SoftwareX 33 (2026) 102536

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Collaborative hybrid intelligence platform *CHIP*: A modular architecture for developing and testing personalized lifestyle support interactions

Floris den Hengst [a,*], Shaad Alaka [b], Bart A. Kamphorst [c]

[a] *Vrije Universiteit Amsterdam, Amsterdam, Netherlands*
[b] *TU Delft, Delft, Netherlands*
[c] *Utrecht University, Utrecht, Netherlands*

ARTICLE INFO

Keywords:
Lifestyle support
Artificial intelligence
Dialogue
Knowledge graphs
Hybrid intelligence

ABSTRACT

The rise of lifestyle-related, non-communicable diseases such as Type II diabetes, cardiovascular diseases, and depression has prompted the development of various behavior change technologies to promote sustained healthy behaviors. User adherence, however, has remained low.

The Collaborative Hybrid Intelligence Platform CHIP is introduced to address adherence challenges by placing the user perspective at the center and facilitating dialogue-based interactions between users and their technical and non-technical support systems—including AI systems, clinicians and caretakers. These interactions aim to uncover barriers to adherence and collaboratively shape personalized lifestyle plans that align with a person's preferences, values, and context.

CHIP is a microservice-based research platform written in Python with modules implemented as Docker containers. Its modularity allows researchers to replace or adapt specific components, such as natural language reasoners, for technical evaluation and domain-specific adaptation.

Metadata

Code metadata.

Nr.	Code metadata description	Metadata
C1	Current code version	0.0.1
C2	Permanent link to code/repository used for this code version	For example: https://github.com/hybrid-intelligence/CHIP
C3	Permanent link to Reproducible Capsule	
C4	Legal Code License	Apache License V2
C5	Code versioning system used	git
C6	Software code languages, tools, and services used	python, javascript, docker engine, redis
C7	Compilation requirements, operating environments & dependencies	git, docker engine
C8	If available Link to developer documentation/manual	
C9	Support email for questions	f.den.hengst@vu.nl

1. Motivation and significance

The global increase in chronic diseases—including Type II diabetes, cardiovascular disease, and depression—presents significant societal challenges [1,2]. Although the specific etiologies and management strategies vary between conditions, behavioral changes to lifestyle factors such as diet, physical activity, and stress management are widely recognized as primary means for prevention, management, and even remission of many of these persistent health conditions [3–5].

Adherence to lifestyle modifications remains challenging in practice [6–10]. Previous research suggests that lifestyle recommendations should be highly personalized, incorporating the patient's unique

* Corresponding author.
Email addresses: f.den.hengst@vu.nl (F. den Hengst), s.alaka@tudelft.nl (S. Alaka), b.a.kamphorst@uu.nl (B.A. Kamphorst).

https://doi.org/10.1016/j.softx.2026.102536
Received 26 October 2025; Received in revised form 10 January 2026; Accepted 22 January 2026
Available online 31 January 2026

