the synergistic development and convergence of computer science, Information and Communication Technologies (ICT) and manufacturing science and technology. In a smart factory, huge amounts of information are collected in real time from the different processes and machines on the factory floor from smart sensors. Afterwards, these data are integrated in the information system, where they are analysed in order to improve the quality and efficiency of the production system controlling the actuators. Production and information systems are constantly synchronized so that the response time to unpredicted events can be minimized.

A robot is a complex machine composed–of different elements which have different functions, such as mechanical architecture, actuators, sensors and a control system. Considering the robot as a black box, an input energy, properly modulated by input commands, is translated into coordinated motion of the mechanical parts.

In order to accomplish this task, the action of an actuator is governed by the control system, that can modify in real-time such an action by combining sensor readings with planned tasks.

A diagram of the process is outlined in Figure 8.1-1.

This chapter is focused on basic and advanced principle of sensing and actuation of robots, since they allow the monitoring and control of the physical system that is the foundations of a smart factory.

G. Fontana, C. Pagano(✉), M. Valori
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, Milan, Italy
e-mail: claudia.pagano@stiima.cnr.it

J.C. Chaplin et al. (eds), *Digital Manufacturing for SMEs*
DOI: 10.17639/J6AM-VP75

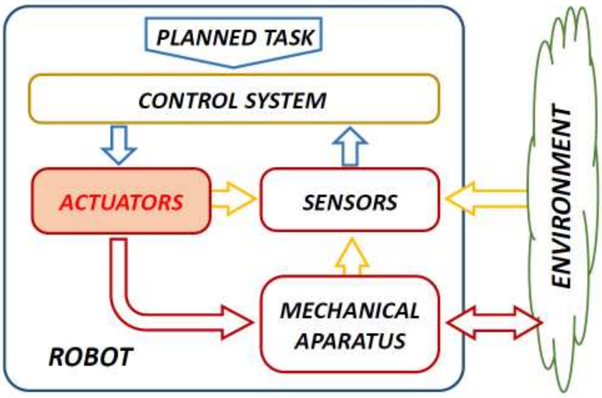


Figure 8.1-1 Block diagram of robot components.

Sensors are also discussed in section 9.2 of Chapter 9 in the context of digital twins of manufacturing sensors. However, they perform an essential role in robotic systems, so a detailed discussion in the context of robotics is provided in the next paragraphs.

8.2 Sensors

Robot sensors are the components of a robotic system in charge of collecting information about its internal state or features of the external environment; the awareness about these aspects is fundamental for the robot to perform its tasks.

Sensors can be thought of as robot senses, as they can provide the robot with different kinds of information depending on the task requirements. In order to achieve this goal, different sensor types are used to measure different physical quantities. Whenever the sensor information is acquired, the robot control system interprets them and acts to execute its task.

Sensors fall into one of two categories:

- *Internal Sensors:* Which measure physical values internal to the system such as the position, velocity, acceleration, forces, torques of robot joints, and inertia of the robot links. Based on these measurements the control system controls the actuators of the robot joints in order to perform the desired motion correctly and accurately.
- *External Sensors:* Which gather information about the environment surrounding of the robot, such as distances to objects, light and temperature measurements, and forces encountered when interacts with the world.

Figure 8.2-1 provides some examples of each of these categories.

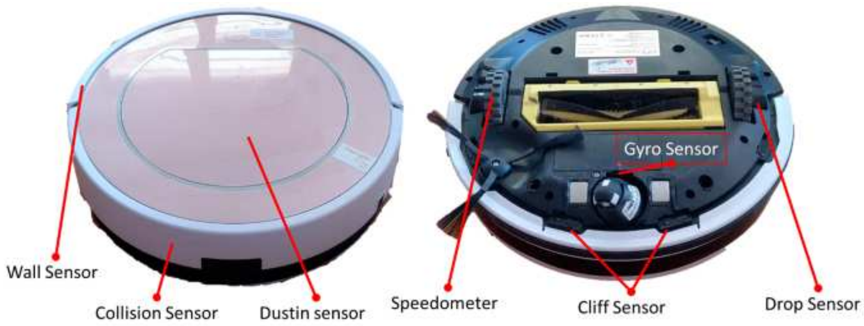


Figure 8.2-1 The top and bottom view of an autonomous vacuum cleaner with internal sensors (gyro sensor, speedometer, suction sensor) and external sensors (dust-in sensor, cliff sensor, infrared sensor receiver, wall sensor, collision sensor, drop sensor).

8.2.1 Sensor transduction principle

Each sensor is based on a *transduction principle*; a conversion of energy from one form to another. Typically, sensors convert to electrical energy to enable the robot to read a signal generated from the measurement of a physical quantity.

Even though sensors might be based on more than one transduction principle, a classification based on the transduction principles [1] can be useful to identify materials, costs and time of response involved and, thus, to choose the most appropriate for the application.

- *Mechanically-based Sensors:* These work as simple switches; the contact closes the circuit inside the sensor and the current is detected. Releasing the contact opens the circuit and no current flows. A minimum force threshold can be set as the requirement for the switching.
- *Resistive-based Sensors:* These detect mechanical changes by detecting resistance changes of the device. In their simplest form they consist of two electrodes separated by a layer of deformable material, whose resistance varies according to the shape it assumes as a reaction to a mechanical stimulus. The simplest resistive sensor is the potentiometer that measure the position or the displacement. Other resistive sensors include strain gauges (to measure the force/torque, strain or acceleration), thermocouples (to measure the temperature), photoresistors (to measure light intensity) and thermistors (to measure the temperature).
- *Capacitive-based Sensors:* These are similar to resistive approaches, where the change of distance between plates or of an area of material due to an external input changes the capacitance of the material. Many types of sensors use capacitive sensing, including sensors to detect and measure proximity, pressure, position and displacement, force, humidity, fluid level, and acceleration.

- *Optical Sensors:* These detect a change in the intensity, phase, or polarization of the transmitted or reflected light and convert it into an electronic signal. They are not affected by electromagnetic interference, are intrinsically safe, and require fewer electrical wires. They are mostly used as position sensors, that activate when an object interrupts a light beam, as photoelectric sensors that detect the distance, absence, or presence of an object, or camera sensors to acquire an image.
- *Optical Fibre-based Sensors:* These use fibres not only as light transmitter but as sensors, as when these fibres deform due to an external stimulus, they produce a variation in the intensity, phase, polarization, wavelength or transit time of light of the transmitted light. Optical fibres are mostly used as sensors to measure strain, temperature, and pressure.
- *Piezoelectric Sensors:* These are based on the property of piezoelectric materials, in that they generate an electrical potential difference when deformed, or vice versa (in this case they are used as actuators). They are used to measure changes in pressure, acceleration, temperature, and strain.

8.2.2 Contact Sensors

External sensors can be classified as either *contact sensors and non-contact sensors*, depending on whether the sensor needs to physically touch the object to function.

Contact sensors are largely used for obstacle avoidance: when a physical contact between the robot and the external environment occurs it triggers the robot to execute a task, such as a movement aimed at avoiding contact.

Force sensors can be *quantitative* where they detect the value of the force, or they can be *qualitative* where the output shows only whether the force overcome a set threshold value.

A common application for these sensors is the need to assess forces or torques that appear during manipulation. This aim can be achieved by directly measuring force/torques on the end-effector wrist (that is the joint holding the end-effector) or sensors within robot joints.

In the first case (see Figure 8.2-2), forces and torques are measured by the use of multi-axes force sensors, mainly consisting of mechanical structures with elastic elements built-in, whose deformation is measured. Since the structural behaviour of these elements is well-known, the applied load can be determined depending on the effective deformation. The deformation can be assessed with different approaches, but the most common relies on the application of *strain gauges;* specific components whose electrical resistance changes depending on the deformation.

Some recent robots equip the gripper joints or directly the gripper tips with force sensor, in order to control the grasping force and compute the contact points. Sensors useful to fulfil this goal include single-axis force sensors as strain gauges, piezoelectric sensors, pressure sensors, or multi-axis force sensors as presented above for the robot wrist.

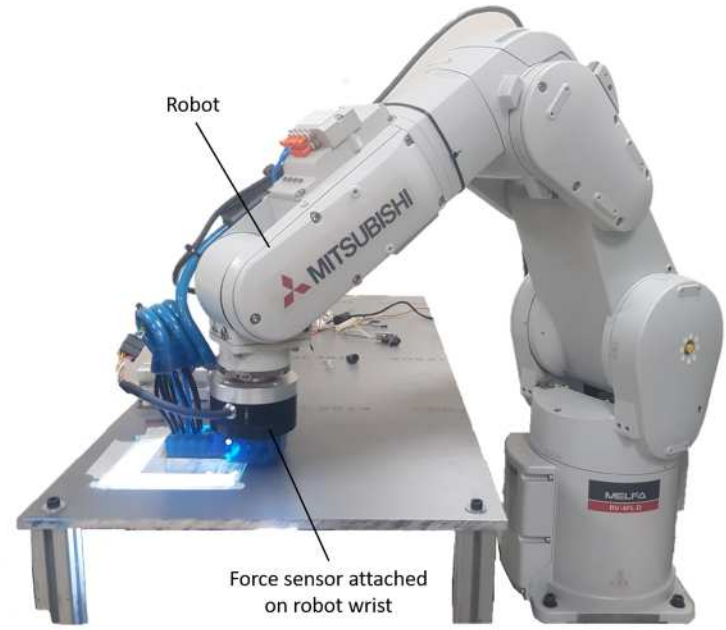


Figure 8.2-2 Anthropomorphic robot with a force/torque sensor attached on wrist.

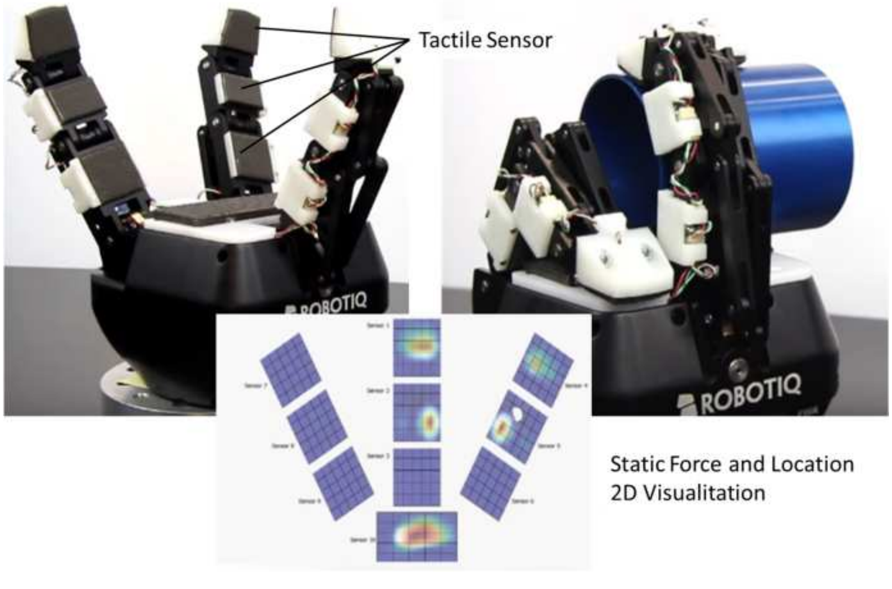


Figure 8.2-3 Tactile sensor on industrial robot gripper – 3-Finger Adaptive Robot Gripper. The array sensor is able to detect force position and intensity, as shown in the 2D visualization. Source: Authors based on Robotiq (www.robotiq.com).

Tactile sensors consist of arrays of sensors (and are hence sometimes called array sensors), whose signals are collected individually to detect the type of physical contact analogous to the human sense of touch.

They can be defined as the continuous sensing of all the various effects originated by contact between the robot (usually the end-effector or the gripper fingertips) and an object. Most significant contact-based effects are contact stresses and forces in the two contacting surfaces. The tactile sensors are able to measure the variable contact forces over an area with a specific spatial resolution.

The tactile sensors on the fingers of the robot gripper allows the robot to detect the presence of the object, the pressure distribution, the object shape, surface texture and roughness, the object pose (i.e. position and orientation), and contact locations, and any slipping information.

The tactile sensors are used on the gripper fingertips, as illustrated in Figure 8.2-3, but they can be assembled on the robot links.

