



An Information Management Framework for Environmental Digital Twins (IMFe) as a concept and pilot

Justin Buck¹, **Andrew Kingdon**², John Siddorn⁴, Gordon Blair³, Alexandra Kokkinaki¹, John Blower⁴, Matt Fry⁵, Ben Marchant², Sam Pepler⁶, John Watkins³, and James Byrne⁷

¹National Oceanography Centre, British Oceanographic Data Centre, Liverpool, United Kingdom

²British Geological Survey, Keyworth, United Kingdom

³UK Centre for Ecology and Hydrology, Lancaster, United Kingdom

⁴National Oceanography Centre, Southampton, United Kingdom

⁵UK Centre for Ecology and Hydrology, Wallingford, United Kingdom

⁶National Centre for Atmospheric Science, Oxford, United Kingdom

⁷British Antarctic Survey, Cambridge, United Kingdom

Environmental science is concerned with assessing the impacts of changing environmental conditions upon the state of the natural world. Environmental Digital Twins (EDT) are a new technology that enable environmental change scenarios for real systems to be modelled and their impacts visualised. They will be particularly effective with delivering understanding of these impacts on the natural environment to non-specialist stakeholders.

The UK Natural Environment Research Council (NERC) recently published its first digital strategy, which sets out a vision for digitally enabled environmental science for the next decade. This strategy places data and digital technologies at the heart of UK environmental science.

EDT have been made possible by the emergence of increasingly large, diverse, static data sources, networks of dynamic environmental data from sensor networks and time-variant process modelling. Once combined with visualisation capabilities these provide the basis of the digital twin technologies to enable the environmental scientists community to make a step-change in understanding of the environment. Components may be developed separately by a network but can be combined to improve understanding provided development follows agreed standards to facilitate data exchange and integration.

Replicating the behaviours of environmental systems is inevitably a multi-disciplinary activity. To enable this, an information management framework for Environmental digital twins (IMFe) is needed that establishes the components for effective information management within and across the EDT ecosystem. This must enable secure, resilient interoperability of data, and is a reference point to facilitate data use in line with security, legal, commercial, privacy and other relevant concerns. We present recommendations for developing an IMFe including the application of concepts such as an asset commons and balanced approach to standards to facilitate minimum interoperability requirements between twins while iteratively implementing an IMFe. Achieving

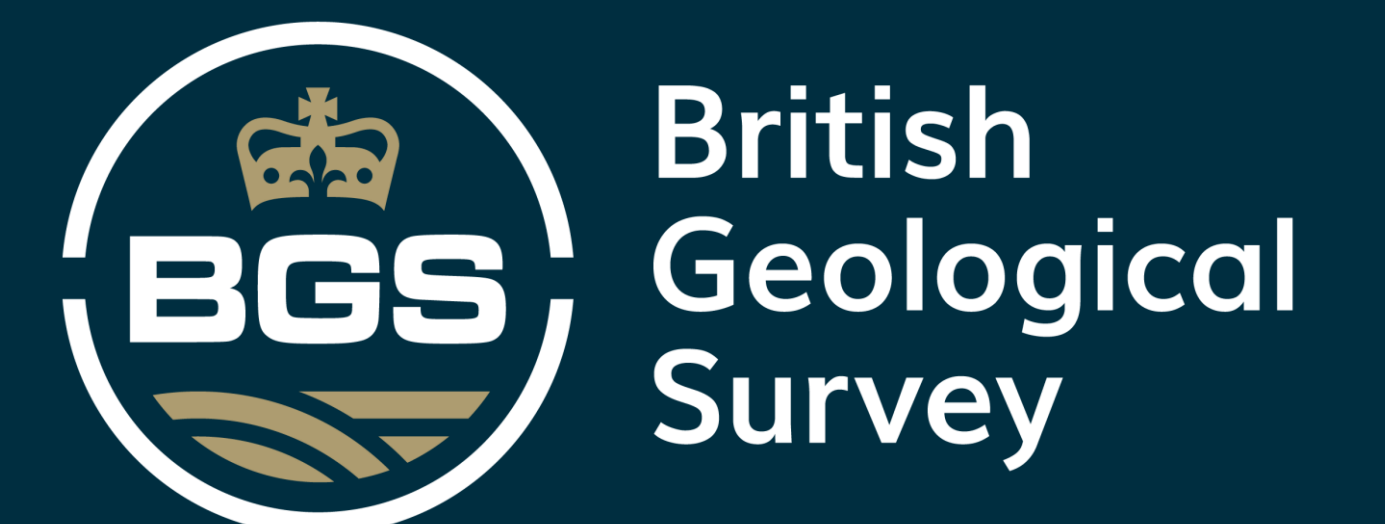
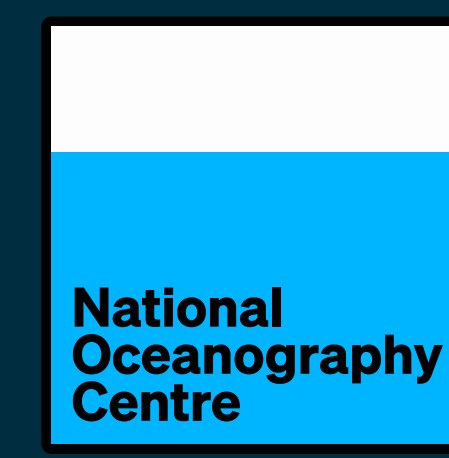
this requires components to be developed that follow agreed standards to ensure that information can be trusted by the user, and that they are semantically interoperable so data can be shared. A digital Asset Register will be defined to provide access to and enable linking of such components.

