

# Data Economy 2.0: From Big Data Value to AI Value and a European Data Space



Sonja Zillner, Jon Ander Gomez, Ana García Robles, Thomas Hahn, Laure Le Bars, Milan Petkovic, and Edward Curry

**Abstract** Artificial intelligence (AI) has a tremendous potential to benefit European citizens, economy, environment and society and already demonstrated its potential to generate value in various applications and domains. From a data economy point of view, AI means algorithm-based and data-driven systems that enable machines with digital capabilities such as perception, reasoning, learning and even autonomous decision making to support people in real scenarios. Data ecosystems are an important driver for AI opportunities as they benefit from the significant growth of data volume and the rates at which it is generated. This chapter explores the opportunities and challenges of big data and AI in exploiting data ecosystems and creating AI value. The chapter describes the European AI framework as a foundation for deploying AI successfully and the critical need for a common European data space to power this vision.

**Keywords** Data ecosystem · Data spaces · Future directions · Big data value · Artificial intelligence

---

S. Zillner (✉)  
Siemens AG, Munich, Germany  
e-mail: [sonja.zillner@siemens.com](mailto:sonja.zillner@siemens.com)

J. A. Gomez  
Universitat Politècnica de València, València, Spain

A. García Robles  
Big Data Value Association, Bruxelles, Belgium

T. Hahn  
Siemens AG, Erlangen, Germany

L. Le Bars  
SAP, Paris, France

M. Petkovic  
Philips and Eindhoven University of Technology, Eindhoven, The Netherlands

E. Curry  
Insight SFI Research Centre for Data Analytics, NUI, Galway, Ireland

# 1 Introduction

Artificial intelligence (AI) has a tremendous potential to benefit European citizens, economy and society and already demonstrated its potential to generate value in various applications and domains. From a data economy point of view, AI means algorithm-based and data-driven systems that enable machines with digital capabilities such as perception, reasoning, learning and even autonomous decision making to support people in real scenarios. AI is based on a portfolio of technologies ranging from technologies for the perception and interpretation of information extracted from vast amounts of information data; software that draws conclusions and learns, adapts or adjusts parameters accordingly; and methods supporting human-based decision making or automated actions.

A critical driver for the emerging AI business opportunities is the significant growth of data volume and the rates at which it is generated. In 2014 the International Data Corporation (IDC) forecasted that in 2020 more than 16 zettabytes of useful data (16 trillion GB) will be made available, reflecting a growth of 236% per year from 2013 to 2020 (Turner et al. 2014). We know today that this forecast was far too low. According to a new update of the IDC Global Data Sphere<sup>1</sup> report, more than 59 zettabytes will be created, captured, copied and consumed. This growth is forecast to continue through 2024 with a 5-year compound annual growth rate (CAGR) of 26%. In consequence, this leads to an exponential growth, i.e. the amount of data being created over the next 3 years will be greater than the amount of data created over the past 30 years. The IDC report revealed that productivity/embedded data will be the fastest growing type of data with a CAGR of 40.3% from 2019 to 2024.

This chapter expands on a recent position paper (Zillner et al. 2018) from the Big Data Value Association community aligning it with recent developments on the European strategies for AI and data. It explores the potential of big data and AI in exploiting data ecosystems and creating new opportunities in AI application domains. It also addresses the ethical challenges associated with AI. It reflects on the need to develop trustworthy AI to mitigate conflicts and to avoid the adverse impact of deploying AI solutions. The European AI framework is described as a foundation for deploying AI successfully. The framework captures the processes and standards to deliver value that is acceptable to the users and citizens based on trust. Finally, the chapter describes the critical role of data and the need for common European data space to strengthen competitiveness across Europe.

---

<sup>1</sup>Worldwide Global DataSphere Forecast, 2020–2024: The COVID-19 Data Bump and the Future of Data Growth (Doc; #US44797920), IDC Report, [https://www.idc.com/getdoc.jsp?containerId=IDC\\_P38353](https://www.idc.com/getdoc.jsp?containerId=IDC_P38353)

## 2 The AI Value Opportunity

The current data explosion, combined with recent advances in computing power and connectivity, allows for an increasing amount of big data to be analysed anytime, anywhere. These technical advances enable addressing industrial relevant challenges and foster developing intelligent industrial application in a shorter time and with higher performance. AI will increase value creation from big data and its use to rapidly emerging B2B, B2G, G2C, G2B and B2C scenarios in many AI application domains. Machines and industrial processes which are supported by AI are augmenting human capacities in decision making and providing digital assistance in highly complex and critical processes.

Established industrial players are starting to implement AI in a wide range of industrial applications, such as complex image recognition, primarily for interpreting computed tomography (CT) and magnetic resonance imaging (MRI); autonomously learning, self-optimising industrial systems such as those used in gas turbines and wind farms; accurate forecasts of copper prices and expected power grid capacity utilisation; physical, autonomous systems for use in collaborative, adaptive, flexible manufacturing as part of Industry 4.0; and many more. At their heart, many of these AI systems are powered by using data-driven AI approaches such as deep learning. Exploiting data ecosystems is essential for AI (Curry and Sheth 2018).

In addition to the above, the EU Big Data Value Public-Private Partnership (BDV PPP) has established 32 projects with their respective experimentation playgrounds for the adoption of big data and AI solutions. In particular, the BDV PPP lighthouse projects play a fundamental role in piloting and showcasing value creation by big data with new data-driven AI applications in relevant sectors of great economic and societal value for Europe (Zillner et al. 2017). These projects demonstrate the essential role of data for AI, a few examples of which are as follows.

**DataBio** Data-Driven Bioeconomy takes on a major global challenge of how to ensure that raw materials for food, energy and biomaterials are sufficient in the era of climate change and population growth. Through big data and AI, DataBio is significantly enhancing raw material production in agriculture, forestry and fishery in a sustainable way. With its 26 pilots, DataBio strives to demonstrate annual increases in productivity ranging from 0.4% in forestry to 3.7% in agriculture and fishery (through savings in vessel costs). This makes up for a productivity gain of 20% over 5 years in agriculture and fishery. Big data pipelines and AI techniques are used in multiple pilots using the DataBio platform deployed in multiple clouds. The platform gathers Earth observation data from satellites and drones as well as IoT sources from in situ sensors in fields and vehicles. It manages and analyses the generated big data and presents it to the end users. These include farmers, foresters, fishers and many other stakeholders, supporting their operational decision making in a user-friendly way by providing them guidance in critical daily questions, such as what and where to grow, crop or fish; how to fight diseases; or when and how to harvest, cut or fish.

