remote sensing

Article

Digital Ecosystems for Developing Digital Twins of the Earth: The Destination Earth Case

Stefano Nativi [1,*,†], Paolo Mazzetti [2] and Max Craglia [1,†]

[1] European Commission–Joint Research Centre (JRC), 21027 Ispra, Italy; massimo.craglia@ec.europa.eu
[2] National Research Centre of Italy, Earth and Space Science Informatics Laboratory (ESSI-Lab), Institute on Atmospheric Pollution Research, 50019 Sesto Fiorentino, Italy; paolo.mazzetti@cnr.it
[*] Correspondence: stefano.nativi@ec.europa.eu; Tel.: +39-0332-78-5075
[†] Disclaimer: The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Abstract: This manuscript discusses the key characteristics of the Digital Ecosystems (DEs) model, which, we argue, is particularly appropriate for connecting and orchestrating the many heterogeneous and autonomous online systems, infrastructures, and platforms that constitute the bedrock of a digitally transformed society. Big Data and AI systems have enabled the implementation of the Digital Twin paradigm (introduced first in the manufacturing sector) in all the sectors of society. DEs promise to be a flexible and operative framework that allow the development of local, national, and international Digital Twins. In particular, the "Digital Twins of the Earth" may generate the actionable intelligence that is necessary to address global change challenges, facilitate the European Green transition, and contribute to realizing the UN Sustainable Development Goals (SDG) agenda. The case of the Destination Earth initiative and system is discussed in the manuscript as an example to address the broader DE concepts. In respect to the more traditional data and information infrastructural philosophy, DE solutions present important advantages as to flexibility and viability. However, designing and implementing an effective collaborative DE is far more difficult than a traditional digital system. DEs require the definition and the governance of a metasystemic level, which is not necessary for a traditional information system. The manuscript discusses the principles, patterns, and architectural viewpoints characterizing a thriving DE supporting the generation and operation of "Digital Twins of the Earth". The conclusions present a set of conditions, best practices, and base capabilities for building a knowledge framework, which makes use of the Digital Twin paradigm and the DE approach to support decision makers with the SDG agenda implementation.

Keywords: digital ecosystem; digital twins; green deal data space; remote sensing; earth observation; system-of-systems engineering; metasystems; SDGs; interoperability science Paolo Tarolli

check for updates

Citation: Nativi, S.; Mazzetti, P.; Craglia, M. Digital Ecosystems for Developing Digital Twins of the Earth: The Destination Earth Case. *Remote Sens.* **2021**, *13*, 2119. https://doi.org/10.3390/rs13112119

Academic Editor: Paolo Tarolli

Received: 7 April 2021
Accepted: 26 May 2021
Published: 28 May 2021

1. Introduction

1.1. Key Challenges for Humankind: Global Changes and the Sustainable Development of Society

Society today is facing unprecedented global environmental challenges in terms of food, water and energy security, resilience to natural hazards, population growth and migrations, pandemics of infectious diseases, sustainability of natural ecosystem services, poverty, and the development of a sustainable economy. Climate change cuts across all of these challenges with the potential to greatly exacerbate them [1]. Global environmental change is even greater today than in 2003 (at the time of the first Earth Observation Summit in Washington DC) when governments and international organizations committed to a vision of a future wherein decisions and actions for the benefit of humankind are informed by coordinated, comprehensive, and sustained Earth observations [1]. Earth observation is essential to provide information on the different systems of our planet and detect its changes; observations can be performed via remote-sensing and in-situ techniques.

In 2012 at the United Nations Conference on Sustainable Development (i.e., the Rio+20 Conference), the international community decided to establish a High-level Political Forum on Sustainable Development. UN Member States also decided to launch a process to develop a set of Sustainable Development Goals (SDGs). This process had broad participation from major groups and other civil society stakeholders. On 25 September 2015, the United Nations General Assembly formally adopted the universal, integrated, and transformative 2030 Agenda for Sustainable Development [2], along with a set of 17 Sustainable Development Goals and 169 associated targets [3].

The implementation of the SDGs agenda requires a significant effort in collecting, analyzing, and sharing relevant information worldwide. For this purpose, the UN calls for the development of a distributed and international knowledge platform to facilitate multi-stakeholder collaboration and partnerships, sharing information, best practices, and policy advice among the United Nations Member States, civil society, the private sector, the scientific community, and other stakeholders [4]. Similarly, the European Green Deal [5] and the "European strategy for data" [6] identify shared infrastructures and EU-wide common, interoperable data spaces as critical to addressing the joint challenges of environmental sustainability and the digital transformation.

In this paper, we introduce a new initiative in Europe to develop such shared infrastructure and data space called Destination Earth. We use it as an example to discuss the features, opportunities, and challenges of digital ecosystems with a focus on their role in enabling the development and use of digital twins of the Earth. Together, digital ecosystems and digital twins of the Earth constitute the new approach to leveraging the characteristics of the digital transformation in addressing environmental change at the global and European scales. In the remainder of Section 1, we review key paradigm changes in the transition from a physical to a cyber-physical society and define the concept of a digital twin of the Earth. In Section 2, we introduce the Destination Earth initiative, and argue in Section 3 that its objectives require a digital ecosystems approach. This novel paradigm is discussed from different perspectives (i.e., viewpoints that address the diverse concerns of different stakeholders) in the central part of the paper. In Section 4, we introduce digital ecosystem key principles and patterns; in Section 5, its constitutive mechanisms and metasystem; in Section 6, the information viewpoint; in Section 7, the functional sub-systems; in Section 8, the engineering architecture; and in Section 9, the virtual cloud layer and orchestration services. Section 10 draws the conclusions and identifies the next steps.

We strongly believe that digital ecosystems are the new model to develop a distributed knowledge system among multiple stakeholders, replacing and bringing up to date the concepts of data infrastructures we saw developing in the 1990s and 2000s [7]. The comprehensive discussion of the multiple dimensions and challenges of digital ecosystems presented in this paper fills an important gap in the current knowledge and marks the path for future developments and implementations.

1.2. *Shifting Paradigms in a Digitally Transformed Society*

In more than 20 years, technological change has substantially transformed industry, the economy, and social behaviors worldwide. Universal connectivity to the "network" and the consequent emergences of innovative and global economic models (Web, Data, and Platform economies) have profoundly changed all the sectors of our society, including the scientific one.

Starting from the 1990s, the diffusion of the Internet beyond the academic and military sectors to the business community and civil society was greatly facilitated by the development of more intuitive navigation tools such as the World Wide Web, and later by the diffusion of mobile-phone based Internet services. The initial paradigm to share data across the Internet was based on search and discover services, with the aim of finding, downloading, and using the data locally. Digital data infrastructures were designed and developed to apply this interoperability paradigm. In the environmental and geospatial domains, valuable examples were the Spatial Data Infrastructures (SDI), such as INSPIRE (https://inspirehep.net/, accessed on 1 May 2021) and the GSDI (http://gsdiassociation.org/, accessed on 1 May 2021). This interoperability paradigm had important merits in pushing metadata formalization and data encoding standardization. However, it also demonstrated its limits in solving semantic, pragmatic, and contextual interoperability—see for example the issues of analysis-ready data and sound data re-usability. Finally, non-technical interoperability issues, such as the diversity of data policies, degree of openness, ownership, and fitness for purpose, hindered the widespread use of this approach.

With the advent of the Digital Transformation, the interconnection between the physical and the digital world has become almost complete: economic, industrial, and social relationships have been moved to the "cyber-physical" world, where all the relevant stakeholders are included more easily and can intensively cooperate in generating the knowledge required for addressing a given purpose. Besides the interaction services and tools, the "cyber-physical" world also provides scalable analytical and interpretation tools and services, enabled by virtualization technologies such as cloud platforms and infrastructures. These capabilities are increasingly interconnected to the physical entities and processes shared on the "cyber-physical" world. As a result, it is now possible to go beyond the interoperability paradigm applied by data exchange infrastructures and address most of its shortcomings. In the "cyber-physical" world, interoperability is pushed from the level of sharing measured/observed data to the one of sharing the information and knowledge generated by analyzing those data. The volume and heterogeneity of data produced daily by a Digitally Transformed society is too high for having stakeholders process and analyze them locally in an effective and sustainable way. Therefore, policy, industrial, and economic organizations do not find the SDI discovery and access paradigm particularly useful any longer. It is up to the "cyber-physical" world (i.e., the digital tools and services provided by the cyber-physical world seen as a distributed system) to aggregate, harmonize, and analyze big data for generating insights and knowledge. There is no longer a distinction between the local digital environment and the network, as most organizations store and process the data on the network itself. We see therefore a new interoperability paradigm that can be defined as the "information and knowledge sharing" paradigm.

An important and related paradigm is that of datafication [8], i.e., the conversion of all aspects of our life into quantified data [9]. This paradigm aims at generating actionable intelligence from data streams and largely builds on three digital processes [10], as depicted in Figure 1:

1. (Big) **Data collection**: the collection, aggregation, and contextualization of digital artefacts and digital "footprints" constantly generated by humans, machines, and real objects connected to the network. This pervasive practice produces what is generically referred to as Big Data. Cloud-based data centers are instrumental to this practice. Noticeable sources of Big Data include social networks, public government and e-commerce procedures, Internet of Things (IoT), and the new generation of remote sensing instruments, e.g., satellite, aircraft, and drone (i.e., remote sensing) sensors.
2. **Deep insights generation**: the recognition of valuable insights by analyzing the big data collected. This is commonly achieved by using big data analytics techniques, i.e., advanced analytic techniques against very large, diverse big data sets that include structured, semi-structured, and unstructured data from different sources and in different sizes in the order of terabytes/zettabytes. Today, these practices make largely use of advanced data management systems and data-driven artificial intelligence (AI) technologies, e.g., machine and deep learning models. For example, scientific methods in remote sensing are changing because of the impact of information technology and data-driven AI to generate insights.
3. **Insights interpretation and actionable intelligence generation**: the interpretation of the generated insights to develop profiled intelligence according to users' needs. This is achieved through specialized online platforms that interact with users and

stakeholders and provide personalized services. This approach offers a richer user experience, applying the principles of the platform economy.

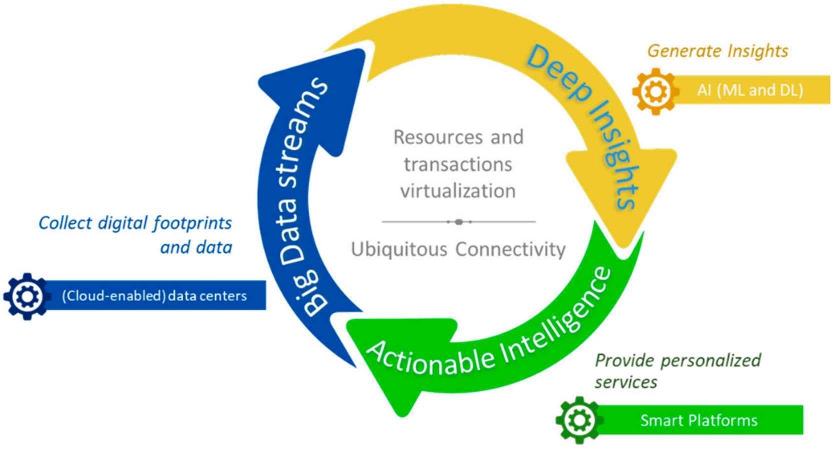

Figure 1. Conceptual "datafication" paradigm to generate actionable intelligence from data streams.