This previously conceptual project has now been enhanced into the Pilot IMFe project aiming to define the architectures, technologies, standards and hardware infrastructure to develop a fully functioned environmental digital twin. During the project lifespan this will be tested with by construction of a pilot EDT for the Haig Fras Marine Conservation Zone (MCZ) that both enables testing of the proposed IMFe concepts and will provide a clear demonstration of the power of EDT to monitor and scenario test a complex environmental system for the benefit of stakeholders.

An Information Management Framework for Environmental Digital Twins (IMFe) as a concept and pilot

Andrew Kingdon (1), Justin Buck (2), John Siddorn (2), Gordon Blair (3), Alexandra Kokkinaki(2), John Blower (2), Matt Fry (3), Ben Marchant (1), Sam Pepler (4), John Watkins (3) and James Byrne (5)

(1) British Geological Survey (2) National Oceanography Centre (3) UK Centre for Ecology & Hydrology (4) STFC NCAS (5) British Antarctic Survey



Abstract

Environmental science is concerned with assessing the impacts of changing environmental conditions upon the state of the natural world. A digital twin (DT) are virtual representations of physical asset enabled through data and simulators for real-time prediction, optimization, monitoring, controlling, and improved decision making. Environmental Digital Twins (EDT) apply these principles to systems replicating natural systems and allow the impacts of environmental change scenarios upon them to be modelled so impacts can be predicted and visualised. They will be particularly effective at explaining these to non-specialist stakeholders.

EDT are now possible due to the emergence of large static environmental data sources, dynamic environmental data from sensor networks and time-variant process modelling. When combined with visualisation these provide the basis of EDT to enable the environmental scientists to make a step-change in understanding of environmental changes. Components may be developed separately, but can be combined to improve understanding provided development follows agreed standards to facilitate data exchange and integration.