**TransformingTransport** Demonstrates in a realistic, measurable and replicable way the transformation that data-driven AI solutions can bring to the mobility and logistics market in Europe. Mobility and logistics are two of the most used industries in the world – contributing to approximately 15% of GDP and employment of over 11 million people in the EU-28 zone, i.e. 5% of the total workforce. The freight transport activities are projected to increase, since 2005, to 40% in 2030 and 80% in 2050. This will transform the current mobility and logistics processes to significantly higher efficiency and more profound impact. Structured into 13 different pilots, which cover areas of significant importance for the mobility and logistics sectors in Europe, TransformingTransport validates the technical and economic viability of big data-driven solutions for reshaping transport processes and services across Europe. To this end, TransformingTransport exploits access to industrial datasets from over 160 data sources, totalling over 164 terabytes of data. Initial evidence from TransformingTransport shows that big data-driven solutions using AI may deliver 13% improvement of operational efficiency<sup>2</sup>. The data-driven solutions in this project entail both traditional AI technology for descriptive analytics (such as support vector machines) and deep learning methods employed for predictive analytics (such as recurrent neural networks). With today's promising results using AI technology (e.g. 40% increase of prediction accuracy), we expect such AI solutions of advanced analytics as enablement to automated decision support for operational systems. These will establish the next level of efficiency and operational improvements in the mobility and transport sectors in Europe.

**BigMedilytics** In 2014, the EU-28 total healthcare expenditure was 1.39 trillion €. Spending is expected to increase to 30% by 2060, primarily due to a rapidly ageing population who typically suffer from chronic diseases. These figures indicate that current trends within the EU's healthcare sector are very unsustainable. The BigMedilytics Healthcare Lighthouse project demonstrates how the application of AI technologies on big data can help disrupt the healthcare sector so that quality, cost and access to care can all be improved. Market reports predict a CAGR of 40–50% for AI in healthcare, with a market size reaching to 22 billion by 2022 €. The project applies data-driven AI technologies over 12 pilots which focus on three main themes: (1) population health, (2) oncology and (3) industrialisation of healthcare. These themes effectively cover major disease groups, which cause 78% in mortality. AI-based methods together with privacy-preserving techniques are deployed to analyse large integrated datasets of more than 11 million patients, which cover a great range of key players in the healthcare sector (i.e. healthcare providers, healthtech companies, pharma and payers). The aim is to derive insights which can ultimately improve the efficiency of care providers while ensuring a high quality of care and protecting patients' privacy.

---

<sup>2</sup>According to the ALICE ETP, a 10% efficiency improvement will lead to EU cost savings of 100 B€.

**Boost 4.0** Roland Berger<sup>3</sup> reveals that big data could see the manufacturing industry add a gross value worth 1.25 T€ or suffer a loss of 605 B€ in lost value if it fails to incorporate new data, connectivity, automation and digital customer interface enablers in their digital manufacturing processes. European Data Market (EDM) Monitoring 2018 reports manufacturing as data market value leader with 14B€. However, the manufacturing industry is losing up to 99% of the data value since evidence cannot be presented at the speed decisions are made. Boost 4.0 reflects on this challenge, leveraging a European industrial data space for connected Smart Factory 4.0 that requires collecting, analysing, transporting and storing vast amounts of data. The Factory 4.0 will use such industrial data spaces to drive efficiencies through the advanced use of data-driven AI capabilities. First, connecting workforce, assets and things to the Internet will enable the leveraging of predictive maintenance to reduce equipment downtime by 50% and increase production by 20%. Second, integration with non-production departments enables new business insights with savings of around 160 B€ only for the top 100 European manufacturers thanks to improved zero-defect manufacturing and the ability to adjust production in real time. Lastly, improved data visibility among companies enables collaborative business models.

**DeepHealth** Healthcare is one of the most important sectors for the EU economy, as previously highlighted by the BigMedilytics project. In order to contribute to the adoption and use of AI and data technologies in the health sector within the EU, the DeepHealth project has two main goals: one at the technological level and the other at the economical level. The objective at the technological level is the development of two software libraries that aim to be at the core of European data-driven AI-based solutions/applications/systems regardless of the sector. In the case of the DeepHealth project, the use of both libraries is focused on healthcare as the 14 use cases are based on medical datasets. These two libraries are the European Deep Learning Library and the European Image Processing Library. Both libraries will make intensive use of hybrid HPC + big data architectures to process data by parallelising algorithms to learn from data and to process digital images. The integration of both libraries into software platforms will considerably reduce the time for training deep learning-based models and contribute to the other objective concerning economy, which is to increase the productivity of IT experts (ML practitioners and data scientists) working in the health sector. IT experts giving support to doctors and other medical personnel are usually faced with the problem of image manipulation (i.e. transformations, segmentation, labelling and extraction of regions of interest) where they need to use a set of different libraries and toolkits from different developers to define a pipeline of operations on images. Installing and configuring different libraries and toolkits is repetitive hard work. The DeepHealth project focuses on facilitating the daily work of IT experts by integrating all the necessary functionalities into a toolkit, including the two libraries and a front-end for using them. The toolkit, one of the

---

<sup>3</sup>[https://www.rolandberger.com/publications/publication\\_pdf/roland\\_berger\\_digital\\_transformation\\_of\\_industry\\_20150315.pdf](https://www.rolandberger.com/publications/publication_pdf/roland_berger_digital_transformation_of_industry_20150315.pdf)

outcomes of this project, will facilitate the definition of pipelines of operations on images and testing distinct Deep Neural Network (DNN) topologies.

### **3 AI Challenges**

The challenges for the adoption of AI range from new business models that need to be developed, trust in AI that needs to be established, ecosystems that are required to ensure that all partners are on board as well as access to the state-of-the-art AI technology. The following subsection will detail all these aspects.

