Navigating the Twilight Zone

Pathways towards digital transformation of food systems

WAGENINGEN UNIVERSITY & RESEARCH

Navigating the Twilight Zone

Pathways towards digital transformation of food systems

Sjaak Wolfert, Lan van Wassenaer, Simone van der Burg, Mark Ryan, Laurens Klerkx, Kelly Rijswijk, Mariette McCampbell, Ioannis Athanasiadis, George Beers

WAGENINGEN
UNIVERSITY & RESEARCH

Index

1 Introduction

1.1 The digital transformation of food systems is in a twilight zone

The food system - from farm to fork - becomes increasingly digital. This is due to the so-called cyber-physical management cycle of smart sensing & monitoring, smart analysis & planning and smart control. Objects, such as crops, animals, or trucks, are monitored by smart sensors, satellites, or drones with cameras. They are smart because their observations go beyond the natural human senses such as sight, smell, and hearing. Moreover, they are capable to communicate and process data 24/7 without becoming tired. Smart sensing and monitoring results in a lot of data that is analysed through smart algorithms in software applications. The outcome results for example in an instruction to fertilize crops or to treat animals. This planning is executed by smart machines that know the local context from the sensors and what is the optimal treatment for your crops or animals. The loop is closed by monitoring again the effect of the control actions. In this way you can see if the growth of animals or crops is going into the right direction. This decision-making cycle can be applied to every object and company in the whole value chain: reality and virtual objects have become digital twins!

Digitalization also means that you can pass the data through the whole chain to address questions on food integrity such as: Where does my food come from? How was it produced? How many pesticides were applied? How long did the animals live before they were slaughtered? However, this requires a technical infrastructure of information systems between companies that exchange the data. Here standardization of data comes into play. If two companies use different product names for pesticide containing the same active ingredient, it becomes a problem to exchange the right information. But beside technical issues, there is also the question if companies are willing to share data? And if they are forced to share them, how do you know that they are not tampering with the data? Even if companies are willing to share data, do they want to invest in these smart sensors, software, and machines? Can these costs be passed on to the rest of the value chain? And if you are not able to do these investments, does it mean that you will be excluded from certain market channels?

The data could also be used for monitoring or control to support public policies on for example food safety, environmental policy, or health-related policies. Currently, most governments have their separate monitoring systems or censuses that are usually lagging behind the actual situation. What if you could plug into these real-time data to monitor the nitrogen emissions or detect a food contamination right at the source? That would be fantastic right? But what happens if data are used for other purposes? What if your company is marked - fairly or unfairly - as fraudulent and the authorities close it down? And more positively, are you also rewarded if your data indicate that you are producing in an environmental-friendly way?

All these developments are driven by innovations from science and technology such as artificial intelligence, blockchain technology, internet of things or big data analytics. But at the same time their development depends on the data that is produced through the application of these technologies. You cannot do big data analytics without access to big data! Traditionally, scientific knowledge is produced by analysing data from controlled laboratory- or field experiments. But the applicability outside that context is always limited. *Your* situation is always different and not as controlled as compared to those experiments. Nowadays, modern data science can analyse substantial amounts of real-time data from various sources. The outcomes can potentially be applied in a much more specific context, for specific fields, crops, animals, etc. Again, the question is if you are willing to share your data with scientists for this purpose? And usually, scientists will also need other data from for example your competitors to create benchmarks and trustful algorithms. Are you able to mobilize your sector or branch to organize this in a way that everybody benefits from it and moreover trust it?

In summary, the digital transformation of food systems is taking place through the introduction of all kinds of (interconnected) smart devices and software. This results in a lot of data that has unprecedented potential for smarter:

1 Decision-making in companies or for consumers
2 Food integrity
3 Public decision making by public authorities

These innovations are driven by new technologies originating from science and other domains than agri-food. Their development is increasingly relying on the data that are derived for their application in real-life environments.

However, from the questions raised, it becomes increasingly visible that this is not only a technological change but implies social, cultural, economic and institutional changes. We should not only consider the individual impact of a digital technology on our own situation, but have to put this in a broader perspective of various types of technologies and IT systems, their functionality, the number and role of the involved stakeholders and the general system complexity of the integration between these various aspects. Additionally, there is also critique on the digital transformation process, showing that it is not only beneficial, sustainable and useful to everyone. Hence, digital transformation has entered a twilight zone where innovations have proven to be promising, but must be up scaled to a higher level of adoption and broader integration. A paradigm shift is needed to navigate properly through this twilight zone involving of multiple aspects such as collaboration, trust, inclusion around topics such as data sharing and new business models. Let’s have a closer look at how the twilight zone can be described more precisely and which paradigm shift is needed then.

1.2 The need for a paradigm shift to navigate through the twilight zone

Digitalization in agri-food in fact has already started in the 1980-ies and many applications of digital technology have become common practice. However, in the last decades the nature of digitalization is evolving and has become more complex along two axes (Figure 1). The IT integration level is shifting from stand-alone applications that target single process operators to systems of systems that target complex business ecosystems in which many different stakeholders are involved. At the left side of the digital transformation ladder, the scope of application widens from a single production process, through the farm and supply chain, to food systems and a data economy. The latter two are in the twilight zone because their definition is currently still under debate and therefore unclear what the pros and cons are and how they should be organized. At the right side of the ladder, the scope of IT systems widens from single apps, through farm and chain information systems, to data platforms and data ecosystems. The latter two are also in the twilight zone because there are only a few preliminary examples of data platforms for food systems known and data ecosystems are only just emerging.

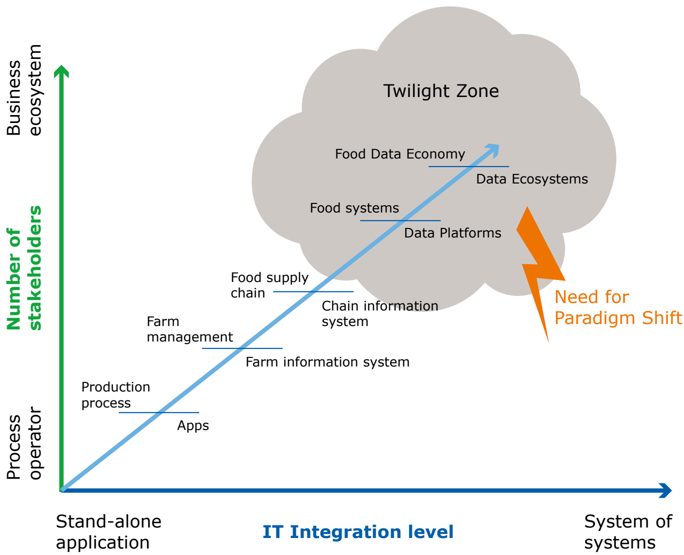


Figure 1 The evolution of the digital transformation in food systems towards a twilight zone where a paradigm shift is needed.

Thus, the early steps in digitalization were focussing on support and automation at process level, shortly followed by the wave of Management Information Systems to support the farmer in managing the farm. A critical success factor was a profound understanding by developers of the users. Strong user involvement in development, test en demonstration of digital innovations is the paradigm here. With the introduction of new technologies such as Internet, cloud and mobile computing, digitalization is now crossing the border of organizations. Data sharing and data exchange across different stakeholders are emerging rapidly. In fact digital transformation is targeting at the optimal use of data available in the whole system. In particular for the food system there are great opportunities to use the vast set of data to combine an efficient and sustainable food production supported by integrated services with well-informed consumers and fact-based policies.

Another characteristic of the new generation information systems is that the user of data is not the same organization as the provider of data. In many cases these are different organizations with different roles and interests. For information systems facilitating different stakeholders, requiring different functionalities, developers are faced with a much more complex challenge. There is the complexity of combining and integrating the different functionalities for a variety of stakeholders and integrating the new application in their existing systems. This is where a new paradigm is needed to illuminate the function of food systems and the data economy and the role of data platforms and data ecosystems. This goes beyond the current paradigm of user-centric software design for single companies or supply chains, but is subjected to a much more complex ecosystem of stakeholders.

In this new paradigm of high stakeholder complexity and highly integrated systems there is a significant risk for investing in digital innovations for food systems. To indicate this risk we mapped the twilight zone on the innovation lifecycle – from invention to market expansion – in relation to the two basic types of funding: public and private funding (Figure 2).

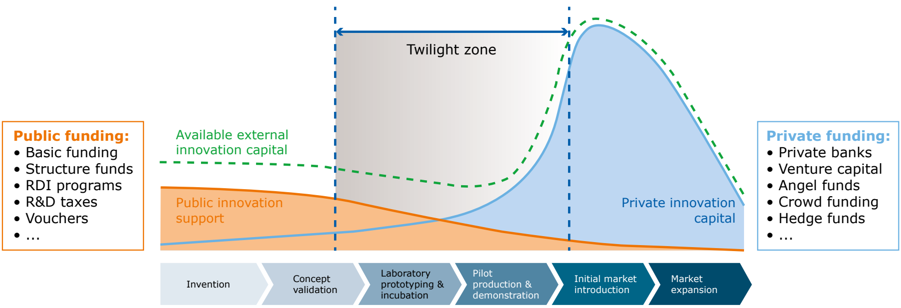
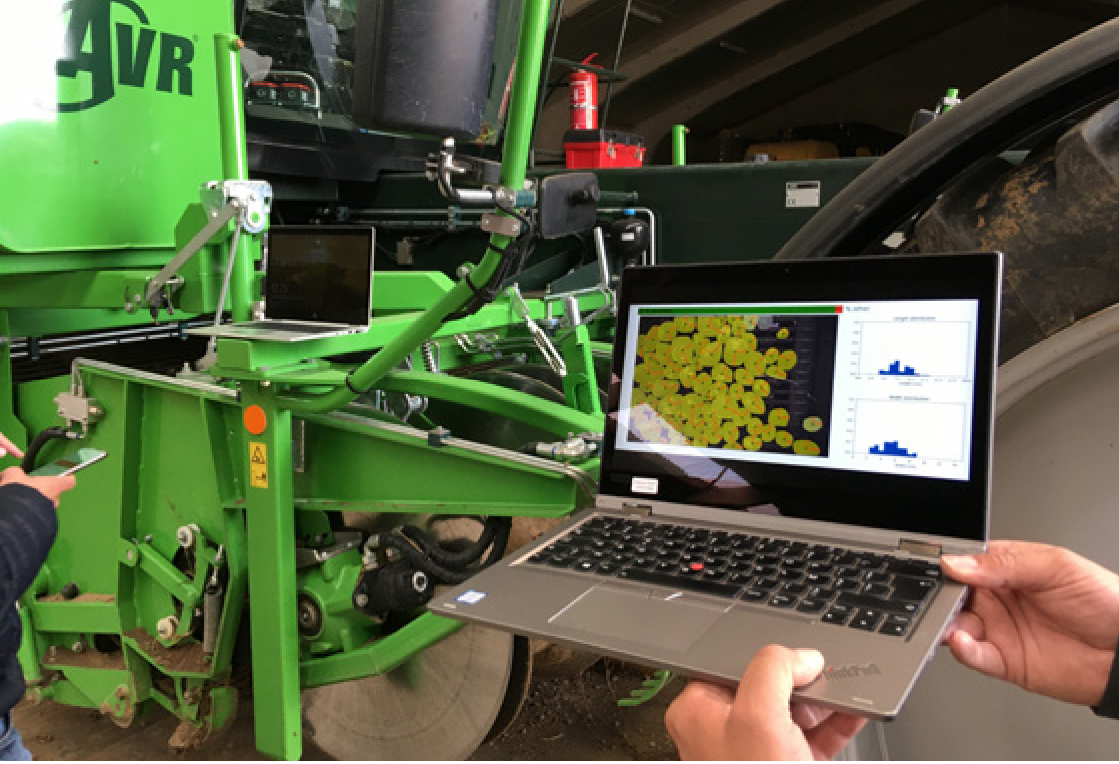


Figure 2 The twilight zone of innovations as mapped on the innovation life cycle from invention to market expansion and the ratio between public and private funding (adapted from TNO)

At the left side, public funding invests in the introduction of novel innovations such as Internet of Things, Blockchain or Artificial Intelligence. Public support decreases when innovations get closer to the market and the risk of inadmissible state support is getting higher. In the twilight zone of digital transformation of food systems, this has become more complex with the high number existing and new stakeholders that are entering the food system.