The application of these two disrupting paradigms to study Global Change and Sustainable Development, according to some researchers [10], introduced a new scientific model that is not only characterized by the collaboration between a large amount of diverse scientific disciplines (e.g., natural sciences, social sciences, and humanities), but also by trans-disciplinary knowledge sharing and co-creation. Scientists must collaborate closely with actors from civil society, industry, and public institutions in order to address socially relevant research questions. This innovative discipline, called Big Earth Data (BED) science [10], aims at studying the set of natural and social phenomena characterizing the Earth system, a class of entities encompassing the local and global changes affecting the natural cycles, the deep under-surface processes, and the interconnections with human society (i.e., our social and economic systems). BED science builds on Data Science and researches the Earth system in a new holistic, multi-disciplinary, and trans-disciplinary way. One of its important objectives is the systematic comprehension, modelling, and implementation of processes to generate information from data and provide the knowledge required by engineers, scientists, and decision makers.

Today, the datafication and the information and knowledge sharing paradigms are at the core of the second generation of IoT platforms [11], which combined with BED science underpin the development of Digital Twins of the Earth [12] discussed in the next section. In particular, the advent of IoT has significantly evolved remote-sensing capacities. IoT sensors enabled fast and cheap acquisition of data from billions of interconnected devices deployed across the globe.

1.3. Digital Twins of the Earth

There are several definitions of Digital Twins (DTs) [13], reflecting the diverse concerns of the industrial, scientific, and standardization sectors, which have been working on their description and realization. According to El Saddik [14], a DT is "a digital replica of a living or non-living physical entity". By bridging the physical and the virtual world, data are transmitted seamlessly, allowing the virtual entity to exist simultaneously with the physical entity. The DT concept (see Figure 2) consists in decoupling a digital system from its physical entity/entities, making it easier to change one without changing the other. It also allows the utilization of advanced data-driven modeling procedures in order to generate those insights that could not be carried out using the traditional observation models [12]. Referring to Figure 2, a DT includes three main components: the physical entity, its digital representation, and the connection between the two [12]. The main interaction features characterizing a DT are interoperability, information model, data exchange, administration, synchronization, and push mode (publish and subscribe model) [15,16].

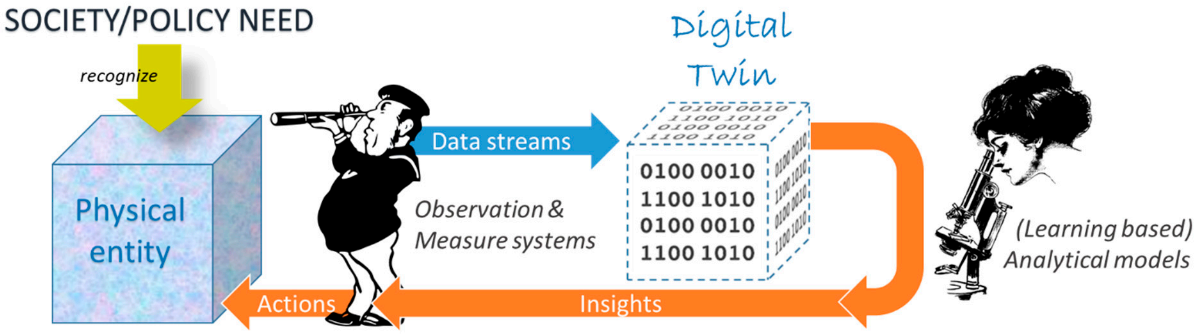

Figure 2. DT philosophy and archetype.

It is useful to distinguish DTs from two other concepts that are often used with an equal meaning: digital model and digital shadow. The three concepts implement different levels of integration between the physical and digital entities [17]: a digital model does not implement any form of automated data exchange between the physical and the digital entities (i.e., in Figure 2, the "Data stream" and "Actions" connections are manual); a digital shadow implements an automated data stream between the state of the physical entity and the digital one (i.e., the "Data stream" connector), while a DT implements automated data exchange in both directions between the physical and the digital entities.

DTs have been around for decades, especially in industrial processes. However, with the recent advent of transformative digital technologies, DTs are changing most of the sectors in society, providing the most advanced pattern to make the physical and the digital worlds interact [18]. This is also true for the scientific sector, and in particular those disciplines that are engaged in understanding and addressing the Global Change effects. Thanks to the development of DTs, it is possible, for the first time, to envision a digital replica of important natural and social phenomena and processes that characterize our planet. It is possible to simulate and predict their behavior by leveraging the large amount of environmental data, available from remote sensing and in situ observations, along with the advancement in AI technology. Academia and research are asked, in addition to improving modelling techniques, to further focus on data optimization and interoperability with modelling platforms [19]. In particular, the diverse scientific sectors must work on four challenges to fully embrace the DT opportunities: (a) to unify data and model standards, (b) to share data and models, (c) to innovate services, and (d) to establish fora to exchange views and knowledge.

Today, the terms "DTs of the Earth" or "Earth DTs" are utilized by several scientific and engineering communities, including: Earth observation [12], space agencies [20], climate research [21], oceanography [22], digital Earth [23], and meteorology and natural disaster risk reduction [24]. Considering the diverse aspects recognized by these communities, it is useful for the scope of this manuscript to define a DT of the Earth as the digital replica of an Earth system component, structure, process, or phenomenon obtained by merging digital modelling and real-world observational continuity –i.e., remote, in-situ, and synthetic data streams. A DT of the Earth must be seen as a living digital simulation model that updates and changes as its physical counterparts change; therefore, a DT of the Earth continuously learns and updates itself.

To move towards sustainability, society must realize a green transition that, according to many, is not achievable without an equally important digital transition. For both transitions, DTs of the Earth play a very important role—as shown in Figure 3. To facilitate the implementation and the evolution of these cyber-physical simulation and projection instruments, it is useful to apply the digital ecosystem paradigm. For this reason, there are several initiatives and programs considering this model at the preset time. At the European and global level, some instances are depicted in Figure 3. In the rest of the manuscript, we use as examples the Destination Earth and GEOSS initiatives.

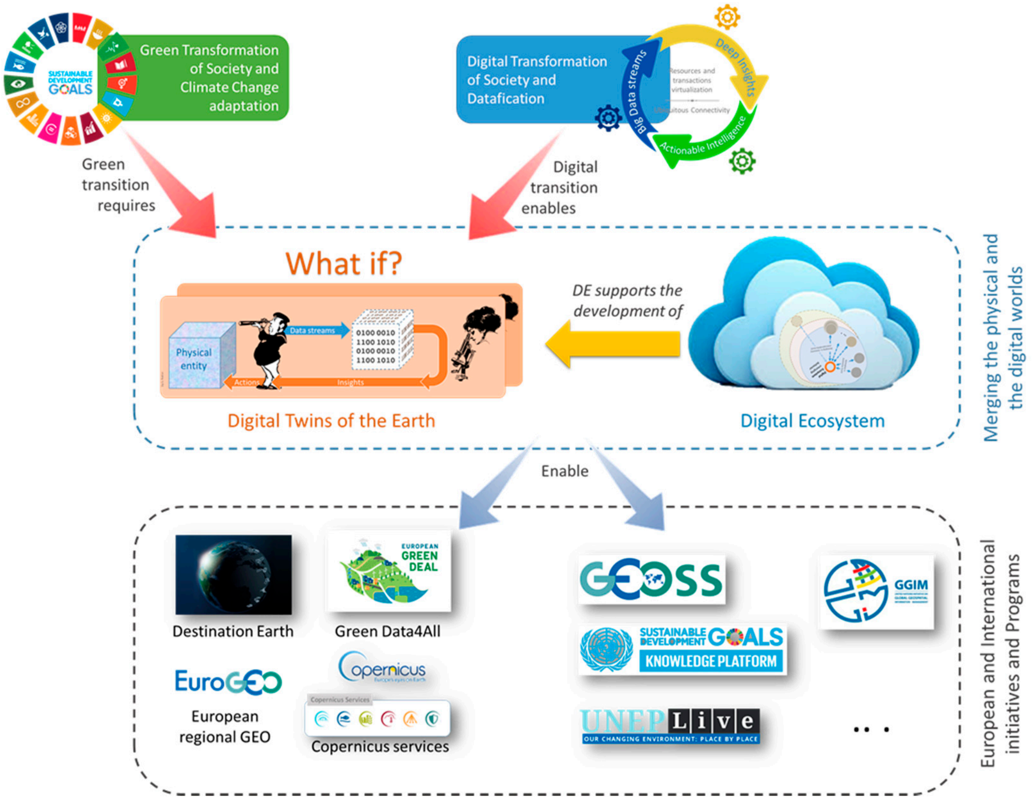

Figure 3. The role of DTs of the Earth in implementing the Green and the Digital transitions of society.

2. The Destination Earth Knowledge Framework and the Green Deal Data Space

The concept of DTs of the Earth is at the heart of the recent and ambitious EC initiative, named Destination Earth [25]—also known as DestinE. This initiative was introduced by the European Strategy for Data [6] as a concrete action contributing to realizing the Common European Green Deal data space, one of nine data spaces envisaged by the Strategy in different thematic domains [6]. The Green Deal data space aims to use the major potential of data in support of the Green Deal priority actions on climate change, circular economy, zero pollution, biodiversity, deforestation, and compliance assurance. Remote sensing and in situ observations are called to provide a comprehensive database of time series to be automatically analyzed by AI and identify changes. According to the European Strategy for Data, the ambition of Destination Earth is to "bring together European scientific and industrial excellence to develop a very high precision digital model of the Earth. . . . offer[ing] a digital modelling platform to visualize, monitor and forecast natural and human activity on the planet in support of sustainable development thus supporting Europe's efforts for a better environment as set out in the Green Deal" [6]. Destination Earth is also seen as contributing to the EU digital strategy [26,27] based on the principles of ethics, democracy, fairness, and open autonomy. Noticeably, Copernicus and other satellite big data streams will contribute to form the core of Destination Earth.

Conceptually, the goal of Destination Earth is to develop a dynamic, interactive, multi-dimensional, and data intensive replica of the Earth (system), which would enable different user groups (public, scientific, private) to interact with vast amounts of natural and socio-economic information. At the heart of Destination Earth is a common infrastructure providing access to data, advanced computing (including high performance computing), software, AI applications, and analytics. This infrastructure then serves different DTs replicating various aspects of the Earth system, such as weather forecasting and climate change, food and water security, global ocean circulation and the biogeochemistry of the oceans, and more yet to be defined [25]. Destination Earth and the Green Deal data space can be major contributors to the SDGs agenda by building on the concept of DTs and applying the datafication concepts depicted in Figure 1.

2.1. Framework Main Requirements and Constraints

An SDG knowledge framework must address important technological and non-technological challenges and constraints, including multi-organizational dimensionality, multi-disciplinary domain, large heterogeneity of contributing components, and a high degree of evolvability and long-term sustainability. To address these challenges, Destination Earth recognized the following main needs:

- To offer a scalable (i.e., cloud-based) core platform providing users with access to:
    - data—characterized by a highly distributed and heterogeneous nature;
    - analytical software—consisting of heterogeneous process and data-driven modeling algorithms, services, and tools;
    - infrastructures—consisting of a mixture of purpose-oriented centralized and distributed components, including HPC (High-Performance-Computing), IoT platforms, and public and private cloud clusters.
- To support the development and utilization of a number of thematic DTs, allowing different types of users (from policy makers to scientific domain experts) to access and interact (at different levels of complexity) with thematic information, services, models, scenarios, simulations, forecasts, and visualizations. DTs must be seen as vertical digital components plugged into the core platform that offers common horizontal services and scalable capacities.
- To enable users to build applications on top of the core platform and integrate their own data and/or analytical software.
- To build on existing (and next coming) operational and scientific infrastructures, platforms, and systems.
- To form the enabling core of a European earth observation, geospatial data, and earth systems applications framework (i.e., the Green Deal Data Space).