#### **3.1 *Business Models***

With the recent technical advances in digitalisation and AI, the real and the virtual worlds are continuously merging, which, again, leads to entire value-added chains being digitalised and integrated. For instance, in the manufacturing domain, all the way from the product design through to on-site customer services is digitalised. The increase in industrial data combined with AI technologies triggers a wide range of new technical applications with new forms of value propositions that shift the logic of how business is done. To capture these new types of value, data-driven AI-based solutions for the industry will require new business models. The design of data-driven AI-based business models needs to incorporate various perspectives ranging from customer and user needs and their willingness to pay for new AI-based solutions to data access and the optimal use of technologies while taking into account the currently established relationships with customers and partners. Successful AI-based business models are often based on strategic partnerships with two or more players establishing the basis for sustainable win-win situations through transparent ways of sharing resources, investments, risks, data and value.

#### **3.2 *Trust in AI***

With AI disruptive potential, there are significant ethical implications on the use of AI and autonomous machines and their applications for decision support. Future AI research needs to be guided by new and established ethical norms. Although the current AI methods have already achieved encouraging results and technical breakthroughs, results in individual cases show some concerning signs of unpredictable behaviour. Recent studies showed that the state-of-the-art deep neural networks are vulnerable to adversarial examples or are unable to cope with new unknown situations. To overcome those shortcomings, for any critical applications (where “critical” needs to be defined with clarity), one should be able to explain how AI

applications came to a specific result (“explainable AI”). Explainability will ensure the commitment of industrial users to measurable ethical values and principles when using AI. One should foster responsible technological development (e.g. avoid bias) and enhance transparency in such exercise. Explainable AI should provide transparency about input data as well as the “rationale” behind the algorithm usage leading to the specific output. The algorithm itself need not necessarily be revealed in this case.

The purpose of AI, data analytics, machine and deep learning algorithms is not only to boost the effectiveness and quality of the services which are delivered to the client but also to ensure that no negative impact is brought as a result of deploying AI solutions in critical applications. For instance, ensuring that AI-powered systems treat different social groups fairly is a matter of growing concern for societies. FAT-ML, i.e. Fairness, Accountability and Transparency in Machine Learning, is an emerging important multidisciplinary field of research (Barocas and Selbst 2016; Carmichael et al. 2016). Related areas including big data for social good, humanistic AI and the broader field of AI ethics have only recently started exploring complex multi-faceted problems, e.g. fostering the creation of social and human-centred values by adding new parameters and enhanced objective functions and restrictions.

Trusted AI involves the simultaneous achievement of objectives that are often in conflict. One critical challenge stems from the ever-increasing collection and analysis of personal data and the crucial requirement for protecting the privacy of all involved data subjects as well as protecting commercially sensitive data of associated organisations and enterprises. There are some approaches attempting to address this issue, including security-oriented (e.g. machine learning on encrypted data with secure computation technologies), privacy-enhancing (e.g. detect privacy risks and alert users) and distributed processing (e.g. federated machine learning) ones. As all privacy approaches add cost and complexity to AI systems, the optimal trade-offs without adding considerable complexity are important research challenges to be addressed. A critical problem is presented by the difficulty to allocate and distribute liabilities and responsibilities across assemblages of continuously evolving autonomous systems with different goals and requirements. While existing risk-based, performance-driven, progressive and proportionate regulatory approaches have promised a more flexible, adaptive regulatory environment, stakeholders are increasingly struggling to deal with the complexities of multi-level, multi-stakeholder and multi-jurisdictional environments within which AI is being developed. Multidisciplinary efforts at both international and regional levels are therefore required to ensure the establishment of an enabling environment where trust and safety of AI are dealt with from a global governance perspective. Existing tools from other domains, such as regulatory sandboxing, testing environments for autonomous vehicles and so forth, could serve as incubators for establishing new policy; legal, ethical and regulatory norms; and measures of trusted AI in Europe.

### 3.3 *Ecosystem*

For developing sustainable data-driven AI businesses, it will be central to consider a value-network perspective, i.e. looking at the entire ecosystem of companies involved in value networks. The ecosystems will be increasingly shaped by platform providers who offer their platform based on open standards to their customers. European economic success and sustainability in AI will be driven by ecosystems which need to have a critical size. Speed is a necessity for the development of these ecosystems.

Data sharing and trading are essential ecosystem enablers in the data economy, although secure and personal data present particular challenges for the free flow of data (OECD 2014; Curry 2016). The EU has made considerable efforts in the direction of defining and building data-sharing platforms. However, there is still a significant way to go to guarantee AI practitioners' access to large volumes of data necessary for them to compete. Further actions must be carried out to develop data for AI platforms, such as awareness campaigns to foster the idea of sharing their data in companies and research centres, and incentives for parties to join data exchange/sharing initiatives. To overcome barriers to data sharing for AI, frameworks for data governance are needed to be established that will enable all parties to retain digital sovereignty over their data assets. Obviously, data sharing must be done, from the legal point of view, by preserving privacy by anonymising all the attributes referring to people, and respecting commercial interests (IPR, competition, ownership) by providing solutions to deal with technical and legal challenges such as data governance and trust-enhancing protocols for data sharing/exchange, decentralised storage and federated machine learning. And from the technical perspective, data sharing is done by (1) designing information systems (i.e. databases) in order to ensure the future use of the datasets with minimal efforts in terms of cleaning data or defining ontologies, by (2) transforming and mapping data sources taking into account the variety and heterogeneity of data in order to gain interoperability and (3) by ensuring the veracity of shared data according to quality standards.

Open AI platforms will play a central role in the data economy at three different levels: (1) definition of protocols and procedures for uploading datasets into data-sharing platforms, (2) definition of standard APIs for different libraries (AI/ML, image processing, etc.) and (3) the design and development of a web-based user interface to allow data scientists to upload data, to define pipelines of transformations to apply to data before training and testing AI models, and to choose among a wide range of AI techniques to run on the same data to carry out comparative studies. Successful European Open AI platforms require the contribution of many agents, such as universities, research centres, large companies and SMEs.

By relying on data-sharing platforms, data innovation spaces, Open AI platforms and digital innovation hubs (DIH), industrial collaborations between large and small players can be supported at different levels: technical, business model and ecosystem while, at the same time, ensuring data and technology access for SMEs and start-ups.



To complement technical and legal infrastructures for the free and controlled flow of industrial data, the building and nurturing of industrial ecosystems fostering data-driven industrial cooperation across value chains and therefore networks will have a critical impact.

