

Digital Ownership in Practice: A Feasibility Study of the Information Bank Model*

Gábor Kusper^{a,b}, Attila Adamkó^b, Imre Baják^c

^aEszterházy Károly Catholic University
kusper.gabor@uni-eszterhazy.hu

^bUniversity of Debrecen
{[kusper.gabor](mailto:kusper.gabor@inf.unideb.hu),[adamko.attila](mailto:adamko.attila@inf.unideb.hu)}@inf.unideb.hu

^cBudapest University of Economics and Business
bajak.imre@uni-bge.hu

Abstract

The notions of *digital ownership* and *data treasure* were introduced in [4]. Digital ownership denotes the user’s ability to control and manage their personal data treasure, which includes personal data, user-generated content, traces of user activity, and data related to owned physical objects. In practice, this data is fragmented across vendor clouds, government portals, messaging platforms, and ad-hoc user archives, making everyday information retrieval tasks unnecessarily costly.

The *Information Bank* (InfoBank) concept was proposed as a user-centric solution to this fragmentation: a secure personal repository that enables collecting, managing, and querying one’s data treasure. At its core, InfoBank operationalizes digital ownership as a practical principle: data describing the user and their owned objects should be accessible, manageable, and transferable by the user, rather than being controlled exclusively by external service providers [4, 5]. This vision aligns with human-centric personal data initiatives such as MyData and Solid, as well as with European regulatory frameworks supporting access, portability, and fair use of personal and product-generated data [1, 7, 9–11]. However, these initiatives do not by themselves constitute a concrete technical blueprint for an ownership-driven InfoBank.

*This research was supported by the GINOP-2.1.2-8-1-4-16-2017-00176 project.

Recent advances in Large Language Models (LLMs) significantly reshape how users can interact with their documents. Users increasingly expect natural-language querying, summarization, and comparison over their private archives. To increase trust and ensure traceability, such interaction must be grounded in authoritative sources. Retrieval-augmented generation (RAG) addresses this requirement by conditioning LLM outputs on retrieved documents, typically using vector-based semantic search [3, 6]. For InfoBank systems, this implies an updated baseline: beyond multi-format storage, they must provide *source-grounded, context-aware assistance* over user-controlled documents.

This work investigates whether the theoretical requirements of the InfoBank concept can be realized in a practical and user-friendly way using current technologies, and whether emerging “chat with your documents” AI tools can be regarded as partial realizations of digital ownership. We structure the feasibility analysis as a component-level mapping from requirements to implementable building blocks, covering data organization, semantic indexing and retrieval, grounded interaction, security and governance, and portability and ownership transfer.

Google NotebookLM exemplifies a source-grounded assistant built around user-provided materials, offering citation-linked responses over selected sources [2, 8], and therefore serves as a useful reference point for our analysis. Rather than proposing a new system, this work provides a structured feasibility assessment of the InfoBank model using existing, widely available technologies. We identify three persistent gaps: (i) ownership transfer for physical objects and their associated digital twins, (ii) interoperability and vendor neutrality beyond proprietary data formats, and (iii) governance mechanisms that extend beyond formal access control toward auditable, user-comprehensible enforcement.

We propose a pragmatic implementation roadmap to narrow the gap between theory and practice: (i) define an open, object-centric data model for digital twins, assigning each physical object a stable identifier and a structured package aligned with emerging product-information initiatives [1]; (ii) standardize a portable package format for transferring object-related subsets of the archive; (iii) integrate RAG-based assistants as the default interface for source-grounded interaction; (iv) connect ownership transfer to decentralized identity standards using registry-independent identifiers (DIDs) and verifiable credentials for interoperable ownership claims [12, 13]; and (v) align governance and portability requirements with regulatory rights and obligations [1, 9, 10].

Although NotebookLM would be a natural choice for our feasibility study, it currently lacks a public API. For this reason, the concrete implementation is based on the OpenAI API, which offers the necessary interfaces for document ingestion, retrieval, and source-grounded interaction.