The requirements and constraints of Destination Earth are based on the analysis of about 30 policy-driven use cases [28] and a survey of similar DT initiatives at national and international levels [12]. The outcome of this analysis identified three priority thematic areas: disaster risk management (from extreme weather-induced natural disasters), climate adaptation (food and water supply security), and digital oceans (around food and energy), with a fourth area including (as yet) less mature applications such as waste management and health and smart cities/urban digital twins.

The Destination Earth high-level architecture [29] builds on the above user requirements and constraints. It also acknowledges that, to address the SDG agenda, it is necessary that the proposed knowledge framework leverages the existing (and heterogeneous) systems—i.e., builds a System-of-Systems. This framework must be able to address evolving policy challenges; for example, through an increasing number of DTs, which then need to be maintained and allowed to evolve with new data, actors, governance challenges, etc. Moreover, the framework must face the challenges of technological, social, and geopolitical landscapes that are constantly changing. For these reasons, the proposed architecture applies to the digital ecosystem paradigm.

A similar conclusion also emerged from a decade of experience in developing the GEOSS (Global Earth Observation System of Systems) infrastructure, which is essentially based on a bottom-up system-of-systems (SoS) approach. Conceived in 2005, the GEOSS infrastructure (managed by the Group on Earth Observation: GEO) became operational around 2015 [30] and continues to develop based on the changing needs of the GEO Community and the evolving global landscape—i.e., the political, economic, technological, and social requirements and conditions.

3. The Digital Ecosystem Nature of the Dynamic Knowledge Framework

The Digital Ecosystem (DE) paradigm seems to fit the constantly evolving needs of system-of-systems like GEOSS and Destination Earth particularly well. Stemming from the concept of natural ecosystems introduced in the biology domain [31], the ecosystem paradigm focuses on a holistic view of diverse and autonomous entities—i.e., the "biotic" component—sharing a common environment—i.e., the "abiotic" component. In search of their own benefit, they interact and evolve, developing new competitive or collaborative strategies, and, in the meantime, modifying the environment. The adoption of the (digital) ecosystem perspective allows the capture of the evolutionary systemic process of DTs and the creation of the necessary conditions for SoS like Destination Earth, GEOSS, or other SDG knowledge frameworks to be sustainable and thriving. The holistic view characterizing the ecosystem paradigm is able to capture the need for the diversity of actors and expertise, directing the attention to synergies and connections, and putting the focus on the capacity to produce common outcomes—what ecologists call the ecosystem services—over time. Due to these features, the ecosystem archetypes have been successfully used to model complex collaborative and competitive social domains, such as business [32] and software realms [33].

In the business environment, co-evolution—as determined by the complex interplay between competitive and cooperative business strategies—is a central concept of business ecosystems. For instance, companies tend to "co-evolve" around a new innovation, working cooperatively and competitively to support new products and satisfy customer needs [34]. Internet and globalization processes made software platforms some of the main engines of innovation, raising an entirely new type of business ecosystems, the software ecosystems, which can be seen as collections of organizations that are related through a software technology or a software related concept [33].

While the common interest in software ecosystems is the evolution and surviving of a keystone technology, for DE, the stakeholders' common interest must be building and operating the datafication value-chain (see Figure 1). It is necessary to them (i.e., the ecosystem "biotic" component) for improving their own utility while contributing to generating value for society (the ecosystem service) in alignment with policy and social targets in a given environment, i.e., the ecosystem "abiotic" component. The strategy should primarily consist of collaborating with each other to extract the maximum value and utility from the available "abiotic" resources (e.g., Earth observation datasets, analytical software, and computing capacities) to survive, thrive, and co-evolve. Therefore, a successful DE (which can be seen as a prominent feature of a data-driven economy) would "bring together data owners, data analytics companies, skilled data professionals, cloud service providers, companies from the user industries, venture capitalists, entrepreneurs, research institutes and universities" [35]. DE should promote, for instance, the necessary actions to allow these stakeholders to interact seamlessly within the digital market, leading to business opportunities, easier access to knowledge, and resources [36]. Policy can significantly contribute to DE development by bringing the relevant players together and by steering the available financial resources that facilitate collaboration among the various stakeholders in the data economy perspective. A DE is enabled by at least three contextual conditions under which actors in the ecosystem operate and which motivate, direct, and/or constrain their actions as providers, intermediaries, or consumers [37]:

1. The regulatory condition—laws, policies, standards, and agreements that have a bearing on how the components of the ecosystem are structured and how they interrelate.
2. The institutional/organizational condition in which the actors operate—each organizational and/or institutional context provides a set of shared social and cultural values, which influence the actors who operate within that particular context [38]. These values inevitably propel and restrain the behaviors of actors in the ecosystem [39].
3. The current ICT capabilities condition—the computational, data storing, analytical software, and network elements, along with the communications protocols that interconnect these elements with the network operators and users [37]. In this context,

for example, the main capabilities are Cloud computing and HPC, AI analytical software, Internet, and the Web. They are all key enabling technologies that respectively introduce new actors to the ecosystem.

Each of these three conditions must be carefully considered and regularly managed by DE operators. As covered in the rest of the manuscript, they can be addressed by defining a metasystemic level that controls the appliance of principles, conditions, and patterns characterizing a DE and its evolution in time.

4. Principles and Patterns for a Thriving Digital Ecosystem

The design and implementation of the architectural framework of a DE must apply a set of principles and related system and software design patterns [40–42].

4.1. Evolutionary Development of a Digital Ecosystem (Evolvability and Resilience)

A DE must be flexible and dynamic because, as indicated in the case of GEOSS in Section 2.1, it operates in an environment in which technology, policies, and use needs are in constant evolution. Therefore, its architecture must implement the following patterns:

- High flexibility and modularity level to effectively de-couple the enabling infrastructure and platform system (i.e., a system-of-systems) from the thematic applications (e.g., implemented through digital twins). The aim is to implement a highly scalable system that implements the elasticity features required by its business goals and objectives.
- Independence from any specific provider, technology, or license—to secure industrial competitiveness and defend the perceived values that characterize the local or regional area where the DE operated and is supposed to create social value.
- Preserve and facilitate the co-evolution of the "digital species" populating the digital environment in which the DE operates. This is important to maximize the DE resilience.
- Equal opportunities of access to the infrastructure and affordability for small organizations and across the ICT value chain.
- Meta-systemic governance of the ecosystem to govern its evolutions, adaptations, mutations, and strains, as well as putting in place the necessary instruments to ensure the maintenance of the DE services offered.

These principles are at the base of the architecture proposed for Destination Earth [29] so that it can develop as a self-sustaining digital ecosystem.

4.2. Emergent Behavior of a Digital Ecosystem as a Whole

A DE is effective only if it generates a value for its member species that is greater than (or different from) the value that they would generate without being part of the DE. Such utility can emerge autonomously as the result of unpredictable interactions, or it can be planned, facilitated, and governed by an internal or external agent. We use here the concept of "emergence" in the weak sense of the so-called epistemological predictive emergence, referring to the practical impossibility of predicting the properties of the ecosystem from the properties of its components. Even natural ecosystems survive and evolve when member species get some utility from being part of them. However, it is well known that disruptive changes can break this dynamic equilibrium and make a natural ecosystem disappear. Therefore, to preserve the value that the natural ecosystem provides to human society (measured by a set of selected ecosystem services), today, it is common to have normative and management intervention. Similarly, for DEs, the value of generating and sharing knowledge to better address societal challenges is potentially created through new and unpredictable ways of interacting among the member species while they co-evolve, maximizing their own utility. However, to favor and maintain such capability and thus preserve the social value of the DE, careful planning and management is necessary. To this aim, the DE design must consider the following aspects:

- Belonging versus autonomy. Like every complex system, a DE is based on the paradox of autonomous and yet cooperative entities, i.e., enterprise systems reduce their autonomy and optimization to precise levels and acceptable risks and thus collaborate [43]. To belong to a DE, for example, the enterprise systems must decide on diverse reductions in their degree of technology and policy autonomy and optimization, which can vary over time. Where useful, in an initial stage, this compromise can be encouraged by financial and normative instruments to demonstrate that a possible initial disadvantage brings several utilities in the medium and long-term.
- Common versus enterprise value. To be effective, a DE is called to preserve the utility of the contributing enterprise systems (i.e., allow them to maintain a good degree of autonomy and diversity) while enabling the emergence of a common value. For instance, Destination Earth has contributions from different systems, each with it its mandate and need to deliver value to its stakeholders. This value has to be maintained whilst the interaction of all the systems needs to create additional value of the initiative as a whole. In other terms, the ecosystem value complements (e.g., augmenting, optimizing, or diversifying) the value created by each enterprise system.
- Critical mass. A critical mass of services and users is needed to guarantee a return-of-investments and facilitate the ecosystem resilience. Otherwise, the risk is that the individual value of belonging to the DE for enterprise systems is too low, and they may prefer complete autonomy, abandoning the DE.
- Cost-effectiveness. It is necessary to avoid duplication of efforts and resources through the use of shared components (built on infrastructure of organizations that can leverage economies of scale) and use of standard and widely adopted technologies, e.g., Internet and Web ones. Partnerships between the public and private sectors can also help in reducing costs and harnessing innovation from a broader stakeholder base.

4.3. Geographical Distribution and Heterogeneity of the Digital Ecosystem Constituent Systems

DE will commonly deal with "Big Data" challenges requiring appropriate strategies to manage volume, velocity, variety, and value of data from multiple sources. In the Earth observation domain, past experiences [30,44] show the importance of moving analytical software where data are located and not vice versa. Network computing (i.e., input/output control and management) is one of the most difficult areas for the cloud. For example, considering the Destination Earth requirements on analytical speed as well as computing and storage scalability, it is essential to optimize the nexus of four (often conflicting) minimization factors: data movement minimization, data replication minimization, analytics time lapse minimization, and energy consumption minimization, whilst preserving distributed data ownership. To realize that, some important patterns are recommended.

4.3.1. Computing Continuum (Osmotic Computing)

The digital transformation computing technologies (i.e., mobile, edge, fog, and cloud computing) introduced the concept of the computing continuum. By applying such a paradigm, at runtime, applications can choose to execute parts of their logic on different infrastructures that constitute the continuum, with the goal of minimizing latency and energy consumption while maximizing availability [45]. The same holistic paradigm to support the revolution of IoT services and applications is also called osmotic computing. This paradigm entails a holistic distributed system abstraction, enabling the deployment of lightweight microservices on resource-constrained IoT platforms at the network edge, coupled with more complex microservices running on large-scale datacenters [46].

4.3.2. Optimizing vs. Satisficing Design Strategy

Important differences distinguish traditional systems and DE engineering. In particular, while the primary objective of system engineering is the optimization of system performance, ecosystem engineering is concerned more with satisficing performance [47]. The term satisficing is a neologism coined by the economist Herbert A. Simon to describe a decision-making strategy aiming at finding a satisfying and sufficing solution instead of the optimal one [48]. Traditionally, rational choice theory assumes that decision makers review all available options and select the best one, according to a set of criteria. However, actually, in complex environments, this strategy can be too expensive and time-consuming, and they simply select the first "good enough" option [49]. For example, the satisficing approach is the preferable (i.e., more rational) choice for decision-making when optimization is not feasible or too expensive to achieve due to complexity or to near-real time requirements. DE design is a typical case of an (architectural) decision-making process in a multi-criteria and multi-objective context. It faces the same challenge of choosing the most rational approach: finding the optimal solution or just accepting a satisficing one.