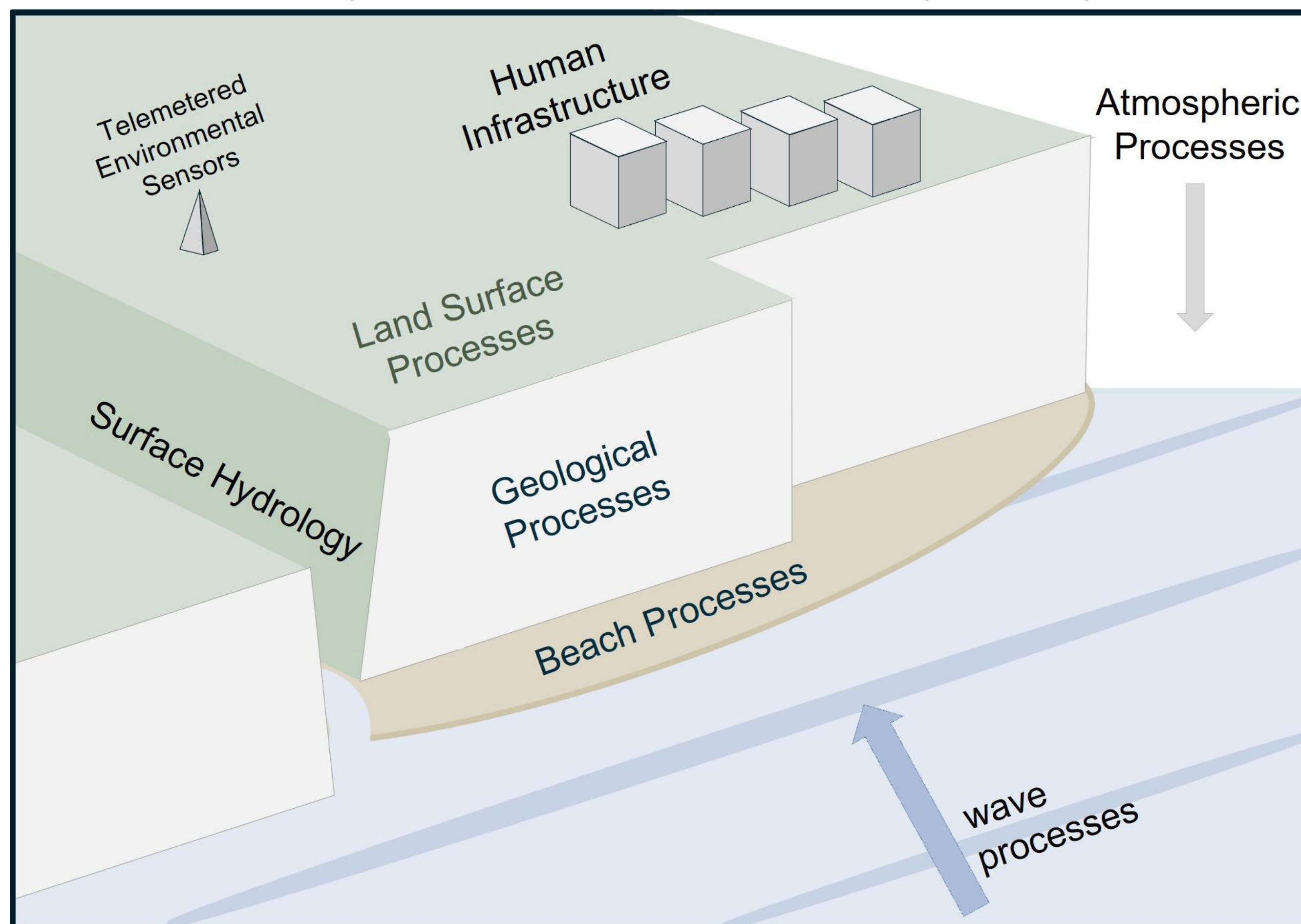


Figure 1: The Coastal Environment: an example of a complex dynamic environmental system with multiple overlapping processes that will benefit from a digital twin to models process interactions

Environmental Digital Twins as tools to deliver interoperable science

EDT could be a transformational technology for explaining complex environmental process interactions using scenario-based predictions. EDT will inform stakeholders, including operators of critical owners of infrastructure. As ever more scientific data is released according to FAIR principals and models components are increasingly being openly shared through code repositories, the barriers to entry are decreasing.

Domain experts can use their expertise to create models and components using common development standards. These can then be shared, both gaining greater value for their development efforts and also community peer-review and iterative improvements. This community of EDT creators need access to common utilities.

The Pilot IMFE project was created to define the facilities and infrastructures needed to enable the environmental science community to build interoperable EDT.

Project Objectives

The overarching objective of this project is to develop an Information Management Framework for digital twins of the environment (IMFe). To build a community of stakeholders around the IMFe work, and an improved understanding of the impact of Digital Research Infrastructure for digital twins.

- 1: To establish an IMFe development framework.
- 2: To conduct pilot studies which test and challenge the framework.
- 3: To deliver an IMFe which addresses all of these elements.
- 4: To define and develop the infrastructures necessary to deliver an IMFE.