References

- [1] EUROPEAN PARLIAMENT AND COUNCIL OF THE EUROPEAN UNION: *Regulation (EU) 2024/1781 establishing a framework for the setting of ecodesign requirements for sustainable products*

- (*ESPR*), EUR-Lex legal text, 2024, URL: <https://eur-lex.europa.eu/eli/reg/2024/1781/oj/eng> (visited on 01/14/2026).
- [2] C. GARTENBERG: *The inside story of building NotebookLM*, Google, July 29, 2025, URL: <https://blog.google/innovation-and-ai/products/developing-notebooklm/> (visited on 01/14/2026).
- [3] J. JOHNSON, M. DOUZE, H. JÉGOU: *Billion-scale similarity search with GPUs*, 2017, arXiv: 1702.08734 [cs.CV], URL: <https://arxiv.org/abs/1702.08734>.
- [4] G. KUSPER: *Definition of Digital Ownership, Digital Data Treasure, Informationbank, some Use Cases and a Business Model*, in: Proceedings of FutureRFID 2014, 2014, pp. 101–113, DOI: [10.17048/FutureRFID.1.2014.101](https://doi.org/10.17048/FutureRFID.1.2014.101).
- [5] G. KUSPER, G. ERDEI, A. SZEKERES: *Requirement Specification of an Information Bank*, in: 2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS), 2024, pp. 1–2, DOI: [10.1109/CITDS62610.2024.10791381](https://doi.org/10.1109/CITDS62610.2024.10791381).
- [6] P. LEWIS, E. PEREZ, A. PIKTUS, F. PETRONI, V. KARPUKHIN, N. GOYAL, H. KÜTTLER, M. LEWIS, W.-T. YIH, T. ROCKTÄSCHEL, S. RIEDEL, D. KIELA: *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*, in: Advances in Neural Information Processing Systems, ed. by H. LAROCHELLE, M. RANZATO, R. HADSELL, M. BALCAN, H. LIN, vol. 33, Curran Associates, Inc., 2020, pp. 9459–9474, URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf.
- [7] E. MANSOUR, A. V. SAMBRA, S. HAWKE, M. ZEREBE, S. CAPADISLI, A. GHANEM, A. ABOULNAGA, T. BERNERS-LEE: *A Demonstration of the Solid Platform for Social Web Applications*, in: Proceedings of the 25th International Conference Companion on World Wide Web, WWW '16 Companion, Montréal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 223–226, ISBN: 9781450341448, DOI: [10.1145/2872518.2890529](https://doi.org/10.1145/2872518.2890529), URL: <https://doi.org/10.1145/2872518.2890529>.
- [8] R. MARTIN, S. JOHNSON: *Introducing NotebookLM*, Google, July 12, 2023, URL: <https://blog.google/innovation-and-ai/technology/ai/notebooklm-google-ai/> (visited on 01/14/2026).
- [9] E. PARLIAMENT, C. OF THE EUROPEAN UNION: *Regulation (EU) 2016/679 (General Data Protection Regulation)*, EUR-Lex legal text, 2016, URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng> (visited on 01/14/2026).
- [10] E. PARLIAMENT, C. OF THE EUROPEAN UNION: *Regulation (EU) 2023/2854 (Data Act): Harmonised rules on fair access to and use of data*, EUR-Lex legal text, 2023, URL: <https://eur-lex.europa.eu/eli/reg/2023/2854/oj/eng> (visited on 01/14/2026).
- [11] A. POIKOLA, K. KUIKKANIEMI, H. HONKO: *MyData: A Nordic Model for Human-Centered Personal Data Management and Processing*, White Paper, Finnish Ministry of Transport and Communications, 2015, URL: <https://julkaisut.valtioneuvosto.fi/bitstreams/e21d52e8-44c0-47c7-b512-21d90abc4661/download> (visited on 01/14/2026).
- [12] M. SPORNY, D. LONGLEY, D. CHADWICK, I. HERMAN: *Verifiable Credentials Data Model v2.0*, W3C Recommendation, May 15, 2025, URL: <https://www.w3.org/TR/vc-data-model-2.0/> (visited on 01/14/2026).
- [13] M. SPORNY, D. LONGLEY, M. SABADELLO, D. REED, O. STEELE, C. ALLEN: *Decentralized Identifiers (DIDs) v1.0: Core architecture, data model, and representations*, W3C Recommendation, July 19, 2022, URL: <https://www.w3.org/TR/did-core/> (visited on 01/14/2026).