

KATY — Knowledge At the Tip of Your fingers: Clinical Knowledge for Humanity

Marta C. Silva¹, Susana Nunes¹, Laura Balbi¹, Ana Filipa Rodrigues¹ and Catia Pesquita^{1,*}

¹LASIGE, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Lisboa 1749-016, Portugal

Abstract

The KATY project aims to develop an AI-empowered personalised medicine system to assist medical professionals and researchers in diagnosing patients more accurately, making predictions about their future health, and recommending better treatments. KATY will tackle the challenge of translating AI-based suggestions into practical decision-making processes and treatment strategies that clinicians can understand and trust by combining high-performing black-box machine learning approaches with a comprehensive knowledge graph. The knowledge graph serves as input to AI methods (e.g. directly, through embeddings) as well as encodes the AI outcomes themselves to create a shared semantic space for data, scientific context and predictions capable of supporting explanation methods.

Keywords

Clinical AI-based system, Biomedical Knowledge Graph, Explainable AI, Ontology Alignment

1. Project information

- Project short name/acronym: KATY
- Project full name: Knowledge At the Tip of Your fingers: Clinical Knowledge for Humanity
- Project website link: <https://katy-project.eu/>
- Start date: 01/01/2021
- End date: 30/06/2025
- Project status: Ending project
- Funding agency: European Commission
- Funding programme: Horizon Europe
- Call identifier: H2020-SC1-FA-DTS-2018-2020
- Project partners:
 - **Academic Partners**
 - * UNITOV – University of Rome Tor Vergata
 - * UEDIN – University of Edinburgh
 - * UG – University of Gdańsk
 - * USTAN – University of St Andrews
 - * UNIVIE – University of Vienna
 - * UNIZAR – University of Zaragoza
 - * ULUND – Lund University
 - * NKUA – National and Kapodistrian University of Athens
 - * NTUU – National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”
 - **Research Institutes**

ESWC 2025: Project Networking track, June 1–5, Portoroz, Slovenia.

*Corresponding author(s).

✉ clpesquita@fc.ul.pt (C. Pesquita)

🌐 <https://www.di.fc.ul.pt/~catiaPesquita/> (C. Pesquita)

🆔 0000-0002-1847-9393 (C. Pesquita)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

- * CEA – The French Atomic and Alternative Energy Commission
- * FC.ID – FCiências.ID – Association for Research and Development in Sciences
- * EURECAT – Fundació Eurecat
- * INT – Fondazione IRCSS Istituto Nazionale dei Tumori
- * CHU – Centre Hospitalier Universitaire de Grenoble
- * HPI – The Health Policy Institute
- **SMEs**
 - * CRT – Caretronic
 - * DST – DSTech
 - * EURICE – European Research and Project Office GmbH
 - * OE – Open Evidence
 - * PG – Personal Genomics
- Project logo: <https://katy-project.eu/user/themes/katy/images/logo.svg>
- Project video (optional): <https://www.youtube.com/watch?v=rCWPj3P1jqk>

2. Project summary

Artificial Intelligence (AI) represents great promise in the field of personalised medicine, with its potential to change how we diagnose and treat serious illnesses, such as cancer, by enabling more precise and effective therapies. However, this is not without its challenges, particularly in how AI-generated insights are translated into practical treatment decisions that clinicians and researchers can understand and apply.

The EU-funded KATY project aims to create an AI-powered personalised medicine system designed to support clinicians and researchers in interpreting and using AI data in their daily work. This next-generation technology will bridge the gap between complex data analysis and clinical application, helping to enhance diagnosis and treatment strategies.

Existing targeted cancer therapies often rely on traditional data analysis methods, and AI can elevate this approach by providing deeper insights and more accurate predictions. However, regardless of how powerful the AI solution is, if clinicians cannot interpret or trust recommendations, it will not be adopted into regular routine care.

In KATY, we tackle this challenge by proposing an AI-empowered personalised medicine system that brings human-interpretable knowledge to the tips of the fingers of clinicians and clinical researchers that they can understand, trust, and effectively use in their everyday working routine. KATY is built around two main components: a Knowledge Graph and a pool of eXplainable AI predictors. As a stress test and due to the lack of personalised clinical responses, KATY targets a low prevalence and complex cancer: clear cell renal cell carcinoma.

The KATY project pursues the following specific objectives:

- Linking omics and clinical data in a Knowledge Graph
- Providing a predictive system to clinicians for AI-based treatment recommendations to support them in their process of selecting the treatment best suited to each patient
- Setting up a proof-of-concept application of AI models and knowledge graphs in the context of a clinical pilot in renal cancer
- Reducing the burden of disease for renal cancer patients by applying existing treatments in a more targeted way
- Enhancing the overall diagnostic capacity for complex diseases by using AI-based models to predict patient response to targeted therapies and the identification of molecular evidence to support these predictions.