The tactile sensors typically used in robotics are array sensors that measure the pressure of contact using the deformation of an elastic skin/overlay. An example of this type of sensor is presented in [2]; it is an elastic capacitive pressure sensor array based on a thin all-elastomeric platform. In particular, the sensor consists of 16 individual capacitively-coupled pressure sensing cells to detect the external tactile information in a 4x4 arrayed configuration [2]. Each capacitive pressure-sensing cell is composed of a pair of CPDMS (Capacitive PolyDiMethylSiloxane) sensing electrodes facing each other across the elastomeric insulating layer (Ecoflex) to construct a parallel-plate capacitor. The applied pressure causes the local change in distance between the CPDMS pairs (since the Ecoflex layer is more easily deformed than the PDMS) and thus the consequent change of the capacitance, which can be detected to obtain pressure and the normal force.

8.2.3 Non-contact Sensors

Non-contact sensors are used to give the robot information about the external environment without the necessity of physical contact.

Proximity sensors detect the presence of objects that are close to the surface of the sensor but not in contact. In robotics these sensors are used to detect possible collisions between a robotic arm and the environment. This could be desirable if the approaching object is to be acted upon, or may cause an undesirable collision so an avoidance plan should be enacted or the robot halted.

Range sensors are devices that can provide precise measurements of object distances, usually measuring the gap between the sensor and the nearest surface. Moreover, they enable the scan of the three dimensional (3D) structure of the environment from the sensor viewpoint. In the robotics field, range sensors are useful for locating objects within the robot working space and also as a feedback in a closed loop control scheme [3]. They are used for robot navigation, obstacle avoidance, and to reconstruct the third dimension from a 2D vision system like a camera.

Range sensing techniques can be based on one of two principles: *triangulation* or *time-of-flight*.

The *laser-based triangulation* working principle is illustrated in Figure 8.2-4: a laser beam is projected from one position onto the surface. The light spot (that can be also a light stripe to acquire simultaneously more points) is observed by a vision sensor from a second position. With the geometrical information of the relative positions and orientations of the laser source and sensor, it is possible to calculate the 3D position of the illuminated surface point or its distance. In order to capture a 3D image, the laser-based triangulation sensor has to move relative to the environment, performing a scanning process. Alternatively, the laser spot can be reshaped with lenses or mirrors to create multiple spots or stripes allowing the measurement of multiple 3D points simultaneously, without movement.

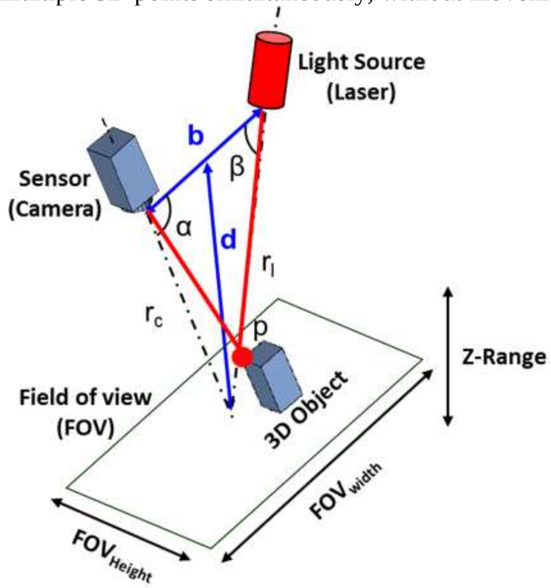


Figure 8.2-4 The physical principle of triangulation sensors: a laser source emits a light, generating a spot on a 3D surface to be measured and the sensor detects the spot on its image plane. Combining this 2D information with the geometrical parameters of the system, it is possible to calculate the 3D position of the projected spot in the environment. A collection of these points (by moving the triangulation system) enables surface reconstruction. The fundamental parameters of the system are the baseline (b) and stand-off distance (d). The former is the distance between the optical centres of the camera and the laser, and it affects to the measurement Z-range, the latter is the distance from the baseline to the focal plane of the camera, and it affects the Z-resolution.

Time-of-flight range sensors compute distance by measuring the travel time of a signal (e.g. light) to cover the path *source* - *object* - *detector*; the detector is usually located close to the source. Laser-based and sonar range sensors are the most common time-of-flight sensors. The former can be considered a radar that is based on light emission and detection. A sonar employs acoustic pulses and their echoes to measure the distance of an object and also its 3D position. As light and sound speed are known, it is possible to obtain the distance from the delay time.

Vision Systems also allow information on the environment to be gathered with no contact with the surrounding. The acquisition of data based on computer vision is a complex sensing process consisting of extraction, characterization and interpretation of the information provided by the images in order to identify objects in the environment. A vision system consists of one or a series of lenses, associated with a vision sensor that converts the visual information into electrical signals. These signals are then analysed by an image digitizer, called a frame grabber, to obtain a digital image.

8.3 Actuators

In the human body, movements are generated in the brain's primary motor cortex, translated and transferred to the muscles, which are able to transform available energy in mechanical motion (see Figure 8.3-1). The muscle appears as the principal contributor to the process, the "*motor*" enabling movement occurrence.

But… what is the muscle of a robot?

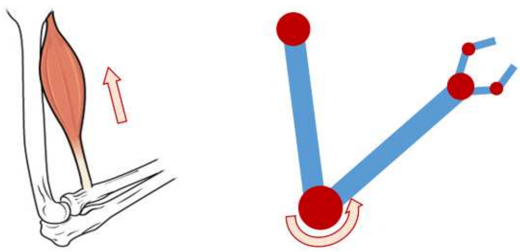


Figure 8.3-1 Human muscles* vs robot actuators. Image rights: OpenStax

In robots this role is played by *actuators*, which are the components responsible of the motion of links, in accordance to desired trajectories. An actuator converts a primary, available form of energy, into mechanical energy to operate the robot.

A robot capable of moving the end-effector in any configuration (limited by its own working volume) is equipped with at least 6 actuators, each one responsible for the motion along/around a single axis. An example is shown in Figure 8.3-2.

* https://commons.wikimedia.org/wiki/File:1015_Types_of_Contraction_new.jpg (partially reported), Link to license

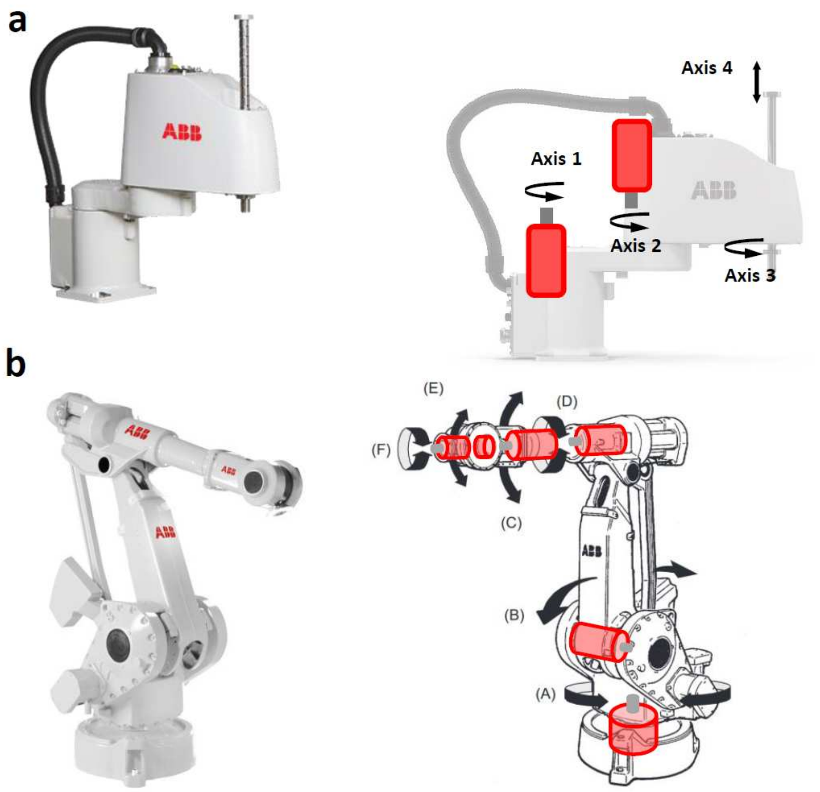


Figure 8.3-2 The four axes of the ABB IRB 910SC (4 axes) VS the six axes of the ABB IRB 4400 (6 axes) with the positions of some actuators highlighted in red*.
Image rights: ABB

An actuator must be properly sized to support loads and inertias, as well as to comply with performance requirements: dynamic modelling of robots allows precise assessment of the performance required for each individual actuator, that can be primarily represented as the power it can exert.

Power is a suitable parameter in order to compare different forms of energy: as already explained, actuators convert power, provided as a certain source, providing mechanical power to links and joints. Even if the effective power is the appropriate parameter to identify the size of an actuator, it is not sufficient for the complete definition of an actuation unit. There are, indeed, other features which characterize the performance required to properly move a mechanical architecture, that accurately specify the most appropriate actuation unit.

* https://new.abb.com/products/robotics/industrial-robots

8.3.1 Type of motion

A first classification of actuators can be defined on the basis of the *type of motion* generated by the action of the actuator:

- *Linear Actuators:* Provide motion along a straight direction and their own geometry limits the stroke; an example outside the robotic world is represented by hydraulic actuators of heavy duty machines (Figure 8.3-3a).
- *Rotary Actuators:* Can rotate, usually spanning 360° and performing more rounds; position limits are instead imposed by the robotic architecture. With appropriate differences, an example in our daily experience is represented by the motor driving a washing machine (Figure 8.3-3b).

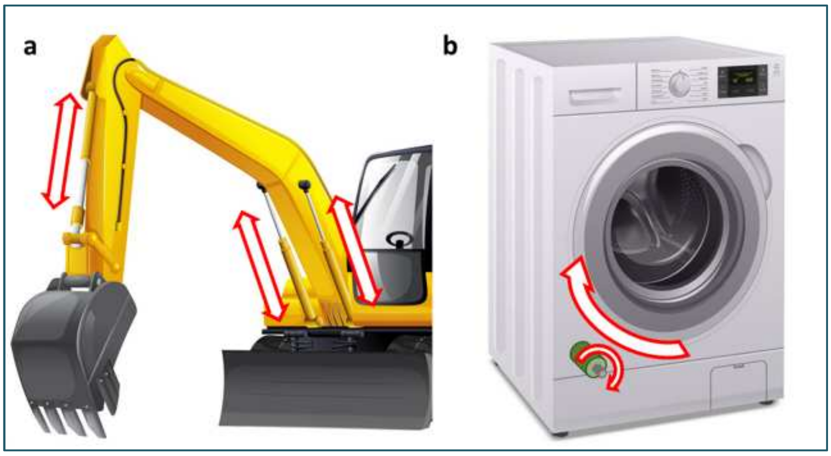


Figure 8.3-3 Examples of actuators in other application fields: a) hydraulic cylinders of overhaul machines (linear) and b) the electrical motor of a washing machine (rotary)*.

Considering the concept of mechanical power, it results as the combination of *force* and *velocity*. This is valid for all mechanical systems, regardless the type of movement. In linear actuators, this is expressed as $P = F{\cdot}V$, whereas in rotary actuators $P = T{\cdot}\omega$, with *torque* (T) representing a form of "rotational force", and ω the angular velocity. It follows that a more detailed specification for actuator performance defines not only the power, but also the maximum force (or torque) and velocity (linear or angular). Due to this, specific speed reduction gears are usually installed in a robot downstream of an actuator in order to properly modulate force and velocity and match the requirements of the particular implementation.

* Images designed by Freepik

Industrial robots sometimes have complex architectures, that are defined by taking into account dynamic features such as dexterity in certain zones of the working envelope and load/inertia distribution. Due to this, actuators, that can be heavy and bulky, often cannot be installed in direct proximity of the joints to be controlled. The need arises for systems able to transfer the power from actuators to joints, leading to the implementation of different motion transmission strategies, based on belts, gears and/or other mechanical systems. According to these considerations, a general actuation unit can be defined as composed of an actuator (the motor unit), reduction gears and a mechanical transmission, as represented in Figure 8.3-4.

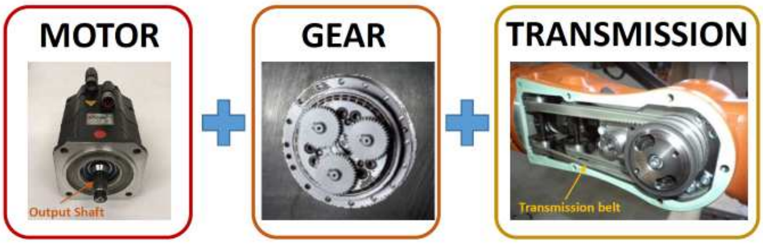


Figure 8.3-4 Examples of robot electrical motor, planetary gearhead and belt transmission*. Image rights: Dr. Arun Dayal Udai.

