Federal Ministry
for the Environment, Nature Conservation
and Nuclear Safety

Digitalisation and Sustainability in the EU: Challenges and utilisation potentials of European environmental data

with practical examples from the European Code4Green Hackathon

Final report

space·tec PARTNERS

terranea

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Acronym	Description
AI	Artificial intelligence
API	Application programming interface (API)
B2B	Business to Business
B2C	Business to Consumer
B2G	Business-to-Government
BfN	German Federal Agency for Nature Conservation
BMU	German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety
EC / COM	European Commission
EIONET	European Environment Information and Observation Network
Engl.	English
EEA	European Environment Agency
FAQ	Frequently Asked Questions
GeoZG	Geodata Access Act
GER	German
GHG	Greenhouse gases (greenhouse gas emissions)
INSPIRE	Infrastructure for Spatial Information in the European Community
MVP	Minimal viable product
(UN) SDGs	United Nations Sustainable Development Goals
SDK	Software development kit
STP	SpaceTec Partners
TFEU	Treaty on the Functioning of the European Union
UBA	German Federal Environment Agency
UIG	German Environmental Information Act (GER. Umweltinformationsgesetz)

INTRODUCTION

The **Digital Policy Agenda for the Environment** of the German **Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU)**[1] proposes concrete measures aimed at implementing digitalisation and use of data in a climate-friendly and sustainable way. Furthermore, it promotes **economic prosperity and competitiveness, social justice and a healthy environment.**[2] Open, trans-European and high-quality environmental data have the potential to provide important insights and solutions to current challenges such as climate change, the decline in biodiversity, or making economic processes more climate-friendly.

To make the provision of environmental data more user-friendly and target group-oriented, an interdisciplinary team composed of the strategy consultancy **SpaceTec Partners**, and the geodata consultancy **Terranea** was tasked with the project **"Digitisation and Sustainability in the EU: Challenges and utilisation potentials of European environmental data".** On behalf of the Federal Ministry for the Environment, the project aimed at identifying challenges in the availability and quality of European environmental data and to propose solutions for increasing their potential use.

The project partners developed and conducted an initial **study on the availability, accessibility and quality of selected European data sets** to assess the possible challenges and potential uses of environmental data. Based on the findings of this study, the partners organised the **Code4Green Environmental Data Hackathon**, in which participants from all over Europe were challenged to create digital solutions in the areas of "climate change", "biodiversity" and "circular economy" using selected open European environmental data. The combination of the results of the initial data study, the solutions developed during the hackathon, and subsequent feedback by hackathon participants and experts resulted in essential new insights.

This report summarises the most important findings regarding the challenges and potential uses of European environmental data. It also formulates a series of **political recommendations** to improve the visibility, availability, accessibility and usability of European environmental data. Furthermore, the report characterises digital and sustainable business approaches as well as the growing innovation ecosystem with its support mechanisms in this field.[3]

[1] Following, Ministry for the Environment / BMU.

[2] BMU (2020): "Digital Policy Agenda for the Environment.

[3] All hyperlinks integrated in this document are online and reachable as of 20.01.2021.

1. USING ENVIRONMENTAL DATA FOR THE DIGITALISATION

Digitalisation constitutes data as an essential resource alongside the entire information, production and value chain of almost all business processes and models today. It is often at the heart of new **opportunities for commercialisation and usage,** as well as being an incremental factor for increasing economic growth, job creation, innovation, and societal progress.

The European Commission's recently published strategy for data *(COM (2020) 66 final)* aims at creating a common market for data that ensures Europe's global competitiveness and data sovereignty. Accordingly, a common European data market should increase the visibility, uptake and usage of open European data both in the economy and its society, while remaining highest standards for privacy and control of data. Besides the **economic potential of a data economy,** its **environmental impact** must however be considered as well.

On the one hand, it is important to organise digitalisation itself sustainably: The operation of data infrastructure and data centres require large amounts of energy; digital devices used for the collection and use of data have a short service life and are not always recycled–all of which leads to a negative environmental impact. On the other hand, digitisation processes also offer great opportunities, of which many are still unused, for a more climate and environmentally friendly design of infrastructure, production, and value chains.

The Digital Policy Agenda for the Environment of the German **Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU)** contains concrete measures aimed at making **digitalisation and data** use **climate-friendly** as well as to promote **economic prosperity and competitiveness, social justice, and a healthy environment**. Furthermore, it aims at contributing to the sustainability goals of the 2030 Agenda. Used favourably, **environmental data can play** a central role to reach these goals: open, trans-European and of high quality, environmental data have the potential to provide important insights and solutions to current challenges such as climate change and the steep decline in biodiversity. Additionally, the data can be used to create more climate-friendly economic processes.

Against this background and at the interface of digitalisation and sustainability, the project partners *SpaceTec Partners* and *Terranea* carried out the project **"Digitalisation and Sustainability in the EU: Challenges and utilisation potentials of European environmental data"** on behalf of the Federal Ministry for the Environment. The focus of the project was to determine current challenges and barriers in accessing and using environmental data, as well as assessing the quality of the data currently available. Additionally, the project focused on identifying solutions to the current challenges and finding ways in which potential data use can be increased. This paper provides a series of **recommendations** which have been derived from internal research, as well as from external input (interviews and participant feedback within the framework of the Code4Green Hackathon).

2. CHARACTERISATION OF ENVIRONMENTAL DATA

An overarching goal of modern environment-oriented legislation is to protect, preserve and improve the environment for present and future generations. Appropriate legislation ensures a high level of environmental protection and in this way preserves the quality of life for EU citizens. Art. 191 (1) of the TFEU defines the objectives of its environmental policy as follows:

- Preserving, protecting and improving the quality of the environment,
- Protecting human health,
- Prudent and rational utilisation of natural resources,
- Promoting measures at an international level to deal with regional or worldwide environmental problems, and in particular combating climate change.

The acquisition and provision of environmental data is a central element for reaching these objectives.[4] This chapter provides an overview of the definition, the scope, and the obligations for the provision of environmental data.

2.1 How is environmental data defined?

Environmental data provide information on the **state of the environment** and include for example measurements on air and water quality, GHG and other pollutant emissions, land cover, noise and radiation pollution, biodiversity, as well as information on energy production and consumption, or waste management and recycling processes. As environmental data usually inherent a geographical reference, environmental data is often available as **spatial data**. Depending on the source, environmental data can be available either freely or commercially. The provision of open and free environmental data is especially relevant as it enables citizens to participate beyond economic value chains and ensures transparency. The **European Green Deal** underlines the need for open and free access to environmental data, not only in the context of European environmental policy but also as an important basis for data-driven innovation in general. Cloud infrastructure and artificial intelligence algorithms are important factors in understanding and addressing environmental challenges and for making evidence-based decisions and policy recommendations.[5]

2.2 The legal framework for the collection, provision, and dissemination of public environmental data in Germany and the EU

Today, free access to public environmental data within Europe is based on European environmental law. This either applies directly (in the form of regulations) or must be transposed into national law (in the form of directives). Common environmental standards serve not only to improve the state of the environment in Europe but also to prevent market distortions within the Union. This ensures that

[4] "In preparing its policy on the environment, the Union shall take account of available scientific and technical data" (Art. 191 para. 3 TFEU).

[5] See e.g. COM (2019) 640 final.

member states can save on production costs if they have different environmental requirements to meet those standards, ultimately leading to competitive advantages.