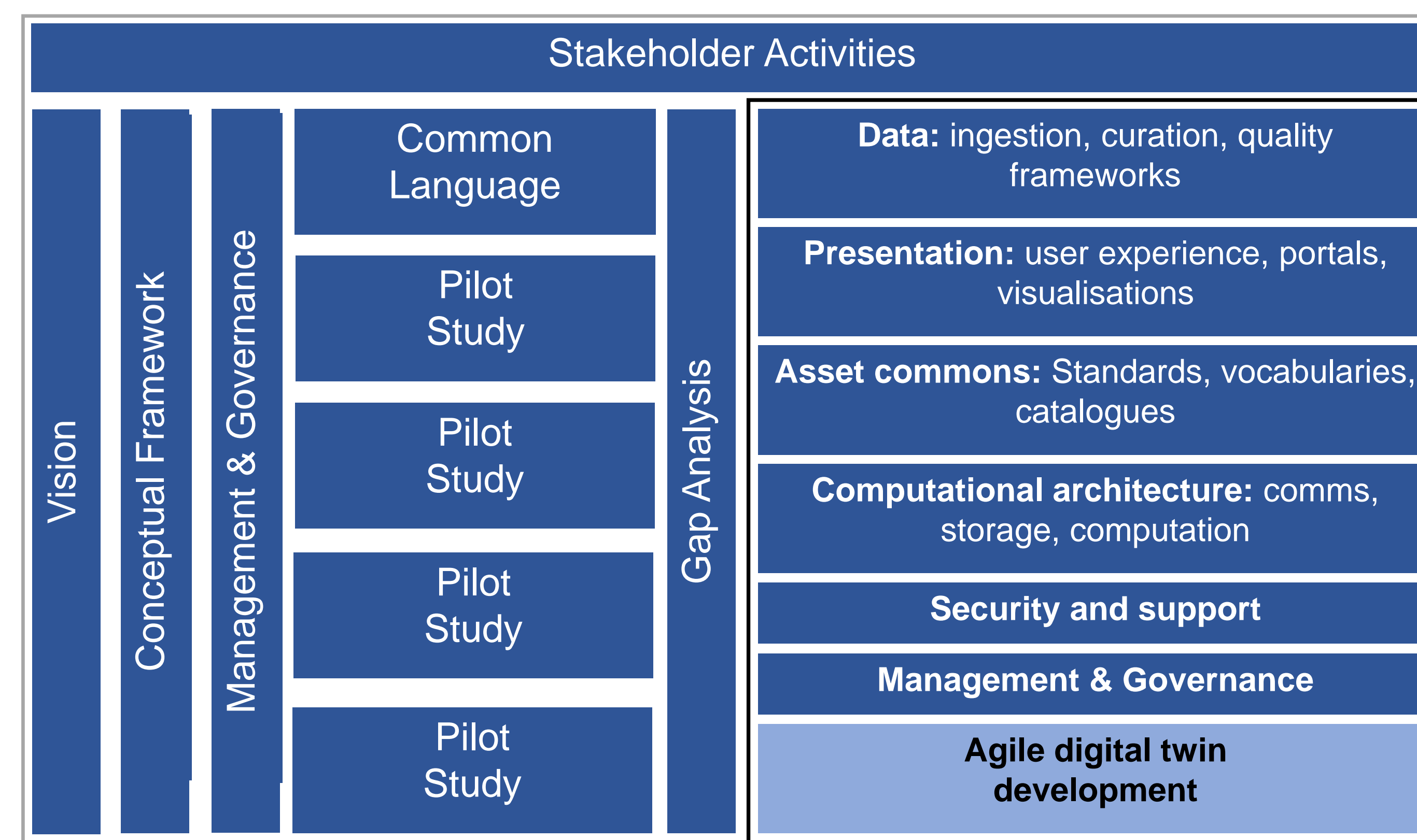


Figure 2: Mapping of the principal elements needed to deliver an Information Management Framework for Environmental Digital Twins

Infrastructures needed to deliver EDT

Synthesising whole environmental systems using Digital Twins will mean sharing specialist knowledge between partners and linking them to understand the complete system. Tools are needed to share data and code for common good, serve them as interoperable components to this community. The proposed solution to this challenge is the "Asset Register" a portal and single access point to federated repositories for NERC environmental science discipline by:

- providing a minimum metadata framework for potential Digital Twin assets
- enabling sharing and reuse of modelling components to maximise their value
- using common principals: Open Science concepts, FAIR data structures and semantics
- enabling interoperable collaborative development of digital twins and their components through sharing of technology, incorporating best practice and making it comprehensible to non-technical users
- ensuring contents and underlying technology can evolve through time
- making this sustainable after the life span of the project