The knowledge graph serves as input to AI methods (e.g. directly, through embeddings) as well as encodes the AI outcomes themselves to create a shared semantic space for data, scientific context and predictions capable of supporting explanation methods. As part of this project we developed a KG for personalised oncology that builds on a network of 33 integrated ontologies to describe transcriptomic and immunopeptidomics data in the context of cancer, stretching the limits of current ontology matching approaches [1, 2]. Given the lack of ontologies able to properly describe the immunopeptidomics domain, we also developed a novel ontology to address this gap [3]. In the spirit of open-science we have focused on re-using and adapting existing tools, in order to demonstrate their flexibility. AgreementMakerLight [4] was adapted to perform the multi-ontology matching tasks required, and was adapted to connect directly to the ontology editor VocBench [5]. In order to generate explanations for the generated predictions, an explanatory framework was developed that focuses on generating paths through knowledge graphs [6].

3. Project available results

- The Immunopeptidomics Ontology [3]: paper, ontology repository and code repository
- The KATY KG: KG repository
- ESWC2022 In-Use paper detailing the alignment of ontologies to build the KG [1]: paper and code repository
- Explanatory framework that generates explanations based on predicted pairs and knowledge graphs[6].

4. Relevance to ESWC conference

The KATY project provides a valuable example of the real-world impact SW technologies can have in the clinical domain. While data management and integration have been popular applications, there are far fewer examples of KGs and ontologies being used to support explainable AI. The complexity of the biomedical domain provides a driving force for the development of novel algorithms and methods capable of addressing complex domain applications.

We will discuss several lessons learned with a particular focus on ontology alignment and explainable AI. Biomedical ontologies are often related but disjoint, requiring more nuanced logical expressions to link concepts through equivalences, leading to the paradigm of Complex Multi-Ontology Matching (CMOM). CMOM has been addressed by the project by establishing a first foray into solving it with a candidate generation strategy that places no limits in the arity of the ontologies and entities used [2]. Additionally, explaining the results provided by the predictive algorithms is essential to increase the trust of the clinicians and clinical researchers, by anchoring the predictions into the biomedical context and providing a logical path from one entity to the other. Our explanatory framework provides multiple paths through the KG context and thus is anchored in scientific knowledge [6]. This approach shows not only an ability to provide scientifically sound explanations, but as also been well accepted by the professionals surveyed.

5. Value brought by the project to ESWC participants

In the networking sessions, we will discuss the experience of developing the KG and associated ontology with deep input from domain experts and in designing explanatory mechanisms that fit with end users' needs. We have made all outputs available to the community and will take this opportunity to explore how these can be used in subsequent endeavours. As the project is nearing its end, we also hope to explore potential collaborations for future funding opportunities, building on this experience.

6. Value gained by the project from ESWC participation

We are looking forward to exploring potential future collaborations and finding partners for future consortia, but also looking to expand the impact of our work by disseminating it more broadly within the SW community.

Acknowledgments

This work was supported by the KATY project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453, and also partially supported by FCT through the LASIGE Research Unit (UID/00408/2025).

References

- [1] M. C. Silva, D. Faria, C. Pesquita, Matching multiple ontologies to build a knowledge graph for personalized medicine, in: European Semantic Web Conference, Springer, 2022, pp. 461–477.
- [2] M. Silva, D. Faria, C. Pesquita, Complex multi-ontology alignment through geometric operations on language embeddings, in: ECAI, 2024, pp. 1333–1340.
- [3] D. Faria, P. Eugénio, M. Contreiras Silva, L. Balbi, G. Bedran, A. A. Kallor, S. Nunes, A. Palkowski, M. Waleron, J. A. Alfaro, et al., The immunopectidomics ontology (impo), Database 2024 (2024) baae014.
- [4] D. Faria, E. Santos, B. S. Balasubramani, M. C. Silva, F. M. Couto, C. Pesquita, Agreementmakerlight, Semantic Web 16 (2025) SW–233304.
- [5] A. Stellato, M. Fiorelli, A. Turbati, T. Lorenzetti, W. Van Gemert, D. Dechandon, C. Laaboudi-Spoiden, A. Gerencsér, A. Waniart, E. Costetchi, et al., Vocbench 3: A collaborative semantic web editor for ontologies, thesauri and lexicons, Semantic Web 11 (2020) 855–881.
- [6] S. Nunes, S. Badreddine, C. Pesquita, Rewarding explainability in drug repurposing with knowledge graphs, in: IJCAI, International Joint Conferences on Artificial Intelligence Organization, 2025.