These can be SMEs, but also big tech companies are looking for chances to apply their technology in the food system. And for the data economy these tech companies are probably also needed.

At the right side, many start-up incubators and accelerators are competing for mostly private funding, currently resulting in a plethora of apps and services, also in the agri-food domain. However, as a farmer cried out: 'we don't need more apps, we need a system!' highlighting the problem of an ever-fragmenting landscape. Referring to Figure 1, this means that they are navigating in the wrong direction (to the lower left) neglecting the need for the paradigm shift. The result is that many promising new technologies are launched but they are hardly adopted at a large scale. Moreover, private investors often like to invest in technology development with a clear prospected return on investment, and tend to pay less attention to research on the ethical, social, and environmental impacts of these technologies on society.

To navigate the twilight zone from a funding perspective, the challenge is to bring together the private and public sector to reap the benefits of both and make optimal use of the totally available innovation capital (see Figure 2): the public sector benefitting from technological advancements in the private sector and the private sector benefitting from the research expertise often found within publicly funded projects. This has been one of the strengths and ambitions of Wageningen over the past number of years, developing dynamic partnerships between technology companies, agricultural organisations, and research institutes, to build a responsive, innovative, and ethical future for all in the agri-food technology sector. This comprehensive approach can be seen in our responsible research and innovation frameworks. This requires a different view on digital innovation representing a holistic approach based on responsible research and innovation.

1.3 Responsible research and innovation looked through five lenses

Responsible Research and Innovation (RRI) is a concept that gained visibility over the last decade. It is predicated on the supposition that technology shapes human (social) lives in profound ways and that therefore it is important to take societal aspects into consideration while the technology is still 'in the making' and it can still be moulded to fit societal values, thus realising technologies whose influence on human (social) life is appreciated and even wanted. To realize technology that is attuned to societal values, RRI fosters an open, inclusive, multi-stakeholder collaboration between researchers and/or innovators and other societal stakeholders, such as citizens, policy makers, businesses, or NGOs. These stakeholders discuss the question of how science and technology should be shaped in the best viable way to not only contribute to solving today's problems, but also create a world that will be desirable and safe for future generations. This means that the focus is not only on achieving outcomes from research and innovation that are socially and economically desired as well as ethically and legally acceptable, but also implies that the process of research and innovation should be inclusive: it should include anticipation of the societal effects and a reflection about their desirability together with societal actors. The result of this anticipation and reflection with societal actors should be shared with the researchers and innovators, to allow them to take it into account during the research and innovation process, including the design process.

With this RRI approach as a basic framework, we will look through five lenses that are important to use while navigating the twilight zone of digital innovation for food systems:

1 **Business models in the data economy**
As data is becoming a key enabler and core of many value creation and distribution processes, economies of scale, scope, and speed are blended into a new generation of business models featured by digital platforms and ‘ecosystem thinking’. New organisational forms (e.g., social enterprises and distributed autonomous organisations (DAO)s) and new currencies (such as digital money, tokens, and cryptocurrencies), challenge traditional business models and governance processes.

2 **Responsible data sharing**
Data sharing is considered to be an important condition for innovation, maintaining a competitive position of the whole food system. However, in practice it is for several reasons difficult to set the rules of the game to support fair sharing of data.

3 **Digital inclusiveness**
Inclusion is a key factor to contemplate in designing digital technologies and infrastructure. Availability, access, and affordability are relatively easy to address, but system complexity and unintended consequences are more difficult to grasp from an inclusion perspective.

4 **Integrative artificial intelligence**
Data science and artificial intelligence are not directly applicable in the domains of sustainability, environment, and agriculture. It requires a change in basic assumptions that involves new techniques, technology, and ways of work.

5 **Cross-sectoral integration**
Agri-food is not a stand-alone system, in fact it is deeply integrated in society and is linked to a lot of other sectors. New opportunities emerge on those topics where data from different sectors and domains can be combined and delivered in new services.

In the next chapter we will take a deep dive into these five topics and describe the implications for navigating the twilight zone. After that, an approach is described in Chapter 3 that addresses the paradigm shift that is needed and integrates the five lenses into an agile navigation pathways to navigate the twilight zone. Finally we will draw the main conclusions and make specific recommendations for various categories of stakeholders.

2 Five lenses to be used for navigating the twilight zone

In this chapter we will look at responsible research and innovation for digital transformation of food systems through five different lenses:

1 Business models in the data economy
2 Responsible data sharing
3 Digital inclusiveness
4 Integrative artificial intelligence
5 Cross-sectoral integration

Each lens will highlight again five sub-topics that we consider most important and provide the implications for navigating the twilight zone.

2.1 Business models in the data economy

Digitalisation has accelerated the production and consumption of data at unprecedented paces. With the explosive growth of data in various forms, data is becoming a key enabler and core of many value creation and business processes. This has given rise to the so-called ‘data economy’ that is still rapidly unfolding but has already shown significant impact on businesses. Why is data economy of interest? How does the data economy differ from previous economies? How are business models being influenced by this development? And how to develop business models in the data economy? This chapter aims to shed light on these questions by addressing the following five sub-topics:

- Data economy and economies of digitalisation
- Impact of data economy on existing business models (the micro questions)
- New values and currencies (the meta questions)
- Ecosystem thinking
- Cultivating data ecosystems

2.1.1 Data economy and economies of digitalisation

The notion of data economy is now trending in policy documents and popular discussions. Data economy is clearly of great interest to both scholars and practitioners. Although the topic is trending, there is not yet a widely accepted definition of data economy, let alone how its impact on businesses can be measured and monitored. Trending concepts tend to diffuse in their meaning. In the case of data economy, this is deeply rooted in different interpretations of the words 'economy' and 'data.' To better understand the implications of data economy for business models, it is important to first unravel these concepts in relation to digitalisation and digital innovation.

An economy can be defined and viewed in many ways. Broadly, an economy refers to a system of organizations and institutions that either facilitate or play a role in the production and distribution of goods and services in a society. This system can have different geographic or thematic scopes, for example, the EU economy, or the 'food economy.' Economy can also refer to careful management of available resources, highlighting concepts such as 'thrift' and 'trade-offs.' Both the systemic and managerial perspectives are conducive to understanding the importance of data economy.

When interpreted as a system, an economy is a system especially of interaction and exchange. This involves the production, distribution, and consumption of goods and services, but also communication and coordination to make the interaction and exchange happen as intended. When interpreted as 'careful management of available resources,' the plural form 'economies' is often used to show advantages in costs or other inputs for the same output. One well-known concept in this regard is the 'economies of scale' – referring to the cost advantages that enterprises obtain due to their scale of operation with cost per unit of output decreasing. In the same vein, the 'economies of scope' refers to advantages gained by producing two or more distinct goods together instead of producing each separately and the 'economies of speed' refer to advantages gained by producing outputs at a higher rate of throughput and through a decrease in the time required. Economies of speed is important both for adjusting existing capabilities (variation) and for developing new products or capabilities (innovation).

Digitalisation is influencing economies in multiple ways. To start with, digitalisation has improved the transmission and reception of audio and visual signals. As a result, communication among people has been much faster and easier as geographical distances no longer form a major obstacle and much information can be recorded and shared for later consultation. This, in combination with improved transport systems, has expanded the boundary of economies and increased the possibilities for optimizing production and consumption processes. Better communication facilitates coordination, although not necessarily improves coordination as the content of communication matters a great deal as well and digitalisation has much less impact on that.

Furthermore, as increased processes in our personal and social lives are being digitized, there is an explosive growth of digital data and information, both in structured and unstructured forms. Structured data can easily translate into the columns and rows of a spreadsheet. Thanks to its well specified format, structured data is highly organized and easily understood by machine language. Unstructured data, on the contrary, usually does not have pre-defined structure to it and comes in all its diversity of forms which makes it difficult to quantify. Examples of unstructured data are YouTube videos, social media interactions, and email messages. With increasing digitalisation, not only the volume, but also the scope of what constitutes data has grown rapidly.

What are then the implications for businesses? Businesses have a long history of using data. In the past, however, companies had no tangible way of analysing unstructured data, so it was discarded while the focus was put on the data that could be easily counted. Nowadays, companies can use artificial intelligence (AI), machine learning opportunities, and advanced analytics to do the tricky unstructured data analysis for them. For example, corporations like Google have made huge advances in image recognition technology by creating AI algorithms that can automatically detect what or who is on a photograph. New businesses have emerged that are specialised in data services (storage, brokerage, and validation) and business intelligence (especially diagnostic and predictive analytics).

Based on these conceptualisation of economy and data, data economy can therefore interpreted in two related but distinct ways: 1) a digital 'economy' that creates and captures the value of data itself, with focus on topics such as 'data ownership', 'data spaces' and 'data market'; 2) a physical economy in which traditional processes are influenced and transformed by better use of data, with focus on topics such as 'big data analytics', 'digital transformation' and 'digital transitions'. In both ways, economies of scale, economies of scope and economies of speed are blended into a new generation of business models featured by digital platforms, ecosystem thinking and new forms of currencies such as cryptographic tokens. The economies of scope and speed are mostly visible in ICT-enabled platforms (e.g., Amazon, AliExpress) that are empowered by modern computing methods and machine-learning algorithms for targeted marketing. The rise of data economy is therefore characterised by the trinity of scale, scope, and speed (Figure 3).

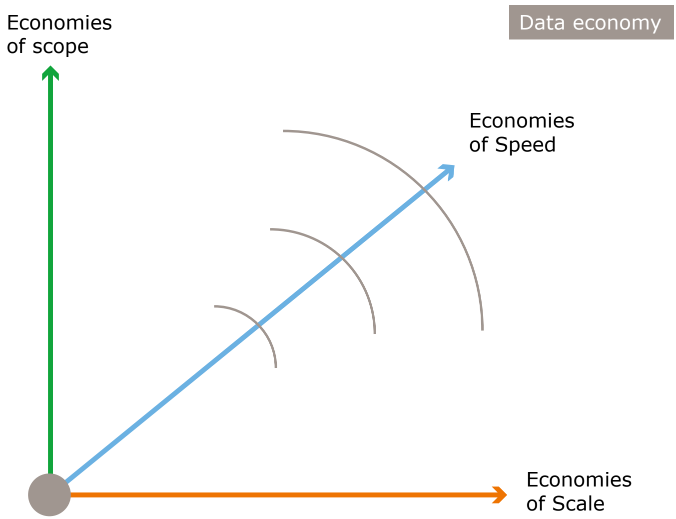


Figure 3 The evolution of business models towards economies of scale, scope, and speed

2.1.2 Digitalisation and business models: the micro questions

Digitalisation has undoubtedly had impacts on various aspects of business models. To understand how specific business models work, it is helpful to address the 'micro-questions' of a business model. These are 'specifications' of the value and revenue streams between the producers and their clients. More specially, the micro-questions are: what product or services is being created, by whom and for whom? What are the inputs and costs? Who is paying for it? How is the payment made? As there are many answers to questions, there is a wide variety of business models, constituting an ever-evolving landscape.

To characterise this evolving landscape requires setting the viewpoints and directions. In view of the central theme of this lecture, two dimensions are relevant to understand the dynamics in the landscape: 1) the level of digitalisation related to the goods and services in the business model; 2) the number of actors or organisations involved in producing the goods and services. This results in a useful framework on which we can plot different business models. Figure 4 shows the most salient features of these business models in four quadrants.

In all four quadrants, new business models are arising, and existing business models are being influenced by digitalisation. Typical considerations are the possibilities to create value (both private and public goods) and economize on inputs (both private and public resources). How can you, as an individual organization, create new revenue streams from data using market intelligence and become part of value networks and business ecosystems? Data can be used for multiple purposes involving multiple stakeholders; multi-sided platforms form the ideal form to facilitate this. Digital currencies are influencing value creation and distribution processes. Data economy is where the economies of scale, scope and speed are blended into a new generation of data-driven business.