Enabling data-driven AI-based business models across value chains and beyond organisational boundaries will significantly maximise the impact of the data economy to power European AI industries. Mechanisms that overcome the lack of data interoperability and foster data sharing and exchange need to be defined and implemented. Notwithstanding, the creation of and compliance with binding international standards is of central importance to the sustainability of solutions, and thus it is a competitive strength. Preferably these standards should be global – because only global standards ultimately lead to success in a world that is more and more networked and where multinational companies make significant contributions to national GDPs.

### 3.4 *Technology*

Success in industrial AI application relies on the combination of a wide range of technologies, such as:

**Advanced Data Analytics:** Many data analytics techniques require adaptation for running more efficiently when working with large datasets. These improvements rely on the development of new algorithms and new ways of transforming data. Additionally, with self-adjusting AI systems, machines will become self-operating by making decisions according to specific contexts to dynamically optimise performance, beyond the level of efficiency the same AI systems can reach when adjusted by humans.

**Hybrid AI:** To derive value from domain knowledge, methods from both symbolic AI and statistical AI need to be combined to give the maximum potential and usability of AI-based applications. This combination of knowledge graphs and statistical AI techniques supports AI solutions concerning (1) data quality issues, (2) better integration and use of training data, (3) explainable AI (no black-box solutions) and, finally, (4) the mutual fertilisation of semantic technologies and AI techniques towards self-optimising machines.

**Distributed AI/Edge Analytics:** The increasing number of intelligent devices at the edge is one of the critical elements for AI. We are now at the point where the collective computing power at the edge (outside data centres) is surpassing the centralised computing capacity. As computing capabilities in the cloud and at the edge are increasingly intertwined, we will see the emergence of distributed AI, i.e. new research approaches that will bring the AI at the core of most future data analytics-based applications.

**Hardware Optimised to AI:** Specialised hardware devices and architectures have an increasingly strong impact both on the AI learning process on applications

with large datasets and on the predicting/inference task, in particular when fast decisions and actuation matter. The designs of powerful and affordable systems on both sides of the AI data flow are an important research topic. Nevertheless, AI algorithms need to be optimised to the specific hardware capabilities.

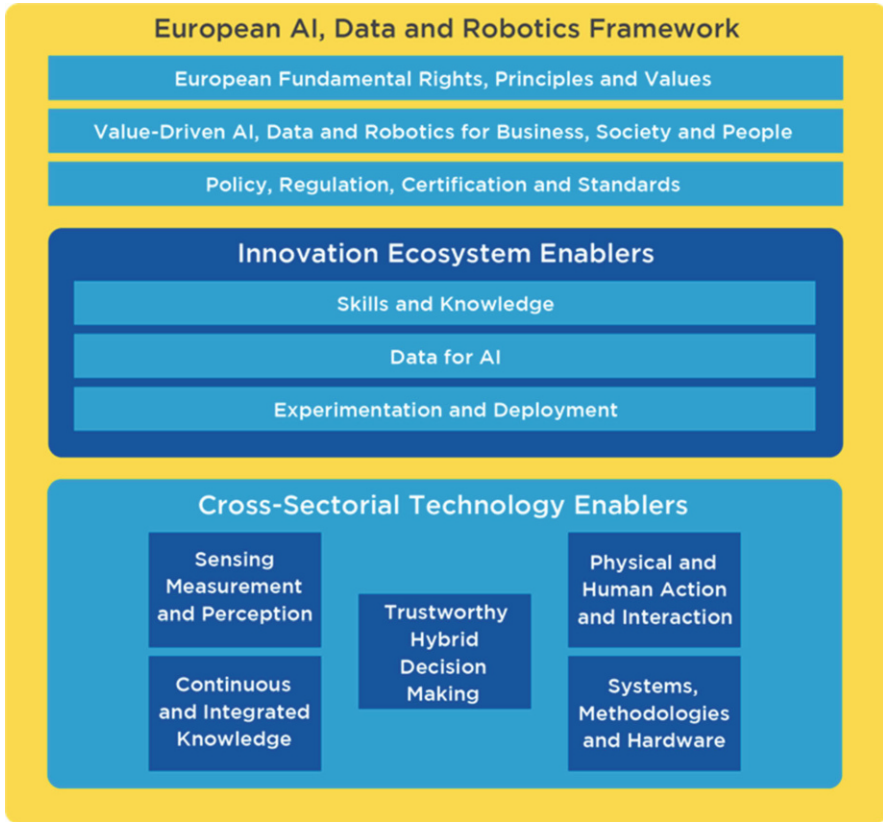
**Multilingual AI:** Humans use language to express, store, learn and exchange information. AI-based multilingual technologies can extract knowledge out of tremendous amounts of written and spoken language data. Processing of multilingual data empowers a new generation of AI-based applications such as question answering systems, high-quality neural machine translation, speech processing in real time and contextually and emotionally aware virtual assistants for human-computer interaction.

## 4 Towards an AI, Data and Robotics Ecosystem



The Big Data Value Association (BDVA) and the European Robotics Association (euRobotics) have developed a joint Strategic Research, Innovation and Deployment Agenda (SRIDA) for an AI, Data and Robotics Partnership in Europe (S Zillner et al. 2019). This is in response to the Commission Communication on AI published in December 2018. Deploying AI successfully in Europe requires an integrated landscape for its adoption and the development of AI based on Europe’s unique characteristics. In September 2020 the BDVA, CLAIR, ELLIS, EurAI and euRobotics are pleased to announce the official release of the joint Strategic Research Innovation and Deployment Agenda (SRIDA) for the AI, Data and Robotics Partnership which unifies the strategic focus of each of the three disciplines engaged in creating the Partnership.

Together these associations have proposed a vision for an AI, Data and Robotics Partnership: “The Vision of the Partnership is to boost European industrial competitiveness, societal wellbeing and environmental aspects to lead the world in developing and deploying value-driven trustworthy AI, Data and Robotics based on fundamental European rights, principles and values”.