8.3.2 Power source

A second fundamental classification of actuators is based on the source of power exploited to generate mechanical motion:

- *Electrical actuators*: These represent the most common choice for industrial robots and are mostly based on direct current (DC), though versions based on alternate current (AC) are available. The most popular are rotary actuators but linear models also exist. Electric actuators (or, electrical motors) are characterized by high compactness and reliability. Due to this, and considering the availability of electricity in all plants, they are actually the most widespread form of actuator in the robotics industry. Electric actuators generate extremely high velocities, and as a result they are always coupled with specific reduction gears.
- *Hydraulic actuators:* These use oil as an incompressible fluid, hydraulic actuators usually provide linear motion, though rotary models are available. They are very powerful, but require maintenance and dedicated circuits for oil

* https://www.youtube.com/watch?v=iRKDfknqtbc (Property of Dr. Arun Dayal Udai, arun_udai@yahoo.com)

circulation. Their implementation in robots is also widespread, especially in "legged" designs.

- *Pneumatic actuators*: They represent the air-powered version of hydraulic actuators. The use of pressurized air does not guarantee sufficient stiffness with high payloads; however, pneumatic actuators do enable a number of manufacturing tasks, to be performed with extreme simplicity (push, pull, lift, positioning, tightening, blend, cut, punching, etc.), even if extreme precision control is not possible. They are mostly linear, they are implemented in stop-to-stop trajectories, such as certain pick-and-place tasks.

The main advantages and disadvantages of the different actuator power sources are presented schematically in Figure 8.3-5.

Power sources	Main advantages	Main disadvantages
Electric actuators	Speed and precision Easy control and, thus, possible complex control algorithms Low cost Reduced dimensions and weight	Reduction gears are a source of mechanical losses and imprecision Limited available power
Hydraulic actuators	High performance (force/speed) Intrinsic backdrivability: does not require energy to maintain a pose	Disadvantageous cost/performance ratio Noise and leakage Bulky envelope
Pneumatic actuators	Low cost High velocity No toxicity in case of leakage	Limited accuracy (due to fluid features) Noise (less than hydraulic) Need for air filters and maintenance

Figure 8.3-5 Main advantages and disadvantages of the most common actuators.

Although the actuators listed above represent the main actuators used in industrial robotics, there are some other actuators which exploit particular physical principles and are worth mentioning.

Piezoelectric actuators use certain crystalline materials, like quartz, which modify their configuration and physically deform when a voltage is applied to the material (see Figure 8.3-6). This deformation, proportional to the electrical field, is very small but also very precise, enabling positioning in the order of µm. It follows that their application can be foreseen in the assembly of products at the microscale. It is worth noting that their working principle is reversible and they can be used as positioning sensors by measuring the induced voltage.

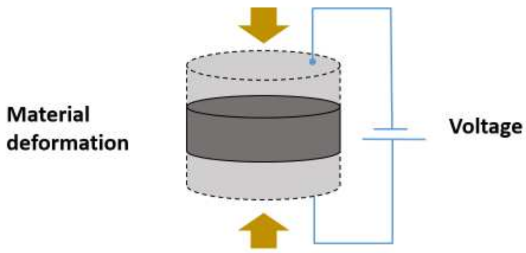


Figure 8.3-6 Working principle of a piezoelectric actuator.

Shape Memory Alloys (SMA) are metallic alloys which are able to “remember” their original position. In order to restore their initial configuration, it is sufficient to heat the material by applying an electrical current (Joule effect). The forces generated are considerable. Currently, their implementation as actuators is still experimental and being investigated in several fields which have demanding geometrical requirements, due to their great advantage of being able to exert a force without the need for bulky actuators.

8.4 Microrobotics

The achievements of miniaturisation are becoming more and more evident in everyday life; miniaturised devices are present in an increasing number of applications, especially electronics devices – such as smart phones, laptops, vehicle sensors and household goods – but also biological probes, medical systems and military devices. In robotics this trend leads to innovative solutions in a large variety of application fields, such as the maintenance and inspection in industry, non-invasive surgery in medicine and micro operation in biology. Due to their reduced dimensions miniaturised robots can perform tasks which are in small spaces that are not accessible to humans. Microrobotics is still a very promising research topic even though several research prototypes have already been developed [4][5][6][7]. Microrobotics is often mentioned as part of the Industry 4.0 technology portfolio - including also sensors, drones, virtual and augmented reality, additive manufacturing – which can address the challenges of the Sustainable Development Goals, set up by the United Nations [8].

In the literature the term *microrobot* refers to a large variety of robotic devices. It is generally applied to *all kind of robots that perform tasks in the microworld*.

However, the term microrobots is more correctly used for *miniature robots whose size is in the order of a few cubic centimetres*. The rationale behind this broad definition of the term is that both categories share some of the main challenges: non-intuitive physics**,** difficult fabrication and observation.

Like conventional robots, microrobots are integrated systems consisting of sensors, actuators and a logic circuit. The most challenging aspect in the development of microrobots is the fabrication of micro-actuators and micro-sensors which can give high efficiency and high stability. Sensors that can be used in small spaces are often not very precise so that the knowledge of the environment is not complete and the control system has to cope with this. Moreover, the *microcontrollers* do not usually have the processing power and memory common at the macroscale. Therefore, simplified control methods are usually preferred at the microscale. Mechanics are intrinsically more robust on the small scale [9]; this is an advantage for microrobots where space can be saved for more critical components.

The assembly of such small components is another relevant issue for microrobots. The scale effects and the consequent predominance of superficial forces affect the operation of a micromanipulator [10]. In particular, the releasing phase of a manipulation is critical and several strategies have been investigated and tested. As for micromanufacturing, each method has advantages and disadvantages which make one method more suitable in some applications than others.

All these issues made not only the fabrication, but also the design of microrobots very complicated. There is a strong dependency between the control and the model and the entire system, cyber and physical parts have to be conceived at the same time and any change in one of them is reflected in a redesign of the other. This makes microrobots complicated cyber-physical systems, more deeply discussed in the next chapter.

8.4.1 Power

One of the main issues related with the miniaturisation of robots is energy; due to the size reduction the space for the power source is reduced and, in addition, leakage currents increase. Therefore, the minimization of power consumption is essential; the operating voltage must be as low as possible and the energy required by all components minimized.

The solutions most commonly used are batteries and supercapacitors. Batteries are suitable in terms of output and durability, but are difficult to miniaturise. Supercapacitors have lower voltage limits and offer lower energy density but higher currents for charging and discharging. Moreover, supercapacitors can be recharged allowing them to be combined with power generators, such as solar cells. Indeed, wireless power sources, such as radio frequency, optical power, and energy harvesting are suitable and often advantageous due to the miniaturisation.

8.4.2 Control

The controller of a robot has to process information and generate suitable actions. Smaller devices tend to deal with less complex and more restricted environments and make smaller and slower movements, thus they might require a simpler and not so fast control system. Nevertheless, due to the very small dimensions of microrobots, onboard processors with enough calculation power are still a challenge.

8.4.3 Sensors

In order to perform its task a robot should be equipped with as many sensors as necessary to perceive the surrounding environment. Mobile robots have to be able to detect obstacles at a great enough distance so they can be avoided; thus, touch sensors, distance sensors and/or proximity sensors are.

Sensors for microrobots have to both be reduced in size, and their power consumption minimized, which can be a major issue. Therefore, passive sensors, which do not supply energy into the environment, or very simple active sensors are the most suitable sensors for microrobots. Cameras and microphones are, thus, very common as passive sensors in microrobots, together with strain gauges which can be easily downsized. Active infrared proximity sensors, which are simple to use, inexpensive and can be found in compact packages, are also very suitable.

8.4.4 Actuators

Actuators are one of the major issues in designing miniature robots. The choice of the actuation principles for the design of a microrobot has to reach a compromise in the range of motion, force, actuation frequency, power consumption, control accuracy, system reliability, robustness, load capacity, etc.

Figure 8.4-1 lists the most common types of actuation and their main characteristics. Each solution offers advantages and disadvantages, but none is perfectly suited to all applications.

Driving principle	Actuator	Driving range	Yield	Response speed	Driving voltage
Static electricity	Electrostatic rotary motor	Large	Small	Medium	High
	Electrostatic linear	Small	Medium	High	High
Electro magnetic induction	Ordinary motor	Large	Small - Medium	Medium - High	Low
Piezo electricity	Bimorph	Small	Medium - High	High	High
Heat	Shape Memory Alloy	Large	Large	Low - Medium	Low

Figure 8.4-1 Typical characteristics of the most common actuator types [11].

Electrostatics, electromagnetics and piezoelectrics are the most commonly used solutions for actuation at small scale.

Electrostatic and electromagnetic fields can be rapidly created and interrupted, allowing very fast actuation. Moreover, electrostatic fields can exert great forces, but in a very short distance, unless high voltage is used. The extremely low current consumption associated with electrostatic devices makes for highly efficient actuation. Electromagnetic fields offer the advantage of converting electrical energy into mechanical work with high efficiency, leading to a low current consumption. The main disadvantage is the poor scalability of the electrical magnets.

Piezoelectric materials (zirconate titanate, PZT, quartz, SiO2, lithium niobite) are also very common. Their molecular crystals show a dimensional change when in an electric field and, vice versa, produce a voltage when deformed. They show a very quick response to the input, great repeatability and high force, but they exhibit very small strokes (under 1%). Thus, they are usually combined in piles or in stick-slip modes to obtain larger strokes.

Other smart materials also become advantageous at the microscale [12]. Shape memory alloys (SMA) are functional materials that have two stable solid crystallographic phases. They can transform reversibly from a crystal structure to the other structure upon heating and generate mechanical work during the phase transformation. They offer large displacement and actuation force within an extremely small volume and a low operating voltage. Some problems, such as low dynamic time response and large hysteresis, limit their use at the macroscale, but these problems are less relevant at the microscale.

Electroactive polymers (EAP) are very promising materials. A very large deformation (10% - 400%) can be achieved when a voltage is applied across the polymeric film, coated with electrodes on both sides. They are lightweight, inexpensive, fracture tolerant and fast, but they usually require a high DC voltage (>150V/μm), which is also very close to the breakdown voltage of the material. Moreover, their applications are limited by low actuation forces.

8.5 Conclusions

Sensors and actuators are the main components of robots and in continuous development in order to improve their performance and meet new challenges. In the context of Industry 4.0 further advancements are required to build smart factories where all the devices are integrated in a network system and the entire production chain is automated. In smart factories, smart sensors serve as the interface between the digital and physical world. They consist of the combination of a sensor, a microprocessor and a communication system so that environmental data is monitored and transmitted, in real time, to the main control system. This data is used to activate the actuators in order to automatically manage the process and maintenance, improving the efficiency and the quality of the manufacturing. A huge amount of data is collected to monitor the production processes, the environment and its modifications. This improves health and safety in the work space, in particular where cooperation between robots and humans occurs.

8.6 References

[1] J. Fraden, Handbook of Modern Sensors: Physics, Designs, and Applications, III Edition, Springer International Publishing, 2016

[2] S-J. Woo et all, "Thin All-Elastomeric Capacitive Pressure Sensor Array Based on Micro Contact Printed Elastic Conductors", *Journal of Materials Chemistry C* vol2, pp. 4415-4422, 2014

[3] R. B. Fisher et al., Range Sensors, Springer Handbook of Robotics, II Edition, Ed. Siciliano et Khatib, Springer Berlin Heidelberg, pp 455-476, 2016.

[4] G. Caprari, T. Estier, and R. Siegwart, “Fascination of down scaling- alice the sugar cube robot”, *Journal of Micro-Mechatronics*, vol. 1, no. 3, pp. 177–189, 2002.

[5] S. Kernbach, R. Thenius, O. Kernbach, and T. Schmickl, “Re-embodiment of Honeybee Aggregation Behavior in an Artificial Micro-Robotic System,” *Adaptive Behavior*, vol. 17, no. 3, pp. 237–259, 2009.

[6] F. Mondada, M. Bonani, X. Raemy, J. Pugh, C. Cianci, A. Klaptocz, S. Magnenat, J.-C. Zufferey, D. Floreano, and A. Martinoli, “The e- puck, a robot designed for education in engineering,” in Proceedings of the 9th conference on autonomous robot systems and competitions, vol. 1, no. 1, pp. 59–65, 2009.