Particularly relevant for environmental data is the **Environmental Information Directive of** the European Union[6], which was adopted in 2003 and transposed into national law in the Federal Republic of Germany in 2005 as the **Environmental Information Act (UIG)**. In addition to the scope and type of data to be collected, the UIG also defines **bodies obliged to provide information.** Accordingly, "the government and other public administration bodies", as well as "natural or legal persons under private law, insofar as they perform public tasks or provide public services (...)" must conduct regular surveys and make them available free of charge (cf. §2 para. 1, 2 UIG). In turn, any person is entitled to access this data without having to demonstrate a legal or economic need (cf. §3 para. 1 UIG). Free access may be restricted under specific circumstances if the environmental data is of vital importance to the public or other interest protection. These include, for example, the disclosure of trade, business and tax secrets, the infringement of copyright or intellectual property, concerns about public safety, defence, international relations, as well as the endangerment of law enforcement (cf. §8f UIG).

Equally important concerning environmental data is the **INSPIRE Directive** adopted by the EU in 2007, which aims to create an infrastructure for spatial data in the European Union. INSPIRE obliges the Member States to make all collected geographic reference data and geospatial data interoperable and freely available via network services. In Germany, the directive was incorporated into national law at the federal level in the form of the **Geodata Access Act (GeoZG)**. It obliges the federal government to make all geodata and services available free of charge for both commercial and non-commercial use.

2.3 Collection and provision of environmental data in Germany and the EU

In Germany, the **collection of environmental data is fragmented**. While the obligation to report and publish data lies with the federal government, both the federal government and the environmental administrations of the respective federal states oversee the underlying analyses and publications. Due to fragmentation in responsibilities, the measurements for these analyses can be carried out by different specialised authorities, which results in the data being published by a multitude of agencies in numerous data portals. Federalism in Germany only reinforces this fragmentation. This can make the targeted retrieval of data difficult in certain cases.

At the European level, the management and analysis of national data are carried out by the European Commission and the **European Environment Agency (EEA)**. The **European Environment Information and Observation Network (EIONET)** represents the network of EEA member states through which reporting takes place. The EEA aggregates the data submitted by the member states and publishes the results in the form of reports, indicators, and maps. As in Germany via the UIG, all environmental

[6] Directive 2003/04/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC. Online at:

data in Europe are treated as public goods and are openly accessible. Restrictions arise, as mentioned above, only from legal provisions, data protection reasons, intellectual property rights or the protection of national or public security.

The obligation to freely distribute environmental data does not apply to private providers and portals, privately funded research projects or citizen science projects.

2.4 Usability of environmental data portals in companies

Today, environmental data is also an important resource for businesses. They make it possible to **develop new services**, **increase the efficiency** of companies and **save costs**. Start-ups complement e.g. public data on the current air quality with data from their sensor networks [7] or provide information on health risks from environmental influences. [8] In agriculture, satellite data is used to determine the ripeness of fruits and create yield potential maps and harvesting plans. [9] Further examples that do not only relate to open environmental data can be found in the literature [10] and on the Internet. [11] For a maximum derived benefit from the available data, however, several prerequisites regarding the data are necessary.

First, the data must be **easy to find and freely accessible**. Ideally, data from different organisations or authorities should be published via a common and central data portal. This makes it easier to find and access data. Searching through various databases from several different providers is an enormous effort for businesses. Standardised metadata description for each data set further helps users in understanding the data and its content, as well as in obtaining further background information on collection methods as well as possibilities and limits of the applicability of the data at hand. They are also a prerequisite for the correct use and interpretation of the data. **INSPIRE ensures that metadata standards are used for public sector environmental data**.

Access to data should be made **possible without barriers** (i.e. without prior registration) and free of discrimination. Even though registration on data portals can support the analysis of user groups and their search and download behaviour (which in turn can lead to an improvement of the data portal and its offer), barrier-free access is preferred from the user's point of view. Discrimination of users could, for example, take place according to gender or place of residence. Compared to other data access challenges, user discrimination in accessing environmental data plays a subordinate role.[12]

For data to be used economically, it must be **machine-readable**. This means that it must be structured and available in a format that can be automatically identified and evaluated by software. Proprietary

[7] https://hawadawa.com/

[8] https://ajuma.eu/

[9] https://www.greenspin.de/

[10] See e.g.: European Data Portal (2020): „Analytical Report 10: Open Data and Entrepreneurship“; Open Data Institute (2017): „Data entrepreneurship: exploring successful business models with open data“.

[11] European Data Portal: „Unlocking value from open data“.

[12] There is no recorded event, where gender, place of residence or other personal attributes hindered the access to publicly available environmental data.

formats usually prevent automated evaluation. Depending on interest and need, data sets can be downloaded and stored individually or in a package. State-of-the-art for doing so today are **Application Programming Interfaces (APIs)**. They enable direct communication between systems and an automated data exchange so that extensive data sets do not have to be stored separately.

If the environmental data is not prepared and processed by the company, but integrated directly into a system via an API, it must be ensured that the data can be retrieved at any time. This requires a robust infrastructure with minimal downtime. An example of this is data on air quality, which is collected via numerous measuring points. For meaningful use, the data must be up to date. Static data, i.e. data that is updated at longer intervals, should also always be retrievable in its most recent version and contain a high degree of accuracy, respectively contain corresponding information about possible inaccuracies within the metadata.

It is also important that the **measurement series** on offer are **sustainable,** i.e. that they are **continued over a long period** for economic exploitation to be worthwhile. If measurements are interrupted or even terminated, a service provider loses the foundation of his business model. Data sustainability requires long-term funding as only the guarantee of long-term funding provides **data and planning security** for companies.

Licences regulate the **use and exploitation possibilities** of **copyrighted data**. They enable users to offer specifically modified data for commercial purposes or to integrate it into their applications. For efficient use, **licences** from different data providers are expected to be as **uniform** as possible so that data can be exploited under the same conditions. There are numerous standard licences, such as those of the Creative Commons (CC) organisation. [13] In Germany, public sector data is often published with a Data Licence Germany 2.0 [14] which allows commercial and non-commercial use with a reference to the source.

[13] https://creativecommons.org/licenses/

[14] https://www.govdata.de/lizenzen

3. SPECIAL FEATURES OF DIGITAL BUSINESS MODELS

Creating a business is easier than ever in the digital age. Contextual knowledge and building blocks are largely available "at one's fingertips". A variety of data storage, processing, shop, CRM, billing, and accounting systems can be integrated by external cloud service providers. Missing connections and networks can be established more easily across country borders and language barriers in the online realm.

Typical data businesses are based on **visualising** contexts or enriching them with **contextual information** for private, commercial, or institutional users. They overcome analogue or digital media discontinuities, use interactive elements to bring data to life, or make data available for **embedding in business processes** via digital protocols (SDK, API, etc.). The identification and development of a successful business model require the willingness to think through multi-layered use cases and to recognise the lowest common denominator where new value creation will come to bear. Ideally, cross-border scaling should also be possible.

Highly scalable data business models are especially attractive as they are self-multiplying ones their use and benefit are proven. The patterns of generic digital business models that go beyond a project or license-based solution businesses are sufficiently described in the literature: Platforms, Freemium, Subscription and Pay per Use are just a few examples.[15] The complexity of implementation along the customer lifecycle is often underestimated. In many digital business models, the **end-user is not necessarily also the paying customer**. The end-user often knowingly or unknowingly contributes to content or pays indirectly by giving the service provider access to his or her user data.