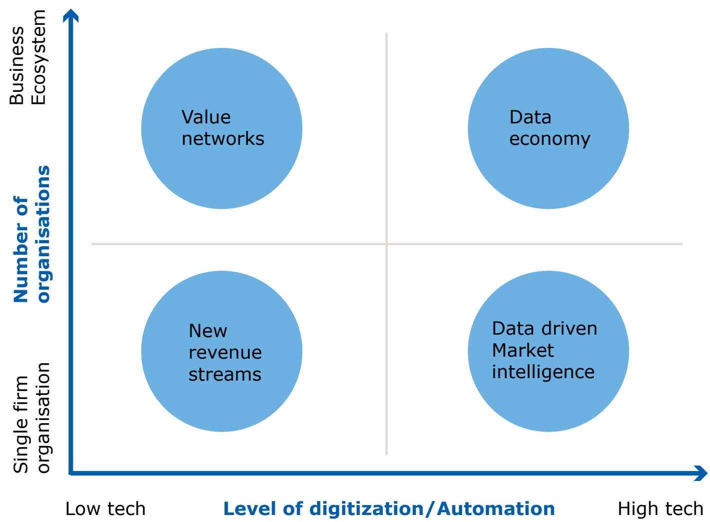


Figure 4 The landscape of business models in relation to digitalisation

2.1.3 The meta questions: values and currencies

The meta questions refer to questions that shape the way micro questions are addressed. For example, how a business chooses the value intends to create, how it defines its client, and how it chooses the distribution channels and ways to capture the values created.

Digitalisation is not a standalone development in society. With increasing environmental and social problems, societies face multiple challenges for which new business models are being developed and organised. These new business models are often characterized by new values, organisational forms, and the use of new currencies in value exchange. Transparency, sustainability, inclusiveness and fairness are new values that businesses seek to create and capture.

Distributing and capturing new values by different stakeholders require the creation of new currencies for tracking and exchanging values. Traditional currencies based on a fiat system have been challenged by a new generation of digital currencies built upon blockchain, more popularly known as cryptocurrencies.

2.1.4 Ecosystem thinking

The term 'ecosystem' was coined in the early 1930s by the British botanist Arthur Roy Clapham at the request of the ecologist Arthur Tansley who used the term to draw attention to the importance of transfers of materials between organisms and their environment (Willis, 1997). In routine use, ecosystem now refers to a complex network or interconnected system. The term ecosystem is now widely used in discussions of software developments, stakeholder management, innovation, and business strategies. As noted by Adner (2017), the rising popularity of 'ecosystem' goes hand in hand with increasing interest and concerns with interdependence across organizations and activities. As a result, ecosystem development has become an important activity in innovative projects where multiple individuals and organisations are involved in diverse ways.

A business ecosystem is a dynamic evolving landscape of stakeholders have multiple interactions with each other. Business ecosystems are evolving as new players take hold in the old field (for example, Amazon and Uber selling food products) and old players taking up new roles (for example, machine producers becoming farming consultant).

The increasing availability of data and digital capabilities post several implications for business models. To start with, governance has become complex in a business ecosystem in which roles and responsibilities and relationships are often not well defined and set in stone.

Business must deal with multiple, sometimes contradicting values (e.g., efficiency vs. resilience). Governance must be arranged at multiple scales: corporate governance, network governance, platform governance, ecosystem governance.

In summary, new digital technologies challenge the boundaries of existing jurisdictions and legal remits (e.g., the rising use of cryptocurrencies and smart contracts). And there is a need to improve data literacy and digital capabilities of the public and farmers who are less favoured/ exploited in the digital economy.

2.1.5 Cultivating data ecosystems

A data ecosystem is a network of databases and stakeholders working together on or being connected by the data they produce or use. A data economy can only exist with a reliable data ecosystem in which different types of data are produced, stored, validated, exchanged and used by different stakeholders.

Although increasingly viewed as a good or commodity, data is not produced and traded in the same way as most physical or conventional products. For example, data, once created, can often be copied or destroyed at neglectable marginal costs – except for very large data sets. Traditionally, the value of economic goods rests on the rivalry and exclusivity of its usage. This is often not the case for data and digital goods. Data market or data-driven markets have therefore distinctive features than traditional markets. Data validation as a service and tokenization (digital money, cryptocurrencies, blockchain) is challenging traditional business models and governance processes.

The way that individuals and organizations have produced, and consumed data has changed with the advent of modern technologies like cloud computing and distributed ledger technology – more popularly known as blockchain. Much of the challenge in understanding and improving the data economy lies in the pluriformity and multiple uses of existing data and co-creation of data that is relevant to the business ecosystem.

Improving data processing and analysing capabilities enable uses of data at various levels of the value ladder and the upward mobility of data (Figure 5). With improved interoperability, data can move from the mere means of recording to being a currency or a means in governance (zero knowledge proof).

Business for impact such as impact investing and other sustainable business models all require new data to be generated that can be aggregated for analytics at the level of supply chain. This creates the challenge of data governance in an ecosystem of stakeholders with different and potentially conflicting interests.

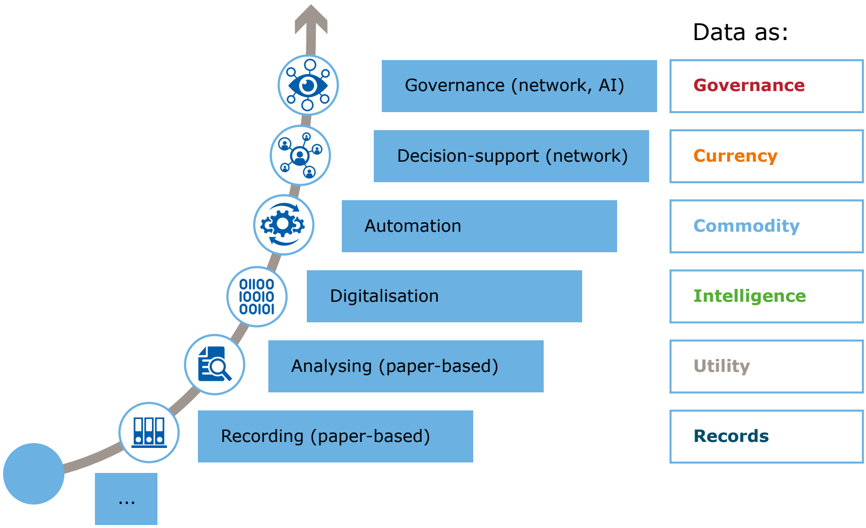


Figure 5 The value ladder of data in different functions and roles

2.1.6 Implications for navigating the twilight zone:

Looking through the lens of business models in de data economy the following implications for navigating the twilight zone can be identified:

- In a data economy, economies of scale, scope and speed need to be considered simultaneously in business models, considering digital possibilities
- Governance has become complex, dealing with multiple contradicting values, multiple scales: corporate governance, network governance, platform governance, ecosystem governance
- New digital technologies challenge or span the boundaries of existing jurisdictions and legal remits (e.g., the rising use of cryptocurrencies and smart contracts).
- There is a need to improve data literacy and digital capabilities of the public and farmers who are less favoured/exploited in the digital economy

2.2 Responsible Data Sharing

Data sharing is an important condition for innovation, maintaining a competitive position of the whole agri-food production complex. However, in practice it is for several reasons incredibly challenging to set the rules of the game to realise trusted data sharing. This chapter further explains these reasons by addressing the following five sub-topics:

- The (un)willingness to share data
- Fair data markets
- Autonomy and control
- Care for the commons
- Trust

2.2.1 The (un)willingness to share data

Europe wants to become a data-driven society for the benefit of its citizens and society (European Commission 2020a). Some kinds of data are protected by regulation, such as personal data which are protected by the General Data Protection Regulation (GDPR), or data whose sharing could jeopardize fair competition are protected by competition law. For non-personal data, often the Free Flow Regulation is fostered (European Commission 2020a).

Digital farming technologies collect data that rarely fall under privacy law and sometimes under competition law. Digital farming technologies collect and process data that usually fall under the Free Flow Regulation, such as, data about chemical components of soil, soil humidity, weather data, emission-data, data about the health and growth of crops or animals, medication data (such as anti-biotics for animals or pesticides for plants), etc. The free flow of these data is fostered in the European Union, as having access to them is thought to empower businesses, research organisations and European, national, and local policy, as they provide better knowledge and enhance innovation activity.

In the past years, however, it has also become abundantly clear that not all stakeholders involved in digital farming networks are eager to share their farm data. The free flow of farm data is therefore far from a reality. Interviews and surveys carried out in various parts of the world, such as Australia, New Zealand, North America, and the EU, have focused on farmers and pointed out that farmers are often unwilling to share their data (Jakku et al. 2019; Wiseman et al. 2019; Regan 2019; Fleming et al. 2018; Carolan 2017; Zhang

et al. 2017). Farmers often distrust the agribusinesses who ask for their data, as they suspect they will re-use these data to build other businesses and services which will benefit them but not the farmer. Furthermore, farmers suspect agribusinesses may sell their data, use them for profiling or as pre-information, prior to their investments on the stock market (Ryan 2020).

In response to farmer's distrust of farm data sharing, different stakeholders have begun shaping guidelines to improve farm data management practices and foster trust in farm data sharing[1]. Principles, codes of conduct and codes of practice provide a valuable start to the discussion about the preconditions for trusted data sharing practices. But these discussions are not completed yet, for distrust in farm data sharing persists.

1 In the US, the American Farm Bureau launched the Privacy and Security Principles for Farm Data in 2014[1]; and in the same year the New Zealand Farm Data Code of Practice [2] was published. In the EU, farmer's representatives from Copa-Cogeca and CEJA (Conseil Européen des Jeunes Agriculteurs) and major agribusinesses presented the EU Code of Conduct for Agricultural Data Sharing by Contractual Agreement in 2018[3]. And in 2020, Australia published the Australian Farm Data Code [4].

2.2.2 Fair data markets

As many stakeholders in the agri-food sector are businesses, they primarily look at farm data spaces as markets. This is what we saw in a lot of our empirical work carried out in our project Internet of Food and Farm 2020 (D7.3; D7.4; Data Sharing action report). Getting farmers to trust others with their data, is perceived by them as an effort to convince them of the added value that this will have for their business.

This *value* is understood in diverse ways. One of the great advantages of sharing data is that it helps to strengthen knowledge and improve farmers' choices. It can help farmers to optimize their products or processes, save time and/or money by making farm processes more effective or reducing the inputs (such as, water, fertiliser, feed for animals, pesticides, etc.) needed to realise a proficient level of production. Eventually this would make farm businesses stronger, as production is enhanced, and the competitive position is strengthened.

Until now, however, there is little evidence available that data sharing will lead to knowledge that brings great business advantages. In the absence of the availability of this evidence, farmers often remain unconvinced that the benefits will outweigh the costs of the investment in digital technologies, or the concerns related to privacy, digital surveillance and profiling. Furthermore, they are often unconvinced that the promised benefits will be worth taking the risk of sharing data with stakeholders that are often distant and faceless newcomers in the sector. Tech companies therefore often suggest that farmers and growers should be informed more about the added value of data sharing, or that incentives should be provided to start sharing more data. In line with the idea that data spaces are like market spaces, many tech companies suggest providing monetary incentives to foster data sharing. According to them, farmers and other stakeholders would not be as hesitant to share data when they would earn money with it, directly or indirectly. This could be done by means of tax incentives or subsidies, or by simply paying money in exchange for data. Other possibilities that respondents mentioned were the provision of cost reductions for specific services such as maintenance of hardware.

This market-line of thinking has however also raised a lot of questions that caused distrust in data sharing. Selling data to a company means a transfer of ownership, but farmers find the negative consequences that this may have sometimes hard to foresee. What will the company who bought their data do with the data? Will these companies develop new products based on the data/ information that is given to them? Will they sell the data to third parties and earn a lot of money with that? Farmers often feel uneasy about these possibilities and want to know where their data will end up. Also, they wonder whether selling data will be a good deal for them, or a better deal for the company buying their data. Weighing the pros and cons of such sales is hard, as the meaning of fairness in the context of selling and buying data is quite unclear. What is a 'fair' deal when it comes to data? What are good arguments in the negotiation when it comes to 'fairness'?