4.4. Effective Governance of the Independent Enterprise Systems of a Digital Ecosystem

A DE is supposed to build on existing building blocks, what we call the enterprise systems, stimulating the creation of new elements to fill possible gaps and realizing a complex SoS (or super-system) [50]. In a DE providing DTs of the Earth to support environmental policies, including the UN SDGs, these blocks are very heterogeneous (e.g., legacy systems with different objectives, technological characteristics, and diverse content) and are managed by different organizations. Consequently, the success of a DE mostly depends on the appropriate governance of the ecosystem as a whole. Such governance must define and apply the set of rules and principles that will help steer the ecosystem evolution and effectiveness through the many changes occurring in the political, social, cultural, scientific, and technological environment where it operates. In the case of the Destination Earth ecosystem, different SoS governance styles were considered [29], encompassing diverse management and control configurations, from fully distributed to fully centralized [51,52]:

- Virtual governance: there is no central management authority and no centrally agreed-upon purpose for the system-of-systems.
- Acknowledged governance: there is a central management organization without coercive power to run the system-of-systems, but constituent systems interact more or less voluntarily to fulfill agreed-upon central purposes.
- Collaborative governance: like in the directed governance below, there are recognized objectives, a designated SoS manager, and resources allocated for the system-of-systems. However, like in the acknowledged governance, the normal operational mode of the constituent systems is not subordinated to the central managed purpose, and they retain their independent ownership, objectives, funding, and development and sustainment approaches.
- Directed governance: an integrated SoS is built to fulfill specific purposes and centrally managed long-term to continue to fulfill those purposes, as well as any new ones the system owners might wish to address.

The 30 policy-driven use cases considered for Destination Earth involved many different actors such as infrastructure managers, data custodians, data scientists, analytical software providers, simulation experts, policy officers, senior decision-makers, and operations managers [28]. In this context, all the stakeholders agreed on importance of designing a multi-level governance [53] for this initiative. The analysis of the use cases suggests that the collaborative and acknowledged governance styles would be the most appropriate to satisfy the business goals and objectives of Destination Earth.

5. Digital Ecosystem Constitutive Mechanism, Metasystem Transition, and Viability

DEs are constantly subject to changes for both internal reasons (e.g., enterprise system changes, new system addition, etc.) and external reasons (for example, changes in the societal and technological environment where the DE operates). Without any control, those changes could (in principle) be very disruptive and make it impossible for the DE to survive and, in particular, to continue the provision of its services for a social utility. Therefore, a DE needs the capability to detect changes and respond to them [54]. Figure 4 shows the dynamic and evolving environment that characterizes a DE—in particular, internal pressures are represented by an enterprise system (managed by Organization #X) that evolves over time, and by a new organization (i.e., Organization #Y) that decides to join the DE at a given time.

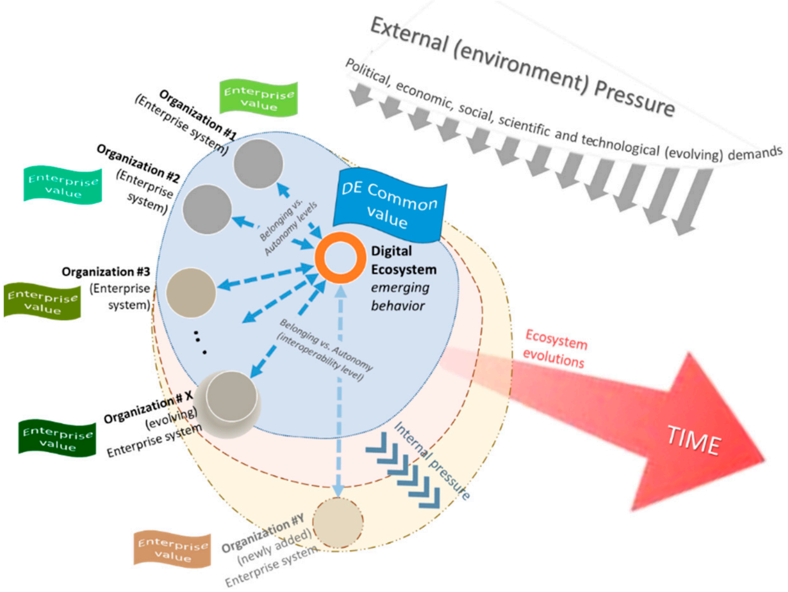

Figure 4. The evolutionary nature of a DE, its main elements, and the contextual pressure it must address.

Referring to Figure 4, a DE constitutive mechanism is achieved when a DE common value (social utility) provides choices for the enterprise systems to fulfill their expected enterprise value (enterprise utility) in a holistic interaction [43]. Therefore, it is essential to define and manage the acceptable behaviors of the constituent parts, their time evolutions, and the consequent communication and interoperability levels that must be supported by the DE. A DE must be able to support the diverse and evolving levels of belonging (to the DE) versus autonomy (from the DE), which are decided by each of its enterprise organizations. These compromises characterize all the different constituent systems and can evolve in time. All these introduced aspects, largely, pertain to the governance sphere.

The strategy to make the DE a viable system, i.e., a system that is able to sustain itself over time [55], consists in recognizing as invariant the process of change, rather than the elements that change [56]. To this aim, the DE should be seen as a collection of parts that are dynamically related, and such dynamicity must be carefully formalized and governed. This can be achieved through the development of efficient communication and control functions. Therefore, to survive, a DE must implement a cybernetic (from the Greek word κυβερνητική meaning "governance" etymologically related to "govern" through Latin gubernare) system, including dedicated components implementing the required communication and control functionalities.The structure of the controlling components varies according to the selected system typology, ranging from the mere expression of emerging properties, as in fully distributed systems (e.g., virtual system-of-systems), to a coercive governing body, as in fully centralized systems (e.g., directed system-of-systems).

The implementation of a cybernetic mechanism is called a metasystem transition [57], since, as depicted in Figure 5, the communication and control structures, i.e., what is called a metasystem [58], constitute a system on their own, which stands "above" and "beyond" the controlled SoS in the DE. In a DE, the enterprise systems work at a first order operation level, while the DE metasystem works at a second order level, to control the SoS evolution and deal with meta-elements and items, e.g., providers, consumers, harmonizers, orchestrators, aggregators, governing boards, interoperability standards, etc. The metasystem communicates through a governance language—a metalanguage [59].

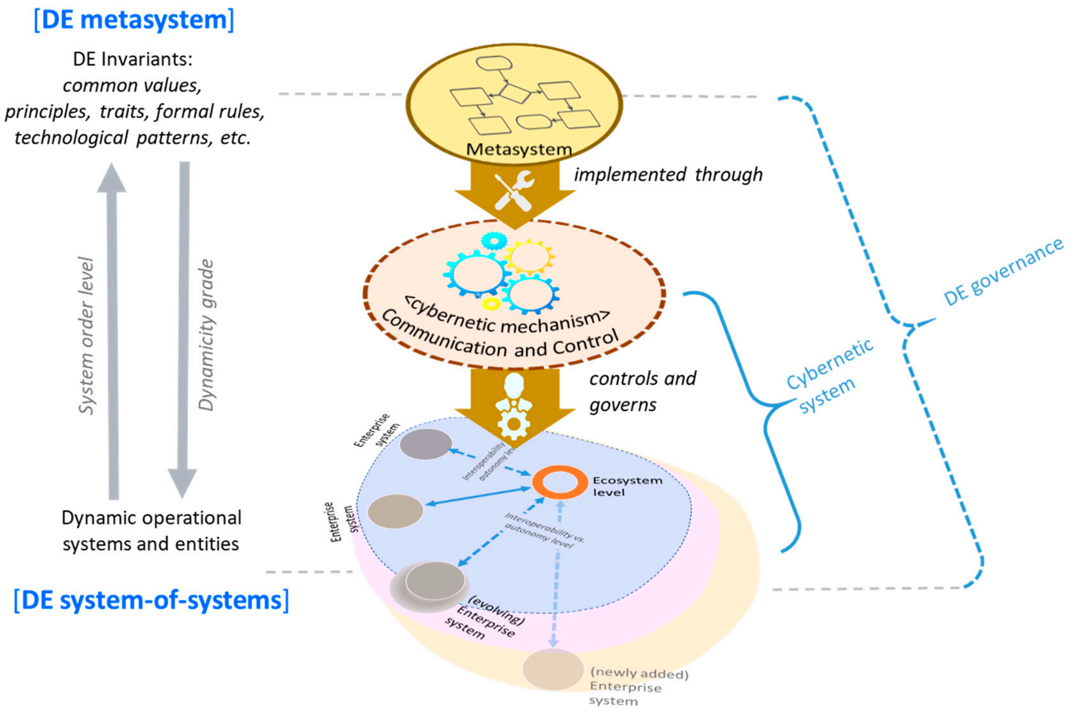

Figure 5. The main (and multi-level) architectural elements implementing a viable DE.

In summary, a DE metasystem is the set of mechanisms that creates, controls, and enables the evolution of the ecosystem to reach its given goals over time and become a cybernetic system. The metasystem assures the SoS invariance (essential for its sustainability and evolution) by introducing a set of DE invariants, including the DE common values, principles, traits, formal rules, technological patterns, etc. Referring to Figure 5, the main architectural elements and their relationships to be introduced are:

- Metasystem schema—including the DE invariants (e.g., DE common value and patterns).
- Governance style and cybernetic mechanisms—to apply the metasystem rules.
- SoS architecture topology—largely depending on and governed by the governance style, the cybernetic mechanisms, and the agreed DE Common Value.

6. Digital Ecosystem Information View

DE dealing with DTs of the Earth (such as the Destination Earth ecosystem), are asked to share a set of main digital entities along with enabling services—a simplified schema is depicted in Figure 6. These entities and their relationships may be grouped into three principal strands:

(a) DE Policy/Business entities;
(b) DE Framework/software entities;
(c) DE constituent infrastructure entities.

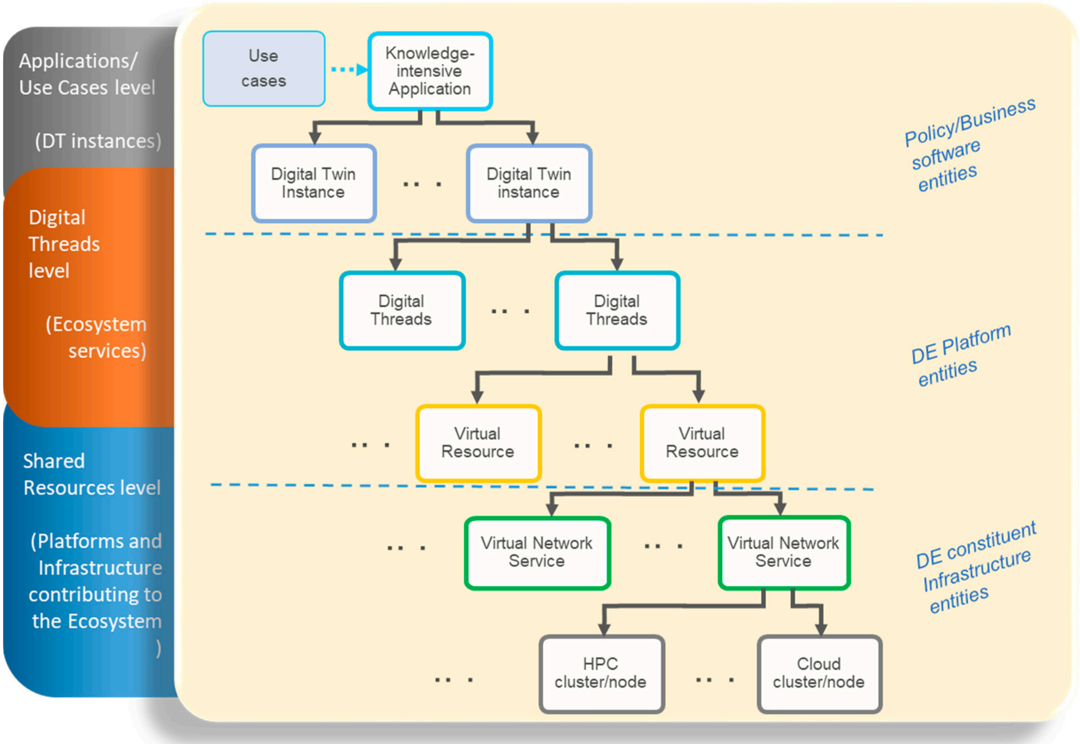

Figure 6. Main high-level entities (and their simplified relationships) shared by a DE providing DTs of Earth.