[7] M. Rubenstein, C. Ahler, and R. Nagpal, “Kilobot: A low cost scalable robot system for collective behaviors,” in *IEEE International Conference on Robotics and Automation*, pp. 3293–3298, 2012.

[8] Kumar A, and Nayyar A, “si3-lndustry: A Sustainable, Intelligent, Innovative, Internet of Things Industry. In: Nayyar A., Kumar A., editors. A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development. Springer; p. 1-21, 2020.

[9] Zhu, TT, Bushby, A.J. and Dunstan, D.J., “Materials mechanical size effects: a review”, *Materials Technology*, 23(4), p. 193-209, 2008.

[10] A. Ghosh, “Scaling laws”, in Mechanics over micro and nano scales, p. 61-94. Springer, New York, NY 2011.

[11] P. Dario et al., “Microactuators for microrobots: a critical survey”, *Journal of Micromechanics and Microengineering*, 2(3), pp. 141-57, 1992..

[12] C. Pagano, I. Fassi, “Introduction to miniaturisation in Micro-manufacturing technologies and their applications”, in I. Fassi, D. Shipley, editors. Micro-Manufacturing Technologies and Their Applications. Springer Cham; p. 1-22, 2017.

Chapter 9
Industry 4.0 in Robotics

Irene Fassi, Claudia Pagano and Serena Ruggeri

9.1 Introduction

The so-called 4th industrial revolution is going to radically change the common idea of the factory, and a great industrial research effort is dedicated to the definition of the “Factory of the Future” models.

The fourth industrial revolution is characterised by the word *automation*, which is achieved in industrial production by addressing four challenges:

1. Intelligent machines control the production in smart factories in which the presence of human is enormously reduced.
2. Intelligent machines optimize the capacity of the production facility, coordinating the movement of the material, analysing the status of the production chain and the stock, and zeroing the downtime.
3. Intelligent machines are essentially self-organized; material planning and the handling of orders are fully automated.
4. Intelligent machines can autonomously reconfigure the production line to respond in a very short time to the personalized customer request.

Intelligent machines are based on a complex cyber part and artificial intelligence algorithm. Indeed, Industry 4.0 mainly has two foundations: the hardware layer, which includes all the physical elements, and a software layer, the cyber part. Robot sensing and actuation, which belong to the hardware layer, were discussed in the previous chapter. In this chapter the cyber part is discussed, focusing on programming, control and artificial intelligence. Moreover, a brief overview of the

---

I. Fassi, C. Pagano(✉), S. Ruggeri
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, Milan, Italy
e-mail: claudia.pagano@stiima.cnr.it

J.C. Chaplin et al (eds), *Digital Manufacturing for SMEs*
DOI: 10.17639/dt0m-a649

ethical issues related with the intelligence implemented in the machines is provided at the end of the chapter.

9.2 Cyber Physical System

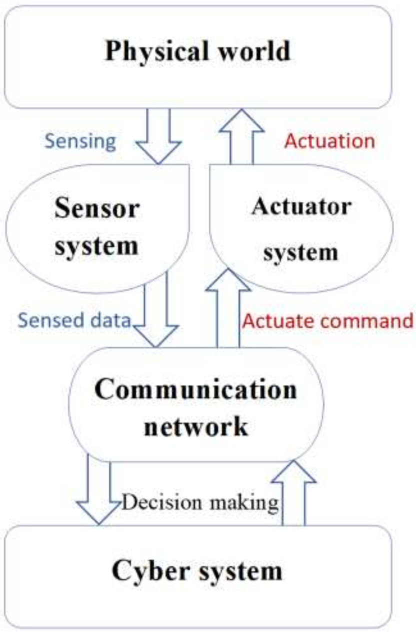


Figure 9.2-1 CPS holistic view.

Smart factories require the synergic deployment of several technologies, and among these the *Cyber Physical System* (CPS) is essential.

The key element for CPS is the *interaction* between physical (real world) and cyber (virtual world) elements, which are cooperating systems, having autonomous behaviours, context awareness through enhanced sensing capabilities, and storage and data processing capabilities from the sensors in the network (Figure 9.2-1).

Some of the practical examples that have already emerged include advanced robotics (e.g. autonomous cars, automatic pilot avionics, micro-robotics, robot-assisted surgery, implanted medical devices), intelligent buildings and the smart electric grid.

Cyber-Physical Systems (CPS) create a link between *physical and digital systems* in order to generate a common infrastructure with advanced capability. They allow integration of the dynamics of the physical processes with those of the software and networking, so that they can be handled as a single entity.

A CPS performs two main tasks:

- *Intense Connection:* In order to guarantee continuous data acquisition from the physical world and information feedback from the cyber space.
- Data Management, Analysis and Computation: In order to create the cyber space.

To fulfil these tasks a CPS is composed of "*collaborating computational entities that connect the cyber world with the surrounding physical environments or processes in an internet environment*" [1]. A CPS includes "*embedded systems (such as equipment, buildings, means of transportation, and medical devices), internet services, logistic, coordination, and management processes*" [1]. A broad range of sensors and actuators are used to connect the elements of the CPS and allow human-machine communication. The CPS stores and processes all the data received from the sensors and communication systems, and controls the physical systems using the actuators.

Research and development in this area is tackling a wide range of issues including inference from empirical data (i.e. elaborate empirical data in order to draw conclusions), sensing and perception, motor learning and control to adapt to different contexts, and the design, implementation, and verification of safe and well-functioning CPS.

A simplified scheme of a CPS and the interaction between humans and machines is shown in Figure 9.2-2.

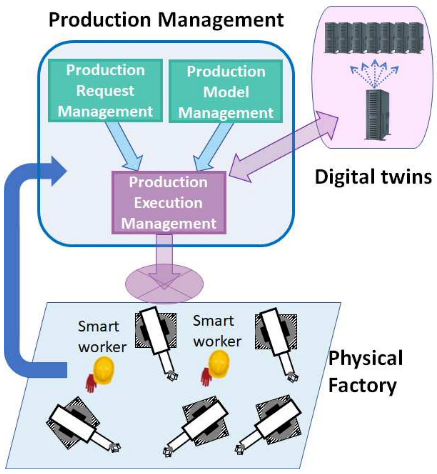


Figure 9.2-2 CPS scheme with interaction connections.

Cyber Physical Production Systems (CPPS) are Cyber Physical Systems for production. They enable and support communication between humans, machines and products. CPPS consist of autonomous and cooperative elements – related to processes, machines, production and logistics – connected across all levels of production, and an information system so that their operation and cooperation activities can be modelled. Thanks to the continuous exchange of data between and within the systems, stored and analysed in real time, the model is always synchronized with the status of the factory and the behaviour of the entire system can be forecast based on past and present situations. This can help in various decision-making processes so that the most appropriate actions can be quickly implemented, improving the productivity of the smart factory [2].

The main characteristics of CPPS are:

- *Intelligence:* The system components are able to acquire information from the environment and act autonomously.
- *Connectedness:* The system elements are able to cooperate and collaborate with each other and are connected to the knowledge and services on the Internet.
- *Responsiveness:* The system is able to react to internal and external changes.

The CPPS concept is strongly linked to other innovative concepts such as Internet of Things (IoT), big data and digital twins. The *IoT* refers to a wireless communication capability integrated with sensors and computing that allows the collection of data related to uniquely identifiable objects through the Internet. *Big data* refers to a new computing paradigm that allows the collection, processing and analysis of massive amounts of data [3].

The Digital Twin (DT) is another technology strictly related with Industry 4.0 and smart manufacturing. As with CPS, the digital twin is associated with the integration between the cyber and physical worlds. By definition a DT creates a virtual model of a physical system in order to predict real behaviour based on the simulation in real time of the system. Therefore, the two concepts have similar goals and approaches. However, while the CPS is implemented with sensors and actuators, which enable the interaction between physical and cyber world, DT is based on models and data. Digital twins are digital replicas of physical systems; they represent evolving digital profiles which are designed using the data collected from the past and present behaviour of a physical object or process. They are mainly applied in tasks such as monitoring, predictive maintenance and optimization of their physical counterparts.

9.2.1 Collaborative CPS

Collaborative CPS (CCPS) are even more complex production systems with respect to other CPS: the concept of “integrality” represents the level of integration of the different modules, aiming at self-learning and reconfiguration properties. In order to improve this integration, that concerns also the Human-Machine Interface (HMI), the CCPS must present a high level of “sociability”, meaning the communication capabilities with the environment, including other CPS [4].

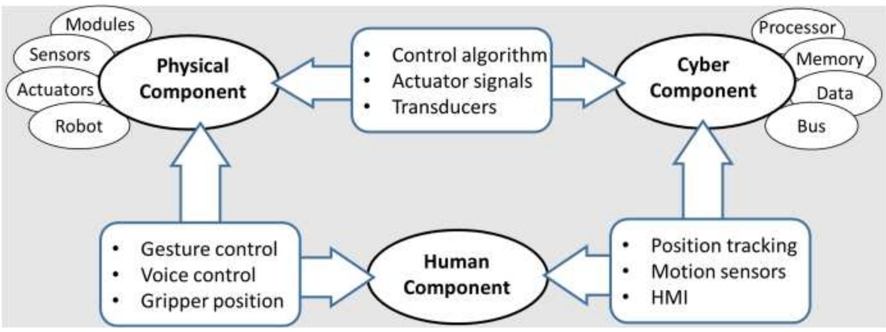


Figure 9.2-3 Structure of a Collaborative Robotics Cyber-Physical System.

CCPS include three main entities: the human component (HC), the physical component (PC) and the computational component (CC), which are interconnected through a variety of technologies (Figure 9.2-3). Being a part of a cyber-physical system, a human and a robot can interact and perform self-organizing tasks. In a future scenario, the integrated sensor network and communication technologies of CPS make the interaction reliable, safe and secure. Several sensors allow the robot to be aware of the human presence and react consequentially, based on the implemented human avoidance scheme. For the interaction to be safe a real-time and accurate human position tracking system is essential. A vision system connected with the robot and monitoring its environment can provide information on the worker location, used as input for speed reduction schemes for the robot. At the same time the monitoring system integrated with a gesture recognition system can be used by the worker to control the robot, together with a voice control. Moreover, force sensors can be integrated in the system to enhance the human robot interaction level. Combining the data from the force sensors and the vision system the robot speed and acceleration can be modulated according to the part of the worker's body in proximity of the end effector. The force sensors can provide additional feature permitting a contact between the human and the robot in order for the worker to train the robot using the hand.

9.2.2 Swarm Robotics and CPS

An early application of the concept of CPS can be found in swarm robotics, a topic largely studied in the last 20 years. Swarm robotics is inspired by nature and particularly the behaviour of swarms able to perform tasks that would be beyond the capabilities of the individuals. It concerns the coordination of distributed robotic teams, interconnected, which can be, thus, considered as simplified CPS.

Several works have been carried out [5], studying and improving the communication among robots and the algorithm managing the group strategy. For practical reason (the available space), analogy with nature (swarm of insects), but also technical reasons, this approach is extensively studied on large numbers of mini and micro robots; indeed, due to space and energy consumption limitation, miniaturized robots have typically a reduced degree of complexity. Therefore, the group strategy is very suitable for such small and simple robots.

A very successful example of swarm robotics has been developed around the e-puck, a mobile minirobot developed mainly for educational purpose [6], and a variety of works to study the group coordination and synchronization [7][8][9] .

The e-puck [10] has a circular section of 75 mm of diameter and variable height, depending on the extensions provided (Figure 9.2-4).

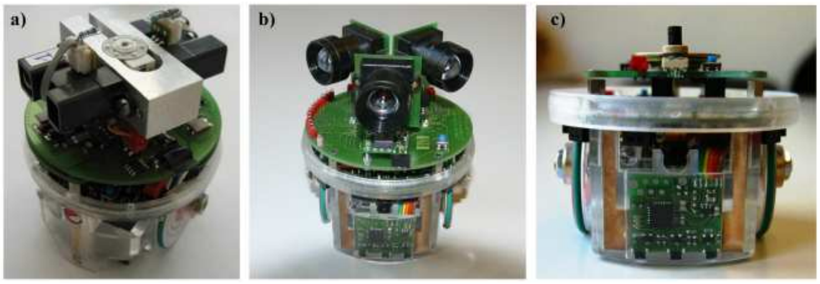


Figure 9.2-4 E-puck robot configuration provided with: a) a IR distance scanner, b) large field of view linear camera and c) ground colour measurement sensor [6].

Among these works some have been focused on the strategy to make a suitable number of e-puck able to move a box and place it in a chosen position. Although the design of the robot allows its customization providing it with the most suitable sensors and features required for the task (Figure 9.2-5), one of the aims of the works is to test several algorithms, not only to achieve the task, but also to do it with the easiest configuration of the robots as possible.