2.2.3 Autonomy and control

In connection to data sharing, many stakeholders talk about the right to decide about data, or to control what happens to them. To offer some clarity about this, the EU Code of Conduct for Agricultural Data Sharing by Contractual Agreement prescribes to shape contracts between stakeholders who start to share data with each other. The tech companies who collect, process, and interpret the data must make explicit in this contract what they are going to do with the data. The farmers or other stakeholders whose data are being collected are provided transparent information and based on that they are enabled to choose whether they want to share data or not.

This contract aims to provide more clarity about what will happen with data and makes the data sharing relationship more transparent. A lot of stakeholders in the data sharing network (farmers and tech companies alike) like this, because they think it will give them more control over data: when they read the terms and conditions of contract formation, they are able to decide whether to share data or not. There is therefore a lot of support for the EU Code of Conduct as it gives greater clarity for all parties involved and their respective responsibilities.

Nevertheless, there is also a lot of doubt about the contracts that are being formed. Signing a contract requires knowledge and expertise about what can be done with data, which not every farmer possesses, or likes to spend time to obtain. Furthermore, as the information provided about what is done with data is often difficult, not every farmer reads the terms and conditions before signing an agreement form for data sharing. This is even more so when the consent procedure is digitalised. Some fear that the farmer will say 'yes' without knowing exactly what he/she is getting into.

Apart from the fact that farmers are often not well informed about the contract they are signing, the contract also simplifies the choice that is being made. As data continually change due to the efforts and technologies of tech providers, it becomes increasingly difficult to discern to whom a certain set of data belongs and who should therefore decide about data, or from what farm data stem and therefore who should be informed about their whereabouts. This fluidity and changing nature of data raises significant challenges to the control of one's data that a contract offers to farmers, as well as the limitations of farmer's right to decide autonomously about 'their' data.

While a contract gives farmers a *sense* of control regarding the question whether to engage in a data sharing relationship or not, it is questionable whether the contract continues to exercise control after it has been signed, or whether something else is needed instead.

2.2.4 Care for the commons

Some suggest that data should be made accessible and usable as a source of knowledge for everyone and benefit society (see Data Governance Act 2020). One way to do this is through a kind of data library that is related to various knowledge centres (van der Burg et al. 2020). As part of a library, data would be considered as public knowledge resources. These data could be used to monitor the effects of food production on the environment, or on the enhancement of the economy, informing public policy, or doing research and enhancing innovation. However, there are concerns that this kind of approach to data sharing could also be harmful for individual farmers. Some farmers are for example afraid that governments will penalize them publicly, when they did something wrong. Or they are concerned that NGO's will publish data to blame farmers, which will have detrimental effects on the reputation of farmers in society. Farmers are afraid people will form opinions based on data, without really understanding the data, as they think one should have farming expertise and know more about the complex contextual story about farms to understand better what the data say.

The care for the commons themes also gives rise to questions regarding who should have access to this data and under what conditions. The question how open data should be is raised by many stakeholders, and for whom and for what purposes. Furthermore, the question is asked whether everyone who provides data to publicly available resources, such as data libraries, should always agree with all the purposes for which data from the library are being used. Questions such as these lead to the further questions, such as:

- Who should oversee publicly available data resources, such as a data library? Who should determine what the preconditions for data sharing and data use should be in a library?
- The degree to which the policy of data libraries ought to be inclusive is a common topic for reflection: should governments be in charge? Tech companies? Or farmers?
- Or should all of them have a role in collaboration?

2.2.5 Trust

All the previously mentioned themes can be considered important constituents of trust. Trust is a theme that is discussed a lot in relation to data sharing. But in these discussions, it is seldomly made clear what is meant with 'trust.' In the development of the various codes of conduct, guidelines and principles developed all over the world to guide data sharing, one approach to trust has been dominant: this is the individualist-contractarian approach to trust. (Coeckelbergh 2012). This is also the approach that underlies the EU Code of Conduct for Agricultural Data Sharing by Contractual Agreement. (See also Van der Burg et al 2020a). This approach to trust considers people first as individuals, who engage in relationships with others only after careful and rational consideration. Formal contracts can be formed to underpin the trust relationships, but that does not have to be the case: contracts can also be a metaphor to describe the rational basis of these trust relationships.

Even if no actual contract is formed, the individualist-contractarian approach takes as a supposition that relationships are always engaged in based on a rational reflection about one's reasons to trust another person, and if there are reasons to retreat from the relationship, this can be done too. There is also the 'social-phenomenological approach' to trust (Coeckelbergh 2012), which starts from the supposition that human stakeholders are always already part of social relationships, before trust arises as a topic for consideration. Trust is part of human relationships from childhood onwards and has an embodied, affective, and social dimension, long before people start reflecting on the reasons, they may have for trusting or distrusting other people. This approach supposes that trust cannot be 'created' after reflecting on the reasons one might have to trust or not to trust another person, but it is primarily a *lived* aspect of human life.

To foster trust in data sharing, we think it is on the one hand important to stick with the individualist-contractarian approach, as in business contracts mark an important beginning to relationships. But contracts will not do enough to bring about trust. After an individual decided to engage in a business data sharing relationship, he or she becomes part of a digital interaction and this interaction should continue to be subjected to norms and values which provide guidance regarding what constitutes acceptable behaviour, but also helps to establish how data sharers can expect to be treated. These norms and values are to help establish and maintain trust in data sharing understood as a social relationship that is shaped and sustained over time. This relationship is mostly a relationship between businesses, which has transformed and became larger and more anonymous when it became digitized.

Next to norms and values (or 'rules of the game') for data markets, we think there is an important trust-building role to play for data libraries. Data libraries serve public goals. As such, data libraries can also help to foster trust in data sharing. Libraries allow stakeholders to engage with data and co-develop governance models for public purposes. Stakeholders could help to shape the data sharing policy of data libraries and establish rules regarding who can use the data, for what (public) purposes, and under what conditions. Next to that, libraries allow to support development of the social-experiential dimension of trust, as libraries can take a role in organising activities which educate stakeholders in the data sharing network and engage them in interactions with data, as well as reflections about what these data can and cannot do for society.

In both business contexts and data libraries, the social dimension of sharing should however receive more attention than it has attracted until now. It should be acknowledged that data sharing is sharing, not simply transferring something from one individual to the next or selling some 'thing.' Even in a market context it makes no sense to speak about data as if they are 'items' that are sold and bought and from which individuals from which these data originate benefit. While we do not deny that there are of course market interests in data, the data economy should be understood as a social interaction. Businesses and organisations who want to benefit from the sharing of data, should therefore spend as much effort to develop the play rules that make the data sharing relationship flourish, as they do to make the technology work.

2.2.6 Implications for navigating the twilight zone

Looking through the lenses of responsible data sharing the following implications for navigating the twilight zone can be identified:

- Code of conducts are a first good attempt, but should be further adapted and expanded to also include guidance for interaction with data in data markets, a new aspect of social life
- Experimental project environments should be created, to stimulate social interactions creating trust
- Trusted environments should be developed to share data 'for the commons,' such as data libraries
- Values and norms that govern interaction with data should be co-created with stakeholders, as this will help protect small players against data misuse and putting them into less powerful positions

2.3 Digital Inclusiveness

A popular assumption is that digitalization is beneficial for everyone, and truly transforms agriculture and increases its sustainability (Klerkx et al., 2019; Basso and Antle, 2020). Other voices state that visions for the role of digital technologies support perpetuation of a status quo that prioritizes maximization of global agricultural production (Lajoie-O'Malley et al., 2020). Regardless of digitalization contributing to sustainable transformation or not, there is a growing realization that digitalization in agriculture may lead to inclusion and exclusion of people in the present or future. Traditionally well-known in- or exclusion factors include being female, disabled, illiterate, indigenous, or (rural) poor. Newer causes for exclusion are for example farm location and size and type of production system. With the increasing importance of (big) data in agriculture, in- and exclusion also become more determined by (access to, ownership of, and power over) data rather than hardware and software. Inclusion and exclusion factors become visible in relation to various aspects of digitalization as will be explained in more detail in this chapter:

- Digital infrastructure – availability and affordability
- Capabilities – awareness, abilities, and agency
- Technological design and power
- System complexity
- Inclusiveness and exclusiveness as a choice?

2.3.1 Digital infrastructure – availability and affordability

Access to digital infrastructure depends on both the availability and affordability of the technologies and their related infrastructures. This involves availability of material (digital hardware, software, and data); infrastructure (required to access and use those hardware, software, and data); institutions (rules and regulations); markets (demand and supply); and a suitable context (i.e., is the digital technology a good and fair fit for the context?).

Affordability relates to economic capacity: capital required to access digital technologies; one off or recurring material investments; and whether the technology delivers profit. Inclusion and exclusion here result from economic inequalities between farmers and farmers and other stakeholders, thus resulting in passive and constitutive exclusion. Affordability challenges may exacerbate with extremely high initial investments, or recurring expenses. Continuous investments become more problematic in case of technological lock-in and path-dependency, tying a farmer to one company or organization due to proprietary software, inability to access farm data without a subscription plan, or inoperability with competitive offers (Bronson, 2018). Additionally, whether investments guarantee profit return or not matters, especially in volatile markets with fluctuating agriculture produce prices (Rotz et al., 2019). As it stands, the two major reasons for exclusion of farmers' access to digital technologies based on their affordability are: by definition (e.g., unable to buy X) and by choice (e.g., unwilling to invest in X).

2.3.2 Capabilities – awareness, abilities, and agency

Another issue for digital agricultural technologies is users' capabilities: are people able to use the technologies and related infrastructure, are they aware of their existence and possibilities, and do they have agency to act upon both their abilities and awareness?

Abilities are about the ease of learning and using a digital technology, and whether farmers can afford investment in additional training and resources (e.g., time, effort, physical strength). Digital literacy is a newer issue relating to skills and knowledge required to use digital technologies, such as using hardware and software, and making sense of data produced or received. In other words, digital technologies need to fit farmers' level of tech savviness to prevent exclusion.

Agency and awareness about the socio-cultural context are fewer tangible issues that are often embedded in the socio-cultural make up of agricultural communities and therefore not directly observable. However, they are critical factors that influence adoption decisions and exclusion, especially in cases of non-adoption or de-adoption, regardless of good availability, affordability, and ability of users.

Reasons for inequalities in access to digital technologies and data are not limited to observable, tangible, or individual factors (like age, gender, and wealth) but also extend to more unobservable, intangible, and aggregated issues which we discuss in the following sections.

2.3.3 Technological design and power

Digital technologies are designed with a specific objective and desired outcomes in mind and often represent the worldview of the designer, thus (re-) distributing power through the design. Design decisions determine the physical, front-end of a digital technology (hardware and software interface) and the underlying system or back-end (programming languages used, location of databases, interoperability with other systems).

Designing a digital technology also requires making decisions about the world that this technology and the data collected by it represents, i.e., whose world is represented, and how is this done? These decisions alter our physical world and how we operate in it, potentially causing unequal opportunities (Cinnamon, 2020). Design decisions are ultimately accompanied by trade-offs: saying ‘yes’ to one design feature usually equals saying ‘no’ to other features. Those trade-offs make exclusion almost inevitable as design-for-all or one-size-fits-all solutions are overly complex and oftentimes simply impossible. An example trade-off is that digitalization leads to (non-deliberate) loss of jobs in traditional manual labour yet may also create demand for skilled employees in newer job-fields like automation.

Designing digital technologies is moreover about distributing power among stakeholders, with some becoming more influential than others. But how are benefits from digital technologies distributed among different stakeholders, such as technology developers, users, data originators, and data owners? Do design choices contribute to reducing inclusion and equal distribution of benefits, or do they create marginalization of individuals or groups? There will be winners and losers; one stakeholder will benefit more from an innovation design than another. Digital agriculture is often associated with high-tech, smart technologies and large-scale, input-intensive farms. Due to those characteristics, wealthier, large-scale, commercial farmers benefit more from digitalization in agriculture (Bronson, 2018). Hence, digitalization may support a limited number of specific agricultural production systems at the expense of others (Bronson and Knezevic, 2016; Klerkx et al., 2019).