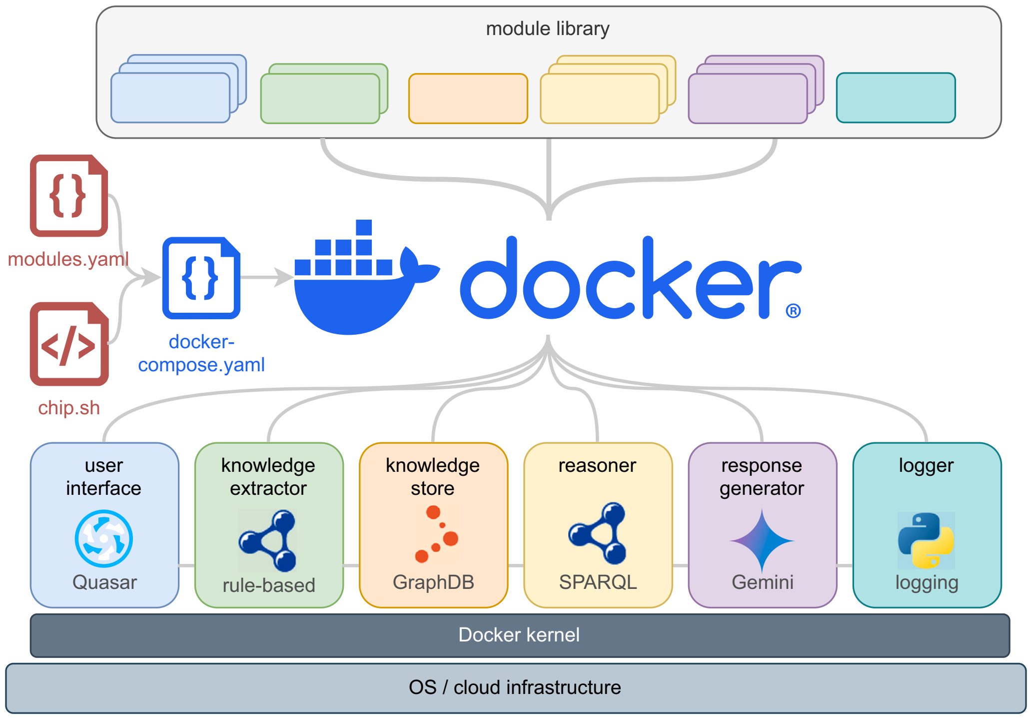

Fig. 1. System architecture. The system consists of six interconnected modules implemented as docker containers. These modules can be selected and configured from a library of existing modules based on the researchers' needs. The user specifies the configuration in `modules.yaml` and runs a command using the `chip.sh` script. This script generates the `docker-compose.yaml` and launches the docker containers.

context, preferences, and values to promote adherence and thus facilitate lasting beneficial health outcomes [11–14].

AI-driven lifestyle support technologies, in particular those based on machine learning, have long been recognized as promising means of offering individuals tailored support in their behavior change trajectories [15,16]. Although promising, most approaches in practice have tended to consider behavior change primarily as an individual enterprise, neglecting environmental factors such as community norms, affordances such as healthy food options, and social support [17]. Failing to incorporate a person's surrounding social and cultural environment, and neglecting the influence of family members, caretakers, and clinicians on the success of someone's behavior change trajectory, has led to low user adherence to technological behavior change support tools, thereby limiting their real-world effectiveness.

Another obstacle of existing technologies has been their lack of transparency. Many existing systems operate as opaque "black boxes" that deliver recommendations without grounded explanations, thus eroding user trust and sustained engagement [18–20]. Some explainable systems exhibit limited adaptability and provide generic advice that does not resonate with diverse patient populations [21]. Furthermore, a common architectural limitation is an insufficient grounding in symbolic reasoning, which complicates the interpretation and clinical justification, and limits compliance with European guidelines such as the Healthier Together NCD initiative [22].

The Collaborative Hybrid Intelligence Platform (CHIP hereafter) aims to address these critical gaps with a hybrid intelligence approach. By combining the scalable adaptivity of machine intelligence with human input from patients, caregivers, and clinicians, the open source software we present here aims for collaboration and patient-centered decision-making. CHIP's core strength lies in interpretable decision- making through the construction and maintenance of knowledge graphs [23]. Specifically, it uses a general domain knowledge graph for health and lifestyle information, along with a user-specific knowledge graph for individual context, values and preferences [24].

When end users interact with CHIP through its chat interface, the user-specific knowledge graph is updated dynamically. This way, CHIP can learn preferences through its user interactions, allowing the reasoner to make recommendations that align with the advice of professionals as well as with the patient's preferences and personal context.

CHIP uniquely serves as a platform for rapid prototyping and evaluation of modular decision-making hybrid human-AI pipelines in research. Related research platforms focus narrowly on either dialogue without any hybrid human-AI decision-making [25–28], or focus on hybrid human-AI decision- making without involving knowledge graphs and without prioritizing interpretability and reasoning correctness [29,30]. CHIP fills this gap by providing a rich and extensible platform for studying the effectiveness of explainable and hybrid AI in lifestyle support with a modular design.

2. Software description

2.1. Software architecture

CHIP is built as a modular and containerized system that facilitates deployment across environments. Fig. 1 provides a visualization of the different modules and their default implementation following a microservice architecture wherein each microservice is an independently deployable, modular service that performs a single function and communicates with other services through well-defined, lightweight APIs [31].

The architecture comprises six core modules, each encapsulated as an independent Docker microservice. Table 1 lists the interfaces for each module. The modules communicate via HTTP endpoints and are orchestrated using an orchestration script called `chip.sh` which interfaces with Docker Compose. Together with a `modules.yaml` configuration file, this script enables the easy start, stop, and reconfiguration of the system with minimal overhead costs. Each module further contains a `compose.yaml` file which specifies its Docker container.