The objects to transport were chosen to be heavy enough to require the efforts of all the group, so that cooperation was essential. The robots are equipped with several IR sensors to detect the obstacles and find the target position, a camera to identify the obstacles (the objects and the other robots) and an optical camera with low image resolution underneath the robot. This camera gives the individual robots the perception of the direction of movement of the object, according to this the robots change the point of application of their pushing, avoiding to neutralize another robot's force. Eventually they are all aligned to successfully transport the object [7].

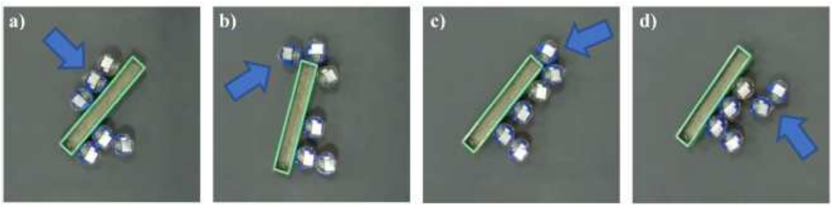


Figure 9.2-5 Alignment of the e-puck robots to successfully transport the object [11].

A more practical application of the typical algorithms of swarm robotics, and specifically path planning and coordination of distributed robotic teams in a CPS, is represented by automated warehouses. As an example, in the warehouses of Amazon and Ocado, the entire system, from the order to the delivery, is automated. A large number of autonomous mobile robots (Figure 9.2-6) are connected among each other and with the system where the orders are placed and an algorithm optimizes the path of the robots according to the positions of the ordered goods so that the displacement of the robots is scheduled to move the goods and pack the order. The connection and coordination among the robots is essential to avoid possible accidents.

In the Amazon warehouse [12] the robots move the shelves containing the required goods toward a worker, who picks the correct goods without the tiring walk amongst the shelves.

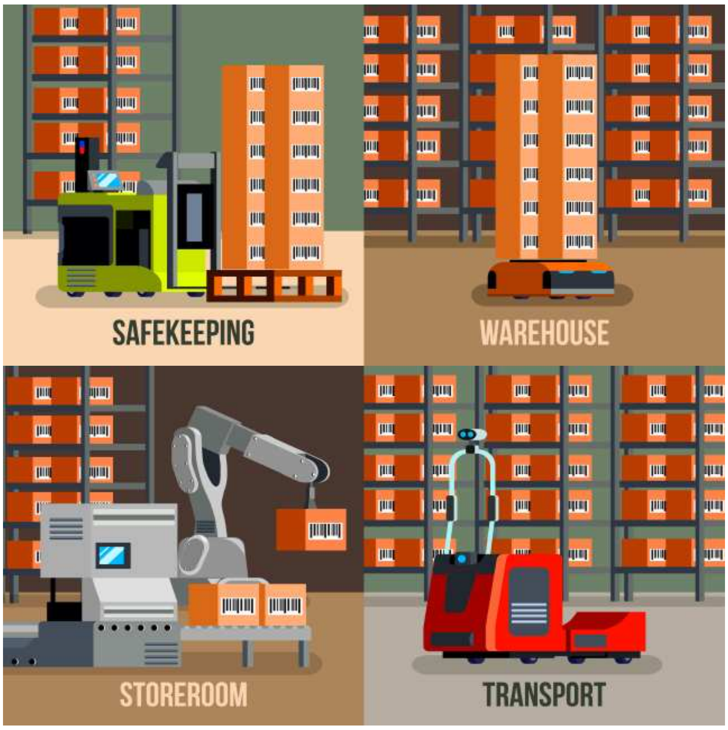


Figure 9.2-6 Automated warehouse*.

In Ocado’s warehouse [13] the robots displace crates and pick goods from them moving along a grid system.

* Image designed by Freepik

9.3 Robot Programming

The execution of a complete task in a robotic application requires the execution of one or more operations by the robot. The robot has to move according to a specific motion that is completely or partially defined by the motion law imposed on its end-effector. The desired motion results from suitable *motion planning and programming of the robot task* and is achieved by a set of commands sent to the robot's actuators in accordance with a predefined control strategy. For the correct execution of the task, it is important to know the kinematic structure of the robot, including the type and the location of the actuators (the motors), the joints, the transmissions, and the dimensions of the mechanical structure. Such knowledge is completed by that of the electronics controlling the actuators and of the electrical and mechanical limits of the components.

9.3.1 Motion Planning

The position and movement of the end-effector or the force it exerts depends on the position, movement and actions that the different parts of the robot exert by means of its actuators. Therefore, the analysis of the robot and its modelling are preliminary steps in the development of a robotic application.

The problem with planning a trajectory for a robot can be broken down into finding a *path* and defining a *timing law* on the path. Indeed, the path identifies the locus of the points that the robot has to follow to perform the desired motion (Figure 9.3-1), whereas the "*trajectory is a path on which a timing law is specified*" [14], e.g. in terms of velocities at each point.

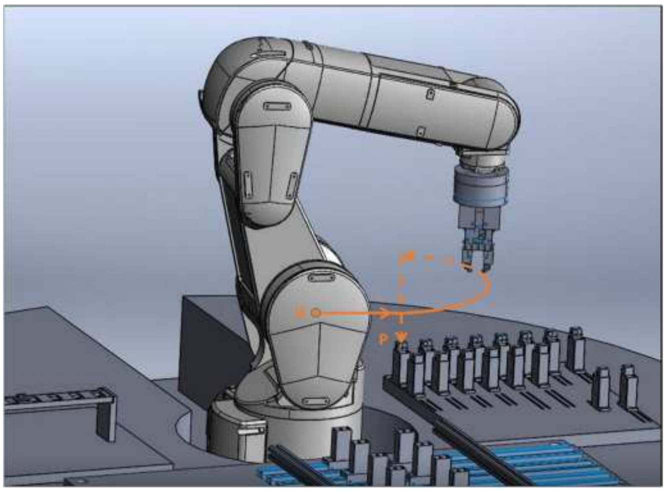


Figure 9.3-1 Example of a robot in simulated motion, moving from an initial point (H – home position) towards point P to pick a component on the table. Image rights: Authors.

Thus, *motion planning* concerns the generation of the timing laws for the coordinates (at the joints or at the end-effector), therefore the inputs to the motion control system [14].

Depending on the operations the robot is required to perform, the robot can adopt one of the following motion strategies [15]: from an initial point to a final point (*point-to-point motion*) with an arbitrary trajectory; through a finite sequence of points assigned along the path (*motion through intermediate points*); or according to a *specific trajectory*.

When defining the timing law, several aspects should be considered, including [14][15]:

- The *constraints* deriving from the type of application.
- The mechanical structure of the robot.
- The *joint forces and torques* should respect the limits of the actuators, the drives and the controller.
- The *trajectories* that interpolate the path points should be *smooth* to minimize undesired effects such as vibrations, the trajectories should not imply a very high *computational load.*
- The *positions and velocities at the joints* should be continuous functions of time (continuity of accelerations is not mandatory although generally welcome).

9.3.2 Programming

Programming a robot means instructing it on the task(s) it has to perform. A programming environment and suitable programming language(s) have to be defined.

Existing *programming methods* include [16][17][18] different approaches:

- *Guiding the robot to the positions of interest* or along the desired paths (teach programming) by using a teach pendant, manually holding the end-effector, or in tele-operation (using a master-slave system, so that the operator moves the master robot and the actual “slave” robot moves accordingly).
- *Writing a program* in a language that the robot can interpret and execute (proprietary language or robotic libraries supporting standard programming languages).
- *Using interactive graphical interfaces*, that is CAD (Computer Aided Design) modelling and system simulation tools to define the task and generate the program automatically.
- *Task-level programming* consists of the user telling the robot what should be done and the robot knows how to do it. This is achieved thorough modelling and sensing of the environment and machine intelligence. Therefore, task-level programming is difficult to achieve at the present time, but future innovative applications may possibly be developed.

Generating robot programs from CAD/CAM software packages is frequently used, since they are very common among manufacturing companies [16]. In the context of the digital factory, digital solutions for the 3D representation and simulation of an individual robot or robotic lines effectively support the design and validation of the manufacturing process. Other approaches to instruct a robot include, for example, using pointing devices or choreographing the task movements [16][18].

Robot programming methods can be separated into *online programming* and *offline programming.*

With *online programming*, the physical robot is used and the operator directly acts on the robot controller. Programming by teaching belongs to this category. The robot is taught or guided as needed and the actions are recorded in the robot's controller memory. The robot then executes the movements in a repetitive way.

Online programming is rather easy and does not require specialist operator skills. It is not sensitive to accuracy errors. However, it does require access to the robot meaning production is shut down during the programming phase. Moreover, the editing of the program (e.g. to add some parameters, correct a line, or add interactions with other devices) can be limited meaning complex tasks can be difficult to set up. Finally, the robot programs are only stored in the robot memory so can be hard to access. This method can be useful if the robot is required to repeatedly execute a task for a long production period. Programming by language online is also possible: in this case the operator writes the commands directly on the teach pendant, but it can be difficult to realise at that moment all the effects, such as possible collisions.

In *offline programming*, software tools (programming by language or interactive graphical interfaces) are used to generate the program (without the physical involvement of the robot) which is then deployed on the robot controller. With offline programming, editing possibilities increase, the real robot can work during the programming phase so production is not shut down, and start-up and product changeover are faster. However, the robot should be accurate (need for a robot calibration) and specialist operator training is required.

A combination of online and offline programming is also possible.

9.3.3 Programming Languages and Environments

There are many *programming languages* used in robotics that have been adopted over the years. Some examples of the first programming languages used were BASIC, Pascal and LISP. While these languages are now outdated, they provided the basis for more recent industrial robotic languages. Other common scripting languages include C, C++, C#, .NET, Python, Matlab and Java, while an example of a visual programming language is Labview.

Moreover, *many proprietary languages* developed by different manufacturers exist and even for similar robots they can be very different. Currently, almost every robot manufacturer has its own proprietary programming language. Some examples are: RAPID by ABB, MELFA-BASIC by Mitsubishi, KRL by Kuka, AS by Kawasaki, VAL3 by Stäubli, PDL2 by Comau, Karel by Fanuc, Inform by Yaskawa and URScript by Universal Robots.

Manufacturers also provide *programming environments for simulation and offline programming*. The robot and the whole work-cell can be represented as a 3D CAD model. The robot movements, arm collision and tool actions can be simulated and the task defined and optimized. Some examples of manufacturers' programming environments are RobotStudio [19] by ABB and RT Toolbox3 Pro [20] by Mitsubishi.

It would clearly be desirable "*to develop either robotic libraries to be used in the context of consolidated standards or new general-purpose languages for industrial automation applications*" [17]. Indeed, one of the current research and development activities aimed at supporting the delivery of Industry 4.0 is the development or use of standards related to programming tools, modelling and simulation, communication protocols and interfaces.

In recent years, third-party offline programming software has been developed that supports all major robotic brands to provide a common framework useful for simulation and programming of different robots in many applications with a single tool. Some examples include RoboDK [21], Robotmaster [22], Delfoi Robotics [23], FASTSUITE [24] and OCTOPUZ [25].

In addition to commercial products, some open-source solutions exist. Among these, *ROS* (Robot Operating System) is becoming very common. ROS provides a common framework for robotics applications. It is a meta-operating system that provides the services expected from an operating system, "*including hardware abstraction, low-level device control, implementation of commonly-used functionality, message-passing between processes, and package management. It also provides packages of tools and libraries for obtaining, building, writing, and running code across multiple computers [26]*."

"ROS-Industrial *is an open-source ROS module that extends the advanced capabilities of ROS to manufacturing automation and robotics. The ROS-Industrial repository includes interfaces for common industrial robots, grippers, sensors, and device networks [27]*." It provides software libraries for e.g. automatic 2D/3D sensor calibration and motion planning.

9.4 Robot Control

At the robot level, a control and supervision architecture needs to be implemented, able to interact with the external environment within the CPPS and transform the external stimuli into actuators commands. Moreover, a device is needed to regulate the actions generated by the actuators so that the robot behaves as desired: this device is called a *controller*. It receives the desired robot behaviour as input and, in order to determine the actions of the actuators, it often requires information about the robot status (position, velocity, exchanged force, etc.) that is obtained by reading the data provided by suitable sensors. When this monitoring function is present, this type of control is called *closed-loop control*, otherwise it is called *open-loop control*.

