

DataBri-X: Expanding Digital Value Creation in European Data Spaces

by Stelios Sartzetakis (ATHENA RC) and Chamanara, Javad (TIB)

In an era when data availability and AI technologies advance rapidly, the European data economy is poised for substantial growth, unlocking new opportunities and innovations. The DataBri-X project [L1] is motivated by the need to foster the development of trustworthy, “made in Europe” AI that embodies European values and ethical standards. DataBri-X focuses on transforming data-sharing ecosystems by advancing data lifecycle practices, tools, and governance frameworks.

Despite the potential of data, data sharing and interoperability are still in their nascent stages. To realise a truly cross-border and cross-sectoral data-sharing environment, DataBri-X employs comprehensive Data Spaces and data processing tools that allow for seamless processing of proprietary, personal, and open public data. Achieving this vision necessitates overcoming various technical, legal, and business challenges throughout the data lifecycle. DataBri-X not only focuses on conventional raw data and its transformations, but also encompasses metadata, models, and processing algorithms.

DataBri-X applies a fundamental rethinking of data lifecycle practices, focusing on the development and implementation of innovative strategies that prioritise transparency, efficiency,

and collaboration in data sharing. By fostering an environment in which data sharing becomes standard practice, the aim is to build trust among stakeholders and establish a robust framework for sustainable data management. DataBri-X advances data tools and services by evaluating current offerings, identifying gaps, and addressing areas for improvement within data-sharing ecosystems. The goal is to create cutting-edge tools that ensure seamless interoperability, accessibility, and usability of data. These tools are designed to align with FAIR principles, reduce energy footprints, and adapt to diverse user needs while supporting innovative business models. FAIR not only in the context of data, but also for the governance process and execution. In DataBri-X, the processes applied on the data as well as the full execution workflow including software systems used, their versions, and the functions applied on data are preserved, can be shared and reused.

Focus is placed on addressing key challenges in data management. Clear frameworks are established to define data ownership rights and responsibilities within decentralised data-sharing landscapes. Mechanisms are implemented to ensure data provenance and verification, enhancing confidence in data quality. Decentralised technologies are explored to enable secure and efficient data sharing, promoting autonomy and privacy without reliance on centralised control. Additionally, robust strategies are employed to safeguard sensitive data through effective confidentiality measures and digital rights management, ensuring that digital rights are respected throughout the data lifecycle. Furthermore, energy-efficient practices are integrated into data processing and sharing to minimise environmental impact and promote sustainability.

Building on the results of relevant past and ongoing initiatives, DataBri-X aims to refine existing data management tools, sys-

