MEDINFO 2025 — Healthcare Smart × Medicine Deep
M.S. Househ et al. (Eds.)

doi:10.3233/SHTI251086

Automated Elicitation of Human and Ecological Health Indicators: An LLM-Based Practical Implementation for One Digital Health

Helit BAUBERG[a], Nimrod TACHNAI[a], Gur HANAN[a], Dana NEHAMA[a], Oscar TAMBURIS[b], Stéfan DARMONI[c,d], Julien GROSJEAN[c,d], Arriel BENIS[a,1]
[a] *Holon Institute of Technology, Israel*
[b] *Institute of Biostructures and Bioimaging, National Research Council, Naples, Italy*
[c] *Department of Digital Health, Rouen University Hospital, France*
[d] *Analyse Intégrée Multimodale en Santé (AIMS) lab,*
ORCiD ID: Oscar TAMBURIS https://orcid.org/0000-0002-0130-7915, Julien GROSJEAN https://orcid.org/0000-0002-7446-644X, Stéfan DARMONI https://orcid.org/0000-0002-7162-318X, Arriel BENIS https://orcid.org/0000-0002-9125-8300

Abstract. This paper presents a new method for automating the identification of human and ecological health indicators using the One Digital Health framework, which combines One Health and Digital Health principles. By applying mainly Large Language Models, we conduct a systematic literature review on urban freshwater environments. This automation streamlines the process of finding and analyzing relevant research, allowing us to extract vital health indicators related to urban aquatic ecosystems and human wellness. The findings support the OneAquaHealth project's goals, enhancing environmental monitoring and linking human, animal, and environmental health in a digital context.

Keywords. Digital Health, One Digital Health, One Health, Large Language Models, Knowledge Discovery, Health Indicators, Ecological Indicators, Review, Automation

1. Introduction

One Digital Health (ODH) is a framework that integrates One Health (OH) and Digital Health concepts to create a unified approach to health ecosystems. [1] It takes a holistic view, addressing human health, animal health, and environmental factors (i.e., the OH umbrella) while leveraging DH-related technologies from different fields, such as artificial intelligence (AI) or the Internet of Things (IoT). ODH promotes proactive health ecosystems, leveraging FAIR (Findable, Accessible, Interoperable, and Reusable) access to data and analytics to improve understanding of complex ecosystems and enhance trend detection. [2]

[1] Corresponding Author: Arriel Benis, Department of Digital Medical Technologies, Holon Institute of Technology, Golomb St. 52, Holon, 5810201, Israel; E-mail: arrielb@hit.ac.il.

In this regard, the European Union-funded OneAquaHealth (OAH, https://www.oneaquahealth.eu/) project aims to improve the sustainability and integrity of freshwater ecosystems in urban areas based on the One Health and ODH approaches. OAH explores the link between ecosystem health and human well-being to enhance environmental monitoring, restore aquatic ecosystem health, promote One Health, and support a better understanding of city health concerns. The process begins with curating collections of scientific papers for indicators elicitation and discovery. Conducting a manual systematic literature review is resource-intensive and often lacks comprehensiveness, especially given the exponential growth of published research. [3] [4] Our main aim was to demonstrate how AI, and more precisely, Natural Language Processing-driven models, such as Large Language Models (LLMs), can support the identification of human and ecological health indicators in urban freshwater stream environments, seen as a medium of interactions with humans and animals that can be monitored by using digital tools. [4]

Our primary objective was to automate a systematic literature review following the PRISMA guidelines. This review aimed to identify health indicators and other relevant topics from the literature. Our goal was to develop a functional platform powered by AI and NLP that could support the objectives of OAH by effectively processing, analyzing, and aligning them with key elements, perspectives, and dimensions of ODH.

2. Methods

The project focused on developing a process that supports a comprehensive automated literature review related to the topics OAH focuses on. This involved building the prompts flow to identify, screen, and assess the eligibility of research papers [5] similarly to the PRISMA flow diagram and guidelines [6], in combination with predefined validity criteria designed to categorize each paper based on its relevance to human or ecological health indicators.

Thus, we automated the classical systematic literature review methodology following the process: (1) Formulating the research question *(i.e., identification of human and ecological health indicators from the literature)*; (2) Planning a detailed protocol to guide the research *(i.e., building a set of prompts allowing categorizing papers by related domain -human, animal, or ecological health alone or combined)*; (3) Collecting relevant papers through PubMed by using a broader set queries running of "All fields" *(i.e., (Urban Freshwater Environments), (Human Health Indicators) AND (Urban Aquatic Ecosystems), (Ecological Health Indicators) AND (Urban Aquatic Ecosystems), (Human Health Indicators) AND (Urban Aquatic Ecosystems), Ecology AND (Urban Aquatic Ecosystems), (Human Health) AND (Urban Aquatic Ecosystems), (Urban Aquatic Ecosystems))*; (4) Preprocessing by removing multiple instances of redundant papers; (5) First screening the collected papers by defining and handling the inclusion or exclusion criteria *(i.e., related or not to ODH or OAH by building a set of specific prompts to select relevant papers based on their title and abstract only)*; (6) Second screening of the remaining papers based on their full content *(i.e., by using a set of prompts designed for considering the relevancy to ODH or OAH and the presence of potentially interesting indicators)*. Extracting pertinent data (*e.g., indicators, values*) from each paper; (7) Synthesizing the extracted data and information across the papers that passed the second screening. In the final stage of this step, all curated data from each paper are compiled

into a structured summary, which is then saved in a text file for domain experts to review for a human-based assessment. (Table 1)

The LLMs handled steps 5 and 6 with specific prompts focused on (a) identifying recurring themes between indicators, (b) detecting disease vectors, and (c) recognizing early warning signs of ecosystem degradation. Experts in each field validated all the prompts used.