The control system is needed to determine the time history of the forces or torques to be developed by the joint actuators to guarantee execution of the commanded task [28]. In simple terms, for example, if y is the actual position of the robot, y_0 the desired robot position, and u the force or torque generated by the robot motor: the controller has to determine u in order to make the robot move so that y matches y_0.

To better comprehend the subject, one can firstly identify the main elements of a control scheme: the "system" to be controlled and the "controller" that provides a control action to the system. The system to be controlled can be seen as a box that responding to an input $u(t)$ produces an output $y(t)$. The objective is to make the system produce an output as the desired controller input $y_0(t)$ (e.g. to make an axis of a robot move in a desired way).

There is an important distinction in control theory between open-loop control, and closed-loop control i.e. control with a feedback loop. In the open-loop control (Figure 9.4-1), the controller calculates the proper control action $u(t)$ to produce an output $y(t)$ equal to $y_0(t)$ on the basis of the known relation between the input and the output of the system.

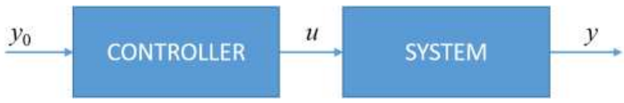


Figure 9.4-1 General scheme of open-loop control.

However, a proper knowledge of the system is often missing, and unpredictable or partially known disturbances (e.g. coupling effects among the joints) can act on the system. Therefore, a common approach consists of closing the control loop with the feedback information on the actual value of the output coming from sensors (Figure 9.4-2). In this way, the controller computes the input to the system in order to reduce as much as possible the error between the desired and actual outputs (e.g. desired position and actual position of the motor).

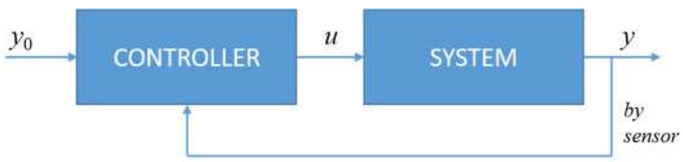


Figure 9.4-2 General scheme of closed-loop control.

In other words, the desired output is compared to the feedback variable provided by a suitable sensor (e.g. encoder), then the controller defines the action on the basis of the error between them. The loop closure allows a better action, such as robustness and reduction of disturbances.

Moving the robot requires the control of different motors (a robot usually has more than one degree of freedom), therefore multi-axis control methods are needed. Depending on the way the different axes are considered and controlled, control strategies can be divided into [28][29]:

- *Decentralized Control*: Each joint is considered independent from other joints and controlled by an independent control loop (Figure 9.4-3).
- *Centralized Control*: In this case the robot is considered as a multivariable system. The dynamic model of the robot is often taken into account (Figure 9.4-4).

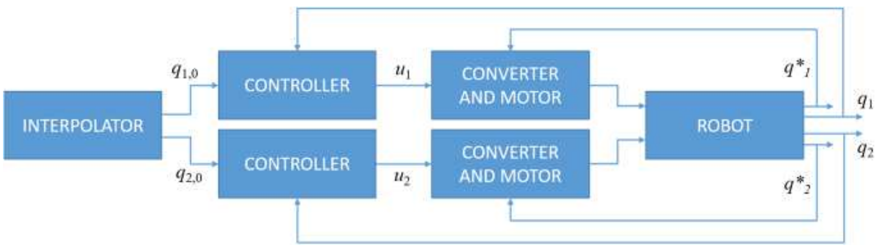


Figure 9.4-3 Decentralized control scheme: example with 2 axes, where q_i (i = 1, 2) represents the joint position and q^*_i the joint velocity. The interpolator block calculates the desired movements required to each axis.

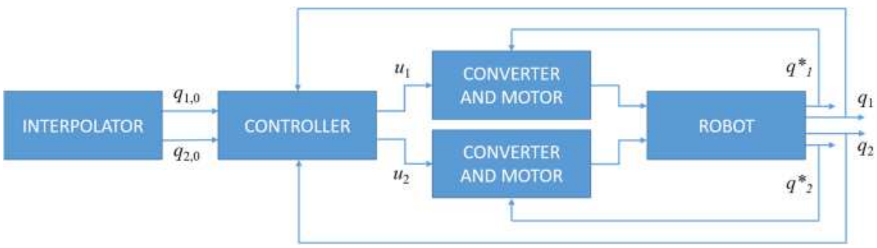


Figure 9.4-4 Centralized control scheme: example with 2 axes, where q_i (i = 1, 2) represents the joint position and q^*_i the joint velocity. The interpolator block calculates the desired movements required to each axis.

In the former case, the actions of each actuator only depend on the actual and desired movements of *that* actuator, while in the case of centralized control the actions of each actuator depend on the desired and actual movements of *all* the actuators.

This also distinguishes the way the dynamic interaction and coupling effects between the joints are considered. In the former case, the dynamic interactions are treated as disturbances, whereas in the latter the coupling effects can be actually taken into account.

Multi-axis control methods often rest on single-axis controllers controlling a motor each, which allows for good performance in many applications. Centralized control strategies are generally implemented when mutual effects between the joints are significant and need to be considered.

Among centralized controllers, different approaches can be considered, such as the *computed torque control* and the *inverse dynamics control*. These controllers are based on the dynamic model of the robot. If this model is not well-known, the performance of the control can scarcely improve. For this reason, a good estimation of the values of the dynamic parameters of the system (e.g. masses, inertias, etc.) is necessary. Such estimation can be achieved by robot dynamic calibration.

Moreover, it is important to note that real robots differ from ideal robots due to the presence of different phenomena including variable inertia, static friction and compliance that have to be taken into account when designing the control scheme.

9.4.1 Force and Vision-based Control

There are cases where a robot is required to exchange forces with the environment, e.g. for pressing a workpiece, for inserting a pin in a hole, or for deburring a workpiece. In these cases, controlling the contact force is more convenient than controlling the motion. Another branch of robotic control is then constituted by *force control* [30][31].

Force control can be divided into:

- *Direct Control*: The force is controlled by closing the force feedback loop. The general scheme of a force controller is illustrated in Figure 9.4-5.
- *Indirect Control*: The force control is achieved by means of motion control (control of a deformation). Impedance control belongs to this category.

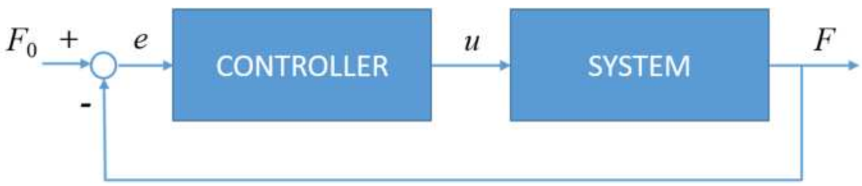
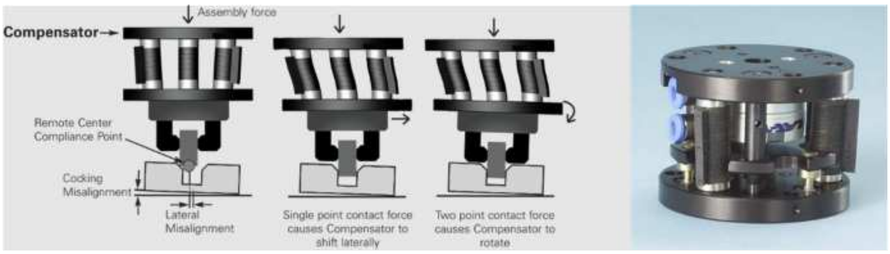


Figure 9.4-5 Direct force control general scheme.

In direct control, the force feedback is provided by a force transducer at the end-effector, at the joints, or at the fingertips of the robot hand when a suitable force is required for grasping and manipulating an object.

It is important to note that it is not possible to control both the motion and the force in the same direction. When the interaction task identifies directions along which the motion has to be controlled and other directions where the force is the controlled variable, *hybrid force/motion control schemes* are used [31][32].

Due to their complexity, sometimes simpler position controllers are combined with a passive compliance on the end-effector, so that small position errors can be compensated by elastic elements. An example is the Remote Center of Compliance (RCC), that is a mechanical device which is commonly mounted between the gripper and the robot's wrist (Figure 9.4-6).



a) b)

Figure 9.4-6 A RCC compensator by ATI Industrial Automation: a) Conceptual scheme†; b) Picture of the device‡. Image rights: ATI Industrial Automation.

The force feedback is also useful for the *teleoperation* of the robot. In teleoperated mode, the robot is remotely commanded by the operator using a master/slave approach. The idea is that the operator moves a device (the master, e.g. a joystick) and the robot (the slave) acts as in the same way. In this case, *haptic interfaces* can be used to provide the operator with the force/torque feedback [33][34].

Nowadays, vision capabilities are often integrated into a robotic work-cell. One or more vision sensors (cameras) can be mounted on the robot (*eye-in-hand configuration*) or in the work-cell (*eye-to-hand configuration*) to make measurements and supervise the scene (Figure 9.4-7). Therefore, they can provide information about the objects to be processed and the surrounding devices and environment.

† Source: https://www.ati-ia.com/Products/Compliance/Compensator_product_desc.aspx

‡ Source: https://www.ati-ia.com/products/compliance/Compensator_ModelDetails.aspx?id=9116-001-A

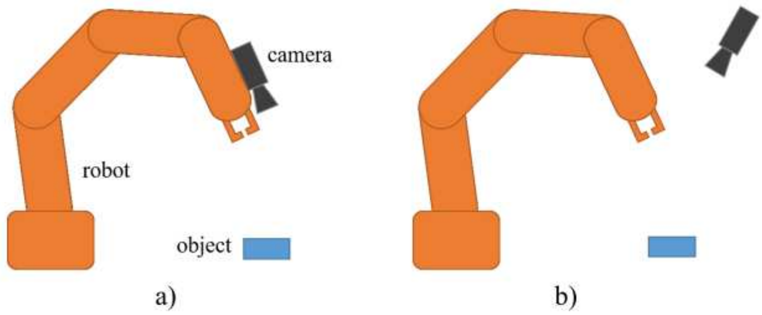


Figure 9.4-7 Configurations of the vision system (simple case of a single camera): a) eye-in-hand; b) eye-to-hand. Image rights: Authors.

In the simplest approach, the robot is commanded in motion control but exploiting the information provided by the camera, implementing the so-called *look-and-move strategy*. As an example, consider a robot that has to pick an object in the working area whose view is provided by a camera. The camera identifies the object and detects its pose which, by means of a proper calibration, can be expressed in the robot coordinate system. This information is then sent to the robot controller that commands the robot to move and pick the object.

When the visual information is used to close the control loop, such a control based on visual feedback is referred to *visual servoing* [35][36]. In other words, "*visual servo control refers to the use of computer vision data to control the motion of a robot*" [35]

The two main approaches to visual servoing are *position-based visual servoing* and *image-based visual servoing* (or hybrid approaches). In both cases the challenge is to move the robot tool into a specific pose relative to the target object. With position-based visual servoing, visual measurements are used to reconstruct the relative pose of the object with respect to the robot, while image-based visual servoing is based on the comparison of the feature parameters of the image of the object between the current and the desired pose [36].

To better understand the two approaches, consider the case of the eye-in-hand configuration, where the robot tool is at a fixed relative pose to the camera and the problem concerns moving the camera.

In position-based visual servoing, a geometric model of the object and camera parameters are used to estimate the object pose relative to the camera. It is then possible to compute the pose error and command the robot to move to correct the pose.

In image-based visual servoing, a model of the object is not needed but it is necessary to detect specific features (e.g. coordinates of points) in the image of the object acquired by the camera. The position of the points in the image depends on the camera pose relative to the object, so that if the camera is moved, the points move in the image. When the camera is in the desired pose, the points appear in known positions in the image, otherwise they appear in different positions. The image-based visual servoing compares the desired and current point features and computes the desired velocity of the camera to move the points where required in the image, therefore the camera to the desired pose.

Image processing algorithms are then essential to extract the image features necessary for controlling the robot.

9.5 Artificial Intelligence

Robotics is one of the fastest growing high technology markets, mainly due to the pervasive introduction of robots in production, and to the growing service and home-care sectors. The actual robotic devices have reached a high level of maturity and reliability in terms of mechanical structures, sensors and actuators, and control techniques to solve basic perception, navigation, and manipulation tasks when working in structured and well-defined environments.

However, to fulfil the I4.0 paradigms robots will need to be more and more versatile, addressing a range of tasks in unstructured and open environments. Therefore, they will need to be equipped with increased cognitive abilities such as knowledge representation, planning, learning, adaptation, and natural human–robot interaction.