**Fig. 1** European AI, Data and Robotics Framework and Enablers (Zillner et al. 2020) (by European Commission licensed under CC BY 4.0)

To deliver on the vision of the AI, Data and Robotic Partnership, it is important to engage with a broad range of stakeholders. Each collaborative stakeholder brings a vital element to the functioning of the Partnership and injects critical capability into the ecosystem created around AI, Data and Robotics by the Partnership. The mobilisation of the European AI, Data and Robotics Ecosystem is one of the core goals of the Partnership. The Partnership needs to form part of a wider ecosystem of collaborations that cover all aspects of the technology application landscape in Europe. Many of these collaborations will rely on AI, Data and Robotics as critical enablers to their endeavours. Both horizontal (technology) and vertical (application) collaborations will intersect within an AI, Data and Robotics Ecosystem.

Figure 1 sets out the context for the operation of the AI, Data and Robotics. It clusters the primary areas of importance for AI, Data and Robotics research, innovation and deployment into three overarching areas of interest. *European AI, Data and Robotics Framework* represents the legal and societal fabric that underpins the impact of AI on stakeholders and users of the products and services that businesses will provide. The AI, Data and Robotics *Innovation Ecosystem Enablers*

represent the essential ingredients for effective innovation and deployment to take place. Finally, the *Cross-Sectorial AI, Data and Robotics Technology Enablers* represent the core technical competencies that are essential for the development of AI, Data and Robotics systems. The remainder of this section offers a summary of the European AI, Data and Robotics Framework, which is the core of the SRIDA (Zillner et al. 2020) developed by the BDVA, euRobotics, ELLIS, EurAI and CLAIRE.

## 4.1 *European AI, Data and Robotics Framework*

AI, Data and Robotics work within a broad framework that sets out boundaries and limitations on their use. In specific sectors, such as healthcare, they operate within the ethical, legal and societal contexts and within regulatory regimes that can vary across Europe. Products and services based on AI, Data and Robotics are shaped by certification processes and standards and impact on users to deliver value compatible with European rights, principles and values. Critical to deploying AI, Data and Robotics is its acceptance by users and citizens, and this acceptance can only come when they can assign trust. This section explores this European AI, Data and Robotics Framework (Zillner et al. 2020) within which research, design, development and deployment must work.

**European Fundamental Rights, Principles and Values** On the one hand, the recent advances in AI, Data and Robotics technology and applications have fundamentally challenged the ethical values, human rights and safety in the EU and globally. On the other hand, AI, Data and Robotics offer enormous possibilities to raise productivity, address societal and environmental challenges and enhance the quality of life for everyone. The public acceptance of AI, Data and Robotics is a prerequisite for it being trustworthy, ethical and secure, and without public acceptance, its full benefit cannot be realised. The European Commission has already taken action and formulated in its recent communications<sup>4</sup> a vision for an ethical, secure and cutting-edge *AI made in Europe* designed to ensure AI, Data and Robotics operate within an appropriate ethical and legal framework that embeds European values. The Partnership (Zillner et al. 2020) will:

- *Facilitate a multi-stakeholder dialogue* and consensus building around the core issue of trustworthiness by guiding and shaping a common AI, Data and Robotics agenda and fostering research and innovation on trustworthy technologies.

---

<sup>4</sup>Communication Artificial Intelligence on 25 April 2018 (see <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>) and Communication Artificial Intelligence on 7 December 2018 (see [https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07\\_en](https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07_en))

- *Seek to promote a common understanding* among stakeholders of the European AI, Data and Robotics ecosystem on the fundamental, rights and values, so that each sector and community are informed and aware of the potential of AI, Data and Robotics as well as the risks and limitations of the current technology and will develop guidance in the responsible implementation of AI, Data and Robotics.
- Establish the basis for identifying and expressing a *European strategic viewpoint* on rights, principles and values by providing clear links to relevant regulation, certification and standardisation.

**Capturing Value for Business, Society and People** Technical advances in AI, Data and Robotics are now enabling real-world applications. These are leading to improved or new value-added chains being developed and integrated. To capture these new forms of value, AI-based solutions may require innovative business models that redefine the way stakeholders share investments, risk, know-how and data and, consequently, value. This alteration of value flow in existing markets is disruptive and requires stakeholders to alter their business models and revenue streams. These adjustments require new skills, infrastructure and knowledge, and organisations may have to buy in expertise or share data and domain know-how to succeed. This may be incredibly difficult if their underlying digitalisation skills, a prerequisite for AI, Data and Robotics adoption, are weak.

Even incremental improvements or more considerable changes carry risks and may create a reluctance to adopt AI, Data and Robotics. There may be little or no support for change within an organisation or value chain, especially when coupled with a lack of expertise. Successful adoption of AI, Data and Robotics solutions requires a dialogue between the different stakeholders to design a well-balanced and sustainable value network incorporating all stakeholder's interests, roles and assets.

To support the adoption of AI, Data and Robotics applications, the Partnership (Zillner et al. 2020) will stimulate discussions to align supply and demand perspectives of the diverse AI, Data and Robotics value-network partners, with the main focus on application areas and sectors that:

- Are crucial for the European economy
- Relate to critical infrastructure
- Have a social or environmental impact
- Can increase European competitiveness in AI

**Policy, Regulation, Certification and Standards (PRCS)** The adoption of AI, Data and Robotics depends on a legal framework of approval built on regulation, partly driven by policy, and an array of certification processes and standards driven by industry. As AI, Data and Robotics are deployed successfully in new market areas, regulation and certification can lag behind, thereby creating barriers to adoption.

Similarly, a lack of standards and associated certification and validation methods can hold back the deployment and the creation of supply chains and therefore slow

market uptake. In some areas of AI, Data and Robotics, the market will move ahead and wait for regulation to react, but in many application areas existing regulation can present a barrier to adoption and deployment – most notably in applications where there is a close interaction with people, either digitally or physically, or where AI, Data and Robotics are operating in safety or privacy critical environments.