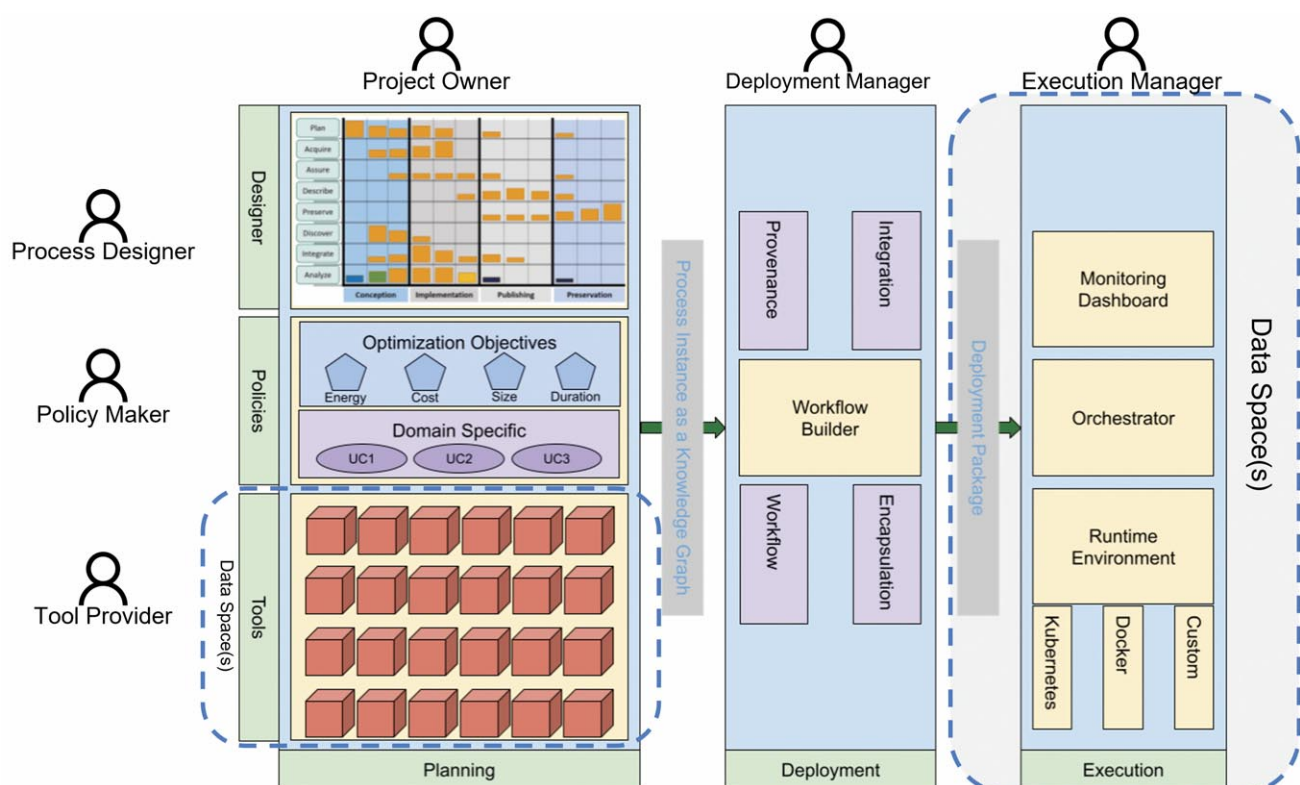


Figure 1: DataBri-X Architecture.

tems, and processes. This includes enabling and automating the creation and maintenance of common ontologies, vocabularies, and data models, as well as supporting automated authoring, co-creation, curation, annotation, and labelling of data. These efforts are geared toward enhancing the usability of data for artificial intelligence and other advanced applications, ensuring that data remains a valuable and adaptable resource for future innovations.

DataBri-X Outcomes

1. A comprehensive management model that redefines data lifecycle practices, fostering a culture of responsible and efficient data sharing.
2. Enhanced maturity of data tools and services that are user-friendly, interoperable, and secure, empowering organisations to leverage data effectively.
3. A framework addressing critical challenges related to data ownership, provenance, confidentiality, and energy efficiency, paving the way for a sustainable and equitable data-sharing ecosystem.
4. Implement IDS-compliant Data Spaces.

DataBri-X Architecture

DataBri-X aims to revolutionise data lifecycle practices with a focus on data-sharing ecosystems. By proposing a comprehensive management model, JenPlane, the project enhances the maturity of data tools and services, ensuring they are equipped to meet the evolving demands of secure and efficient data sharing.

JenPlane is a governance model that advocates a process-based approach to data management. It reimagines traditional data lifecycle practices by introducing a dynamic, non-linear framework tailored to the complexities of modern data-sharing ecosystems. By conceptualising the data lifecycle as an interactive plane rather than a sequential cycle, JenPlane offers a flexible model that adapts to the diverse needs of users engaged in data management activities.

JenPlane consists of multiple elements, namely, the processes, the designer, the composer, and the builder. At its core, JenPlane empowers users to design their data lifecycle by customising one of the available process templates and then selecting the most energy-efficient and complementary tools tailored to their specific data management tasks with the help of an LLM-based recommendation engine. This process designer and composer address the challenges of tool selection, collaboration, and orchestration, enabling semi-automatic deployment, execution, and orchestration of data-driven projects. By creating structured working areas that encompass various disciplines such as planning, data collection, validation, semantic annotation, preservation, discovery, and integration, JenPlane facilitates a comprehensive approach to managing data.

While JenPlane serves as the governance backbone, it operates within a broader ecosystem of the DataBri-X project. This ecosystem comprises various software tools that provide essential services across different segments of data-intensive projects. These tools are designed to collaborate seamlessly, ensuring smooth data flows and workflows, ultimately enabling users to meet their project requirements effectively.

Figure 1 “DataBri-X Architecture” illustrates the overall JenPlane Data Governance. Key components include the JenPlane Process Designer and the JenPlane Composer. These components allow users to specify project requirements, select appropriate tools, and assemble an efficient workflow for data governance. JenPlane’s unique structure, which represents project phases on a two-dimensional axis, ensures flexibility and allows multiple activities to proceed in parallel, enhancing adaptability to different data-centric projects. The toolbox will also include a Policy Centre that stores customisable policies, ensuring compliance with security and privacy regulations, such as GDPR, while facilitating sustainability and energy-efficient data processing.

In total, 11 DataBri-X partners are bringing 25 data tools and services together for different areas of the data lifecycle. The tools are improved on the TRL level and integrated into the DataBri-X toolbox that can be configured along the project governance components for easy deployment in Data Spaces.

The effectiveness and practicality of the DataBri-X toolbox are validated through its implementation in three distinct pilot use cases spanning telecommunications, energy, and legal sectors, designed to showcase the toolset’s potential in driving innovation and improving decision-making across diverse domains.

DataBri-X project has received funding from the EU’s Horizon Europe research and innovation programme under grant agreement no. 101070069.



Link:

[L1] <https://databri-x.eu>

Please contact:

Stelios Sartzetakis, ATHENA Research Centre
stelios@athenarc.gr

Chamanara, Javad, TIB
Javad.Chamanara@tib.eu