Referring to Figure 6, a DE dealing with DTs of the Earth will support the following conceptual and digital entities:

- Digital Twins (DTs) are the DTs recognized by the use cases that the ecosystem is called to support. In the case of Destination Earth, the initial DTs will address disaster risk management from extreme weather-induced natural disasters, climate adaptation (food and water supply security), and digital oceans including food and energy [28].
- Digital Threads are the “the flow of data fueling the digital insights behind customer-centric experiences” [60]. A digital thread can be defined as a record of the entity/system lifetime, from its design to its cancelation [61]. For the DE, the digital thread concept is crucial to conceive, implement, (re-)use, and assess a DT. For example, a digital thread should deal with the instantiation and operation of the DT showed in Figure 2. Digital thread conceptual entities are commonly implemented by artifacts like workflows and services orchestration processes.
- Digital/Virtual Resources, including:
    - datasets, along with relevant metadata schemas and the associated data streams, e.g., remote sensing imagery, IoT data, social/economic time series;
    - process-based analytical software, along with relevant scientific algorithms, models, and workflows, e.g., physics-based models;
    - learning-based AI analytical software, along with relevant neural network models and configurations, and workflows, e.g., Machine Learning and Deep Learning models;
    - complex DTs, which may result in a composition of two or more elementary DT instances. In this advanced context, DTs will be managed as another kind of digital resources.
- Virtual Network Services provide functions deployed on cloud-based virtual machines in the hosted network services environment, in the public cloud or premise-based Virtual Machines subject to availability. An important role for them is to simplify the ecosystem network by bringing together distant and disparate resources in a more

efficient way. In a DE, they are used to implement multiple combinations of network functions and/or multiple vendor services at remote and cloud locations, optimize the automation and orchestration services, and implement rapid service scaling.

- HPC and cloud clusters/nodes are physical or virtual machines providing computer, storage, and networking services. They are the key entities to implementing a multi-cloud approach. They can be also seen as digital/virtual resources.

7. Digital Ecosystem Functional Sub-Systems

Building on the architecture layers and applying the recognized ecosystem and software patterns, a DE dealing with DTs of the Earth must implement five functional sub-systems—three of them cover the specific DE scope, while the other two implement the utilities required to implement a secure cybernetic system—see Figure 7:

- A knowledge-intensive application development system to implement/use applications that address policy needs using DTs instances.
- A distributed analytics system that generates intelligence from the appliance of analytical models to data streams (e.g., datasets from operational satellite constellations and IoT infrastructures), making use of the multi-cloud scalable capacities—they are all virtual shared resources.
- A multi HPC and Cloud system that provides the storage/access, network, and computational scalable resources.
- A cybernetic mechanisms system that governs and controls the correct behavior of the whole DE, i.e., of the other systems and of their interactions.
- Multi-level security system that applies to all the previously introduced systems and implements the DE security, privacy, and trust requirements.

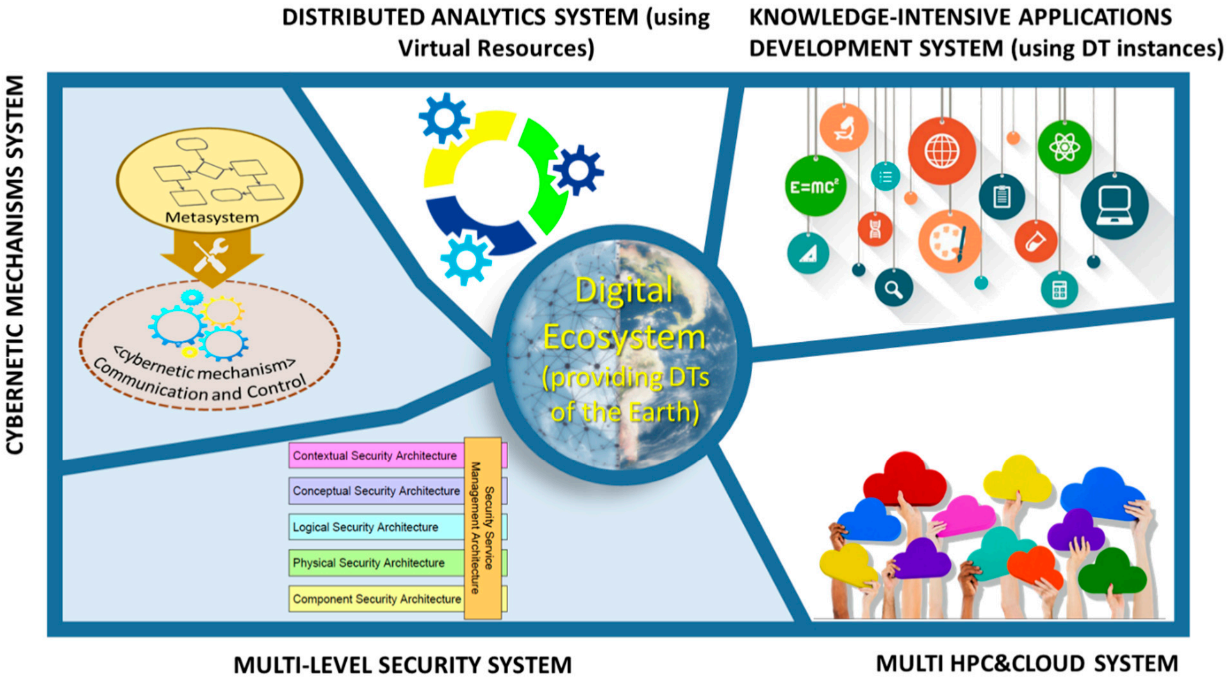

Figure 7. DE functional sub-systems: business related (white background) and secure cybernetic ones (pale blue background).

8. Digital Ecosystem Engineering Architecture

To implement a DE providing DTs of the Earth, the technological reference framework should consider the following engineering and interoperability paradigms and models.

8.1. Multi-Cloud Approach and Virtual Cloud Paradigm

The multi-cloud approach commonly refers to the use of two or more clouds from different cloud providers [62]. This can be any mix of Infrastructure-, Platform-, or Software-as-a Service (IaaS, PaaS, or SaaS) [63]. Differently from cloud federations, in the multi-cloud approach, cloud services providers do not have to agree to share resources for a given application [64]. The multi-cloud approach also includes using private clouds and hybrid clouds with multiple public cloud components. Adopting a multi-cloud approach requires, to provide a rich user experience, the implementation of the virtual cloud paradigm [65], i.e., the development of a virtual layer between the heterogeneous cloud systems and the common DTs/applications development platform. According to several international surveys on cloud usage [66,67], more than 90% of interviewed organizations are already using more than one infrastructure cloud. By applying a multi-cloud strategy, organizations and applications are able to increase their efficiency and effectiveness, choosing the right service and provider for each different use case. On the other hand, it is important to understand and manage some security and performance challenges in multi-cloud architectures, such as networking between clouds, scalability limits, and multiple clouds monitoring [68]. Users and clients interact with the DE multi-cloud-based platform via a virtual (private) cloud.

8.2. IoT–Edge/Fog–Cloud

While cloud data centers are large facilities deployed in a limited number of locations (due to special infrastructure and management needs), in a digitally transformed society, cloud users are spread everywhere—IoT and 5G enabled applications are significant examples. Commonly, clients and users are far from the cloud data centers that are managed by their preferred providers. Edge and/or fog computing infrastructure are likely to be closer to those devices and applications to bring computing capacity with lower response time [69]. Whereas cloud computing effectively supports non-real-time and long-period data driven scenarios, edge computing is effective for real-time and short-period data driven scenarios, such as local decision-making.

8.3. Serverless Computing (Supporting Near Real-Time Services)

Understanding which services should execute on a cloud data center and which ones on the edge devices is a challenge due to unpredictable factors, including the vast heterogeneity of client/user devices (i.e., their computational capabilities) and the network systems' diversity (i.e., their latency times) [69]. This is an explanation for the success of the serverless model, which focuses on the provision of computational functions, with limited resource requirements, that can be deployed closer to user devices; commonly, serverless functions are triggered by user-defined events. This model enables real-time data streams processing and services. Serverless computing fully supports the multi-cloud paradigm by allowing organizations to build services as necessary across different clouds with different approaches [70].

8.4. HPC-as-a-Service

HPC is important to enable extreme digital twin applications; however, access to supercomputers is out of the range of the majority of users and applications. Recently, cloud computing introduced the concept of providing on-demand resources that can be services (i.e., IaaS and PaaS), resolving existing issues such as hardware maintenance and networking expertise. The HPC community is making a great effort to apply these HPC computing concepts to obtain the benefits of the cloud [71]. The main target is to create an HPC system that can provide computing power as a service, i.e., HPC-as-a-Service (HPCaaS). Therefore, HPCaaS is the provision of high-level processing capacity to customers through the cloud [72]. This paradigm provides the resources required to process complex calculations, avoiding the investments of skilled staff and demanding software refactoring.

9. Digital Ecosystem Virtual Cloud Layer and Orchestration Services

The utilization of a virtual cloud services layer (i.e., IaaS, PaaS, and SaaS) allows the utilization of a distributed multi-cloud solution in a dynamic fashion and in a transparent way for the users. This, ultimately, obeys the flexibility and scalability requirements acknowledged for a DE (see Section 4).

A key concept and component of a virtual cloud is represented by its orchestration strategy and system. In a DE, virtual cloud orchestration utilizes software instruments to manage the interconnections and interactions among the disparate systems that constitute the multi-cloud infrastructure. Cloud orchestration connects automated tasks into a cohesive workflow to achieve a goal, in keeping with the DE security and policy constraints. The orchestration results in a set of virtual machines (organized in clusters, connected according to a specific topology), providing the necessary resources to implement the requested services scalability and high availability—see Figure 8.