The framework was written in Python 3. Input and output files were accessed from Google Drive. Full-text PDF files of the papers were manually downloaded to Drive from the provided sources. The LLM versions used were Gemini-1.5-flash and GPT-4o-mini (OpenAI for ChatGPT).

3. Results

Figure 1 presents the PRISMA flow diagram for the first step of automated human and ecological health indicators elicitation.

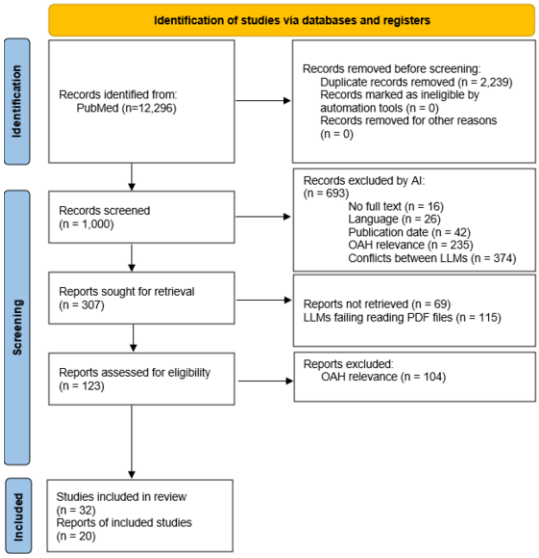

Figure 1. PRISMA flow diagram for the Automated Elicitation of Human and Ecological Health Indicators.

The large number of papers initially identified as potentially of interest was 12,269. 10,057 were unique (82.0%). A subset of 1,000 papers (9.94% of the unique ones) was randomly selected due to limited technical resources allocated to this proof-of-concept experiment. Considering exclusion criteria (i.e., no full text available, non-English papers, Publication date before January 2013 or after May 2024, non OAH relevance - i.e., not related at all to “urban freshwater streams”-, disagreement between the LLMs regarding previous inclusions/exclusion criteria -e.g., not rightly handling the “non ODH relevance”-), 307 (30.7%) were sought for retrieval. 69 (22.5%) were not retrieved due to the non-availability of the PDFs through the access provided by the institutional electronic library used; moreover, 115 (37.5%) of the PDF versions were not readable

by the LLMs. 123 (40.1%) were finally evaluated for eligibility with specific prompts; 32 (26.0%) were included in the review process and forwarded to the next step.

The papers that were finally selected were compared with those chosen in a similar PRISMA review handled by human researchers end-to-end. The overall result was around 75% overlapping, even though the automated process ran on a subset of only 1,000 papers. The initial number of papers identified as potentially engaging in the human-based review was around 4,500 (44.7%).

As this paper aims to mainly present the methodological process, one example of a paper included in this automated systematic review is presented below. It relies on an article entitled "*Protecting and restoring freshwater biodiversity across urban areas in Aotearoa, New Zealand: Citizens' reporting of pollution in stormwater drains and waterways*." [7] OpenAI processed it in 102 seconds, compared to a human domain expert who needed at least 1 hour. [8]

Table 1. Abstract of the key findings' summary related to ODH and OAH in GPT-4o-mini

Prompt	OpenAI answer
Ecological and Human Health Indicators […]. Ecological indicators measure ecosystem health and the impact of human activities, while human health indicators assess population well-being and public health effectiveness. […]. What indicators are mentioned in the paper? How were they measured?	The paper mentions several indicators […]: 1. Ecological Indicators: - Biological: ***Species diversity***. - Physical: ***Water quality***. - Chemical: ***Pollutant levels***. 2. Human Health Indicators: - While specific human health indicators are not detailed in the provided context, the ***focus is on behaviors that can impact both ecological health and community well-being***. […]
Express the key findings of the paper in terms of the relationship between the indicators discussed in the paper. Reply in plain text.	The paper's key findings reveal a ***complex interplay between various indicators influencing urban residents' behaviors related to freshwater biodiversity and pollution reporting***. Perceived responsibility also plays a vital role, […] highlighting the need for tailored interventions that address specific barriers faced by different audience segments. […] is essential for ***promoting sustainable behaviors and enhancing freshwater conservation efforts***.
What One Digital Health keys are addressed in this paper?	The keys addressed in this paper are ***One Health*** and ***Digital Health***.
What One Digital Health perspectives are addressed in this paper?	The perspectives addressed in this paper are ***individual health and well-being***, ***population and society***, and ***ecosystem***.
What One Digital Health dimensions are addressed in this paper?	The dimensions addressed in this paper are ***environment*** and ***human and veterinary health care***.