Many successful digital business models combine business-to-government (B2G) and business-to-business (B2B) aspects and are often overlaid with a variety of partnerships. Digital entrepreneurs often fail to define **a simple comprehensible business model**, solve a key customer problem, or to validate market interest for business partners and investors with a so-called "Minimal Viable Products" (MVPs).

When implementing digital business models, **a multitude of legal requirements** must be taken into account, from data protection to copyright and liability issues. If public data is used, it is not easy to offer a "service level agreement" based on it, as the service provider is dependent on public sources. These are generally inexpensive or even free of charge, but also do not include a sustainable guarantee of supply.

Entrepreneurs who want to attract venture capital investors must focus on scalable business models that can serve a sufficiently large market in a relatively short period. In the data business, the creation of intellectual capital (patents, property rights) is of little relevance; what matters most is **implementation speed and user growth.**

[15] For a comprehensive overview see e.g.: Aagaard, Annabeth (2019): "Digital Business Models. Driving Transformation and Innovation.", Springer International Publishing, Basel, Switzerland.

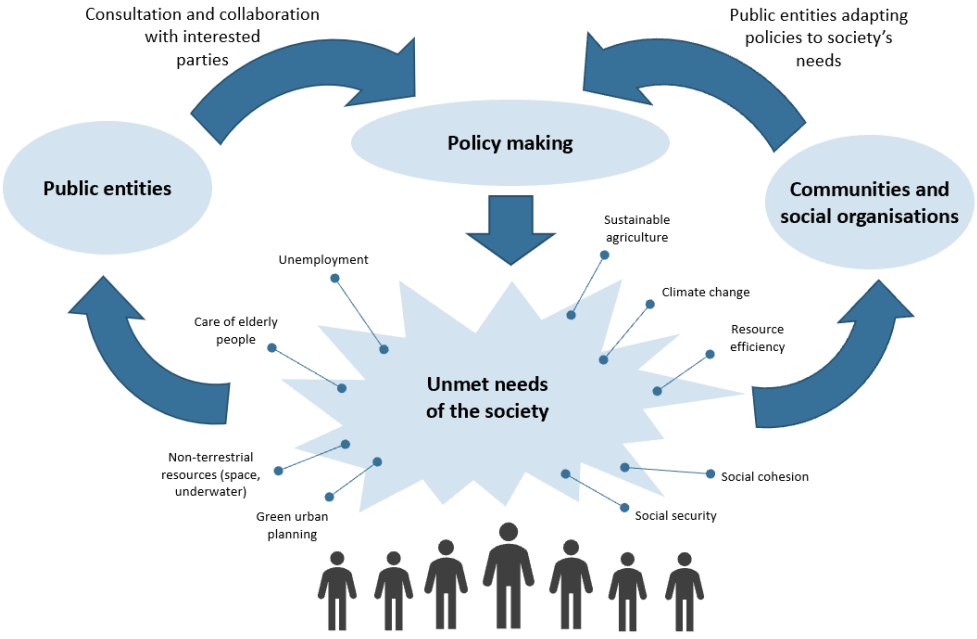
4. SUPPORT FOR PUBLIC BENEFIT-ORIENTED FOUNDERS

Entrepreneurs and innovators are characterised by a high passion for their ideas. They have an intrinsic drive to create value. Innovators want to make a difference, bringing their knowledge, skills, and contextual understanding to the table. Increasingly, there is a growing awareness of sustainability in business among established entrepreneurs and start-ups. Issues such as global climate change or creating and maintaining a liveable environment are gaining traction. Many new business plans refer to the fulfilment of the UN Sustainable Development Goals (SDGs).

Figure 1: The United Nations Sustainable Development Goals (SDGs)

Whilst their idealistic end benefits are usually desirable for a clear majority, applications-oriented towards the common good are often difficult to address through market forces. Often, these applications aim at improving developments and conditions that meet public interest but are dependent on the goodwill of state institutions or the cooperation with social organisations for progress.

Figure 2: Illustration of a possible policy cycle to achieve social needs

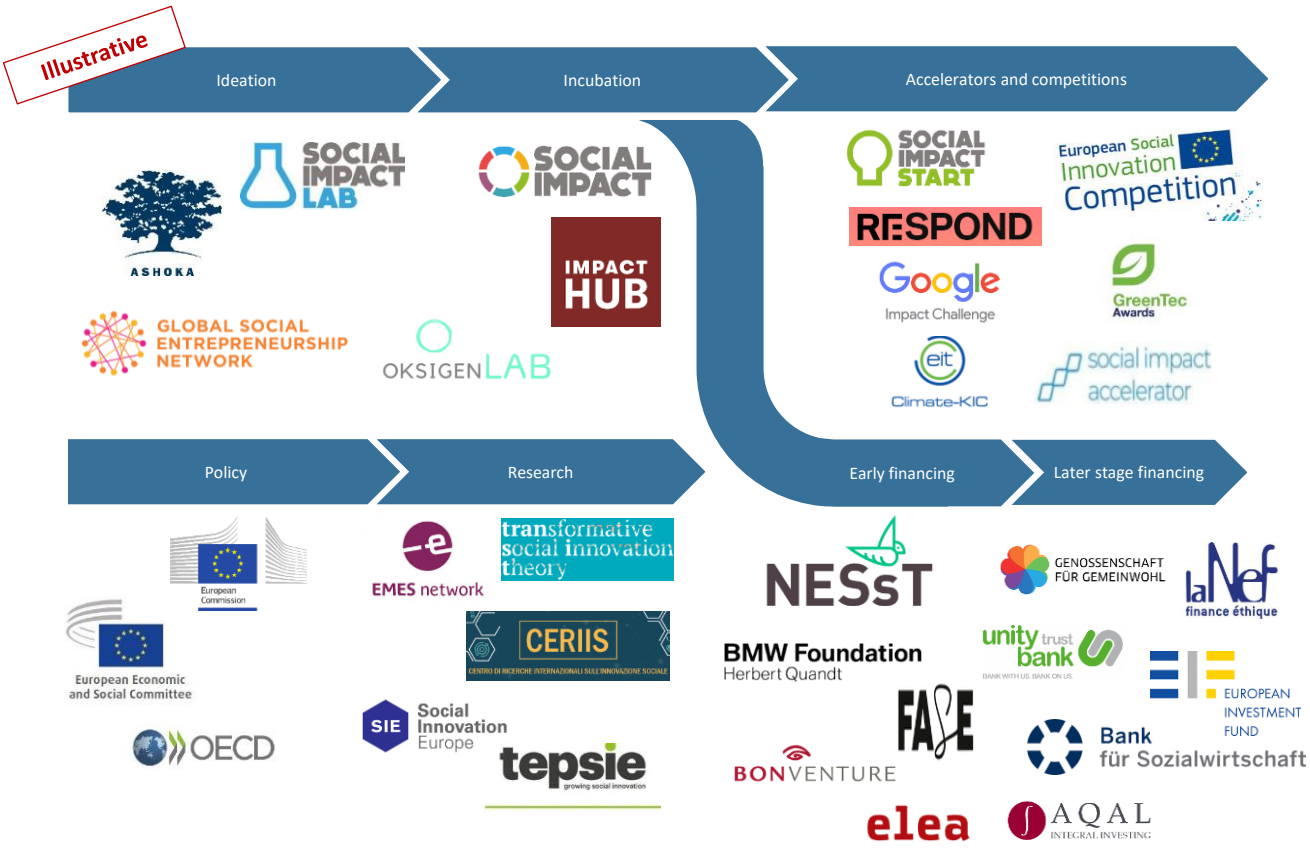


In addition to idealism and professional competence, entrepreneurial skills such as enthusiasm, the ability to work in a team and perseverance are of particular importance. Since the **desired goals can often only be achieved indirectly**, openness to new topics, the ability to compromise and a high degree of flexibility are necessary.