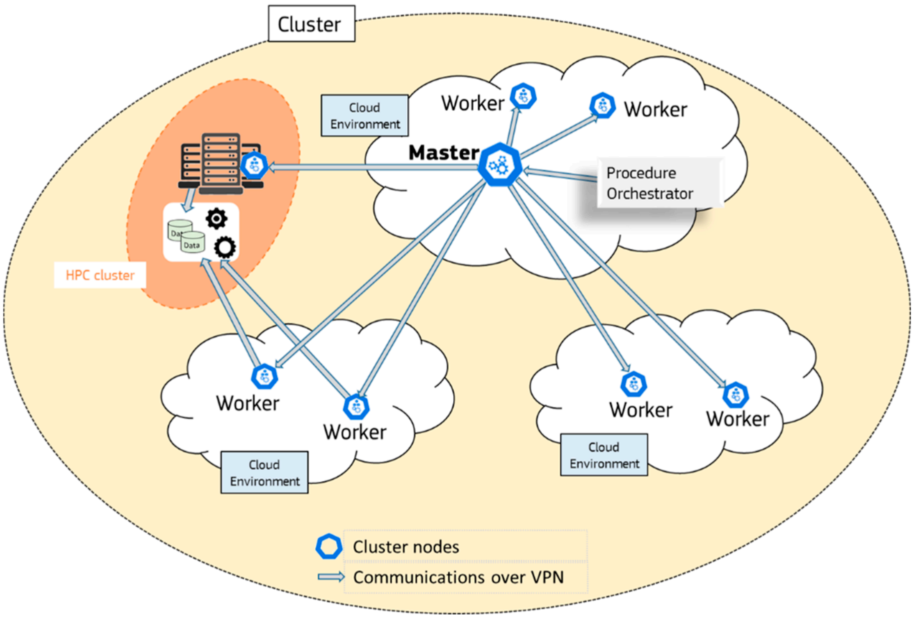

Figure 8. Centralized orchestration architecture for implementing a multi-cloud DE.

For example, in Destination Earth, the virtual cloud orchestration should be designed to allow a highly dynamic scalability and facilitate the DE evolvability (i.e., SoS evolutions) in a transparent way. In agreement with the possible DE government styles (discussed in Section 4.4), few orchestration architectures are possible for the ecosystem technological implementation, realizing either a centralized or distributed (e.g., meshed or opportunistic) networking approach [29]. In the case of Destination Earth, a full discussion about the orchestration solutions is in [29]. Figure 8 depicts the implementation architecture of the centralized approach, which can be considered propaedeutic for the other two (more complex) solutions.

9.1. Technological Framework

The engineering and technology framework characterizing a DE virtual-cloud framework is shown in Figure 9. The framework makes use of microservices and API technologies to implement elastic interoperability between software components. This framework fully supports the DE traits of flexibility, modularity, and evolvability. That is achieved by de-coupling DTs and software applications from the virtual cloud layer—for example, in Destination Earth, this is characterized as the horizontal common infrastructure. In turn, the virtual cloud layer virtualizes a set of operational infrastructures (i.e., the multi-cloud environment) by implementing the necessary mediation and brokering services and publishing a set of open, common, and consistent network services and related APIs to clients. The virtual cloud services allow the multi-cloud infrastructures (i.e., the ecosystem enterprise components) to remain substantially autonomous and evolve freely in time, avoiding the propagation of changes to the DTs and the client applications. Moreover, new operational infrastructures can be added to enrich the ecosystem in a transparent way. It is noticeable how de-coupling the layers supports the separation of the different concerns, expressed by the three principal DE stakeholders: the ecosystem users and clients, the organization(s) steering and funding the ecosystem as a whole, and the enterprises managing the infrastructures that contribute to the ecosystem SoS (i.e., the enterprise systems). As depicted in Figure 9, their needs and constraints clearly influence the different services layers of the framework; thus, the more these layers are decoupled, the more the stakeholders' concerns are separated, applying the "separation of concerns" computer science pattern [73]. Finally, this technological framework supports both a Service-Oriented Approach (SOA) and a Resource-Oriented Approach (ROA), satisfying another important DE principle.

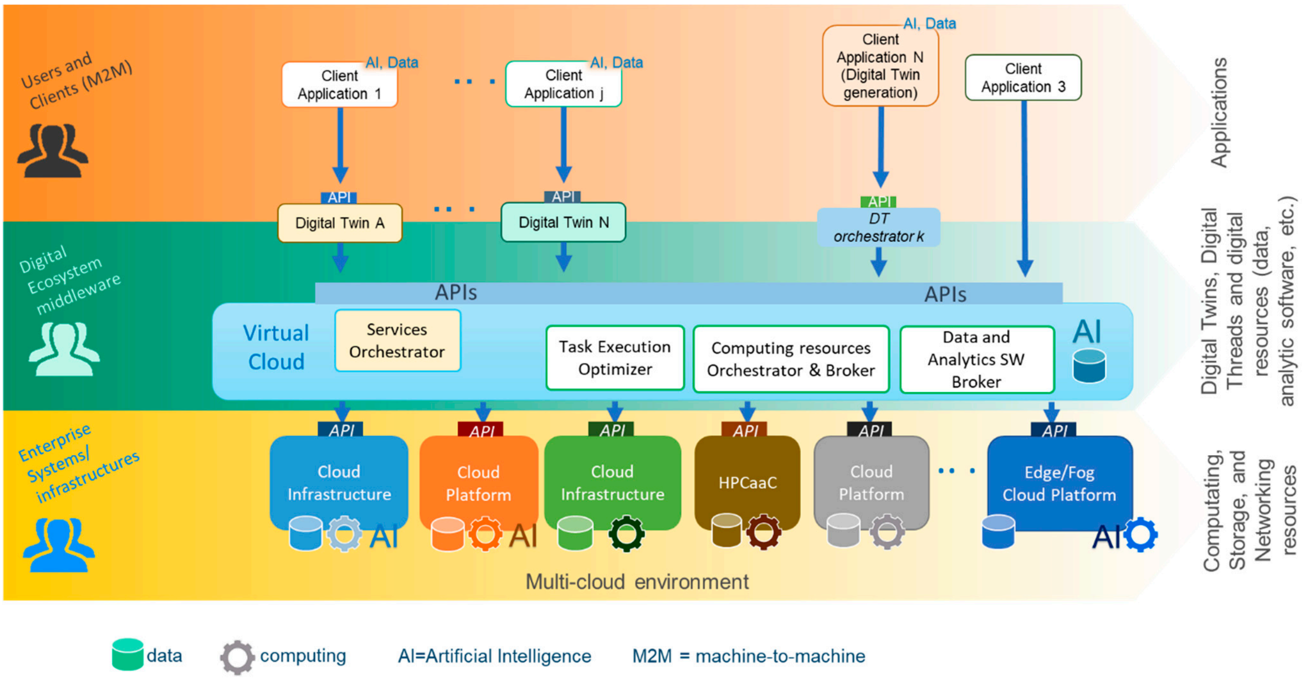

Figure 9. DE technological framework based on the Virtual-Cloud paradigm.

9.1.1. Client (Software) Application

It is a software application (Machine-to-Machine or GUI-enabled) that makes use of the DE services, by accessing macro- and micro-service APIs. In particular, Figure 9 shows applications utilizing the DT instances to develop a business logic, as well as applications using the virtual cloud services to discover, benchmark, and orchestrate shared digital resources, e.g., datasets and analytical software.

9.1.2. DT (Service) Orchestrator

It is a software component that implements the business logic for orchestrating the components and services that are necessary to implement a DT instance, e.g., a sound operational workflow made up of data streams and analytical software. This component can be provided either by a third party (i.e., a client application) or by the DE; in this case, it consists of a more general Internet services orchestrator, is deployed in the virtual cloud, and is named "Service Orchestrator"—see Figure 9. In this case, Service Orchestrator orchestrates shared components to finalize the necessary tasks and generate a DT, including (i) analytical software retrieval/access/provision, (ii) software container(s) instantiation,

(iii) data ingestion requests management, (iv) container(s) execution request management, (v) output(s) storage request management, and (vi) security aspects implementation. This component exposes a well-known API to be invoked by client applications.

9.1.3. Virtual Cloud Middleware

DE virtual cloud is a middleware framework that allows the execution of tasks and applications (noticeably, workflows orchestration for DT instance generation) on a multi-cloud environment (including HPC-as-a-Cloud), exposing it as a unique and consistent virtual capacity. This middleware implements the necessary interoperability arrangements to finalize application execution and infrastructural scalability. In addition, the middleware is capable of querying the underlying multi-cloud infrastructures for accessing and using the digital resources they provide, e.g., computing, data, and analytical software. When an orchestrator requests a workflow execution, the Virtual Cloud, on the basis of a set of predefined criteria (e.g., data and analytical models availability, computing capacities, and energy saving) decides where to execute the job and satisfy a set of agreed policies (e.g., to minimize data movement and time lapse), sending the request to the appropriate computing infrastructure(s). The virtual cloud exposes a set of APIs to discover, access, and benchmark the available virtual resources (e.g., data, analytical models, computing resources, etc.) that are provided by the connected infrastructure, supporting a resource-oriented approach (ROA). To build on a multi-cloud environment and support ROA, the virtual cloud leverages a set of mediation, brokering, and orchestration services implemented by the following components:

- Infrastructure resources Orchestrator and Broker implementing the resources orchestration and scalability as well as the required interoperability arrangements to execute jobs on different computing infrastructures, making it transparent for the orchestrator.
- Data and analytics SW Broker implementing the discoverability and access of the data and analytical software resources that are available on the federated infrastructures. It also implements the mediation and harmonization tasks required. This component is detailed further below.
- Task execution Optimizer implementing the optimization of a given task execution, on the basis of a set of parameters to be minimized or maximized, recognizing the infrastructure(s) where to execute the job and sending the request.

Data and Analytics Software Brokering

In the era of Big Data (i.a., big satellite image time series), data movements must be minimized as much as possible (see Section 4), building distributed solutions and moving analytical software around the ecosystem. To implement an effective and usable distributed data system, metadata sharing and data system microservices must be particularly curated. Considering the multi-disciplinary domain charactering a DE providing DTs of the Earth, and the consequent high heterogeneity of datasets to be processed, data mediation and brokering services (along with their APIs) must be considered and developed [30,74].

In a distributed system, where analytics software is the main resource to be moved around, it is crucial to fully understand the level of interoperability implemented by that software, e.g., learning-based AI models for remote sensing. It is possible to distinguish among three diverse levels of interoperability, according to the openness, digital portability, and client-interaction style that characterize a given analytics tool, i.e., a data/process-driven analytical model provided as a digital software or service:

- Model-as-a-Tool (MaaT): interoperability consists of user's interaction with a software tool (developed to utilize a processing/analytical model) and not with the model itself or a service API. A given implementation of the analytical model runs on a specific server, and a user interface is exposed to interact with the software. It is not possible to move the model and make it run on a different machine. Benefits include a strong control of the model use and execution. There are limitations on the usability and

flexibility of the model, as well as its scalability due to the limitation of the specific server. Machine-to-machine interoperability (chaining capabilities) is not allowed.
- Model-as-a-Service (MaaS): as for the previous case, a given implementation of the analytical model runs on a specific server, but this time, APIs are exposed to interacting with the model. Therefore, interoperability consists of machine-to-machine interaction through a published API, e.g., for a run configuration and execution. Nevertheless, it is not possible to move the model and make it run on a different machine. Also in this case, it is not possible to move the analytical software that run on the model server. Concerns deal with a still limited flexibility and possible scalability issues (depending on the server capacities). To note, this time, the existence of possible concerns for less control on the model (re-)use.
- Model-as-a-Resource (MaaR): the interoperability level resamples the same patterns used for any other shared digital resource, like a dataset. This time, the analytical model itself (and not a given implementation) is accessed through a resource-oriented interface, i.e., API. That allows to effectively move the model and make it run on the machine that best performs for a specific use case. There are clear benefits in terms of flexibility, scalability, and interoperability. The main concerns are about the model sound utilization.