Designing a digital technology therefore inevitably is accompanied by risks. These design-related impacts may not always be intended; unintended consequences are likely, which in turn can lead to exclusion or even unfavourable inclusion. In combination with uncertainty about emerging effects of digitalization, various concerns exist about misuse of data and blurring roles and responsibilities in the digital agriculture system. Currently, roles, stakeholders and data-owners are not clearly defined; neither are governance models, establishing who is accountable for what. Design choices should ideally anticipate unintended consequences that could become design-related risks

(Rijswijk et al., 2020) including answering the question 'who is/are responsible for the consequences of digitalization?

2.3.4 System complexity

Digitalizing agriculture is complex and multi-faceted, involving many heterogeneous (non-human) stakeholders that all need to be (digitally) connected. The digital agricultural system is complex in multiple ways: variations in crop production systems and value chains; national and international jurisdictions; the multitude of stakeholders involved; and the ever-growing diversity of digital technologies and technological packages which may or may not be interconnected or interoperable. The complexity and motions of digital systems make prediction and visibility of different forms of inclusion and exclusion challenging.

System complexity also increases uncertainty about issues such as the quality of data and information as input and output of digital systems. A response is more technological integration. Integration offers opportunities for synergies and reduced complexity, yet a lack of integration can become a digital trap (Rijswijk et al., 2020). For example, a user may become stuck with a particular piece of hardware or software that is not interoperable with other items or cannot be updated. Interoperability and coupling of systems are critical. In contrast, too tight coupling of systems leads to vulnerability and potential domino effects, i.e., if one system fails all fail. How do digital traps and domino effects relate to inclusion and exclusion? The first can result in perpetuating inclusion or exclusion: those included remain included, those excluded remain excluded. Instrumental exclusion may be the outcome of the latter because of the causal linkages between systems.

The presence of digital technologies and data-based decision-making inherently affects real-life interactions, such as between people or between people and animals. Traditional human-to-human interactions become moderated or replaced by machines, changing relationships between humans and their natural, technical, and social environments, and allowing for less empathy, trust building, and judgement of intentions and preferences (Scholz et al., 2018). In cultures where human-to-human interaction has important cultural value, trust is important for acceptance of (digital) technologies (Aker et al., 2016). According to Scholz et al. (2018), data can be a disturbing variable and distractor for sharing experiences and knowledge, taking away agency from the human individual.

More concretely, digital systems rely on data input to operate. However, data inconsistency is a known problem, especially with large datasets from heterogeneous sources, needing investment in rigorous efforts to reduce data noise and correct inconsistencies (Philip Chen and Zhang, 2014). Another challenge with data aggregation is the need to consider variances in how data is interpreted. Although mainstreaming interpretations enhances interoperability, it also raises the question of whether 'hybrid' interpretations are trustworthy or provide a new form of interpretative doubt (Mansour et al., 2016), and whether they support or undermine equality. For example, the outcome of interpretational mistakes may be that people are passively included or excluded, which is hard to control for and may have unforeseen consequences.

2.3.5 Inclusiveness and exclusiveness as a choice?

While inclusion is important, also in the technology design and system complexity, to avoid digital divides, it may not always be favourable to those included. It also raises questions of who should be involved in digitalisation processes.

Negative socio-economic impacts of digitalization, e.g., lack of access, problematic technology design or incalculable system complexity, have often been summarized under the umbrella of the so-called 'digital divide.' This divide is often linked to the more tangible in- or exclusion factors like location, age, gender, ethnicity, wealth status, and education level, resulting in the rural digital divides, gender digital divides, etc. Additionally, the increasing importance of data has led to the emergence of a specific new type of divide: the data divide. The data divide refers to asymmetries between the 'data haves' and 'have-nots' (Scholz et al., 2018). This leads to social and economic marginalisation and uneven socio-economic development (Rijswijk et al., 2020; Rotz et al., 2019; Salemink et al., 2017), and increasingly lead to in- and exclusion mechanisms that are intangible in nature (e.g., algorithmic bias, user profiling). Intangible factors, resulting from design choices and system complexity, are powerful determinants of who is included or excluded and whether inclusion and exclusion is beneficial or harmful due to e.g., expanding access to data, aggregation of data, and capacity for data computation and manipulation. Thus, inclusion is a crucial factor to contemplate in designing digital technologies and infrastructure.

However, to date, design choices and system complexity are rarely considered and as a result policy making lacks behind in terms of offering solutions to the emerging challenges. Although designers and implementers of digital technologies may (and have a responsibility to) anticipate many unintended consequences, some fall into the category of unknown consequences and simply cannot be predicted beforehand. However, designers cannot account for all unintended consequences, especially when they require transformations beyond the technological design such as in the institutional or socio-cultural environment. The latter is an important responsibility for policymakers.

Additionally, digital inclusion is not always favourable. Not everyone always has to be (directly) included and the right to be excluded by choice should exist. This requires thinking about options to opt out. Too often lock-in is inevitable. Furthermore, the perception that technology and technological progress are inherently good and needed for growth is fundamentally flawed when it comes to digital technologies. The trade-offs and unintended consequences that come with digitalization and datafication should receive more recognition and consideration.

Stakeholders across the agricultural sector need to redevelop their identity and build new capacity and expertise. This includes different skillsets and expertise of classical stakeholders in the agricultural sector working on e.g., policy making, crop production, advisory services, or value-chain development, as well as involving different types of stakeholders and expertise (e.g., designers, operators, and regulators of digital platforms and systems).

2.3.6 Implications for navigating the twilight zone

Looking through the lens of digital inclusiveness the following implications for navigating the twilight zone can be identified:

- Digital inclusion is a crucial factor in designing digital technologies and the technical and organizational infrastructure.
- The trade-offs and unintended consequences that come with digitalization and datafication should receive more (upfront) anticipation, recognition, and consideration.
- When designing digital solutions attention should be paid to the right to be excluded, not only by the designers but also by policy makers.
- Digitalization of agriculture requires new skillsets and expertise of classical stakeholders and involvement of new types of stakeholders.
- Continuous monitoring and evaluation of digital systems and their design is necessary to observe how in- and exclusion unfold to act upon it.

2.4 Integrative Artificial Intelligence

Today, big data is ubiquitous, machine learning applications are thriving, artificial intelligence (AI) appears in everyday conversations, and Internet of Things is present in more appliances, machinery, and technology. Agri-businesses and organizations are increasingly employing cloud computing, and high-performance facilities are progressively accessible as a service. Opportunities and benefits are becoming omnipresent, as operations can be organized more effectively with more computational power, and huge amounts of data can be analysed using tailored machine learning algorithms to discover new insightful patterns and better decisions. However, translating generic-purpose big data and artificial intelligence technology into meaningful applications in agri-food still requires further development. Recently, there have been several success stories that transferred mature technological advancements from other domains to agriculture. For example, computer vision has been successfully applied in a variety of domains before starting being picked up in agricultural applications, from plant phenotyping to robotic harvesters. Corresponding innovations have emerged from the twilight zone and reached market expansion, changing the landscape of precision agriculture and livestock farming, among others.

However, this is not always the case. A use-inspired, integrative approach is needed to extend and adapt new big data technologies and make them useful in food systems. It requires a paradigm shift to harness the power of digital technologies and enable data-intensive scientific discoveries in environmental and agri-food systems. This emerging science era, named *the fourth paradigm for research* by Turing Award winner Jim Gray is founded on a *new paradigm, beyond experimental and theoretical research and computer simulations of natural phenomena—one that requires new tools, techniques, and ways of working*. Data-driven discoveries will be enabled by data-intensive research collaborations that go beyond disciplinary boundaries, and involve big data infrastructures, new sensing technologies, data science, machine learning and artificial intelligence. Recently, a Nature Food editorial pointed out that we can do more with what is already available: computational capacity, experience and big data are already there and open opportunities for tackling food security problems more accurately and inclusively. The question is how? Below we outline five topics for further accelerating integrative artificial intelligence in agriculture for further accelerating data-driven discoveries:

- The big data gap
- Privacy-preserving AI
- Knowledge re-use
- Transparency and explainability
- Iterative and continuous learning

2.4.1 The big data gap

As high-fidelity sensory devices are becoming more accessible, data records are accumulated and open opportunities for tackling food security problems more accurately and inclusively. However this process is not always done in a cooperative manner. Unimportant irregularities may become practical obstacles that stand in the way of theoretical success. Despite the increasing volume and variety of sources of data becoming available, more data are not necessarily useful for improving our understanding of the underlying phenomena, to improve decision-making on the farm or support new policy choices. Data in our domains are too coarse, incomplete, or noisy, with high spatial and temporal variability, as they inherit the measurement instruments' failures, biases, and noise. But more importantly, they suffer from the fragmentation of case studies, lack of standardization, and data sovereignty issues, including ownership and licensing, that impair our capacity to benefit from artificial intelligence and big data advancements.

There is an avalanche of data that became available in the past decades: on the one side recent advancements in satellite imagery offer unprecedent opportunities for monitoring earth processes in higher-than ever resolutions, able to penetrate clouds and even report topsoil conditions. Several satellite data products are openly accessible through open data infrastructures. Similarly, as unmanned aerial vehicles become more affordable, aerial data collection has also gained its place as a data collection instrument in agriculture. On the other side, advances in DNA sequencing devices have allowed to collect massive datasets, including full genomes and population-wide genotype collections of several species. An explosion of biological data is also generated by high-throughput screening technologies that routinely measure genome-wide differences at the biochemical level (van Dijk, er al 2021) yielding massive "-omics" data sets, that are also commonly shared.

Agri-food is flooded with data at the two ends of the scales (i.e., genomic and remote sensing data), but few data are systematically available in the middle: at a farm/field level. And those available, typically, are incomplete, noisy, in data silos and not properly linked to each other. Such data are extremely important for our understanding of food systems, as they document the management aspects across the Genetics-Management-Environment (GxMxE) continuum. This is a major drawback for composing a complete view of GxMxE interactions, hindering the development of next generation digital twins, able to allow for holistic, integrative views across scales. We need to further advance our field in data standardization curation and sharing. The FAIR (Findable, Accessible, Interoperable, Reusable) data principles offer a framework for sharing data (Wilkinson et al, 2016), but still, we miss good practices and standards for annotating and sharing field- and farm- level information that may serve as reference (ground truth) for future machine learning applications.

2.4.2 Privacy-preserving AI

Data privacy has been identified as a major issue for making Europe fit for the digital age. The General Data Protection Regulation (GDPR) has changed the landscape in Europe and internationally, and “privacy by design” is a key component for its successful implementation. Data privacy and industrial secrecy are major reasons for firms and governments not exchanging information, to the extend it may hinder optimal decision making. Across food value chains data are massively collected about several processes, and include sensitive information about individuals, both producers and consumers, who use appliances, services, or even simply consume food. Geo-privacy is also of particular concern, as farmers may be individually identified, for example through the location of their tractors.

A new generation of privacy-preserving, data-driven technologies can enable better decision-making across food value chains, without exposing an individual’s sensitive information. From cloud data infrastructures and blockchain, to federated machine learning there is a challenging research agenda lying ahead that can allow sharing data, while still accounting for privacy. Cryptography research techniques as homomorphic encryption allow performing computations, and eventually learning, from encrypted data without requiring access to the original. Such advancements, along with federated machine learning, a form of collaborated learning where AI learns from decentralized data, offer new avenues for learning from sensitive data, including socio-economic and geo-referenced information sources, and address emerging ethical issues about data privacy and sovereignty.

2.4.3 Knowledge re-use

A lot of the knowledge about agri-food systems lies with experts, farmers, extension officers, reports, scientific models, and several other tacit forms of knowledge. Agri-food systems require novel methods for learning from data, while reusing the knowledge that is already available: from farmer experience and good practices; to extension officer handbooks and crop calendars; to complex process-based simulation or statistical models. A major challenge ahead is how to employ artificial intelligence for learning seamlessly from both mathematical models, historical observations, and expert knowledge, encoding prior knowledge into reusable AI models, with applications across various scales in food systems.