To implement applications orientated towards the common good, the entrepreneur must develop a sense for the interests, concerns and constraints of institutions, non-governmental organisations,

and lobby groups. It is difficult to develop sustainable business models based on voluntary work or donation funding. Ideally, **public funding should serve as initial start-up funding** to enable a team to fully take on the issue. Acquiring funding is a way for many entrepreneurial teams to articulate their ideas and gain experience with the public sector as well.

Figure 3: Representation of the European funding and financing ecosystem.



In recent years, numerous innovation ecosystems focused on sustainability and social impact have emerged. Incubator and accelerator programmes serve to bring entrepreneurial ideas to fruition, create efficient teams and expand networks. Furthermore, **accelerators serve as a showcase for interested investors** who are also influenced by the recent trend towards sustainable solutions: A growing number of investors want to finance companies that do not only generate revenues but also bring along a positive societal benefit.

5. EUROPEAN PERSPECTIVES

Environmental data play an increasingly important role in current European legislation, as is evidenced by the publication of the **European Strategy for Data** or the **European Green Deal**, which both reference environmental data as an important source for further developments. In the following chapter, the key legislative points of the European Green Deal and the European Strategy for Data are presented. The chapter also summarizes the most important findings of the expert panel discussion led by the German Federal Minister for the Environment, Nature Conservation and Nuclear Safety Svenja Schulze on the **European Model of Digitalisation**. It was broadcast live as part of the Code4Green Hackathon on 12 July 2020.

5.1 Environmental data in the European Green Deal and the European Strategy for Data

The European Green Deal

On 11 December 2019, EU Commission President Ursula von der Leyen officially presented the European Green Deal for the first time. The goals are:[16]

- A resource-efficient and competitive economy,
- Carbon neutrality of the European Union by 2050,
- The restoration of biodiversity,
- The EU as a pioneer in the fields of environment and climate.

The Green Deal creates an important basis for future environmental policy measures within the EU and formulates a roadmap to achieve the adopted goals–regarding also the handling of European environmental data.[17]

According to the Commission, only free access for private and business entities would allow **stakeholders at all levels to be involved in future climate protection measures** and, especially at the business level, to be able to develop better risk management strategies regarding climate change.

In addition to the open provision of publicly collected data (see chapter 2.2 of the final report), the Commission also considers business enterprises themselves to have obligations: In the future, they should make their climate impact and environmental data more visible and accessible to make it possible for public and private investors to see comprehensive information on the sustainability aspects of their investments.

The Commission considers (environmental) data to be especially relevant for innovation. According to the Commission, "accessible and interoperable data (...) are at the heart of data-driven innovation".[18] Data, in combination with digital infrastructure and artificial intelligence applications,

[16] https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

[17] COM (2019) 640 final. Aviailable online: https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1&format=PDF

[18] COM (2019) 640 final: 23.

would play a crucial role in **evidence-based decision-making** and in **addressing environmental challenges.** The Commission particularly emphasises the possibility of using data to improve the **prediction and management of environmental disasters**. [19]

The European Strategy for Data & INSPIRE

Following the Green Deal, the Commission published "A European Strategy for Data" in February 2020.[20] Among other things, it promotes the creation of a **Common European Data Space for the European Green Deal**. This would be used to support measures set forth by the Green Deal regarding climate change, the circular economy, carbon neutrality, biodiversity, and deforestation, as well as helping to ensure compliance with its regulations.

The Commission created the **GreenData4AI** initiative, which aims to review in detail and, if necessary, revise the INSPIRE spatial data infrastructure and the Access to Environmental Information Directive (see chapter 2 of this report). This should ultimately make it easier for public authorities, businesses, and civilians to "support the transition to a greener and carbon-neutral economy" with less administrative effort.[21]

5.2 Environmental policy goals of the German Presidency of the Council of the European Union

Based on the European Green Deal, and the many opportunities that arise from the collection, provision, and its use, environmental data also play an important role in the context of the German Presidency of the Council of the European Union. [22] **The German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, therefore, put the entwined challenges of digitalisation and sustainability in the focus of the German EU Council Presidency**.

Accordingly, the BMU highlights to use digitalisation and digital data (including environmental data) in a targeted and efficient way for environmental protection issues, whilst making digitalisation itself more environmentally friendly.

"For example, a European Data Space for environmental data can support adaptation to climate change and the implementation of environmental policy (...) but exploding data volumes and computationally intensive applications [also] have negative effects on the environment."[23] Therefore, during the German EU Council Presidency, the BMU is actively engaged in making the use of European environmental data **more target-oriented** and **user-friendly**, as well as more **ecological.**[24]

[19] COM (2019) 640 final: 23.

[20] COM (2020) 66 final. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=DE

[21] COM (2020) 66 final: 31.

[22] Following: German EU Council Presidency.

[23] https://www.bundesregierung.de/breg-de/themen/europa/ziele-umwelt-1791978

[24] https://www.bmu.de/eu-ratspraesidentschaft-2020/unsere-ziele-fuer-die-deutsche-eu-ratspraesidentschaft/

5.3 Environmental data and digitalisation in the European Model

> "*It is important to have our own way. **Europe is different from China and from the US** and we could do digitalisation in a way that is **more sustainable**, that is **more open**, that is **democratic** and uses public data differently to how the **U.S. or China** does.*"
>
> - German Federal Environment Minister Svenja Schulze.

In the frame of the Code4Green Hackathon, the organisers conducted a panel discussion on the topic "**Sustainable, innovative, social - shaping the European way of digitalisation**". The panel was joined by the German Federal Environment Minister Svenja Schulze, Hans Bruyninckx (Executive Director of the European Environment Agency EEA), Francesca Bria (President of the Italian Innovation Fund) and Lukasz Gadowski (venture investor, founder, and head of Team Europe), who discussed the future path of European digitalisation.[25]

There was consensus among the panellists that the digital future of Europe should set itself apart from global role models such as the USA or China in terms of accessibility, privacy and sustainability.

Francesca Bria emphasised a digital future for Europe based on freely accessible data, the protection of individual personal and fundamental rights, democratic values, and sustainability. Based on these values, Europe's technological digital transformation should bring a special global competitive advantage and contribute to solving current problems such as climate change.

Figure 4: Federal Environment Minister Svenja Schulze during the panel discussion on the digital future of Europe.

[25] A video compilation of the panel discussion is available online: https://www.bmu.de/service/veranstaltungen/wettbewerbe/code4green/#c48030

Federal Environment Minister Svenja Schulze shared these views and, referring to the German EU Council Presidency, emphasised the role of EU member states in particular: the digital future could best be shaped at the EU level - cooperation between the EU member states would therefore be of vital importance. Hans Bruyninckx spoke of a genuine EU brand of digitalisation which, unlike the global competition, is based on more democratic values and more sustainability. If, according to Bruyninkcx, Europe succeeds in becoming more independent in the areas of AI and big data in the future, Europe and the EU could take on a special leadership role in global digitalisation.