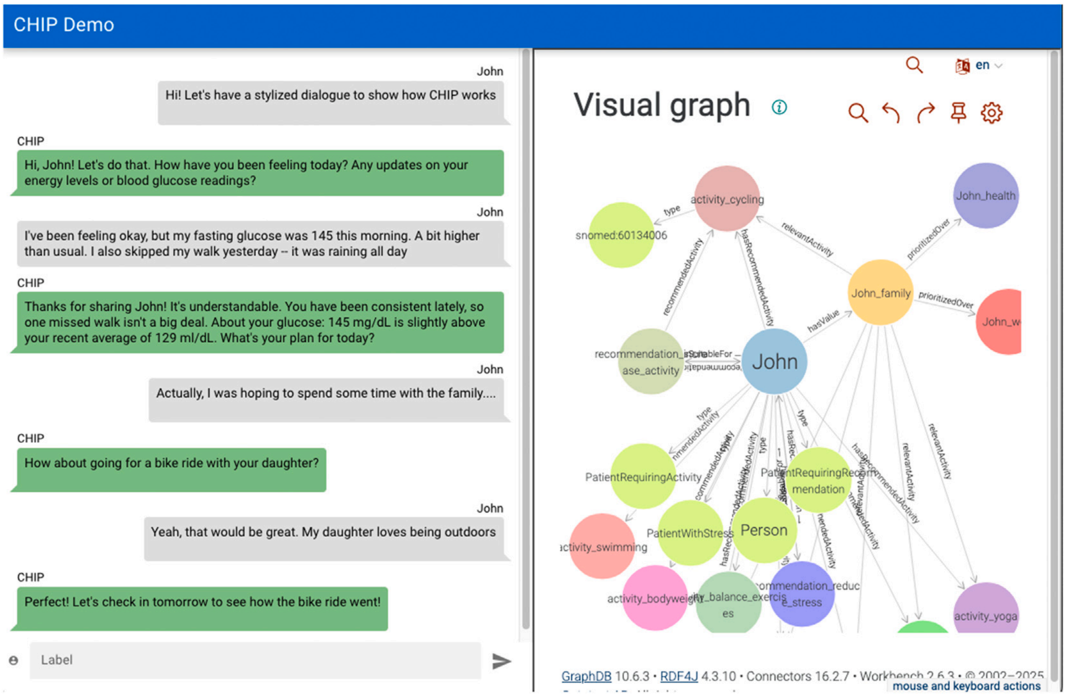

Fig. 2. Quasar-powered default implementation of the user-facing front-end with a dialogue capability (left) showing a constructed, illustrative dialogue and an interactive visualization of the current state of the knowledge graph (right).

Table 1
Interfaces for each component.

User Interface			
Accepts	**Description**	**Returns**	**Description**
`str`	patient input	`JSON`	patient name, sentence, timestamp
Knowledge Extractor			
Accepts	**Description**	**Returns**	**Description**
`JSON`	patient name, sentence, timestamp	`JSON`	sentence data, SPO triple
Reasoner			
Accepts	**Description**	**Returns**	**Description**
`JSON`	sentence data, RDF triple	`JSON`	sentence data, query result data, query result type
Knowledge Store			
Accepts	**Description**	**Returns**	**Description**
`SPARQL`	query	`str`	query result
Accepts	**Description**	**Updates**	**Description**
`SPARQL`	RDF triple insertion	`knowledge`	inserts a new RDF triple
Response Generator			
Accepts	**Description**	**Returns**	**Description**
`JSON`	sentence data, query result data, query result type	`str`	natural language response
Logger			
Accepts	**Description**	**Updates**	**Description**
`log entry`	message, severity	`file`	centralized log

CHIP runs on any OS or cloud infrastructure that supports Docker Compose and bash, including all major cloud providers and Linux distributions, macOS and Windows OS (via WSL2), ensuring portability and scalability. Each core module encapsulates a distinct functional component and comes with an open source default implementation to ensure that the system can be used out-of-the-box without proprietary dependencies.