PRCS issues are likely to become a primary area of activity for the AI, Data and Robotics Partnership. Increasingly it is regulation that is the primary lever for the adoption of AI/Data/Robotics systems, particularly when physical interactions are involved or where privacy is a concern. Similarly, the development of standards, particularly around data exchange and interoperability, will be key to the creation of a European AI, Data and Robotics marketplace. Establishing ways that ensure conformity assessments of AI, Data and Robotics will underpin the development of trust that is essential for acceptance and therefore adoption. In addition, the Partnership also has a role to advise on regulation that creates or has the potential to create unnecessary barriers to innovation in AI, Data and Robotics. The Partnership (Zillner et al. 2020) will need to carry out the following activities to progress PRCS issues:

- Identify key stakeholders in each area of PRCS and ensure there is good connectivity between them and to the AI, Data and Robotics Ecosystem.
- Work with stakeholders and the emerging AI, Data and Robotics Ecosystem infrastructure (digital innovation hubs, pilots and data spaces) to identify key issues that impact on adoption and deployment in each major sector.
- Promote best practice in deployment regarding PRCS issues and provide signposts to demonstrators and processes that can accelerate uptake.
- Support and collaborate in standardisation initiatives and the harmonisation of regulation across Europe to create a level AI, Data and Robotics single marketplace and connect with European and global standards and regulatory bodies.
- Foster the responsible testing of AI, Data and Robotics innovation in regulatory sandbox environments.
- Consolidate recommendations towards policy changes and provide support for related impact assessment processes.
- Drive European thinking and needs towards international standardisation bodies.

## ***4.2 Innovation Ecosystem Enablers***

The Innovation Ecosystem Enablers are essential ingredients for success in the innovation system. They represent resources that underlie all innovation activities across the sectors and along the innovation chain from research to deployment. Each represents a key area of interest and activity for the Partnership (Zillner et al. 2020), and each presents unique challenges to the rapid development of European AI, Data and Robotics.

**Skills and Knowledge** As traditional industry sectors undergo an AI, Data and Robotics transformation, so too must their workforces. There is a clear skills gap when it comes to AI, Data and Robotics. However, while there are shortages of people with specific technical skills or domain knowledge, there is also the need to train interdisciplinary experts. AI, Data and Robotics experts need insight into the ethical consequences posed by AI, by machine autonomy and by big data automated processes and services; they need a good understanding of the legal and regulatory landscape, for example, GDPR, and the need to develop and embed trustworthiness, dependability, safety and privacy through the development of appropriate technology.

The Partnership will work through its network to ensure that all stakeholders along the value chain, including citizens and users, have the understanding and skills to work with AI-enabled systems, in the workplace, in the home and online. The Partnership has a critical role to play in bringing together the key stakeholders: academia, industry, professional trainers, formal and informal education networks and policymakers. These collaborations will need to examine regional strengths and needs in terms of skills across the skill spectrum, both technical and non-technical. It is critical to ensure that the skill pipeline is maintained to ensure the AI, Data and Robotics transformation of Europe is not held back. Some concrete actions the Partnership (Zillner et al. 2020) will focus on are as follows:

- Promote equality and diversity within the current and future workforce and ensure diversity and balance in the educational opportunities that drive the skill pipeline.
- Ensure the alignment of curricula and training programmes for AI, Data and Robotics professionals with industry needs.
- Establish AI, Data and Robotics skills, both technical and non-technical, through certification mechanisms for university courses, professional and vocational training, and informal learning.
- Development of complementary short courses related to artificial intelligence aimed at decision makers in industry and public administration and those wishing to upgrade, enhance or acquire AI-based skills.
- Support for secondary education and adult learning to cover STEM skills including the ethical, social and business aspects of AI together with the changing nature of work as well as support for vocational training.

**Data for AI** In order to further develop AI, Data and Robotics technologies and meet expectations, large volumes of cross-sectoral, unbiased, high-quality and trustworthy data need to be made available. Data spaces, platforms and marketplaces are enablers, the key to unleashing the potential of such data. There are however important business, organisational and legal constraints that can block this scenario such as the lack of motivation to share data due to ownership concerns, loss of control, lack of trust, the lack of foresight in not understanding the value of data or its sharing potential, the lack of data valuation standards in marketplaces, the legal blocks to the free flow of data and the uncertainty around data policies. Additionally,

significant technical challenges such as interoperability, data verification and provenance support, quality and accuracy, decentralised data sharing and processing architectures, and maturity and uptake of privacy-preserving technologies for big data have a direct impact on the data made available for sharing. The Partnership (Zillner et al. 2020) will:

- Create the conditions for the development of **trusted European data-sharing frameworks** to enable new data value chain opportunities, building upon existing initiatives and investments (data platforms, i-spaces, big data innovation hubs). Data value chains handling a mix of personal, non-personal, proprietary, closed and open research data need to be supported. The Partnership would promote open datasets and new open benchmarks for AI algorithms, subject to quality validation from both software engineering and functional viewpoints.
- Define specific measures to **incorporate data sharing at the core of the data lifecycle** for greater access to data, encouraging collaboration between data value chain actors in both directions along the chain and across different sectors. Additionally, the Partnership will provide **supportive measures for European businesses** to safely embrace new technologies, practices and policies.
- Facilitate **coordination and harmonisation of member states efforts** and realise the potential of European-wide AI-digital services in the face of global competition. It would guide and influence standards concerning tools for data sharing, privacy preservation, quality verification, collaboration and interaction. Promote standardisation at European level but maintain collaboration with international initiatives for made-in-Europe AI to be adopted worldwide.

**Experimentation and Deployment** They are central levers for AI/Data/Robotics-based innovation because of the need to deploy in complex physical and digital environments. This includes safe environments for experimentation to explore the data value as well as to test the operation of autonomous actors. AI/Data/Robotics-driven innovations rely on the interplay of different assets, such as data, robotics, algorithms and infrastructure. For that reason, *cooperation* with other partners is central to gaining access to complementary assets. This includes access to the AI, Data and Robotics Ecosystem covering AI platform providers, data scientists, data owners, providers, consumers, specialised consultancy, etc. The Partnership (Zillner et al. 2020) will:

- Stimulate cooperation between all stakeholders in the AI, Data and Robotics value chain around experimentation and deployment.
- Enable access to infrastructure and tools in combination with datasets covering the whole value chain as a basis for doing experiments to support development and deployment.
- Support the creation and linking of DIHs, centres of excellence and all other EC initiatives.
- Support AI/Data/Robotics-based incubators as well as testbed developments as well as promote initiatives that enable SME access to infrastructure and tools at low cost.



- Foster set-ups that bring together industrial users with research excellence and domain experts with data science skills, aiming to fill the gaps between domain/business and technical expertise.

### ***4.3 Cross-Sectorial AI, Data and Robotics Technology Enablers***

The last part of the framework is the *technology enablers* for building successful AI products and services. Each embodies the concept that AI, Data and Robotics need to work in unison to achieve optimal function and performance. They represent the fundamental building blocks needed to create AI, Data and Robotics systems of all types.