4. Discussion and conclusion

The automated elicitation of human and ecological health indicators through AI and NLP significantly supports ODH integration in challenging multidisciplinary and interdisciplinary complex projects involving the area. By leveraging LLMs, our study demonstrates that comprehensive literature reviews can be streamlined, making it feasible to manage the ever-expanding body of research in broader fields, such as the one investigated in the OAH project, which examines the interactions between human health and behavior and urban freshwater ecosystems. The ability to sift through thousands of papers in a scalable manner and in a short amount of time facilitates the elicitation and enhances the comprehension of existing health—and ecology-related topics. This research presents an LLM-based literature review tool that cuts the review process from

months to a few days. Domain experts found the tool's outputs accurate and relevant, closely matching human reviews and proving valuable for timely insights. However, reliance on automated systems must be approached with caution. The predefined validity criteria and prompts used to direct the LLMs were essential in maintaining relevance. Yet, due to LLM hallucinations, there is a risk of overlooking nuanced findings that may not conform to established categories. [9] Thus, reducing ethical issues involving domain experts in validation seems critical to mitigate this risk. Our research shows that automating systematic literature reviews can effectively extract key health indicators in urban freshwater streams. Next, we plan to cluster papers by domains identified by LLMs and generate hypotheses on emerging topics. This work aims to promote healthier, sustainable urban ecosystems and highlights the transformative potential of technology in tackling current health and ecological issues.

Acknowledgments

This work was supported by the European Union's Horizon Europe Research and Innovation Program (HORIZONCL6-2022-GOVERNANCE-01) under grant agreement No. 101086521 - OneAquaHealth (Protecting urban aquatic ecosystems to promote One Health).

References

[1]. A. Benis, O. Tamburis, C. Chronaki, and A. Moen, One Digital Health: A Unified Framework for Future Health Ecosystems, *J Med Internet Res*. 23 (2021) e22189. doi:10.2196/22189.

[2]. O. Tamburis, and A. Benis, One Digital Health for more FAIRness, *Methods Inf Med*. (2022) a-1938-0533. doi:10.1055/a-1938-0533.

[3]. B. Nussbaumer-Streit, M. Ellen, I. Klerings, R. Sfetcu, N. Riva, M. Mahmić-Kaknjo, G. Poulentzas, P. Martinez, E. Baladia, L.E. Ziganshina, M.E. Marqués, L. Aguilar, A.P. Kassianos, G. Frampton, A.G. Silva, L. Affengruber, R. Spjker, J. Thomas, R.C. Berg, M. Kontogiani, M. Sousa, C. Kontogiorgis, and G. Gartlehner, Resource use during systematic review production varies widely: a scoping review, *Journal of Clinical Epidemiology*. 139 (2021) 287–296. doi:10.1016/j.jclinepi.2021.05.019.

[4]. F.M. Delgado-Chaves, M.J. Jennings, A. Atalaia, J. Wolff, R. Horvath, Z.M. Mamdouh, J. Baumbach, and L. Baumbach, Transforming literature screening: The emerging role of large language models in systematic reviews, *Proc Natl Acad Sci U S A*. 122 (2025) e2411962122. doi:10.1073/pnas.2411962122.

[5]. B. Meskó, Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial, *J Med Internet Res*. 25 (2023) e50638. doi:10.2196/50638.

[6]. M.J. Page, D. Moher, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, R. Chou, J. Glanville, J.M. Grimshaw, A. Hróbjartsson, M.M. Lalu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, and J.E. McKenzie, PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews, *BMJ*. (2021) n160. doi:10.1136/bmj.n160.

[7]. L.J. McLeod, D.W. Hine, T.L. Milfont, Z. Dorner, N.A. Tassell-Matamua, R.D.C. Maris, J.C. Kitson, and P. Stahlmann-Brown, Protecting and restoring freshwater biodiversity across urban areas in Aotearoa New Zealand: Citizens' reporting of pollution in stormwater drains and waterways, *J Environ Manage*. 351 (2024) 120019. doi:10.1016/j.jenvman.2024.120019.

[8]. J. Clark, P. Glasziou, C. Del Mar, A. Bannach-Brown, P. Stehlik, and A.M. Scott, A full systematic review was completed in 2 weeks using automation tools: a case study, *Journal of Clinical Epidemiology*. 121 (2020) 81–90. doi:10.1016/j.jclinepi.2020.01.008.

[9]. F. Fatehi, M. Samadbeik, and A. Kazemi, What is Digital Health? Review of Definitions, in: A. Värri, J. Delgado, P. Gallos, M. Hägglund, K. Häyrinen, U.-M. Kinnunen, L.B. Pape-Haugaard, L.-M. Peltonen, K. Saranto, and P. Scott (Eds.), Studies in Health Technology and Informatics, IOS Press, 2020. doi:10.3233/SHTI200696.