For instance, in the Destination Earth case, all the three levels of analytical software interoperability are likely to be supported. Metadata describing analytical models is an important challenge [75] and must be carefully considered by any DE dealing with digital twins of the Earth.

9.2. Developed Prototypes

In the framework of a JRC study for the description of the Destination Earth ecosystem architecture [29], the engineering schema and virtual cloud solutions discussed above were successfully tested by developing two proofs-of-concept. A first development tested the realization of the Destination Earth DE, which was based on a virtual multi-cloud environment consisting of a set of heterogeneous scalable infrastructures managed or utilized by the European Commission, ESA, ECMWF, and EUMETSAT. Most of the discussed DE principles and patterns were fruitfully tested (e.g., flexibility and usability).

This prototype makes use of well adopted open-source technologies. As for the virtual cloud implementation, it built on Kubernetes (https://kubernetes.io/, accessed on 1 May 2021) and OpenStack (https://www.openstack.org/, accessed on 1 May 2021) technologies. By utilizing ClusterAPI (https://cluster-api.sigs.k8s.io/, accessed on 1 May 2021) solution, the virtual cloud supports different cloud infrastructure environments, including AWS, GCP, Azure, and OpenStack.

A second proof-of-concept was then developed to implement an DT of the Earth instance, following the philosophy and archetype depicted in Figure 2. For the orchestration solution, it makes use of the VLab technology [76]. The two proof-of-concepts will be discussed in a forthcoming manuscript.

9.3. GAIA-X Initiative

GAIA-X [77] is a European initiative aiming at developing a digital ecosystem regulated by its members. “The initiative is working to create an environment in which data can be shared and stored under the control of data owners and users; and where rules are defined and respected, so that data and services can be made easily available, compiled and exchanged” [77]. This project aims to develop an open, transparent digital ecosystem, where data and services can be made available, collated, and shared in an environment of trust. Being aligned with the European Data Strategy, GAIA-X already declared the adoption of most of the principles and patterns introduced in this manuscript for DEs. The project is working on the full specification of its technological architecture [78], which promises to provide an open framework to implement secure DE in Europe. That would serve different application domains, including the Green Deal data space and the generation of DTs of the Earth.

10. Discussion and Conclusions

Humankind is called to face unprecedented challenges to try and preserve our planet and achieve a more sustainable and just society. This is primarily a political challenge, but needs to be supported by a sound and robust multi-disciplinary knowledge platform that can provide the scientific evidence required by policy and decision makers based on environmental observation from remote-sensing and in-situ measurements. To be operative, this platform must go beyond the traditional data exchange interoperability pattern and apply the information and knowledge sharing paradigm that characterizes the "cyber-physical" world created by the Digital Transformation of society. This paradigm is enabled by another innovative model, called datafication, which generates actionable intelligence from data streams. To study the disrupting effect of applying both these archetypes to Global Change and sustainable development, a new scientific discipline was proposed: Big Earth Data (BED) science. This new discipline can leverage the emergence of Digital Twins of the Earth, seen as living digital simulation models that update and change as their physical counterparts change. The concept of DTs of the Earth is at the heart of the recent and ambitious EC initiative, named Destination Earth.

In this paper, we introduced the concept of a digital ecosystem as a new approach to harness the characteristics of the digital transformation and develop the required knowledge framework for Global Change and SDG agenda implementation. The discussed framework is a general one, and it effectively supports the generation and use of Digital Twins of the Earth. The presented framework has emerged from several scientific and development activities that we have performed in the past fifteen years at the European and international level, such as the INSPIRE Directive implementation, the Copernicus DIAS evaluation, the GEOSS design and development, the WMO Hydrological Observing System (WHOS) design and implementation, the AI Watch (https://knowledge4policy.ec.europa.eu/ai-watch_en, accessed on 1 May 2021) implementation, and, more recently, the Destination Earth architecture definition for developing the Green Deal data space.

We referred to the Destination Earth and GEOSS as examples to introduce the concept of a digital ecosystem as a new approach to harness the characteristics of the digital transformation and develop the required knowledge framework for Global Change and SDG agenda implementation. In the paper, we have discussed digital ecosystems from different perspectives, including key principles and patterns, metasystemic governance, the information viewpoint, and engineering architecture. Based on the multi-year experience of developing a global Earth Observation System of Systems and an initial pilot of the Destination Earth architecture, the paper has identified a set of important contextual conditions and good practices to put in pace a sustainable digital ecosystem:

Contextual conditions:

1. The existence of regulatory conditions (e.g., laws, policies, standards, and agreements).
2. The existence of institutional/organizational condition (e.g., a set of shared social and cultural values).
3. The existence of ICT capabilities condition (e.g., computational, data storing, analytical software, and network elements).

Good practices:

1. To apply the data value-chain ecosystem model to better support flexibility and evolvability.
2. To establish the datafication value-chain as the common interest shared among the ecosystem stakeholders
3. To adopt a collaborative style of governance.
4. To recognize three macro-categories of stakeholders (or elements): data/knowledge providers, intermediaries, and data/knowledge consumers.

5. To aim at implementing a collaborative consumption approach to the economic model of sharing.
6. To aim at a satisficing behavior instead of an optimization approach in the DE design and implementation to favor a quick response to changes.
7. To carefully design and implement a meta-systemic level of the DE by defining what should remain invariant in an ever-changing context.
8. To exploit innovative paradigms for information processing (e.g., mobile code, cloud continuum), preserving technology neutrality (in compliance with the metasystemic invariance vs. technology evolution).
9. To utilize a virtual cloud building on a multi-cloud environment.

The knowledge framework must be able to implement and manage a set of concepts and functional systems, including:

Concepts

○ Digital Twins;
○ Digital Threads;
○ Digital/Virtual Resources (e.g., datasets, analytical software, composite DTs);
○ Virtual Network Services;
○ HPC and cloud clusters/nodes.

Functional systems

○ Knowledge-intensive application development system;
○ Distributed analytics system;
○ Multi HPC and Cloud system;
○ Cybernetic mechanisms system;
○ Multi-level security system.

From our experiences, we have learned that to realize a DE, one must think in a systematic way, not focusing on one specific action only. Moreover, it is important to first design a good process and then recognize the useful technology to leverage the process in an effective way. Good technologies are unlikely to correct a poorly designed process. Finally, the human factor is an important factor to be considered too.

We strongly believe that digital ecosystems are the new model to develop a distributed knowledge system among multiple stakeholders, replacing and bringing up to date the concepts of data infrastructures we saw developing in the 1990s and 2000s. The comprehensive discussion of the multiple dimensions and challenges of digital ecosystems presented in this paper fills an important gap in the current knowledge and marks the path for future developments and implementations. The discussed DE framework is a general one and can be utilized to implement any use case, making use of a DT of the Earth.

Future work includes the further development of the initial prototype implemented, extending the number and typology of enterprise systems contributing to the experimented DE. Moreover, on the top of the extended DE, advanced digital threads will be implemented to generate and operate DTs that are significant for the Green Deal data space. GAIA-X promises to provide a technological framework to implement an open and secure DE, in Europe, which serves different application domains, including the green deal data space and the generation of DTs of the Earth. A future study may analyze the ecosystem services to develop DTs and their suitability to implement Destination Earth and other knowledge platforms to realize the SDG agenda.

Author Contributions: All authors contributed to the conceptualization, methodology, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by European Commission DG RTD: 34538, European Commission DG CNECT: 35713, and European Space Agency: 4000123005/18/IT/CGD.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: With special thanks to Nicholas Spadaro, Mattia Santoro, and Massimiliano Olivieri. The authors recognize the important contribution provided by the JRC team on Destination Earth, and the colleagues from ECMWF, EUMETSAT, ESA, and DG CNECT engaged in Destination Earth. A final thank goes to the reviewers for their comments that helped to improve the manuscript. The research leading to these results has been conducted in the framework of the following scientific and innovation activities: the "EO-Value" project (an Administrative Agreement between EC, DG, RTD, and DG JRC), the "Mission Earth study" project (an Administrative Agreement between EC, DG, CNECT, and DG JRC), the "Copernicus 2" project (an Administrative Arrangement between EC, DG, DEFIS, and DG JRC), and the "DAB4EDGE" project (ESA-ESRIN and CNR collaboration).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the writing of the manuscript or in the decision to publish the results.

References

1. GEO. *GEO Strategic Plan 2016–2025: Implementing GEOSS*; GEO: Geneva, Switzerland, 2016.
2. United Nations. Resolution Adopted by the General Assembly on 25 September 2015, 21 October 2015. Available online: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (accessed on 2 October 2016).
3. United Nations. Sustainable Development Goals, 21 August 2016. Available online: http://www.un.org/sustainabledevelopment/sustainable-development-goals/ (accessed on 27 May 2021).
4. United Nations. Sustainable Development knowledge Platform, 2016. Available online: https://sustainabledevelopment.un.org/index.html (accessed on 27 May 2021).
5. European Commission. A European Green Deal, September 2020. Available online: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en (accessed on 2 January 2020).
6. European Commission. *A European Strategy for Data -COM(2020) 66 Final*; European Commission: Brussels, Belgium, 2020.
7. Nativi, S.; Santoro, M.; Giuliani, G.; Mazzetti, P. Towards a knowledge base to support global change policy goals. *Int. J. Digit. Earth* **2019**, *13*, 188–216. [CrossRef]
8. Mayer-Schönberger, V.; Cukier, K. *Big Data: A Revolution that will Transform how We Live, Work, and Think*; John Murray: London, UK, 2014; p. 242.
9. Ruckenstein, M.; Schüll, N.D. The Datafication of Health. *Annu. Rev. Anthr.* **2017**, *46*, 261–278. [CrossRef]
10. Guo, H.; Nativi, S.; Liang, D.; Craglia, M.; Wang, L.; Schade, S.; Corban, C.; He, G.; Pesaresi, M.; Li, J.; et al. Big Earth Data science: An information framework for a sustainable planet. *Int. J. Digit. Earth* **2020**, *13*, 743–767. [CrossRef]
11. Nativi, S.; Kotsev, A.; Scudo, P.; Pogorzelska, K.; Vakalis, I.; Benetta, A.D.; Perego, A. *IoT 2.0 and the Internet of Transformation (Web of Things and Digital Twins)*; European Union: Luxembourgh, 2020.
12. Nativi, S.; Delipetrev, B.; Craglia, M. *Destination Earth: Survey on Digital Twins Technologies and Activities, in the Green Deal Area*; European Commission: Luxembourgh, 2020.
13. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
14. El Saddik, A. Digital Twins: The Convergence of Multimedia Technologies. *IEEE Multimed.* **2018**, *25*, 87–92. [CrossRef]
15. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. [CrossRef]
16. ISO TC184. *ISO/DIS 23247 Automation Systems and Integration-Digital Twin Framework for Manufacturing*; ISO: Geneva, Switzerland, 2020.
17. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-Pap.* **2018**, *51*, 1016–1022. [CrossRef]
18. Jacoby, M.; Usländer, T. Digital Twin and Internet of Things—Current Standards Landscape. *Appl. Sci.* **2020**, *10*, 6519. [CrossRef]
19. Tao, F.; Qi, Q. Make more digital twins. *Nat. Cell Biol.* **2019**, *573*, 490–491. [CrossRef] [PubMed]
20. ESA. ESA Digital Twin Earth Challenge, 2020. Available online: https://copernicus-masters.com/prize/esa-challenge/ (accessed on 1 January 2021).
21. *World Climate Research Programme*; Report of the 41st Session of the Joint Scientific Committee; WMO: Geneva, Switzerland, 2020.
22. European Commission. A Transparent & Accessible Ocean: Towards a Digital Twin of the Ocean, 2020. Available online: https://ec.europa.eu/info/sites/info/files/research_and_innovation/green_deal/gdc_stakeholder_engagement_topic_09-3_digital_ocean.pdf (accessed on 2 January 2020).
23. Van Genderen, J.; Goodchild, M.F.; Guo, H.; Yang, C.; Nativi, S.; Wang, L.; Wang, C. *Digital Earth Challenges and Future Trends, in Manual of Digital Earth*; Springer: Singapore, 2020; pp. 811–827.
24. Australian Government-Bureau of Meteorology. *A Land of Storms, Floods and Bushfires: Seamless and Integrated Forecasting*; Australian Government: Canberra, Australian, 2020.
25. European Commission-DG CNECT. Destination Earth (DestinE), 30 November 2020. Available online: https://ec.europa.eu/digital-single-market/en/destination-earth-destine (accessed on 2 January 2021).