The sensing and perception and knowledge and learning technology enablers create the data and knowledge on which decisions are made. These are used by the reasoning and decision-making technologies to deliver: edge and cloud based decision making, planning, search and optimisation in systems and the multi-layered decision making necessary for AI, Data and Robotic systems operating in complex environments.

Action and interaction cover the challenges of human interaction, machine to machine interoperation and machine interaction with the human environment. These multiple forms of action and interaction create complex challenges that range from the optimisation of performance to physical safety and social interaction with humans in unstructured and multi-faceted environments.

Systems, hardware, methods and tools provide the technologies that enable the construction and configuring of systems, whether they are built purely on data or on autonomous robots. These tools, methods and processes integrate AI, Data and Robotics technologies into systems and are responsible for ensuring that core system properties and characteristics such as safety, robustness, dependability and trustworthiness can be integrated into the design cycle and tested, validated and ultimately certified for use.

Each technical area overlaps with the other; there are no clear boundaries. Indeed, exciting advances are most often made in the intersections between these five areas and the system-level synergies that emerge from the interconnections between them.

## **5 A Common European Data Space**

For European data economy to develop further and meet expectations, large volumes of cross-sectoral, unbiased, high-quality and trustworthy data need to be made available. The exploration of ethical, secure and trustworthy legal, regulatory and

governance frameworks is needed. European values, e.g. democracy, privacy safeguards and equal opportunities, can become the trademark of European data economy technologies, products and practices. Rather than be seen as restrictive, these values enforced by legislation should be considered as a unique competitive advantage in the global data marketplace.

To reflect this new reality, the European data strategy was revised in 2020 to set out a vision for the EU to become a role model for a data-driven society and to create a single market for data to ensure Europe's global competitiveness and data sovereignty. As highlighted by EU Commissioner Thierry Breton<sup>5</sup>: "To be ahead of the curve, we need to develop suitable European infrastructures allowing the storage, the use, and the creation of data-based applications or Artificial Intelligence services. I consider this as a major issue of Europe's digital sovereignty".

Alignment and integration of established data-sharing technologies and solutions, and further developments in architectures and governance models aiming to unlock data silos, would enable data analytics across a European data-sharing ecosystem. This will enable AI-enhanced digital services to make analysis and predictions on European-wide data, thereby combining data and service economies. New business models will help to exploit the value of those data assets through the implementation of AI among participating stakeholders including industry; local, national and European authorities and institutions; research entities; and even private individuals.

As part of the revised data strategy, common European data spaces will ensure that more data becomes available for use in the economy and society while keeping companies and individuals who generate the data in control (*Communication: A European strategy for data* 2020). Platform approaches have proved successful in many areas of technology (Gawer and Cusumano 2014), from supporting transactions among buyers and sellers in marketplaces (e.g. Amazon), innovation platforms that provide a foundation on which to develop complementary products or services (e.g. Windows), to integrated platforms which are a combined transaction and innovation platform (e.g. Android and the Play Store). The idea of large-scale "data" platforms has been touted as a possible next step to support data ecosystems (Curry and Sheth 2018). An ecosystem data platform would have to support continuous, coordinated data flows, seamlessly moving data among systems (Curry and Ojo 2020). Data spaces, platforms and marketplaces are enablers, the key to unleashing the potential of such data. Significant technical challenges such as interoperability, data verification and provenance support, quality and accuracy, decentralised data sharing and processing architectures, and maturity and uptake of privacy-preserving technologies for big data have a direct impact on the data made available for sharing.

The nine initial common European data spaces (Fig. 2) will be the following:

- An **industrial** data space, to support the competitiveness and performance of the EU's industry

---

<sup>5</sup>15 July 2020: [https://ec.europa.eu/commission/presscorner/detail/en/SPEECH\\_20\\_1362](https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_20_1362)

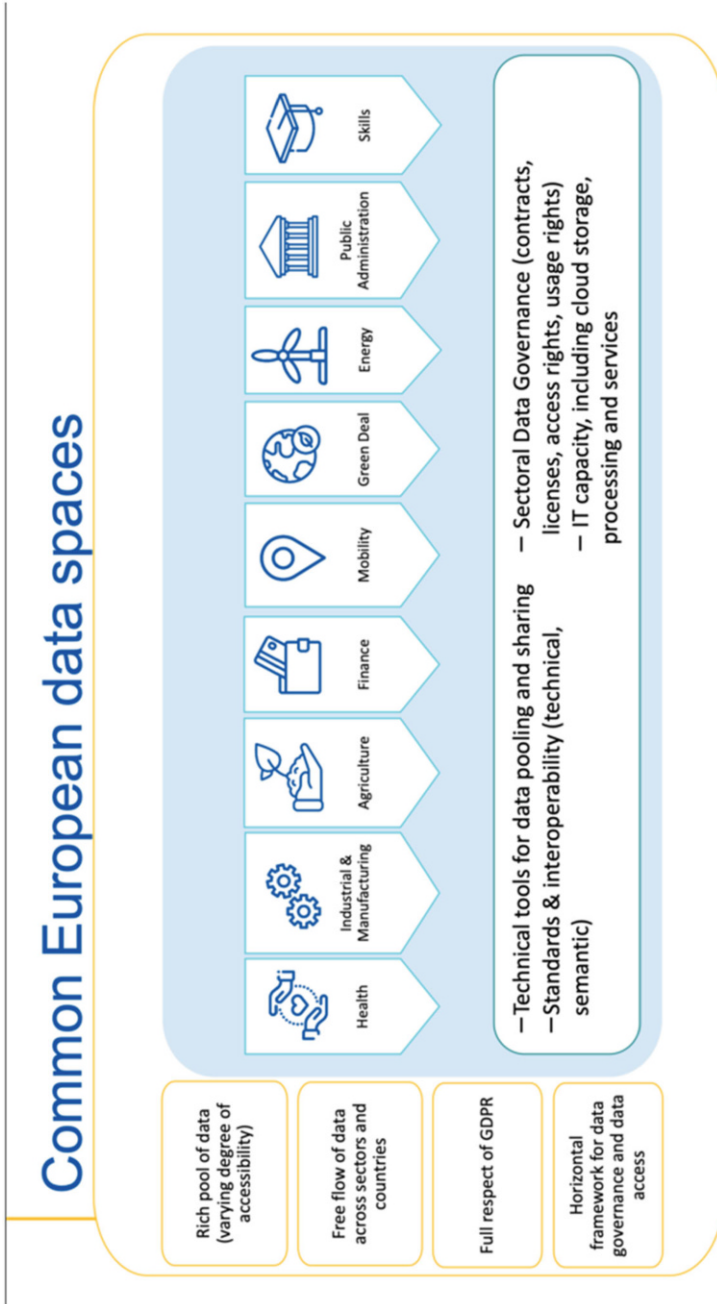


Fig. 2 Common European data spaces (Communication: A European strategy for data 2020) (by European Commission licensed under CC BY 4.0)

