Category	Report
Title and Version	The Global Health Network Ontology and AI-powered search – Version 1.0
Valid from	1st July 2024
Date of next review	1st July 2025
Approval date	1st July 2024
Authors	Driver S, Boante L, Segrt A, Lang T
Languages	English

GRANT INFORMATION

Wellcome Grant Number	226688/Z/22/Z
Host Organization	Oswaldo Cruz Foundation (host organisation)
Project Title	The Global Health Network Latin America & the Caribbean: Creating Equity in Health Research by Connecting Excellence and Sharing Know-how.
Type of Grant	Discretionary Award - Data for Science and Health application

TECHNICAL INFORMATION

BACKGROUND

The Global Health Network is a collaborative and open facility that connects researchers, healthcare professionals, and organizations working towards the improvement of global health. Established in 2010, The Global Health Network has played a pivotal role in fostering collaborations, disseminating knowledge, and facilitating capacity building in global health research.

The Global Health Network has an immense volume of resources that has been created by and is available to a vast and trusted community. Currently there are 65 Knowledge Sharing Hubs, each hosting their own set of resources such as tools, templates, SOPs, guidance documents, workshop and webinar recordings, and learning materials. Many of these are resources that cut across various topics and knowledge hubs. The success of this body of resources has created a challenge in making sure they are discoverable from any part of The Global Health Network platform. The aim is to ensure that any user, whether they are experienced in the Network or new to it, can find what they are looking for, discover what they need, and keep up with news they follow and might be interested in.

In order to achieve this, all existing content on The Network needs to be catalogued and indexed more granularly so that resource metadata is more comprehensive, relationships between resources can be identified, and resources can be ranked in terms of importance. Existing medical ontologies were reviewed but found to be unsuitable on their own for the type of social science, qualitative content that is common across The Global Health Network. Therefore, it was decided to create a The Global Health Network ontology and index resources against that to create a more structured and relevant database.

Several years ago, The Global Health Network and the World Health Organization created a Professional Development Scheme for researchers and research teams. The Professional Development Scheme provides a comprehensive, high quality mechanism for recording, tracking

and guiding professional development in health research by capturing core competencies, qualifications and training. A competency dictionary was created as part of this work, with a hierarchical structure and keywords for topics relevant to researchers in global health and ideally suited to become the foundation for The Global Health Network's ontology.

The Global Health Network team has scoped various technology solutions to improve discoverability including leveraging Google Cloud indexing and natural language processing APIs, University of Sheffield's GATE proprietary text analysis and language processing application, and an ontology builder powered by a graph database by Cognitive City. These avenues were explored and were found to be costly in time and resource and largely superseded by advances in machine learning and artificial intelligence.

Advances in Artificial Intelligence (AI) presents a judicious opportunity to enable transformational access to information and knowledge to research teams across the globe – if applied responsibly, expertly and ethically. If applied with these principles at the forefront, this stands to enhance the efficiency and effectiveness of knowledge discovery across the field. The opportunity that AI technology brings is about providing curated, focussed and trusted access to this knowledge, and that of connected partners and organisations, to deliver precisely the information required by the research community relevant to the exact context they are working in.

PROCESS FOR REVIEW THE APPROPRIATE ONTOLOGY

Step 1 - Define Objectives

1. **Implement AI Technologies:** Integrate state-of-the-art AI technologies into the existing infrastructure of The Global Health Network and their connected partners to standardise global health research terminology, and optimize and automate knowledge organization, categorization, discovery and retrieval.
2. **Create a proof of concept around outbreaks research.** Build Enhanced Search and Recommendation System. Create a global health research ontology using the existing competency dictionary as a foundation to index the current database. Feed the index and associated database of resources into machine learning algorithms for continuous training and leverage a retrieval-augmented generation workflow to surface contextually relevant resources and recommendations, accessed by semantic search with generative AI capabilities, and further filterable by employing a faceted search feature.
3. **Facilitate Multilingual Accessibility:** The semantic search empowered by large language models (LLMs) will allow a user to find precise and relevant recommended resources using natural language in a conversational way in multiple languages, ensuring inclusivity and accessibility for a diverse global audience.

EXPECTED RESULTS

1. **Improved Knowledge Discovery for better studies:** Enhanced recommendations and discoverability will lead to more equitable, efficient and accurate discovery of relevant methods, standards, training and tools improving the generation of quality evidence to drive improvements in health, everywhere.
2. **Increased Platform Usage:** The integration of AI is expected to attract a wider audience, fostering increased collaboration and knowledge-sharing among global health research teams and organisations and other networks

3. **Diverse and Inclusive Access:** Multilingual support will make the platform accessible to a broader and more diverse use across roles and regions, promoting inclusivity and opportunity in global health research.

SUSTAINABILITY PLAN

Upon completion of the project, a sustainability plan will be implemented to ensure the ongoing effectiveness and relevance of the AI-driven enhancements. This plan includes:

1. **Continuous Monitoring and Optimization:** Regular monitoring of machine learning algorithms and outputs to ensure accuracy, responsibleness and relevance, with a commitment to ongoing optimization based on user feedback.

CONCLUSION

The integration of Artificial Intelligence into The Global Health Network and our wider partners organisations represents a ground-breaking opportunity to understand the potential to revolutionize the way global health research knowledge is discovered, accessed, and utilized. The Global Health Network is the most appropriate facility in which to test and learn about this rapidly evolving field because it has the critical mass of information needed, it has a trusted and connected community, it is open access and will benefit all, and it works across all diseases, all types of research, and all settings. This can enable unparalleled curated access to trusted methods, approaches, skills and standards from across The Global Health Network's community and the wider connected organisations. This will make knowledge available to teams that they did not know existed, removing barriers and raising standards. We can also measure how these changes and increases the ability to undertake research in many settings and in teams where a lack of access to know-how, training and methods has prevented engagement in research. This project could serve as a key indicator of how to best harness this innovation responsibly in applications across the field of research in all settings in global health.