26. European Commission. The European Digital Strategy, 2020. Available online: https://ec.europa.eu/digital-single-market/en/content/european-digital-strategy (accessed on 2 January 2020).
27. European Commission. *Shaping Europe's Digital Future*; European Commission: Luxembourgh, 2020.
28. Nativi, S.; Craglia, M. *Destination Earth: Use Cases Analysis*; Publications Office of the European Union: Luxembourg, 2020.
29. Nativi, S.; Craglia, M. *Destination Earth: Ecosystem Architecture Description*; European Commission: Luxembourgh, 2021.
30. Nativi, S.; Mazzetti, P.; Santoro, M.; Papeschi, F.; Craglia, M.; Ochiai, O. Big Data challenges in building the Global Earth Observation System of Systems. *Environ. Model. Softw.* **2015**, *68*, 1–26. [CrossRef]
31. Blew, R. On the Definition of Ecosystem. *Bull. Ecol. Soc. Am.* **1996**, *77*, 171–173.
32. Moore, J.F. *The Death of Competition: Leadership & Strategy in the Age of Business Ecosystems*; Harper Business: New York, NY, USA, 1996.
33. Jansen, S.; Brinkkemper, S.; Cusumano, M. *Software Ecosystems: Analyzing and Managing Business Networks in the Software Industry*; Edward Elgar Publishing Limited: Cheltenham, UK, 2013.
34. Moore, J.F. Predators and prey: A new ecology of competition. *Harv. Bus. Rev.* **1993**, *71*, 75–86. [PubMed]
35. European Commission-DG Connect. A European Strategy on the Data Value Chain, 2013. Available online: http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3488 (accessed on 17 August 2019).
36. European Commission. *A Digital Single Market Strategy for Europe -COM(2015) 192 Final*; European Commission: Luxembourgh, 2015.
37. Cavanillas, J.; Curry, E.; Wahlster, W. *New Horizons for a Data-Driven Economy*; Springer Nature: Berlin, Germany, 2016; p. 303.
38. Scott, W.R. *Institutions and Organizations: Ideas, Interests, and Identities*; SAGE Publications: Thousand Oaks, CA, USA, 2013; p. 360.
39. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]
40. Cuppan, A.P.; Sage, C.D. On the Systems Engineering and Management of Systems and Federation of Systems. *Inf. Knowl. Syst. Manag.* **2001**, *2*, 325–345.
41. Maier, M.W. Architecting Principles for System of Systems. *Syst. Eng.* **1998**, *1*, 1998. [CrossRef]
42. De Laurentis, D. *Understanding Transportation As a System of Systems Problem, in System of Systems Engineering: Innovations for the 21st Century*; Jamshidi, M., Ed.; Wiley: Hoboken, NJ, USA, 2009; pp. 520–541.
43. DiMario, M.J.; Boardman, J.T.; Sauser, B.J. System of Systems Collaborative Formation. *IEEE Syst. J.* **2009**, *3*, 360–368. [CrossRef]
44. Nativi, S.; Spadaro, N.; Pogorzelska, K.; Craglia, M. *DIAS Assement*; European Commission: Luxembourgh, under publication.
45. Baresi, L.; Mendonça, D.F.; Garriga, M.; Guinea, S.; Quattrocchi, G. A Unified Model for the Mobile-Edge-Cloud Continuum. *ACM Trans. Internet Technol.* **2019**, *19*, 1–21. [CrossRef]
46. Villari, M.; Fazio, M.; Dustdar, S.; Rana, O.; Ranjan, R. Osmotic Computing: A New Paradigm for Edge/Cloud Integration. *IEEE Cloud Comput.* **2016**, *3*, 76–83. [CrossRef]
47. Keating, C.; Rogers, R.; Unal, R.; Dryer, D.; Sousa-Poza, A.; Safford, R.; Peterson, W.; Rabadi, G. System of Systems Engineering. *Eng. Manag. J.* **2003**, *15*, 36–45. [CrossRef]
48. Simon, H.A. Rational choice and the structure of the environment. *Psychol. Rev.* **1956**, *63*, 129–138. [CrossRef] [PubMed]
49. Simon, H.A. A Behavioural Model of Rational Choice. *Q. J. Econ.* **1955**, *69*, 99–118. [CrossRef]
50. Jamshidi, M. *System-of-Systems Engineering-A Definition*; IEEE SMC: Big Island, HI, USA, 2005.
51. Baldwin, J.; Dahmann, K. Understanding the current state of US defense systems of systems and the implications for systems engineering. In Proceedings of the 2008 2nd Annual IEEE Systems Conference, Montreal, QC, Canada, 7–10 April 2008; pp. 1–7. [CrossRef]
52. Holt, J.; Perry, S.; Payne, R.; Bryans, J.; Hallerstede, S.; Hansen, F.O. A Model-Based Approach for Requirements Engineering for Systems of Systems. *IEEE Syst. J.* **2015**, *9*, 252–262. [CrossRef]
53. European Commission-DG CNECT. Destination Earth (DestinE) Architecture Validation Workshop-Summary Report, 2 February 2021. Available online: File:///C:/Users/Stefano/Downloads/DestinE-ArchitectureValidationworkshopreportpdf%20(1).pdf (accessed on 22 February 2021).
54. Austin, S.A.; Selberg, M.A. Toward an Evolutionary System of Systems. In Proceedings of the Eighteenth Annual International Symposium of The International Council on Systems Engineering (INCOSE 2008), Utrecht, The Netherland, 15–19 June 2008.
55. Beer, S. *Cybernetics and Management*; The English Universities Press Ltd.: London, UK, 1959.
56. Klir, G.J.; Elias, D. *Architecture of Systems Problem Solving*; Springer Science + Business Media: New York, NY, USA, 2003.
57. Turchin, V.F. A dialogue on Metasystem transition. *World Futures J. Glob. Educ.* **1995**, *45*, 5–57. [CrossRef]
58. Beer, S. *Brain of the Firm-Second Edition*; John Wiley & Sons: Hoboken, NJ, USA, 1981.
59. Flood, R.L.; Carson, E. *Dealing with Complexity: An Introduction to the Theory and Application of Systems Science*; Springer Science and Business Media LLC: Berlin, Germany, 2013.
60. Accenture Consulting. The Digital Thread Imperative, 2017. Available online: https://www.accenture.com/_acnmedia/PDF-67/Accenture-Digital-Thread-Aerospace-And-Defense.pdf (accessed on 27 May 2021).
61. Miskinis, C. What does A Digital Thread Mean and How It Differs from Digital Twin, November 2018. Available online: https://www.challenge.org/insights/digital-twin-and-digital-thread/ (accessed on 27 May 2021).
62. IBM. Multicloud, 05 September 2019. Available online: https://www.ibm.com/cloud/learn/multicloud#:~{}:text=Multicloud%20is%20the%20use%20of,%2C%20PaaS%2C%20or%20SaaS).&text=With%20multicloud%2C%20you%20can%20decide,based%20on%20your%20unique%20requirements (accessed on 27 May 2021).

63. Mell, P.; Grance, T. *The NIST Definition of Cloud*; US Dept. of Commerce: Washington, DC, USA, 2011.
64. Hong, J.; Dreibholz, T.; Schenkel, J.A.; Hu, J.A. An Overview of Multi-cloud Computing. In *Advances in Human Factors, Business Management, Training and Education*; Metzler, J.B., Ed.; Springer: Matsue, Japan, 2019; pp. 1055–1068.
65. McKee, D.W.; Clement, S.J.; Almutairi, J.; Xu, J. Survey of advances and challenges in intelligent autonomy for distributed cyber-physical systems. *CAAI Trans. Intell. Technol.* **2018**, *3*, 75–82. [CrossRef]
66. IDC White Paper. Automation, Analytics, and Governance Power, August 2019. Available online: https://go.cloudhealthtech.com/rs/933-ZUR-080/images/IDC-whitepaper-multicloud-management.pdf (accessed on 1 August 2020).
67. Flexera. 2020 State of the Cloud Report, 2020. Available online: https://resources.flexera.com/web/pdf/report-state-of-the-cloud-2020.pdf?elqTrackId=8e710f56b1c44fe39130b0d6168944ed&elqaid=5770&elqat=2 (accessed on 27 May 2021).
68. Tozzi, C. Multicloud Architecture: 3 Common Performance Challenges and Solutions, 25 November 2019. Available online: https://www.itprotoday.com/performance-management/multicloud-architecture-3-common-performance-challenges-and-solutions (accessed on 27 May 2021).
69. Bittencourt, L.; Immich, R.; Sakellariou, R.; Fonseca, N.; Madeira, E.; Curado, M.; Villas, L.; DaSilva, L.; Lee, C.; Rana, O. The Internet of Things, Fog and Cloud continuum: Integration and challenges. *Internet Things* **2018**, *3–4*, 134–155. [CrossRef]
70. Harper, J. 2021 Trends in Cloud Computing: The Omni Present Multi-Cloud Phenomenon, 19 October 2020. Available online: https://insidebigdata.com/2020/10/19/2021-trends-in-cloud-computing-the-omni-present-multi-cloud-phenomenon/ (accessed on 27 May 2021).
71. Imran, H.A.; Wazir, S.; Ikram, A.J.; Ikram, A.A.; Ullah, H.; Ehsan, M. HPC as A Service: A Naive Model, 1 March 2020. Available online: https://arxiv.org/abs/2003.00465 (accessed on 27 May 2021).
72. Rouse, M. HPCaaS (High-Performance Computing as a Service), 28 December 2020. Available online: https://whatis.techtarget.com/definition/HPCaaS-high-performance-computing-as-a-service (accessed on 27 May 2021).
73. Dijkstra, E.W. *On the Role of Scientific Thought. Selected Writings on Computing: A Personal Perspective*; Springer: New York, NY, USA, 1982.
74. Nativi, S.; Bigagli, L. Discovery, Mediation, and Access Services for Earth Observation Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 233–240. [CrossRef]
75. Nativi, S.; Mazzetti, P.; Geller, G.N. Environmental model access and interoperability: The GEO Model Web initiative. *Environ. Model. Softw.* **2013**, *39*, 214–228. [CrossRef]
76. Santoro, M.; Mazzetti, P.; Nativi, S. The VLab Framework: An Orchestrator Component to Support Data to Knowledge Transition. *Remote Sens.* **2020**, *12*, 1795. [CrossRef]
77. GAIA-X Consortium. GAIA_X: A Federated Data Infrastructure for Europe, April 2021. Available online: https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html (accessed on 14 May 2021).
78. GAIA-X Consortium. GAIA-X: Technical Architecture, June 2020. Available online: https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf?__blob=publicationFile&v=5 (accessed on 14 May 2021).