Artificial intelligence is a branch of computer science related to the programming and design of hardware and software systems, allowing the machines to have abilities typical of human and animal beings, such as visual, space-time and decisional perceptions.

The concept of intelligence is therefore defined not only as the ability to compute abstract data, but it includes also all those different forms recognized by the Gardner's theory of multiple intelligences [37][36], ranging from spatial to social intelligence, from kinesthetic to introspective.

An intelligent system is created by trying to recreate one or more of these different forms of intelligence. Although often defined as exclusively human or natural, they can be simplified to particular behaviors reproducible by some machines.

Artificial intelligence is already largely exploited in everyday life. For example, the various voice recognition tools regularly used in many applications from smartphones to security systems, are based on machine learning algorithms, which are a form of AI.

Another popular application of machine learning and artificial intelligence paradigm is in the automotive sector, where advanced speed change systems in semi-autonomous driving cars make use of AI algorithms based on fuzzy logic [38]. Moreover, AI is playing a key role in the development of autonomous driving cars, although many technical and ethical challenges still need to be addressed in order to have a reliable and safe self-driving vehicle.

In the medical field, artificial intelligence mainly uses neural networks, with applications in the analysis of the heartbeat, in the diagnosis of some cancer forms, and in the creation of new therapeutic drugs. Recently, AI techniques have been exploited to develop tools for predicting the coronavirus pandemic, and the early detection and diagnosis of infections [39].

Further sectors in which artificial intelligence is used on a regular basis are the stock market, image recognition (including facial recognition), social media marketing, and robotics. In addition, intelligent systems are also used to further improve many sectors of information technology itself, including proving the correctness of algorithms and creating new learning methods.

9.5.1 Historical Overview

In 1954, George Devol built the first programmable robot, called Unimate. It was a hydraulic manipulator arm that could perform repetitive tasks. It was bought by General Motors in 1961 for use in automobile assembly lines to automate metalworking and welding processes. Although in light of today's AI performance its degree of intelligence is very rudimentary, at the time of its invention it was considered by many as one of the few AI products with commercial value.

Since these early developments, innovative works have been done, specifically improving the robotic capabilities for sensing and perception (mostly through machine vision, but also adding force feedback, tactile abilities, proximity sensors and so on) thus increasing the capabilities of the machines to interact with the surrounding environment. Already in the late 1970s humanoid robots were developed that could stack blocks. The further improvements in ICT hardware and AI technologies boosted the development of more and more advanced robotic systems.

From the coining of the term AI in 1956 until the early1970s, the enthusiasm about AI gave birth to several research subfields, which constitute the foundation of the modern theory of AI: rules-based systems, machine learning, Natural Language Processing (NLP), shallow and deep neural networks, image processing and computer vision, Speaker Recognition and Speech to Text Processing. However, practical applications of AI based programs were still very uncommon and largely limited to solving rudimentary problems, mainly due to limitations in the available computing power.

Moreover, most of the early predictions about the future of strong artificially intelligent machines were not yet realized, making investors skeptical and reluctant to invest more funding. This resulted in a bust phase of AI. Most of the limited number of inventions from this period, such as backpropagation and recurrent neural networks, went largely overlooked and substantial effort was spent to rediscover them in the subsequent decades.

The recent advancements in information and communication technologies gave new life to research into AI, as hardware and network connectivity became cheaper and faster; parallel and distributed systems became practical, and huge amounts of data (Big Data) were easily stored and then available for training AI systems.

9.5.2 Techniques

The use of algorithms capable of reproducing the reasoning of human beings or animals in different situations allowed intelligent systems to further improve their behavioral skillset. Research efforts focused on the development of algorithms which could imitate different behaviors depending on environmental stimuli (the skill often called context-awareness).

These complex algorithms implemented within intelligent systems are therefore able to make decisions or choices according to the context where they are applied. In the case of intelligent vehicle systems, for example, a car without a driver can decide in the case of danger whether to steer or brake according to the situation. For example, depending on whether the information sent by the various sensors such as car speed and direction, size and position of other objects, and ground and weather conditions, it calculates if braking or steering has a higher percentage of safety for the driver, passengers, and pedestrians. The basic knowledge originally embedded in the system is than expanded, created through experience.

A specific sector has grown, focused on the representation of knowledge, which studies all the mechanisms of human reasoning and defines methods to make this knowledge comprehensible to machines, through language and increasingly precise and detailed controls. Transferring existing human knowledge to machines requires transferring not only domain knowledge, but also experiences. This adds the possibility of understanding new information by exploiting the knowledge already present in the starting system.

This information is provided to the machine in various ways, the most important of which are those based on the *Theory of Formal Languages* (which studies the syntactical aspects of structured languages), and the *Theory of Decision*, to identify optimal decisions under specific circumstances.

One of the main steps forward in AI development was made when algorithms were created that were capable of learning through experience just like humans. Learning makes the system capable of improving the ability of the machine to act and make decisions.

Developing algorithms that can learn from their experiences and mistakes is essential to creating intelligent systems that operate in situations for which programmers cannot foresee all the future possibilities.

Through *machine learning,* also discussed in Chapter 5, a machine can learn to perform a certain action even if this action has never been programmed among the possible actions. For example, a robot can perform a peg in hole assembly operation even if the shape and position of the hole differ from the ones it was trained with, or can recognize an object to pick even if that exact image was not previously stored in its memory.

Machine learning is probably the most common branch of artificial intelligence considered in mainstream culture, and constituted the basis of several successful literary works. However, the research is still far from the scenarios represented in science fiction movies and literature, and develops along both theoretical and practical paths in computational learning theory and pattern recognition.

Machine learning was enabled by the development of *artificial neural networks*, i.e. a particular mathematical model inspired by brains and neurons. The name *neural network* derives from the fact that this mathematical model is characterized by a series of interconnections between artificial neurons, and these connections are necessary for the different calculations. Just like biological neural networks, an artificial neural network also has the characteristic of being adaptive to the varying needs deriving from the different information obtained in the different learning phases.

9.6 Ethical Issues

As the degree of autonomy and decision making capability of automatic systems increases, it is essential to define a code of conduct, a list of rules that have to be respected in order to avoid unpredicted behaviours. A typical case is that of autonomous guided vehicles, such as industrial AGVs used for logistic applications, drones and autonomous cars. It is very likely that, in a near future, these concepts will be extended also to robots acting in the industrial environment.

Robot ethics is defined as the set of rules and the code of conduct which is embedded in the robot control system by the designers.

The first set of rules (the well-known *three laws of robotics*) was defined by the science fiction writer Isaac Asimov [40] and the scientific community considers that they are still valid today:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

He later added also a *Law 0*: No robot may harm humanity or, through inaction, allow humanity to come to harm.

However, these laws are not sufficient and there might be scenarios where they cannot be fully obeyed.

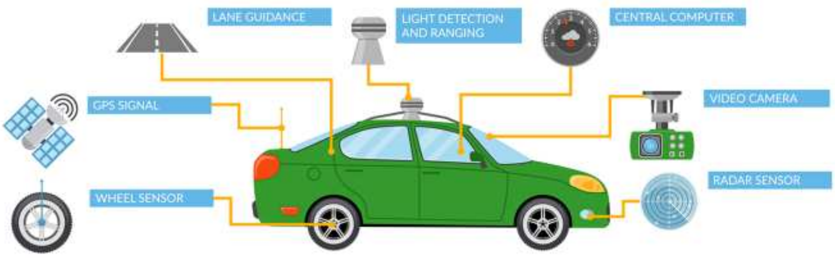


Figure 9.6-1 Autonomous car set-up§.

The three laws imply that a robot has sufficient intelligence (in terms of perception and cognitive abilities) to also make moral decisions in complex situations, which is still not the case. Therefore, they cannot provide a practical basis for writing a robot code of conduct. An example of a set of rules easily embeddable in a robotic system which was proposed by robo-ethical scientists is:

1. Do not kill
2. Do not cause pain
3. Do not disable
4. Do not deprive of freedom
5. Do not deceive
6. Keep your promise
7. Do not cheat
8. Obey the law
9. Do your duty

On the ethics side, several moral dilemmas need be addressed.

Figure 9.6-2 shows a typical moral dilemma, which could be hardly embedded in the machine's software code: knowing that there are only two possible events – crashing into the people crossing the street and killing three of them with two injured or crashing the car into the barrier and killing four passengers – which alternative should be chosen? Obviously, this would be a hard decision also for a human driver, with many different factors to be considered.

Several research projects are on-going in the field and many other examples of moral dilemmas can be found at http://moralmachine.mit.edu/, an online experimental platform designed to explore the moral dilemmas faced by autonomous vehicles and to collect human opinions on how machines should make decisions when faced with these moral dilemmas.

§ Image design by Freepik.

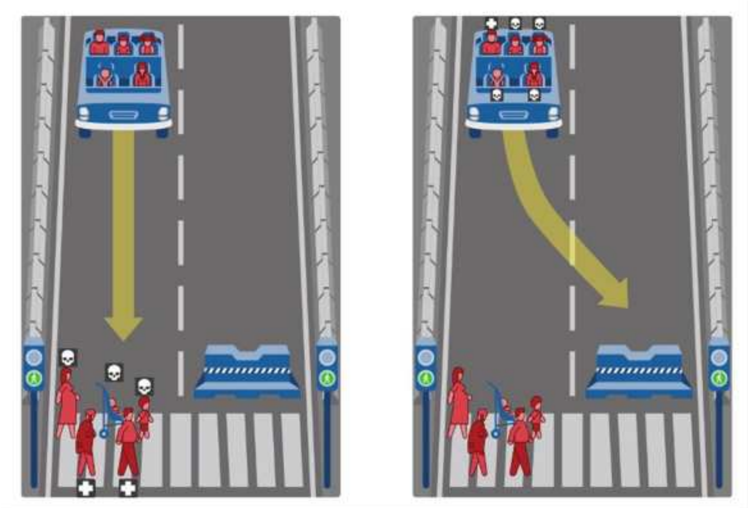


Figure 9.6-2 Driverless car moral dilemma [41].

These topics have also been used as a background principle in the context of the study commissioned by the European Parliament's Legal Affairs Committee "*to evaluate and analyse, from a legal and ethical perspective, a number of future European civil law rules in robotics*". In February 2017 the subsequent European Parliament Resolution "Civil Laws in Robotics" was adopted, covering a wide range of different areas, including liability rules, ethical questions, standardisation, safety and security, data protection and intellectual property rights, autonomous vehicles, drones, care robots, medical robots, human enhancement and education and employment and a code of ethical conduct for robotics engineers and a code for research ethics committees. Moreover, it proposes an EU Agency for Robotics and Artificial Intelligence.

The reaction from the robot manufacturing companies was quite fast, showing anxiety that this regulation could hamper the progress and threaten the competitiveness of the entire manufacturing industry. Several social studies have also been carried out to assess the impact of these rapidly evolving technologies on the job market, and a strong debate is still ongoing in newspapers, social media and international organizations.

Indeed, on May 22, 2019 the Organisation for Economic Cooperation and Development adopted the OECD Council Recommendation on Artificial Intelligence [42], which sets principles for responsible stewardship of trustworthy Artificial Intelligence systems and recommendations to national governments for national policies and international cooperation. These principles are based on: inclusive growth, sustainable development and well-being, human centred values and fairness, transparency and explainability, robustness, security and safety; accountability.

9.7 Conclusions

The innovative technologies of Industry 4.0 enable a new generation of factories where industrial robots perform autonomously all tasks while the workers have mainly a supervisory role. Thanks to the improved sensing capabilities and real time control the cooperation between human and robot become easier and safer, enhancing the productivity, production quality and flexibility. The cyber physical systems provided with artificial intelligence not only monitor the entire production process, but make predictions and decisions in order to improve the efficiency of the production.