From an entrepreneurial point of view, Lukasz Gadowski pointed out that capital and financial markets still need to be improved and facilitated to support a successful digital EU model. As they currently stay underdeveloped in international comparison, (social) entrepreneurs would have a hard time developing and implementing sustainable digital products. The fragmentation of the EU market is an additional issue when compared to the more closed US and Asian markets. Despite these drawbacks, Gadowski considers the more open, transparent, and sustainable European data economy as a significant advantage over other international competitors.

Figure 5: View on the mixing desk during the panel discussion with Federal Environment Minister Svenja Schulze, EEA Director Hans Bruyninckx, Francesca Bria and Lukas Gadowski

6. DATA IN ACTION: THE CODE4GREEN HACKATHON

As part of the project " Digitalisation and Sustainability in the EU: Challenges and utilisation potentials of European environmental data", the **Code4Green Hackathon** was the public core of the project. The second edition of Code4Green took place under the patronage of the **German EU Council Presidency.** It aimed to actively promote **sustainability and digitalisation in the EU**, to attract public attention to digital sustainability and to gather, promote and facilitate sustainable project ideas of its participants. During the hackathon, approximately 70 participants from all over Europe had 36 hours to come up with and develop a sustainable, digital project idea in one of the following thematic areas (challenges):

- **Climate Change**
- **Biodiversity**
- **Circular Economy**

The participants were supported by ten mentors and nine subject-specific data coaches from the organisation team and partner organisations. To ensure easy access to open environmental data, topic-specific **data catalogues** were prepared in advance. The participants were expressly asked not to limit themselves to the data sets offered and were encouraged to find additional data sources themselves. Due to the excellent support and high motivation of the teams, twelve team projects were published on the Code4Green developer webpage.[26] The winning teams (one team per challenge, plus an additional audience prize) were awarded a total of 10,000 € prize money. Additionally, the winners received three-month mentoring support from SpaceTec Partners and Terranea to further develop their project ideas.

Figure 6: German Federal Environment Minister Svenja Schulze and BMU Head of Department for administration, budget, research and digitalisation Dirk Meyer congratulate the winning team of the biodiversity challenge

[26] https://code4green.devpost.com

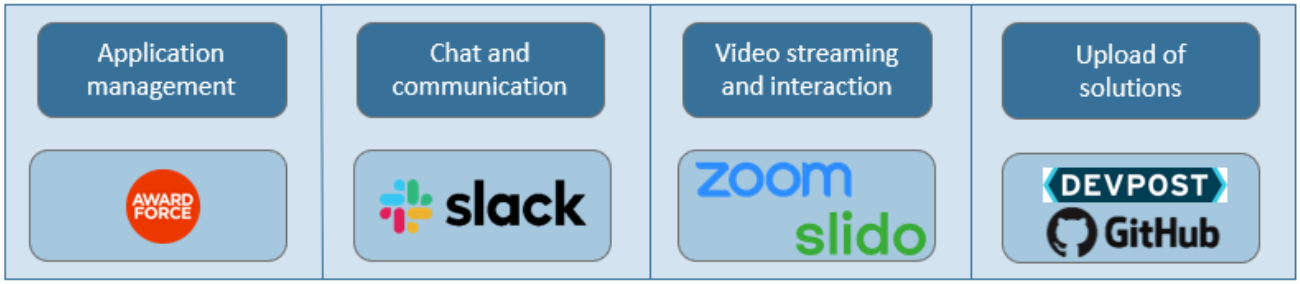
6.1 The virtual event concept

Originally, the Code4Green hackathon was planned as a physical event in Mainz from 12 to 14 July 2020. Due to the corona pandemic, the event concept had to be changed to a **virtual format**. Several concepts and software solutions for the realisation of the hackathon were evaluated in advance. Ultimately, Code4Green was virtually implemented with physical "headquarters" at the BMU's premises in Berlin. While all hackathon participants, mentors and data coaches joined the event virtually, the organisers guided through the event from a physical stage during several live streams.

To be in line with the **European Green Deal Initiative** and the **EU Digital Strategy**, the conceptualisation of the Code4Green hackathon aimed at maximum sustainability and inclusiveness from the very start. The virtual design of the hackathon however offered several additional advantages. By **avoiding or minimising physical contact,** the **safety of participants** and organisers was ensured during the corona pandemic. As no physical travel of the participants was required and all materials were provided digitally, **carbon emissions were reduced** even further, whilst **inclusion was even higher**. As such, the Code4Green Hackathon incorporated goals from the **European Green Deal Initiative** as well as the **EU Digital Strategy**.

All communication between participants took place via the web-based instant messaging service Slack. The live programme, taking place on the physical stage in Berlin, was broadcasted in a live stream via the platforms Zoom and Periscope on the BMU website.

Figure 7: Software used for the Code4Green Hackathon



6.2 Recommendations and impressions of a virtual hackathon

The virtual concept of a multi-day hackathon was a novelty for the organisers involved. Therefore, as a follow-up to the hackathon, 30 participants were contacted via video calls to gather feedback on the process and their personal experience of and during the event. Based on these results and experiences gathered by the organizers before and during the hackathon, several recommendations can be formulated for the future conception of virtual hackathons or virtual events in general.

One source of truth: The organisers of the Code4Green hackathon had agreed to keep all participants up to date via one central news channel. All questions that were expected to arise or were raised before the event were collected in a **FAQ** document. The document was always accessible to all participants but could only be edited by organisers–the document was expanded continuously during the hackathon with new questions and answers. By only having this one official source of information, misunderstandings regarding deadlines, requirements or software (for example by having contrary information in two different documents) could be avoided.

Drop-outs and no-shows: A big difference between physical and virtual events is that in virtual events it is easier for participants to depart early or not show up at all. Virtual events do not have reimbursable travel expenses and do not share the same social setting as physical events.[27] As the registration for the hackathon is designed to be quick and easy, organisers have to expect several registrations by undecided participants. As a result, a higher number of no-shows and drop-outs can be expected at virtual events.[28] Organisers should therefore take this into account if they require a minimum number of participants. For virtual events, it is advisable to increase the maximum number of participants and to design the event programme in such a way that it can still take place as planned with a reduced number of participants (scalability of the programme).

Team building and complementarity: For virtual events, the following aspects should be taken into account concerning team building to maintain the complementarity of teams effectively.

a) Due to the more informal nature of the event and the higher risk of "no-shows" (see above), teams defined in advance may lose important competencies (for example necessary programmers or data scientists). Therefore, the division of the team should be designed in such a way that the team can continue working on the project even if certain team members are no longer available.

b) In principle, different approaches can be chosen for the allocation and building of teams. For example, a team can be created by the organizers in advance or created freely by the participants themselves.

   - **Teams created by the organisers**: In this variant, also used for the Code4Green hackathon, teams are built in advance by the organisers based on their interests, skills and professional backgrounds - initial project ideas presented by applicants can be also be taken into account here. This allows organisers to ensure a minimum number of teams.
   - **Team creation by the participants**: An alternative variant allows participants to present their project ideas before the event and recruit corresponding and interested team members. Participants thus freely come together based on their personal preferences, ideas and skillsets. This does not ensure a minimum number of teams. However, the free choice of participants to choose their team members can be motivating for participants and increases the likelihood that team members harmonise.[29]

For Code4Green, the first variant was chosen, and care was taken to assemble the teams as internationally and diverse as possible, based on their interests, skillsets and professional backgrounds. During the event, however, this variant showed some weaknesses: above mentioned "no-shows" had a significantly negative impact on the workflow and rendered some teams incapable of competing already at the beginning of the event (missing programmers and data scientists). The

[27] In a physical hackathon, interaction between team members and organizers is more direct and intimate. The eventcharacter and direct social interaction usually prevents participants from giving up and leaving the event early. If travel expenses came up, participants will endeavor to be able to claim reimbursement through their presence.