Most of the recent advances in agri-food digitalization still suffer from the compartmentalization of the agricultural domain itself, as manifested from datasets, models, and practice of the several (sub-)disciplines involved in agri-food. It hinders the widespread exploration of various big data and artificial intelligence advancements and prohibits the knowledge transfer across applications with similar requirements. Several factors of variation are typically hardwired into location-, time- or firm-specific settings that hinder systems' capacity for generalization and reuse. A major challenge ahead is to make headway with hybrid intelligent systems that can incorporate extensive knowledge about the socio-environmental factors that drive agri-food systems, but also take advantage of massive datasets available for data-driven research.

2.4.4 Transparency and explainability

While artificial intelligence is expected to play a more significant role than it has done so far, there are rising concerns about the lack of transparency in these technologies (cf. Villani report). Many of the recent advances in artificial intelligence are not able to explain how they conclude to certain decisions. While for some applications such explanations may not be necessarily useful, they are essential in food value chain applications for users to understand, trust and eventually adopt smart solutions. Moving forward from black-box automata, we need intelligible systems, able to offer meaningful explanations of how decisions about certain individuals are taken, in compliance with the GDPR.

Going a step ahead, explainable, and transparent artificial intelligence in food systems need not only embrace the fragmented, multi-modal knowledge of the domain, but also be able to offer new insights about food systems, on how to respectfully feed the world using one planet. This brings forward explainability aspects that go beyond justification and control, but requires smart systems able to discover and encode new knowledge. In this respect, AI technologies need to be explainable, auditable, inclusive, and reusable, if we are to entrust them with co-shaping our future. In the food systems, this task is more challenging as we deal with living systems, tipping points and cross-scale effects, but also with a wide variety of stakeholders.

2.4.5 An "integration first" approach to AI

While data, computational capacity and new artificial intelligence methods are available for driving the next wave of business opportunities and tackling food security problems more accurately and inclusively, we still lag in fit-for-purpose methods, and good practices on how to enable new data-intensive scientific discoveries. An iterative, interdisciplinary approach is needed to co-develop the next generation of digital solutions in agriculture.

An "integration-first" approach is required where various knowledge assets are brought together, and new artificial intelligence models are developed and evaluated in a synthesis space for knowledge integration and reuse across disciplinary silos. To this end, co-shaping the right problems, and sharing benchmarking problems is key for future progress. For example, it took several years for the plant phenotyping community to start benefiting from advances in machine vision, as appropriate data formulations and data were lacking (Tsaftaris and Sharr 2019). Once the right data along with a challenging problem were available, they attracted the interest of both plant phenotyping and machine vision communities, leading to an interdisciplinary symbiosis that accelerated performance, increased awareness, and provided the foundations for new data-driven discoveries.

2.4.6 Implications for navigating the twilight zone

Looking through the lens of integrative artificial intelligence the following implications for navigating the twilight zone can be identified:

- Narrowing the agri-food big data gap, i.e. ground truth data at farm- and field- level, is key for accelerating future developments towards agricultural AI.
- A new generation of privacy-preserving, data-driven technologies can enable better decision-making across food value chains, without exposing an individual's sensitive information
- Agri-food AI needs to build on top of knowledge that is already available in the sector: from farmers practices to science
- Explainable agri-food AI is needed if we are to entrust them with co-shaping our food future.
- Artificial intelligence may assist in addressing food security challenges and shape the future of farming, but requires new ways of working together, and shifting the research agenda from "technology push" to "integration first."

2.5 Cross-Sectoral Integration

Food systems are not stand-alone systems. In fact, it is deeply integrated in society and is linked to a lot of other sectors. Most of the digital technologies have their origin in other sectors (e.g., military, manufacturing, automotive, medical) and are adjusted and modified to be used in the agri-food sector. Other sectors like banking, governments, insurance, transport, and retail are pushing digitalization in agri-food by imposing digital transaction systems.

When navigating in the twilight zone we see new opportunities emerging on those topics where data from different sectors and domains can be combined and delivered in new services. These new services will also enable digitalization in agri-food. This section described five more specific topics with opportunities that cross-sectoral integrated digitalization can offer:

- Rural development
- Food and Health
- Circularity and logistics
- Water management
- Citizen dialogue platforms

2.5.1 Rural development

Most of the agricultural activities and forestry are by nature land-bound to rural areas. One of the factors hampering a decent IT infrastructure (e.g., broadband, mobile coverage) is the business model for the internet providers. When more companies and citizens are using the internet, the rural area will be more attractive for internet providers.

A lively rural area is an important policy objective where rural development is targeting on economic development and employment with a challenge to make the rural area attractive for other economic activities. IT infrastructure will not only serve farmers to improve their business, but also enables other people to develop online services and generate other business. These will not be limited to rural activities like ecosystem services and rural tourism. A high quality and reliable IT infrastructure will be a key facilitator for attracting footloose companies to establish their business in the rural area. A crucial condition for making this happen is related to livelihood of the rural areas. For entrepreneurs, workers, and their family, it is important that e-supported basic services are available. These services are required to make the rural area an attractive environment for families to live in. For this education, health care, shops and mobility are important to be accessible by e.g., distant education, distant health care, internet shopping, etc. For users, business, or citizens, it is important to have access to an integrated IT infrastructure that facilitates services required in the rural and remote areas.

2.5.2 Food and Health

There is a growing awareness that healthy food and nutrition is a basis for personal health. As food diet is a key factor in human health, food becomes increasingly integrated with health care. At the preventive side, consumers can use data for buying and consuming food that keeps them healthy. Well-informed consumers can avoid unhealthy habits. This also holds for the curative side where a growing number of people are on specific diets and must avoid specific ingredients e.g., because of allergies. For those consumers, the information on food can be critical to life saving. Other diets are targeting on providing on extra nutrients, also for those diets it is important to have clear and reliable data on food.

Cooperation between health advisors and food producers will generate smart devices and enable two-way communication between consumers and food producers will open new opportunities for digital services for healthy living.

2.5.3 Circularity and logistics

Usually, logistics require data on operational level, related to shipping. However, more strategic exchange of data between logistics and agricultural processes with e.g., prediction and steering of crops create opportunities for savings on transport kilometres. By preventing half loaded trucks and empty return freights significant reduction of carbon footprints can be realized.

The agri-food sector produces and uses waste products, waste from one company can be input for another which is not limited to agri-food companies, but also other sectors can produce waste that can be used in agri-food and the other way round (e.g., fibres for clothes). The latest information-based services will arise from matching the supply and demand of waste products. Waste products however must be transported and are characterized by high volumes and low prices; thus logistics will be a crucial factor in circularity of waste and needs to be very efficient. Matching services will link the agri-food sector companies with logistics and waste producers and users in other sectors.

2.5.4 Water management

In most regions agriculture is a dominant consumer of water. In increasingly number of regions, the availability of water becomes a challenge to meet the demands. Water management needs to balance the interest of different sectors, households, tourism, agriculture, and industry. Data on when and how much water is needed for each of those sectors would support the optimization of the water supply. From agriculture side, crop prediction models and weather forecasting models will provide important information for the agricultural demand for water. Integration of data and collaboration between e.g. farmers and water boards could significantly enhance water management at local, regional and national levels.

2.5.5 Citizen dialogue platforms

Agri-food is quite often subject in societal debates, e.g., on environmental impact, animal welfare and biodiversity. For the quality of the debate, it is crucial that all participants are well informed. Agri-food needs to be transparent and provide facts about the sector and the behaviour of its companies. Integration with citizens-oriented platforms as appearing in the smart cities community can be supportive in informing citizens about the origins of their food.

2.5.6 Implications for navigating the twilight zone:

Looking through the lens of cross-sectoral integration the following implications for navigating the twilight zone can be identified:

- IT infrastructure in rural areas should be improved for integrated use by agri-food business and other rural inhabitants.
- Development of cross-domain applications and services should be stimulated in projects and programmes; Digital Innovation Hubs, and Digital Earth Twins could play a crucial role in this.
- More advanced methods (e.g. on data anonymization) are needed to support opening reliable public data to facilitate the public debate protecting the individual data suppliers.
- Active involvement of agri-food sector in cross sectoral standardization.

So far, we have looked through five different lenses that are needed to navigate the twilight zone where innovations have proven to be promising but must be up scaled to a higher level of adoption. The next chapter will integrate this into an approach to navigate the twilight zone.

3 A multidisciplinary, collaborative, agile approach for navigating the twilight zone

3.1 Requirements for a multidisciplinary, collaborative, agile approach

Considering the implications of the five lenses to navigate the twilight zone of the previous chapter it can be concluded that it is needed to design digital innovation programmes for food systems that address the following topics:

- **Experimental real-life environments** in which all relevant stakeholders are collaborating on developing digital solutions, continuously interacting on technical as well as organizational aspects. It should provide a safe and trusted environment in which stakeholders can share data and experiences.
- **Agile, iterative design** in which fit-for-purpose and user acceptance are leading and that can early detect risks and unintended consequences
- **Multidisciplinary support and interaction** that concurrently deals with:
  - business modelling embedding the digital solutions in the data economy
  - data science and AI enabling evaluation of policy options and management actions
  - governance looking at multiple scales: from corporate to ecosystem level
  - ethics co-creating values and norms to protect less powerful actors and prevent data misuse
- **Ecosystem development** to upscale digital solutions stimulating cross-domain applications and monitor projects for in- or exclusion of stakeholders

3.2 A multidisciplinary, collaborative, agile project approach

From these requirements, a responsible research and innovation project approach can be derived as visualized in Figure 6. The approach starts by defining use case projects in which multiple stakeholders are going to develop a certain digital solution. A use case means that you already start to use the solution in the project. It is tested in a real-life environment in which user involvement is a key success factor. Development of the solution consists of an iterative cycle of four steps:

1. Design, based on a set of clear objectives or challenges (e.g. increase yield, reduce pesticide use or better transparency for consumers)
2. Implementation and Integration, building the solution in the real-life environment
3. Testing and Demonstration, to see if it meets the objectives and openly communicate about this
4. Evaluation, considering performance, fit-for-purpose and the extent to which the objectives were met

The spiral in this picture indicates that development goes through the 4-step cycle, but each time trying to end at a next level. This next level is determined by so-called minimum viable products (MVP). A minimum viable product is a version of a product, (or service,) with just enough features that can be evaluated by the users. So, an MVP is more than a technical prototype to see if it works. Features should also include aspects of practical use, costs, and benefits.

Dependent on the outcome of the evaluation phase, the design is adapted, and the process goes through the cycle again. It can mean that objectives are changed because of the insights gained. In some cases the expectations must be lowered but it can also happen that they were not ambitious enough. In the worst case, you must start with a complete redesign or you conclude that a certain solution simply does not work. In any case, lessons learned will always be valuable and need to be shared in the ecosystem.

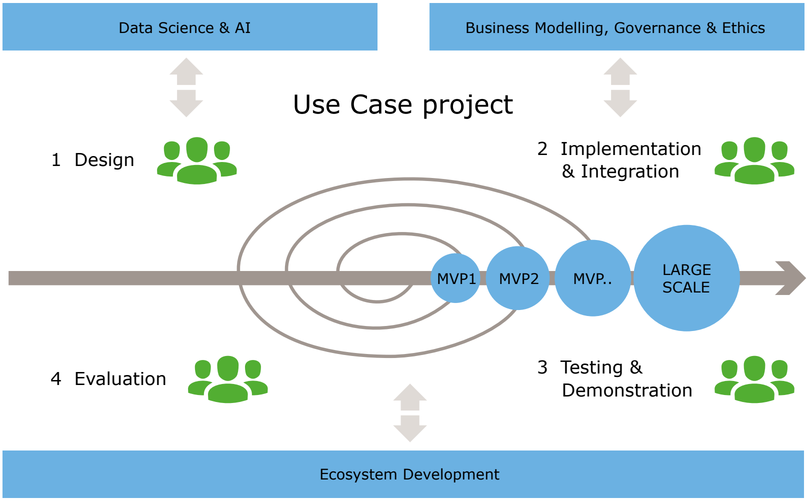


Figure 6 A multidisciplinary, collaborative, agile approach for digital innovation

Usability and user acceptance remain the key principles to guide this development. Therefore, it is important that all relevant stakeholders are involved throughout this whole development cycle. Stakeholders can be for example technology providers, farmers, logistic providers, consultants, or researchers. The goal is to reach a large-scale implementation and adoption of the innovation. Depending on the specific context and scope of an innovation, a use case project can last from several months up to several years. Also, the period of the development cycles can vary from weeks to months or years.