- A **Green Deal** data space, to use the major potential of data in support of the Green Deal priority actions on issues such as climate change, circular economy, pollution, biodiversity and deforestation
- A **mobility** data space, to position Europe at the forefront of the development of an intelligent transport system
- A **health** data space, essential for advances in preventing, detecting and treating diseases as well as for informed, evidence-based decisions to improve the healthcare systems
- A **financial** data space, to stimulate innovation, market transparency, sustainable finance as well as access to finance for European businesses and a more integrated market
- An **energy** data space, to promote a more substantial availability and cross-sector sharing of data, in a customer-centric, secure and trustworthy manner
- An **agriculture** data space, to enhance the sustainability performance and competitiveness of the agricultural sector through the processing and analysis of data
- Data spaces for **public administrations**, to improve transparency, accountability and efficiency of public spending, fighting corruption, both at EU and national levels
- A **skills** data space, to reduce the skills mismatches between the education and training systems and the labour market needs

## 6 Summary

AI, Data and Robotics have a tremendous potential to benefit citizens, economy, environment and society. AI, Data and Robotics techniques can extract new value from data to enable data-driven systems with digital capabilities such as perception, reasoning, learning and even autonomous decision making. Data ecosystems are an important driver for data-driven AI to exploit the continued growth of data. We need to establish a solid European AI, Data and Robotics framework as a foundation for deploying AI, Data and Robotics successfully and a common European data space to power this vision. Developing both of these elements together is critical to maximising the future potential of AI and data in Europe.

**Acknowledgements** Editor and contributors to the BDVA position paper on data-driven AI: Andreas Metzger (paluno, University of Duisburg-Essen), Zoheir Sabeur (University of Southampton), Martin Kaltenböck (Semantic Web Company), Marija Despenic (Philips), Cai Södergard (VTT), Natalie Bertels/Ivo Emanuilov (imec-CiTiP-KU Leuven), Simon Scerri (Fraunhofer), Andrejs Vasiljevs/Tatjana Gornosttaja (Tilde), Axel Ngongo (Technical University of Paderborn), Freek Bomhof (TNO), Yiannis Kompatasiaris and Symeon Papadopoulos (ITI Greece), Nozha Boujemaa (Inria), Juan-Carlos Perez-Cortes (ITI Valencia), Oscar Lazaro (Innovalia Association).

## References

- Barocas, S., & Selbst, A. (2016). Big data's disparate impact. *California Law Review*. <https://doi.org/10.15779/Z38BG31>
- Carmichael, L., Stalla-Bourdillon, S., & Staab, S. (2016). Data mining and automated discrimination: A mixed legal/technical perspective. *IEEE Intelligent Systems*, 31(6), 51–55. <https://doi.org/10.1109/MIS.2016.96>
- Communication: A European strategy for data. (2020). Retrieved from [https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020\\_en.pdf](https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf)
- Curry, E. (2016). The big data value chain: Definitions, concepts, and theoretical approaches. In J. M. Cavanillas, E. Curry, & W. Wahlster (Eds.), *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe*. [https://doi.org/10.1007/978-3-319-21569-3\\_3](https://doi.org/10.1007/978-3-319-21569-3_3)
- Curry, E., & Ojo, A. (2020). Enabling knowledge flows in an intelligent systems data ecosystem. In *Real-time Linked Dataspaces* (pp. 15–43). Berlin: Springer. [https://doi.org/10.1007/978-3-030-29665-0\\_2](https://doi.org/10.1007/978-3-030-29665-0_2)
- Curry, E., & Sheth, A. (2018). Next-generation smart environments: From system of systems to data ecosystems. *IEEE Intelligent Systems*, 33(3), 69–76. <https://doi.org/10.1109/MIS.2018.033001418>
- Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management*, 31(3), 417–433. <https://doi.org/10.1111/jpim.12105>
- OECD. (2014). *Data-driven Innovation for Growth and Well-being*.
- Turner, V., Gantz, J. F., Reinsel, D., & Minton, S. (2014). The digital universe of opportunities: Rich data and the increasing value of the internet of things. *Report from IDC for EMC*.
- Zillner, S., Curry, E., Metzger, A., Auer, S., & Seidl, R. (Eds.). (2017). *European big data value strategic research & innovation agenda*. Retrieved from Big Data Value Association website: [www.bdva.eu](http://www.bdva.eu)
- Zillner, S., Gomez, J. A., García Robles, A., & Curry, E. (Eds.). (2018). *Data-driven artificial intelligence for european economic competitiveness and societal progress: BDVA position statement*. Retrieved from [www.bdva.eu/sites/default/files/AI-Position-Statement-BDVA-Final-12112018.pdf](http://www.bdva.eu/sites/default/files/AI-Position-Statement-BDVA-Final-12112018.pdf)
- Zillner, S., Bisset, D., García Robles, A., Hahn, T., Lafrenz, R., Liepert, B., & Curry, E. (2019). *Strategic research, innovation and deployment agenda for an AI PPP: A focal point for collaboration on artificial intelligence, data and robotics*. Retrieved from <http://www.bdva.eu/sites/default/files/AIPPPSRIDA-ConsultationVersion-June2019-Onlineversion.pdf>
- Zillner, S., Bisset, D., Milano, M., Curry, E., Hahn, T., Lafrenz, R., et al. (2020). *Strategic research, innovation and deployment agenda - AI, data and robotics partnership. Third Release (3rd)*. Brussels: BDVA, euRobotics, ELLIS, EurAI and CLAIRE.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