[28] Of originally 94 registered participants for the Code4Green hackathon, 27 participants did not show up to the virtual event. During the hackathon 18 participants decided to leave the event early due to a variety of reasons.

[29] During feedback-calls, several participants noted that they would have preferred to pick their respective team members themselves, since personalities and personal interests sometimes did not match each other ideally.

organisers were able to react quickly and distributed the team members to new teams or combined individual teams. This way all teams and team members could continue working, however, this also led to a drop in motivation from some of the team members.

For hackathons, it might therefore make more sense to make participants create their teams by themselves ahead of the event, to ensure the creation of compatible, harmonising and motivated teams. During the event, organisers need to be in constant contact with the teams to be able to react quickly if a team loses an important member.

More time to get to know each other: A virtual event format makes it harder for participants to get to know each other before the event. During the Code4Green, some participants complained that they did not have sufficient time to get to know their team members well ahead of the event. We accordingly recommend opening the communication channels between participants two days before the event, so team members may already get in touch more personally. We also recommend organisers to encourage all participants to proactively approach their respective team members.

Space for exchange: A hackathon is, in its essence, a competition. For this reason, the organisers limited the interaction between the individual teams to a minimum to avoid free-riding or the duplication of ideas. However, innovative project results and a social component are also an important aspect of these events. Several Code4Green participants felt that the interaction and exchange of ideas between teams came up short during the event. Future events need to find a good balance between these aspects of competitiveness, social interaction and inspiration.

6.3 Solutions and impressions of the Hackathon winners

Besides evaluating the visibility, availability and usability of European environmental data, the Code4Green Hackathon also aimed to generate and promote digital and sustainable solutions for environmental data. At the end of the hackathon, a total of **three main prizes and one audience prize** were awarded to projects that demonstrated the practical usability of European environmental data creatively and sustainably.

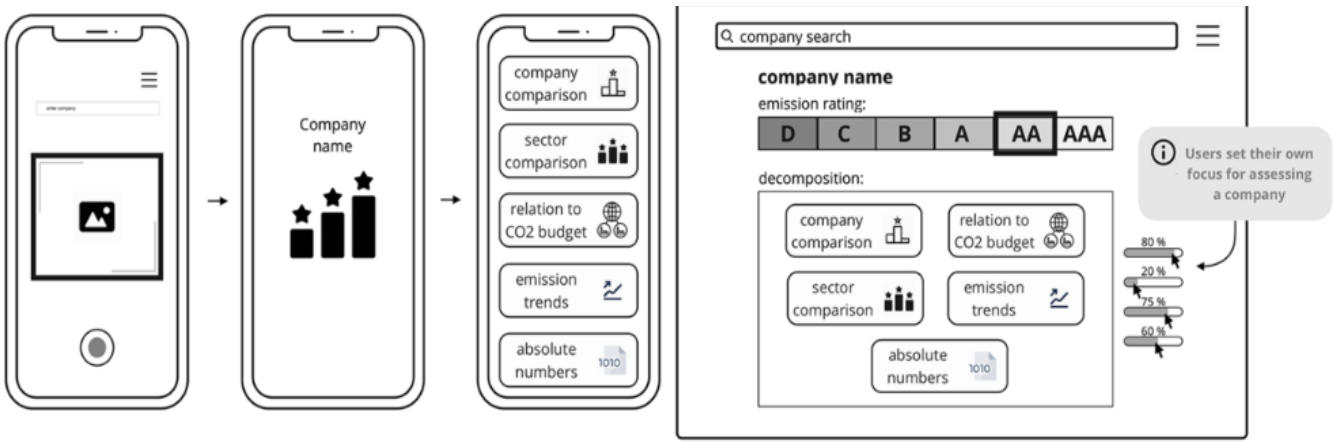
The three main winning teams each prize money adding up to a total value of 10,000 €. Additionally, they received a three-month mentoring programme from SpaceTec Partners, Terranea and other project partners such as the Leaders for Climate Action (LFCA) to help to develop their project ideas further. The winners of the audience award similarly received three months of mentoring. All four winning teams finally presented their ideas live to the 27 Environment Ministers of the EU member states during the EU Environment Ministers' Conference in Berlin at the end of September 2020.

Figure 8: The Code4Green winning teams with Federal Environment Minister Svenja Schulze and European Commissioner for the Environment & Oceans and Fisheries Virginijus Sinkevičius (© BMU / Toni Kretschmer)

The following section summarizes the project solutions of the winning teams and their feedback on the usability of environmental data.

Main Winner - Climate Protection: The Greenspect Project

Aiming to increase the transparency of corporate sustainability efforts and to tackle "greenwashing", the Greenspect Project utilized information that had been made publicly available by companies and had been subsequently collected and published by initiatives such as Wikirate.[30]



[30] https://wikirate.org/

Users of the Greenspect Project app can scan or search for products and see the carbon footprint by the respective companies to help them make a buyers' decision based on the environmental impact of companies. Additional sources such as the European Pollutant Emissions Register (E-PRTR)[31] and transaction data from the European Emissions Trading System (EU ETS)[32] help evaluating the carbon footprint and the sustainability of companies in comparison to their claims. The team expressed the need for better metadata, as well as use-case examples for datasets available. Other problems emerged in the actuality of data. E-PRTR for example only summarizes information from all European member states for previous periods but does not provide up-to-date data. The project currently focuses only on carbon emissions and aims at integrating other sustainability parameters in the future.[33]

Main winner - Biodiversity: WellBeeing

The WellBeeing team developed a prototype that allows beekeepers, farmers, and local authorities to connect via an interactive map to share information about the health of bee colonies. Unfortunately, the team was unable to locate open data containing the status of bee populations, their threats and related location information. Such data is collected largely through research projects and non-profit initiatives[34] that are not necessarily made available to the public as open data. Also, no information on locally applied pesticide amounts could be found. Information on current agricultural use had to be collected using remote sensing data from satellites.

Figure 10: Satellite map designed by WellBeeing to show bee-friendly ecosystems.

[31] https://prtr.eea.europa.eu/#/home

[32] https://www.eea.europa.eu/data-and-maps/dashboards/emissions-trading-viewer-1

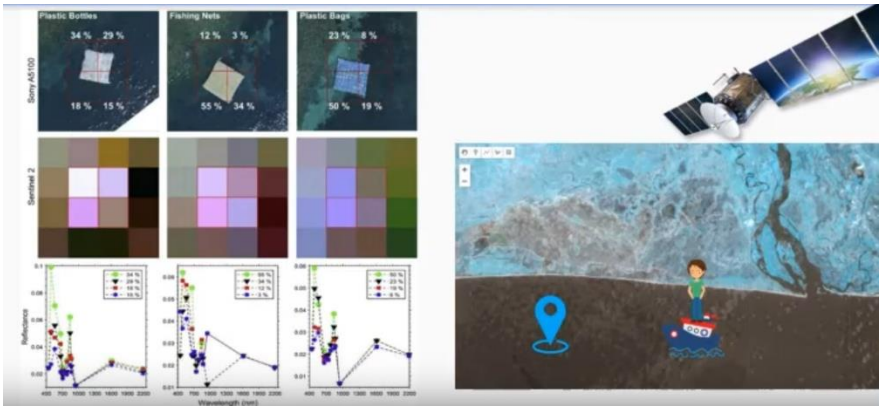
[33] For more information on the project, please refer to their website: https://www.greenspect.org/

[34] https://coloss.org/ / http://www.wildbienen-kataster.de/

Main Winner - Circuit Economy: RosAthene

The RosAthene team combined satellite imagery from the European earth observation programme Copernicus with locally collected data to have artificial intelligence (AI) automatically detect plastic waste in the ocean. The project aims to help clean-up initiatives better coordinate their efforts as well as to predict and prevent coastal pollution before it occurs.