A use case project usually takes place at a small scale: a few organizations and persons form the core. This is inherent to the character of a use case in which you want to create a safe and trustful environment. The involved stakeholders want to learn by making mistakes without being watched by the too many people. However, the risk is that a use case is too isolated and lacks input from state-of-the-art knowledge. It is also important that digital technologies build on existing standards to be able to scale-up afterwards. Besides, there can be external factors, such as laws and regulations that determine the success of the innovation. Therefore, it is important that a use case is supported from three different groups of disciplines, which are based on the five lenses of Chapter 2:

- Data science and AI
- Business Modelling, Governance and Ethics
- Ecosystem Development.

The first discipline, Data Science and AI, helps to analyse data according to the latest techniques and scientific insights. Information management helps to organize data exchange and integration between various devices and systems that are involved. It is important to think about the information architecture and data standards to be used. This is not only a technical thing – it interacts with business processes and users of the various organizations involved.

The second discipline is Business Modelling, Governance & Ethics could be approached as separate disciplines, but they are usually very much intertwined. Business modelling helps to define what is the added value of the digital solutions at stake. What are the costs and benefits? How are they shared between the involved stakeholders? Through the development cycle it can appear that many data get into the hands of a particular stakeholder. How do you want to deal with that? That is where the governance and ethics come in. What are the underlying values that determine who you want to do business with and what kind of agreements do you make about sharing data?

The third discipline Ecosystem Development starts at the preparation phase of the use case project. With whom are you going to work on developing the digital solution? During the development cycle it can be necessary to involve other players. And if you want to scale-up the innovation you must find new investors and the right communication and dissemination channels.

Although these three groups of disciplines are presented here in separate blocks, there is much interaction and overlap between them. For example, the choice of a particular business model will influence the way you build your information architecture. It could be incredibly open or closed. And for example, ethical choices can lead to inclusion or exclusion of certain types of organizations and thus influences the way the ecosystem develops.

The involvement of these different disciplines can be organized in many ways. If you have an individual use case project you can include researchers or consultants that bring in the knowledge that is needed. Another possibility is that a use case is embedded in a larger project with multiple use cases. In that situation, dedicated expert teams are supporting the use cases from these different disciplines then. Like that, an extra advantage is that you can also learn from other use cases. This can also be further extended by embedding projects in larger programmes in collaboration with other domains such as health, logistics or manufacturing.

3.3 Connecting the dots by Digital Innovation Hubs

The approach as described in the previous section has basically already been used in the IoF2020 project by 33 use case projects[2]. In this way a large, coherent ecosystem and collaboration space was formed that is expected to sustain after a project. In several cases, use case projects successfully navigated through the twilight zone and digital solutions are now introduced into the market and adopted at a larger scale. However, still other promising innovations need more time to take these steps. And although use case projects were organized at a local, regional level they were dependent on the multidisciplinary support from the central project, also for the funding. To breakthrough this deadlock the SmartAgriHubs project was launched that is based on the concept of Digital Innovation Hubs. Another relevant project for food systems in relation to this is the agROBOfood project[3] that specifically focusses on the application of robotics in agri-food.

2 https://www.iof2020.eu/use-case-catalogue
3 https://agrobofood.eu/

A Digital Innovation Hub (DIH)[4] refers to an ecosystem through which any business can get access to the latest knowledge, expertise and technology to test and experiment with digital technology relevant to its products, processes, or business models. The DIH also provides the connections with investors, facilitates access to financing and helps to connect users and suppliers of digital solutions across the value chain (Figure 7).

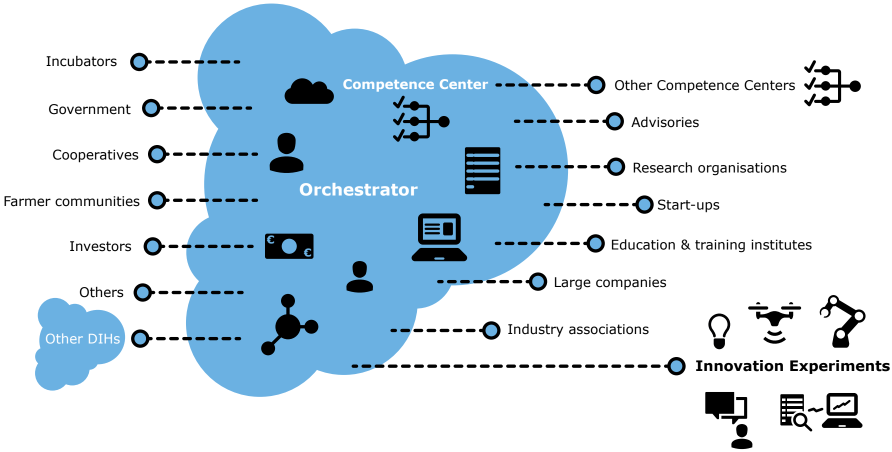


Figure 7 The Digital Innovation Hub as an orchestrator to connect various stakeholders creating and supporting innovation experiments

Such an ecosystem will accelerate digital innovation because it makes the connection between technology, business, and the market. A DIH offers all required innovation services acquiring full representation of the local ecosystem.

In SmartAgriHubs the focus is on agricultural DIHs, although DIHs can also target multiple industries or sectors. DIHs for agriculture are different from other DIHs, because activities are usually land-bound and more specific for a local region while they should provide a local one-stop shop in the proximity of their stakeholders (i.e. farmers, technology providers, etc.). The main aim of the SmartAgriHubs project is to consolidate and foster a European-wide network of Digital Innovation Hubs for Agriculture, to enhance the Digital Transformation for Sustainable Farming and Food Production. This overall aim

4 https://ec.europa.eu/futurium/en/system/files/ged/dei_working_group1_report_june2017_0.pdf

will be achieved by accomplishing the following specific objectives:

1 Build a network of Digital Innovation Hubs, covering all regions in Europe and ensuring a broad coverage in terms of relevant players and technological, business and sector expertise.
2 Support a critical mass of dedicated pan-European “Innovation Experiments” that bring together the farming sector and technology suppliers. Innovation experiments are the same as use case project in the approach that is described in Section 3.2.
3 Provide structural financial support to third parties through open calls supported by European and regional public and private funds.
4 Ensure the long-term sustainability of the network, including a business plan for the DIHs, to attract investors and address the needs of the agri-food sector.
5 Enable and promote the expansion of the DIHs by including new DIHs in the network and through capacity building measures, ensuring that DIHs reach their full innovation-accelerating potential.

Beside DIHs, a network of Competence Centers (CCs) that provide the multi-disciplinary support and interaction is build-up that can be used by DIHs and included in innovation experiments. In this way, it is foreseen that the drivers for digital innovation and the support from and interaction with multiple disciplines are organized in a structural manner, embedded in a local-specific context. At the same time, an overarching network - as formed by projects such as IoF2020 – provides access to state-of-the-art knowledge and technology.

The SmartAgriHubs network is currently running in Europe and a vast network of Digital Innovation Hubs and Competence Centers that are coming together in the Innovation Portal[5]. However, the concept and approach as described in this document can be easily extended to other parts of the world, including developing countries. Starting from a local-specific context of e.g., smallholder farms new or existing innovation experiments can be conducted supported by local DIHs and CCs. By participating in the larger network of SmartAgriHubs it is easier to get access to state-of-the-art knowledge and technology. It can be expected that it also becomes attractive for funding bodies to co-invest in these innovations because it is embedded in a larger, sustainable network.

5 https://smartagrihubs.eu/portal/home

4 Conclusions and recommendations

4.1 Conclusions

The digital transformation of food systems is happening and will be pushed further by novel technologies such as Artificial Intelligence. However, the success and adoption rate and of digital technologies in agri-food is lagging behind in comparison to other sectors and domains. Digital transformation has entered a twilight zone where innovations have proven to be promising, but must be up scaled to a higher level of adoption and broader integration. There are unexpected and unintended effects that emerge during the innovation process which challenge issues such as business models, trust relationships and inclusiveness. The nature of digitalization is evolving and has become more complex. The IT integration level is shifting from stand-alone applications that target single process operators to systems of systems that target complex business ecosystems in which many different stakeholders are involved. From a funding perspective, the challenge to navigate the twilight zone is to bring together the private and public sector to reap the benefits of both and make optimal use of the totally available innovation capital: the public sector benefitting from technological advancements in the private sector and the private sector benefitting from the research expertise often found within publicly funded projects.

For further progress in the digital transformation of the agri-food sector a paradigm shift is needed to navigate properly through this twilight zone involving multiple aspects such as collaboration, trust, inclusion around topics such as data sharing and new business models. A Responsible Research and Innovation approach, implemented in practice, can help to design better and more accepted digital solutions and consequently navigate successfully through the twilight zone and improve uptake. More specifically we provided five lenses in Chapter 2 to look through, and that can help to navigate digital innovations through the twilight zone:

1 Business models in the data economy
2 Responsible data sharing
3 Digital inclusiveness
4 Integrative artificial intelligence
5 Cross-sectoral integration

The new paradigm to navigate the twilight zone was practically translated into a multidisciplinary, collaborative, agile approach in Chapter 3 that is based on:

- Experimental real-life environments - in which all relevant stakeholders are collaborating on developing digital solutions, continuously interacting on technical as well as organizational aspects.
- Agile, iterative design - in which fit-for-purpose and user acceptance are leading and that can early detect risks and unintended consequences
- Multidisciplinary support and interaction - that concurrently deals with business modelling, data science and AI, governance at multiple scales, and ethics co-creating values and norms
- Ecosystem development - to upscale digital solutions stimulating cross-domain applications

By embedding this approach in large projects such as IoF2020, agROBOfood and SmartAgriHubs, sustainable innovation ecosystems are created that will be better enabled to navigate through the twilight zone towards sustainable food systems.

Against this background, we provide a number of specific recommendations in the following section that can help to turn the required paradigm shift into actions.

4.2 Recommendations for a responsible and sustainable digital transformation of food systems

In the following subsections, we will provide specific recommendations to navigate the twilight zone. The categories of topics is based on a final report of the IoF2020 project and some of the recommendations overlap with that[6].

6 https://www.iof2020.eu/deliverables/iof2020-policy-recommendations.pdf

4.2.1 Sustainability recommendations:

1 Sustainability should become integral part of the business model and made explicitly in the objectives of digital innovations. Policymakers need to provide incentives to implement technologies with sustainability benefits. Subsidising the investments or paying for the reduced pollution are options.
2 There is often a high focus on the sustainability of the on-farm production, but the entire life cycle of food products and strategies should be taken into account to see if the total food system is sustainable or not.

4.2.2 Trust and data-sharing recommendations

1 Responsible data sharing should be facilitated by co-creation of values and norms leading to more trust between stakeholders.
2 There should be a clear unified data-sharing security standard applied within the agri-food sector to ensure the protection of data, while ensuring greater trust.
3 There should be an active promotion of systems in which stakeholders can manage their data by their consent with authorisations and a data locker system.
4 The EU Code of Conduct for Agricultural Data Sharing by Contractual Agreement (EUCC): it should be clear that it is a text that guides (legal) contract formation, it should be understandable (shorter, more practical), and include ‘example contracts’ and ‘checklists’ for agribusinesses and farmers.
5 A new code of ethics. There should not be an overreliance on contracts for ethical conduct. A new code of ethics should be developed for tech developers and tech service providers, to help realizing trusted data sharing practices, and which includes a richer set of values then the EUCC as well as clear guidance on how to implement and administer it in concrete companies, in order to make it part of the data sharing culture.
6 The creation of a kind of digital data libraries may help to (a) foster standardisation, (b) foster the re-use of data for public purposes such as monitoring it to track impacts on environment or food safety etc, (c) allow to develop common (shared) policy regarding the way data stored in ‘libraries’ should be used, which allows for more democratic governance by well-informed stakeholders.
7 Standardisation for farm data sharing should be developed to protect those sharing their data and also those using it. This helps to reduce conflict, false expectations, and to avoid legal issues from data misuse. It should be clear, implementable, and enforceable.