A lot of progress has been achieved in the last decades, but several challenges have still to be faced both in the physical (hardware) and cyber (software) world, improving sensors, actuators, controllers, data storage and analysis, and artificial intelligence algorithms in order to achieve the new generation of factories envisioned by the Industry 4.0 paradigm. The interconnection requires a *standard communication system*, proprietary languages, thus, are proposed to be replaced by open source solutions; human-robot collaboration needs *advanced perception skills* to work safely in unstructured spaces and *psycho-social aspects* of this collaboration from the point of view of the workers have to be wisely considered; intelligent machines able to make decisions introduce serious *ethical issues*, so appropriate laws have to be defined and implemented in the robots; CPSs are exposed to security threats and the transmission and sharing of data necessitates new levels of *cyber security and privacy*; the modularity and configurability needed for fully customized products requires *new business models*; all the necessary transformations require corresponding changes in workforce skills and organizational structures, which can be achieved with a transformation in the *educational system,* offering not only new contents but also new methodologies for the new generation of employees. Industry 4.0 is an industrial revolution evolving at exponential pace with a huge impact on the quality of life for populations all over the world. A conscious awareness of such impact is essential to benefit the most from this breakthrough since it could also yield to greater inequality. The immediate benefits will be mainly for innovators and investors, the job market will be strongly affected increasing the demand of highly skilled workers and providing few offers to workers with less education and lower skills. Digital interconnections will also shorten the distances between people providing new opportunities for cross-cultural exchanges, improving understanding and cohesion. However, also unrealistic expectation and extreme ideologies will spread faster and with no control accessing more vulnerable people, who could also be victims of privacy violations and stolen data. Therefore, the opportunity offered by the fourth industrial revolution has to be globally shared and shaped toward a future with common objectives and values, to benefit powerless people rather than increasing the gap between the social classes.

9.8 References

[1] N. Jazdi, "Cyber physical systems in the context of Industry 4.0", in Proceedings of the IEEE International Conference on Automation, Quality and Testing, Robotics, Cluj-Napoca, Romania, 22–24 May, pp. 1- 4, 2014.

[2] E. Geisberger, & M. Broy, (Eds.). agendaCPS: Integrierte Forschungsagenda Cyber-Physical Systems Vol. 1. Springer-Verlag.; ISBN 978-3-642-29098-5, 2012.

[3] D. Zuehlke, "SmartFactory Towards a factory-of-things", *Annual Reviews in Control*, vol.34, pp.129-138, 2010.

[4] A. Khalid et al. "A methodology to develop collaborative robotic cyber-physical systems for production environments", *Logistic Research* 9:23, 2016.

[5] I. Navarro and F. Matía, "An introduction to swarm robotics". *International Scholarly Research Notices*, 2013.

[6] F. Mondada, M. Bonani, X. Raemy, J. Pugh, C. Cianci, A. Klaptocz, S. Magnenat, J.C. Zufferey, D. Floreano, and A. Martinoli, "The e-puck, a robot designed for education in engineering", in Proceedings of the 9th conference on autonomous robot systems and competitions Vol. 1, No. CONF, pp. 59-65, IPCB: Instituto Politécnico de Castelo Branco, 2009.

[7] M.H.M. Alkilabi, A. Narayan, C. Lu, and E. Tuci, Evolving group transport strategies for e-puck robots: moving objects towards a target area. In Distributed Autonomous Robotic Systems, pp. 503-516 Springer, Cham. 2018.

[8] M.H.M. Alkilabi, A. Narayan, and E. Tuci, "Cooperative object transport with a swarm of e-puck robots: robustness and scalability of evolved collective strategies", *Swarm intelligence*, 11(3-4), pp.185-209, 2017.

[9] F. Perez-Diaz, R. Zillmer, and R. Groß, "Control of synchronization regimes in networks of mobile interacting agents",. *Physical Review Applied*, 7(5), p.054002, 2017.

[10] http://www.e-puck.org/

[11] https://www.youtube.com/watch?v=t8wbicHdRFk

[12] https://www.youtube.com/watch?v=E7jRYvEFytE

[13] https://www.youtube.com/watch?v=4DKrcpa8Z_E

[14] B. Siciliano, L. Sciavicco, L. Villani, G. Oriolo, Robotics - Modelling, Planning and Control, Chapter 4, Springer-Verlag London Limited DOI 10.1007/978-1-84628-642-1, 2010.
[15] G. Legnani, I. Fassi, Robotica Industriale, Chapter 9, CittàStudi, 2019.
[16] M. Hägele, K. Nilsson, J.N. Pires, R. Bishoff, Industrial Robotics, in Springer Handbook of Robotics, Chapters 54, Siciliano, B. Khatib, O. (Eds.), 2nd edition, DOI 10.1007/978-3-319-32552-1, 2016.
[17] B. Siciliano, L. Sciavicco, L. Villani, G. Oriolo, Robotics - Modelling, Planning and Control, Chapter 6, Springer-Verlag London Limited DOI 10.1007/978-1-84628-642-1, 2010.
[18] G. Legnani, I. Fassi, Robotica Industriale, Chapter 11, CittàStudi, 2019.
[19] https://new.abb.com/products/robotics/robotstudio
[20] https://us.mitsubishielectric.com/fa/en/products/industrial-robots-melfa/engineering-software/rt-toolbox3-pro/rt-toolbox3#
[21] https://robodk.com/
[22] https://www.robotmaster.com/en/
[23] https://www.delfoi.com/web/products/delfoirobotics/en_GB/off-line_ENG/
[24] http://www.fastsuite.com/en_EN/fastsuite/why-fastsuite.html
[25] http://octopuz.com/
[26] http://wiki.ros.org/ROS/Introduction
[27] https://rosindustrial.org/about/description/
[28] B. Siciliano, et al. *Robotics - Modelling, Planning and Control*, Chapter 8, Springer-Verlag London Limited 2010. DOI 10.1007/978-1-84628-642-1
[29] G. Legnani, I. Fassi, Robotica Industriale, Chapter 10, CittàStudi, 2019.
[30] L. Villani, et al. *Force Control*, Chapter 9, in *Springer Handbook of Robotics*, Siciliano, B. Khatib, O. (Eds.), 2nd edition, DOI 10.1007/978-3-319-32552-1, 2016.
[31] B. Siciliano, et al., Robotics - Modelling, Planning and Control, Chapter 9, Springer-Verlag London Limited DOI 10.1007/978-1-84628-642-1, 2010.
[32] Kevin M. Lynch and Frank C. Park, Modern Robotics: Mechanics, Planning, and Control, Chapter 11, Cambridge University Press, ISBN 9781107156302, 2017.
[33] B. Hannaford et al., Haptics, Chapter 42, in Springer Handbook of Robotics, Siciliano, B. Khatib, O. (Eds.), 2nd edition, DOI 10.1007/978-3-319-32552-1, 2016.
[34] G. Niemeyer et al. Telerobotics, Chapter 43, in Springer Handbook of Robotics, Siciliano, B. Khatib, O. (Eds.), 2nd edition, DOI 10.1007/978-3-319-32552-1, 2016.
[35] F. Chaumette et al. Visual Servoing, Chapter 34, in Springer Handbook of Robotics, Siciliano, B. Khatib, O. (Eds.), 2nd edition, DOI 10.1007/978-3-319-32552-1, 2016.
[36] B. Siciliano, et al. Robotics - Modelling, Planning and Control, Chapter 10, Springer-Verlag London Limited DOI 10.1007/978-1-84628-642-1, 2010.
[37] H. Gardner, Frames of mind: The Theory of Multiple Intelligences, 1983.

[38] Fuzzy Logic (FL) is a method of reasoning that resembles human decision making process, working on the levels of possibilities of input to achieve the definite output, helping to deal with uncertainties in engineering. It can be implemented at both hardware and software level.
[39] C. Menni, et al., "Real-time tracking of self-reported symptoms to predict potential COVID-19", Nature *medicine*, *26*(7), 1037-1040. doi.org/10.1038/s41591-020-0916-2, 2020.
[40] I. Asimov, Runaround, 1942.
[41] E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J.-F. Bonnefon, I. Rahwan, "The Moral Machine experiment", *Nature*, 563, pp.59–64, 2018.
[42] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449

Acknowledgements

Many people have contributed to the development of the Digit-T online training course and e-book, and the editors and authors would like to thank everyone involved. We hope we have listed everyone within this acknowledgements sections and their main areas of contribution; we apologise for anyone we have unknowingly missed.

The learning material authors are listed at the start of the sessions in the training course and the chapters of the e-book. However, many additional people have contributed to the discussions concerning the content to be included who have not been listed there. These include *David Bainbridge, Panorios Bernardos, David Branson III, Sergio Carol Llopart, Elkin Castro, Mireia Dilmé i Martínez de Huete, Felip Esteve Oró, Victor Garcia Fernandez, Konstantinos Kampouropoulos, Laura Lopez Calvo, Jordi Palmiola Creus, David Sanderson, Svetan Ratchev, Laia Subirats Mate, Emma Shires, German Terrazas* and *Alison Turner*. Several of these people have also contributed to the peer review of the sessions in the online course which has helped us to ensure the quality of the materials developed.

Particular thanks goes to the team at EURECAT who developed the Moodle platform for the online training course. The team includes *Álvaro Martín Sancho, Magalí Lescano Correa, Lara Santana* and *Laia Subirats Mate*. We all appreciate your patience while you waited for material from us and our many requests for minor edits to the platform.

Our thanks also go to the people who helped with the translations of the materials into Italian and Spanish. There was considerably more translation required then we anticipated at the start of the project, and ensuring the technical content was translated correctly was no small task. For the Italian translations we thank *Sarah Behnam, Vincenzo Bellantone, Andrea Mazzoleni, Valeria Marrocco, Francesco Modica, Rossella Surace* and *Gianluca Trotta*. For the Spanish translations we thank *Ariadna Ros Carque* and *Francisco Gomez Villalba*.

We thank everyone who has helped to upload the sessions to the training platform, proof-read the sessions in the training course (in English, Italian or Spanish), worked to ensure consistency in the overall look and feel of the training course, and ensure the e-book is correctly formatted. Contributors here include *Jose*

A. Mulet Alberola, Mireia Dilmé i Martínez de Huete, Nancy Martin, Trunal Patil, Kryssa Roycroft and *Ruth Strickland.*

We give a special thanks to all the manufacturing companies that have been involved since the beginning in the Digit-T project. They have provided valuable suggestions for the definition and content of the training by participating in the user needs survey undertaken to understand the needs of SMEs and of different local regions, and by providing useful case studies that have been integrated in the training materials. We also thank everyone who has helped during the testing phase of the online training course.

A final thanks goes to *Helena Arrand*, for her project management and leadership which has been essential to the successful delivery of the project.

About the Digit-T project

The Digital Manufacturing Training System for SMEs (Digit-T) project was created to provide people, in particular SMEs, with an easily accessible introduction to Digital Manufacturing. The project has developed a free online training course and associated e-book – both can be accessed from the Digit-T website at digit-t.eu.

The project itself was a collaboration led by the University of Nottingham (UK) in conjunction with STIIMA-CNR (Italy), EURECAT (Spain) and Associazione Fabbrica Inteligente Lombardia (AFIL, Italy). It ran from 2017 – 2020 and was co-funded by the Erasmus+ programme of the European Union under grant 2017-1-UK01-KA202-036807.

University of Nottingham
UK | CHINA | MALAYSIA

The University of Nottingham is a world top 100 university and a world leader in pioneering research. The Institute for Advanced Manufacturing is driving the development of technology with the aim of improving all aspects of high-value manufacturing. The Institute pulls together internationally renowned academics in a diverse range of fields including digital manufacturing, advanced manufacturing technologies, additive manufacturing, composites manufacturing, manufacturing metrology, operations management, and human factors in manufacturing.

www.nottingham.ac.uk/manufacturing

Eurecat is the leading technology centre of Catalonia, providing the industrial and business sector with differential technology and advanced expertise to boost their competitiveness in a fast-paced environment. The centre brings together the expertise of more than 650 professionals who generate a volume of income of €51 million per year. Serving more than 1,600 companies, Eurecat is involved in more than 200 national and international research and development projects with high strategic value.

www.eurecat.org/en/field-of-knowledge/robotics-automation

The Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA) of the National Research Council (CNR) acts as a strategic player in the continuous process of definition of the European, national and regional visions on the manufacturing sector, through the participation in initiatives and platforms devoted to the formulation of new paradigms. The research and development activities of the Institute aim at innovating products, processes, and organisations in critical research areas of advanced manufacturing by focusing on mechatronic knowledge-based systems, robotics, multilayer adaptive control systems, microsystems, virtual prototyping and integrated simulation of mechatronic manufacturing systems, supply chain and new business models.

www.stiima.cnr.it/en

AFIL is an Italian private association, recognized by Lombardy Region as the regional technological cluster for Advanced Manufacturing. The cluster promotes and facilitates research and innovation actions by creating and animating communities of stakeholders with the final goal to improve Lombardy manufacturing system sustaining its leadership and competitiveness. To accomplish its mission, AFIL is involved in several interregional networks fostering the connection and exchange between local and foreign stakeholders operating in the field of Advanced Manufacturing.

www.afil.it