Figure 11: Design of the AI-based algorithm for automatic image recognition of marine pollution.



Extensive training data is required for the successful automatic classification of satellite image data by AI. Since no suitable training data could be identified during the hackathon, the team set up their training data using Google Earth. The team plans to include local groups and citizens in the collection of in-situ data in the future to collect data more directly - this was of course not possible to initiate during the hackathon.

Audience Award Winner: Wisely

Team Wisely developed a prototype that allows private households to adjust their energy consumption according to the availability of renewable energy. Initially, only data from the Open System Power Data Portal[35] was used, which collects and makes various data sets available needed for energy modelling. However, the team emphasised that, despite extensive documentation, this data was difficult to use and exploit without prior experience.

Figure 12: Illustration of the user interface designed by Wisely.

[35] https://open-power-system-data.org/

7. DATA PREPARATION FOR THE CODE4GREEN HACKATHON

Hackathons generally offer data sets that have either been processed or are at least documented. The Code4Green hackathon also adhered to this rule selected and documented European environmental data available for use during the hackathon.

7.1 Data preparation and cataloguing

In preparation of the hackathon, the project partners Terranea and SpaceTec Partners prepared data catalogues for the participants: A total of 52 data sets were identified, evaluated, and catalogued for the individual challenges. To find suiting datasets, the project partners carried out extensive internet research and utilised the vast expert knowledge from the BMU and other partner organisations. All identified potential datasets were analysed and evaluated in terms of their available **metadata**, their **(open) accessibility** and **usability** (licences/formats/API availability). In individual cases, data publishers were contacted directly to check whether data could be made accessible for the hackathon.[36]

There was no separate provision and pre-processing of data on a dedicated hackathon server. Instead, the respective URLs to the corresponding download pages together with the evaluation and information on the data were collected and presented in three challenge-respective catalogues. For some downloads, prior registration was required – these were checked to ensure that the registration process is automated and could be carried out at short notice during the event. For the ENTSOE service, API keys were requested in advance from the provider, as they conduct key generation manually. EU-ECO label data was made available in the form of Excel files by the European Commission.

7.2 Recommendations for future data preparation

Given how extensive and time-consuming data research is, it is advisable to define the overarching work topics and specific challenges at an early stage to limit the necessary selection of data sources. This very early specification helped the Code4Green organising team to search precisely and target-oriented. It is also advisable to **contact experts from relevant authorities early** to get a better sense of further existing data portals and data sets and possible restrictions in access. Early contact with external data providers ensures timely access to data from the private sector, associations, and European institutions. In addition to potential **licensing issues**, **technical questions** must also be clarified in advance. These include for example, can one access the data directly via a provider's platform or are data sets transmitted by email? Does one have to apply for access data or are API keys manually generated? How can datasets, which are transmitted by email, be stored and what should be done with them at the end of a hackathon?

[36] The catalogs included data sets from, for example, the Copernicus Land Monitoring Service (CLMS), the Copernicus Atmosphere Monitoring Service (CAMS), the Copernicus Climate Change Service (C3S), the German Weather Service (DWD), the European Environment Agency (EUA) and the European Soil Data Center (ESDAC).

The processing of data sets before a hackathon is not necessary from the organiser's point of view. Firstly, it can be assumed that participants of a hackathon are experienced enough to carry out the necessary processing steps themselves. Secondly, the necessary processing steps also depend on how the data will be used and may differ from any initial data processing.

In times of **big data**, organisers should check in advance whether access to a **cloud computing environment** will be necessary during project implementation and if so, how this can be most effectively implemented. It is also important to remember that cloud computing represents an extra cost factor. It is possible to incorporate cloud computing service providers as sponsors in the event, but this requires enough lead time to plan and coordinate with them.

8. DATA FEEDBACK FROM THE CODE4GREEN PARTICIPANTS

Participant feedback and **best practice experience** generated important new insights on challenges regarding the use of European environmental data. Following the hackathon, 30 participants in addition to the winning teams were consulted about their experience with European (environmental) data.[37] Feedback was collected on the **data sets provided as well as on their general experience with European (environmental) data**. Among other things, the participants were asked to rate the access to (free) European environmental data, concerning:

- Data visibility,
- Data availability/data gaps,
- Availability of metadata/data information,
- Data access (licence required? Registration required?),
- Data usability (standardisation, availability of APIs, machine readability, etc.).

Furthermore, the participants shared suggestions on how to improve access and usability of European environmental data. The following challenges and suggestions were derived from the interviews:

- *"Finding suitable data is unnecessarily burdensome as it is published across numerous portals from different providers and in different countries. "*

Complementary datasets are often published in country-specific portals. The participants would like to see **a central European contact point** to acquire **cross-border complementary data sets.**

- *"Data portals usually provide diverse and detailed data sets. However, the user guide is often complex and only a targeted search leads to success. "*

In general, the participants praised the high quality and diversity of European environmental data. However, problems often arise in the **complexity of user guides**: outdated or unnecessarily complicated user interfaces of websites are unattractive, outdated search algorithms of integrated databases would furthermore require too specific search queries.

- *"Access to data is not always possible without restrictions. The provision of data in different formats and a straightforward download function would be desirable. "*

The hackathon participants were generally satisfied with the data itself. However, many providers require users to **register** to access the data–for many users this is not very attractive and **is also not in the spirit of free data**. Some providers must approve registrations to their databases manually. This can lead to considerable delays in accessing some of the data sets.

- *"To reduce the need for downloading data, it would be good to have interfaces that facilitate the easy integration of data into one's applications. "*

[37] The feedback process was carried out by doing qualitative interviews in an open interview format with key questions. The qualitative data was not \quantified due to the limited number of participants.

Participants refer here to the need for **standard integrations of APIs** in order to integrate the data more easily into their solutions.

- *"Data must be up to date and regularly updated. "*

Some participants complained that **there are too many outdated data sets** available. Outdated data sets should be better marked and easier to distinguish from current data sets.

- *"Use-case examples should be used to showcase possible uses of the data. "*

Better documentation of the metadata makes it easier to identify the areas of application of the data, especially for users not familiar with the topic. A minority of participants would also like to see a clear description of what the data can be used for, for example by providing exemplary use-cases.

- *"Licences should be clear and easy to understand. "*

Even open environmental data can subject to certain conditions of use.[38] Participants would like to see **clear and easily understandable documentation on the terms of use**. According to feedback from some participants, the terms of use are often inadequately presented - as a result, potential users of the data do not dare to use it.

Feedback from the hackathon participants hereby again highlights the barriers to easy, target-group-oriented access to environmental data.

Figure 14: The participant's feedback at a glance

[38] In some cases, the data provider must be referred to in the illustrations or in the imprint. In some cases, the consent of the provider must be obtained before commercial use.