8 Policymakers need to protect those sharing data and their intellectual property by regulations against unwanted distribution of data, prevent monopolies or a market dominated by a few big players, which would be in breach of the European competition law.

4.2.3 Technological limitations and connectivity recommendations:

1 Digital solutions should be integrated in the existing setting of companies and equipment, use of standards needs to be promoted for this.
2 Transnational and national policymakers need to find greater convergence of regulations to ensure easy, effective, and mutually beneficial transitions between borders, allowing easier adoption of new digital technologies.
3 Implement sufficient and affordable internet connections in rural areas. Ensure that rural areas have sufficient connectivity, promote an awareness of this availability, and education of how the sector can benefit from cloud-based services and online business channels.
4 Make specific policy efforts for fair access to, and education of, technology to avoid the "digital divide" and information asymmetries. Policymakers need to ensure there is a level playing field and nobody gets left behind.
5 Technology standards and the use of performance standards need to be set, along with policy decisions including holistic consideration of benefits, costs, effects of digital technologies, climate, re-use, and recycling. The technology itself must be robust and reliable, achieved through independent testing facilities to ensure they are effective.

4.2.4 Common Agricultural Policy (CAP) eco-schemes and conditionality recommendations:

1 Policy should encourage data sharing in agri-food, as it would allow for better comparison between technologies, and better baselines for sustainability.
2 Digitized data should be encouraged to streamline the process for farmers and policymakers. Eco-schemes hold the potential to lead to a bigger administrative burden if farmers must prove that they have fulfilled the clauses of their eco-scheme contract. Blockchain technology with smart contracts could be beneficial here.
3 Policymakers should stimulate open data platforms and technology accessibility. Digitally accessible or open data (animal registers, cadastre, pesticide register, etc.) are important to help farmers easily show they abide by the clauses in their eco-schemes.
4 Policymakers can make use of the policy of the Performance Monitoring and

Evaluation Framework (PMEF) and the Farm Accountancy Data Network (FADN). They provide an excellent opportunity to test digital solutions, such as providing farmers with a digital dashboard and key performance indicators. This could go hand in hand with the promotion and development of the Agricultural Data Spaces that is foreseen in the Digital Europe programme by scaling the data space concept down from EU level to Member State- and regional level.

4.2.5 Agricultural Knowledge and Innovation System (AKIS) recommendations:

1. Data sharing should be embedded in topics where some trust is available or in topics where farmers have direct benefits from the innovation (e.g., sharing data on soil and water use could help farmers to investigate strategies to cope with climate change).
2. There should be continuous monitoring and evaluation of inclusion of stakeholder groups as an important parameter in the design and implementation of digital solutions; transparency of the consequences of in/exclusion is necessary.
3. An R&D ecosystem with a well-established network of farms as testing & demonstration infrastructure should be established.
4. Education should include the benefits of data sharing. This would allow future workers to increase their awareness of digital technologies and acquire hands-on experience through training.

5 New technologies should be brought closer to their potential users, in settings where they can see them in action and understand their benefits without the need of big spending ("test before invest").
6 Data from nationally funded projects and initiatives related to infrastructure like weather stations, soil sampling data, should be collected in a common web portal and be accessible to new projects. This would lead to alignment between the EU and Member State policies for the promotion and support of Digital Innovation Hubs for agri-food, and with the Testing and Experimentation Facilities instrument in the Digital Europe Programme.
7 Governments should have a strong open data policy, explained in their National Strategic Plan with clear initiatives to share their data, such as animal health passports and logs of pesticide use. There are many beneficial data sources that are not being utilized because there are insufficient incentives to share them.
8 Wide demonstration of good practices to stimulate the use of previous experiences and results; to stimulate building on existing knowledge and experience.

4.2.6 Artificial Intelligence (AI) recommendations

1 The availability of comprehensive datasets should be stimulated, along with showing the benefits of using these datasets. It should be made clear how to use these datasets. Satellite data should be supplemented with other types of data, e.g., from farm information systems, farm accounting, or sensor networks, to perform on issues like antibiotics, soil management or pesticides, that cannot be monitored from the outside of the farm by satellites or other devices.
2 AI - as the latest development in digital technologies - making digital innovation even more complex should pay even more attention to transparency, privacy, inclusion and human involvement.
3 AI standardisation in the agri-food sector should be improved to facilitate growth in the industry and overcome blockages, while ensuring best practices. This standardisation can be adapted from other industrial sectors or applied to the agri-food sector, rather than investing in timely, costly, standardisation procedures.
4 Ensure that sufficient data and an effective label-train-validation process are critical to keeping the AI systems updated and ready to use by farmers. AI requires data, but also labelling the data to train the AI systems properly.

4.2.7 Data marketplace recommendations

1. Data marketplaces and data sharing initiatives need to be encouraged and supported by policy.
2. There needs to be support to achieve both scale and innovation. Alternative and novel business models based on data- and digital infrastructure sharing should be explored and incentivised accounting for fair distribution of added value.
3. Policy should facilitate interoperability between the public and private domain. For example, share a common “template” for data sharing or storage, instead of having many different database structures.
4. Open source and open data should be encouraged to push towards greater interoperability. Data marketplaces are not just about selling data, but also, sharing data. Stakeholders can benefit from data solutions, but not necessarily, only economically.

4.2.8 Research policy recommendations

1. Further research needs to be conducted around attitudes to sharing data to understand the reasons for distrust in data sharing and how to overcome this.
2. Research agendas need to help establish more concrete business models for agri-food innovation processes. The more concrete the business model in use case projects, the more successful the innovation process.
3. Investigate ways of synergies between different research projects across multiple domains. Exchange between projects to bring the same topic and work groups closer together.
4. Apply innovation funding schemes for replication of good practices.
5. Cross-collaboration and –fertilization between agri-food and other domains and sectors must be stimulated in new innovation programmes and Digital Innovation Hubs.

We hope that these recommendations are taken into account in policy adaptations both at governmental and business level or that they can help to design future projects that address the challenges of the rising data economy in food systems creating competitive, sustainable and fair data ecosystems.

References

Aker, J. C., Ghosh, I., & Burrell, J. (2016). The promise (and pitfalls) of ICT for agriculture initiatives. Agricultural Economics, 47, 35–48.

Basso, B., Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. Nature Sustainability 3, 254-256.

Bronson, K. (2018). Smart farming: including rights holders for responsible agricultural innovation. Technology Innovation Management Review, 8(2), 7–14.

Bronson, K. and Knezevic, I. (2016). Big Data in food and agriculture. Big Data & Society, 3(1), 2053951716648174.

Cinnamon, J. (2020). Data inequalities and why they matter for development. Information Technology for Development, 26(2), 214–233.

Coeckelbergh, M. (2012). Can we trust robots? Ethics and information technology, 14(1), 53–60.

ECPA (2018). 'Code of Conduct on Agricultural Data Sharing Signing.' ECPA. 23 April 2018. https://www.ecpa.eu/media/news/code-conduct-agricultural-data-sharing-signing.

Klerkx, L., Jakku, E. and Labarthe, P. (2019). A review of social science on digital agriculture, smart farming, and agriculture 4.0: New contributions and a future research agenda. NJAS—Wageningen Journal of Life Sciences, 90-91, 100315.

Lajoie-O'Malley, A., Bronson, K., Van Der Burg, S. and Klerkx, L. (2020). The future(s) of digital agriculture and sustainable food systems: An analysis of high-level policy documents. Ecosystem Services, 45, 101183.

Mansour, I., Sahandi, R., Cooper, K. and Warman, A. (2016). Interoperability in the heterogeneous cloud environment: a survey of recent user-centric approaches. In: ICC 2016 proceedings of the International Conference on Internet of Things and Cloud Computing, 62, ACM.

Philip Chen, P. C. L. and Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques, and technologies: A survey on Big Data. Information Sciences, 275, 314 –347.

Rijswijk, K., Bulten, E., Klerkx, L., Dessein, J., Debruyne, L., Brunori, G., . . . Metta, M. (2020). Digital Transformation of Agriculture, Forestry and Rural Areas: Developing a futureproof Socio-Cyber-Physical System. Retrieved from http://desira2020.eu/wp-content/uploads/2020/07/D1.1_CAF-report_I.pdf

Rotz, S., Duncan, E., Small, M., Botschner, J., Dara, R., Mosby, I., … Fraser, E. D. G. (2019). The politics of digital agricultural technologies: A preliminary review. Sociologia Ruralis, 59(2), 203–229.

Ryan, M. (2020). Agricultural big data analytics and the ethics of power. Journal of Agricultural and Environmental Ethics, 33(1), 49–69.

Salemink, K., Strijker, D. and Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. Journal of Rural Studies, 54, 360–371.

Scholz, R. W., Bartelsman, E. J., Diefenbach, S., Franke, L., Grunwald, A., Helbing, D., ... Montag, C. (2018). Unintended side effects of the digital transition: European scientists' messages from a proposition-based expert round table. Sustainability, 10(6), 2001.

Taddeo, M. (2010a). Modelling trust in artificial agents, a first step toward the analysis of e-trust. Minds and Machines, 20(2), 243–257.

Taddeo, M. (2010b). Trust in technology: A distinctive and a problematic relation. *Knowledge, Technology and Policy*, 23(3–4), 283–286.

Tsaftaris, S. A., & Scharr, H. (2019). Sharing the Right Data Right: A Symbiosis with Machine Learning. Trends in Plant Science, 24(2), 99–102. https://doi.org/10.1016/j.tplants.2018.10.016

Van der Burg, S., Bogaardt, M.-J., & Wolfert, S. (2019). Ethics of smart farming: current questions and directions for responsible innovation towards the future. NJAS Wageningen Journal of Life Sciences. https://doi.org/10.1016/j.njas.2019.01.001.

Van der Burg, Simone, Elsje Oosterkamp, Marc-Jeroen Bogaardt, Aine Regan, Eugen Octav Popa, Ewa Tabeau, Elena Favilli, Gianluca Brunori and Cor Wattel (2020). 'Futures of farm data sharing practices; perspectives of European farmers, researchers and agri-tech businesses' (report IOF2020, D7.4), October. https://doi.org/10.13140/RG.2.2.21562.41924

Van der Burg, Simone, Wiseman, L., & Krkeljas, J. (2020). 'Trust in farm data sharing: reflections on the EU code of conduct for agricultural data sharing.' Ethics and Information Technology, 1-14.

Van Dijk, A. D. J., Kootstra, G., Kruijer, W., & de Ridder, D. (2021). Machine learning in plant science and plant breeding. iScience, 24(1), 101890. https://doi.org/10.1016/j.isci.2020.101890

Villani, C., Schoenauer, M., Bonnet, Y., Berthet, C., Cornut, A.-C., Levin, F., & Rondepierre, B. (2018). For a meaningful artificial intelligence: Towards a French and European strategy. https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Colofon

Authors
Sjaak Wolfert, Lan van Wassenaer, Simone van der Burg, Mark Ryan, Laurens Klerkx, Kelly Rijswijk, Mariette McCampbell, Ioannis Athanasiadis, George Beers

Photography
Schuttelaar & Partners
Shutterstock

DOI
https://doi.org/10.18174/552346

To explore the potential of nature to improve the quality of life

The mission of Wageningen University & Research is "To explore the potential of nature to improve the quality of life". Under the banner Wageningen University & Research, Wageningen University and the specialised research institutes of the Wageningen Research Foundation have joined forces in contributing to finding solutions to important questions in the domain of healthy food and living environment. With its roughly 30 branches, 6,500 employees (5,500 fte) and 12,500 students, Wageningen University & Research is one of the leading organisations in its domain. The unique Wageningen approach lies in its integrated approach to issues and the collaboration between different disciplines.

www.wur.eu