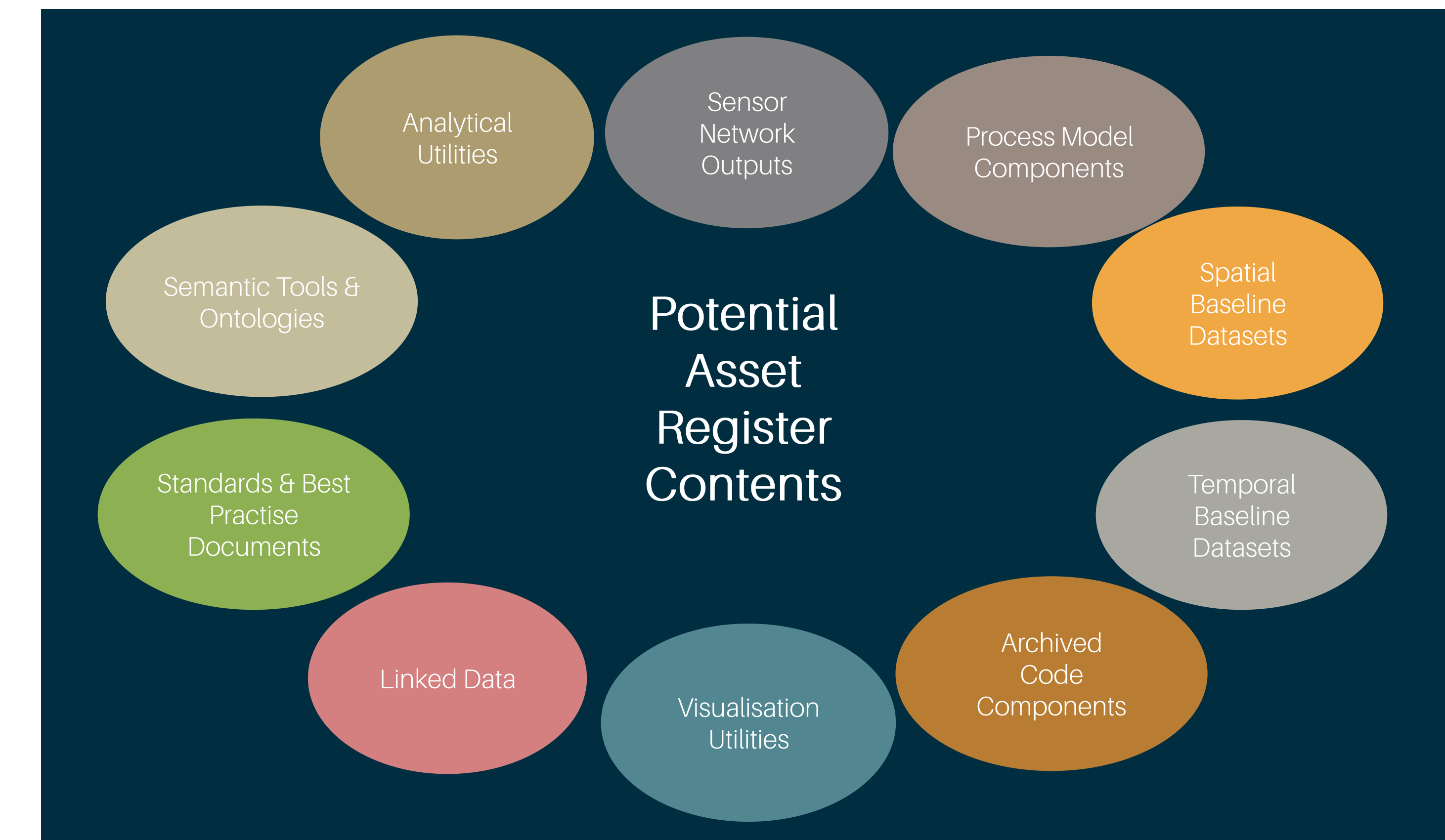


Figure 3: Mapping of the potential components necessary to build an EDT

Next steps

Full roll out of the Asset Register to a wider user community and its population with:

- Relevant code components
- Standards
- FAIR datasets
- other utilities to enable widespread EDT development

Delivery of the Haig Fras MPZ pilot EDT which will test these infrastructures and demonstrate the practicality of the mechanism for creating and delivery of EDT

Creation of the Community of Practise for the development of EDT

This work was supported by the UKRI Natural Environmental Research Council Grant number NE/X016765/1]