9. RECOMMENDATIONS FOR ACTION REGARDING FURTHER DIGITALISATION WITH ENVIRONMENTAL DATA

9.1 Data innovators and start-ups

Innovators should live out their entrepreneurial spirit but keep in mind the proven patterns of data business models.

The free availability of environmental data enables comparatively simple planning for businesses. Ideally, the public sector should make data easily available via central portals which can be then integrated into all types of applications. As the data is being collected to monitor and enforce environmental regulations, it can be assumed that respective data will be collected over a long period. The lengthy and often cost-intensive development of own sensor networks, as well as their operation and maintenance, are thus no longer necessary. If a digital business model is based on freely available data, start-ups should nevertheless still be sure to go through **the associated licensing model in detail**. This is the only way to ensure that copyright is respected. Furthermore, innovators should consider that future end customers will assume the continuous availability of services and with it, the associated data.

When using external data, a possible dependency on data providers must be considered. Data providers in the public sector can hardly be expected to engage in so-called "service level agreements" and legally commit themselves to minimise downtimes and even pay compensation in the event of system failures. It should also be considered that, due to changing legal framework conditions, measurement series can be terminated or changed. As such, it would be advisable to have **multiple sources of data**, i.e. obtaining data from several sources that are independent of each other. Additionally, a healthy distrust towards data produced in research projects is advised, as a regular updating of datasets is not possible in the case of missing follow-up financing.

9.2 Environmental data providers

To enable optimal use of data, **data providers should take the perspective of the user**. Considering how one's own data could be used, but also studying different use cases, helps to optimise the provision of data. In particular, the **development and provision of application programming interfaces (APIs)**, as also required by the recent Directive on open data and the re-use of public sector information (Directive (EU) 2019/1024)[39], are state of the art and enable easy communication between systems. They should therefore be developed and provided as a standard feature.

The development of interfaces also offers the opportunity to promote innovation within the authorities and to increase their efficiency. As a result, optimised interfaces can improve interactions between authorities, businesses, and citizens and accelerate work processes. However, the

[39] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=DE

necessary structures and resources to ensure the long-term provision of the data must also be considered.

Collecting environmental data is required by European environmental legislation. In this context, cross-border use of the data should also be considered and strongly supported. For this reason, **standardised and multilingual metadata** are highly advised. Integrated European data portals simplify finding relevant data. The European Data Portal is a positive example in this context. This metadata portal accesses public sector information from different countries. However, due to its extended scope, a targeted and detailed search is needed to find suitable data. **Thematic data portals, such as those provided by the European Environment Agency, make it easier to find data more quickly**.

Interoperability of data from different providers is also necessary for cross-border and cross-sectoral use of data. Environmental and climate issues must be considered in terms of society as a whole and coordination with e.g. agriculture, forestry, transport, and health is vital for long-term solutions. Interoperability ensures the flow of data between the systems of different institutions and enables resources from different agencies to be accessed more effectively and efficiently. **Interoperability involves not only data formats but also processes, guidelines and licences**. The European Commission already provides guidelines for this and countries are free to implement them.[40]

Uniform licensing is also desirable to facilitate the integration of data from different organisations from different member states. Accordingly, it should be checked whether national licences are compatible for example with the European Union Public Licence (EUPL). [41] Additionally, standardisation through common licences such as those of the Creative Commons organisation[42] should be promoted at the European level.

Open data should be offered free of charge. The release or transmission of open data by public companies may be subject to charges. The INSPIRE Directive (Directive 2007/2/EC) allows fees when they are needed to ensure the financial viability of public authorities, especially those that are obliged to secure their revenue. However, fees levied on data can hinder the evolution of innovation.

9.3 Organisers of innovation competitions

Innovation competitions present a valuable opportunity to test new data sources regarding their international usability and how they are perceived by digitally savvy users. The processing and cataloguing of data should be prioritised according to thematic clusters and explained by showcasing direct applications.

When dividing the working groups, attention should be paid to the **complementarity of the members.** Above all, programming skills should be present in all teams. Furthermore, time should be

[40] https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf

[41] https://joinup.ec.europa.eu/collection/eupl/eupl-text-eupl-12

[42] https://creativecommons.org/licenses/

allocated for team members to get to know each other so that motivated teams can be formed in a short timeframe.

Organisers should not overestimate the commercial value of the innovation projects. Many participants simply do it as an intellectual pastime to connect with new people, broaden their horizon, practise their language skills, and to experience intercultural social interaction.

The **mentors‘ complementary business experience** broadens the teams' perspective and expands their horizons of ideas. **Data coaches** have proven to be valuable elements, acting also as indirect communication channels for the organisers to monitor the teams' activity and offer appropriate support when needed. At the same time, the coaches are also able to learn where data sources have gaps that need to be addressed.

Through **targeted mentoring,** particularly promising ideas (usually the winning teams) can be further refined and developed after the event. This way, even projects not initially constructed to be viable in the long-term can turn into promising, sustainable, and scalable product solutions. In the mentoring phase, topics such as the identification of suitable business models, short and long-term financing, possible company forms and pitch-training are vital. In addition to administrative and financial challenges, the different and often cross-border residences of European teams must of course also be taken into account. Before completing the mentoring, it should be ensured that the teams have sufficient networking and have made initial contacts in a possible follow-up financing. Three of the Code4Green winning teams have announced that they intend to pursue their project ideas further as companies.

9.4 Environmental institutions & politics

The project ideas of the Code4Green profited highly from the direct contact with data coaches from national environmental agencies and experts at the BMU. The great interest on the part of the innovation partners in participating in Code4Green shows that **entrepreneurial activities** are justified as a **field of dialogue** and may even provide approaches for interdepartmental cooperation.

Ministry staff and those from other authorities learn how policy fields are perceived and understood by the public through dialogue with teams of innovators. It also helps them to understand where technical language could be simplified and how environmental challenges can be met through entrepreneurial solutions.

Sustainable entrepreneurship is already a **growing field** that will continue to gain in importance in the coming years. It has the potential to become a significant pillar of the economy and has the potential to become a role model internationally.

International innovation competitions offer the opportunity to develop business ideas across borders, even outside the EU Council Presidency. Broad policy fields such as the European Green Deal will continue to give rise to innovations in the field of environmental data processing and networking for years to come.

Authors

SpaceTec Partners bundles unique expertise in the fields of **strategy, management, and innovation consulting** at its locations in Munich and Brussels and is involved in a wide range of data application fields. SpaceTec Partners' client base includes public authorities and agencies in Germany (BMWi, BMVI, DLR, etc.) as well as at the European level (European Commission, European Space Agency, European GNSS Agency, European Environment Agency, etc.). Furthermore, SpaceTec Partners serves a variety of clients in the private sector - start-ups, SMEs, and large international companies.

www.spacetec.partners

As a service provider in the field of **geospatial data analytics & services**, Terranea specialises in the automated processing of large data volumes and the development of location-based services. In addition to classic **geospatial data processing,** Terranea's team increasingly works with algorithms from the field of artificial intelligence to evaluate heterogeneous data. Terranea integrates data from different sources, e.g. social media, IoT sensors, drones and satellites. Smart cities, agriculture and forestry, natural resource management and natural hazard management are some of the application areas. Terranea works for the European Commission, European Space Agency, European Environment Agency, Federal Environmental Agency, as well as for companies from the private sector.

www.terranea.de