2.2. Software functionalities

User interface. A web-based front-end supports chat-based dialogue with end users and real-time visualization of the agent's knowledge graph as exemplified in Fig. 2.[1] The intuitive chat interface is currently targeted towards the person receiving lifestyle support, but can in the future also be used to interact directly with clinicians, caretakers, and the social network of the patient. The default implementation is based on Quasar [32], which offers efficient cross-platform deployment, a rich set of responsive UI components, and seamless integration with real-time data sources.

Knowledge extractor. This module converts free-form dialogue text into a knowledge graph format i.e., subject-predicate-object triples in RDF [33]. The default implementation uses predefined rules and Natural Language Processing (NLP) techniques.

Knowledge store. A repository for all data within CHIP, including the domain and user-specific knowledge graphs in an RDF database. The default implementation is based on GraphDB [34] and comes with knowledge about Type II diabetes.

Reasoner. This component determines whether to ask a question or provide lifestyle advice. We provide a basic reasoner based on SPARQL

[1] The knowledge graph visualization is for testing and demonstration purposes only and can be hidden during experiments.

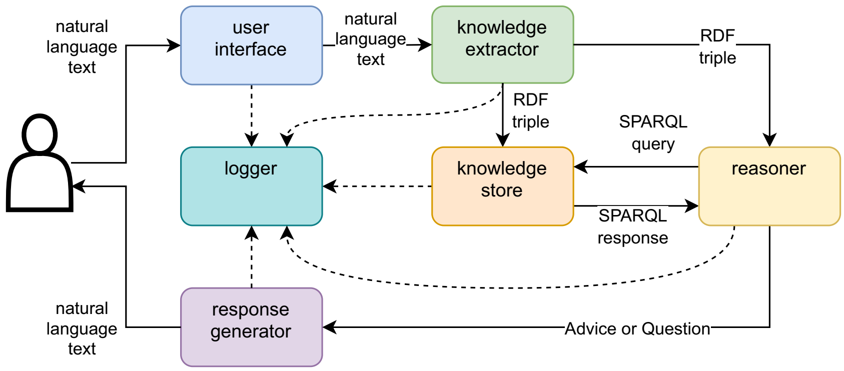

Fig. 3. Overview of system modules.

queries with SWRL rules by default [35,36]. More advanced approaches, for example those that actively assess and improve the quality of their knowledge through conversation by either reinforcement learning or logic-based reasoning, are possible as well [37,38].

Response generator. This module synthesizes a coherent and natural-sounding response based on the structured response provided by the Reasoner. Responses are either questions for clarification or preference elicitation, or statements such as answers to user questions, explanations, and suggestions. The default implementation consists of a conversion of the structured response into a basic sentence structure using a set of templates. Alternatives are based on large language models (LLMs) Gemini[2] or the Creative Commons-licensed MedAlpaca [39,40]. These LLM alternatives take into account the conversational context and produces sentences that are contextually aligned with the flow of the conversation.

Logger. This module can be used to analyze user engagement, evaluate the effectiveness of different dialogue strategies, and study how the system's reasoning affects adherence. The default implementation is based on Python's `logging` module.

3. Illustrative example

In this section, we demonstrate the usage of the system from the perspective of a researcher aiming to study a particular research question. The researcher begins by executing the `chip.sh` utility to bootstrap the experimental environment: `./chip`. This script generates the necessary configuration infrastructure (Fig. 3).

First, the script generates a `core-modules.yaml` file, which defines the active microservice implementations for each component listed in Fig. 1 and Table 1 as follows:

```
logger_module: logger-default
frontend_module: front-end-quasar
response_generator_module: response-generator-default
reasoner_module: reasoning-default
triple_extractor_module: text-to-triples-rule-based
```

The system then generates a `setup.env` file for each module, linking to its respective network endpoint using naming conventions. This ensures that services can resolve each other within the containerized network. Thirdly, the script creates local configuration files for each module based on a version-controlled configuration template. This enables the researcher to change local parameters such as API credentials without modifying the global source code.

The researcher can now launch the services by executing the following command: `./chip start`. This command generates and updates the required configuration files for all modules, ensuring that the system's components can interact as required. It also launches the interlinked Docker containers based on the generated configuration, and serves the UI so that the researcher can visit the specified URL to be presented with the UI in Fig. 2.[3] Depending on the specifics of the research questions involved, one or more modules from Fig. 1 can be altered while the other modules remain intact.

Consider an example research question about the performance of an LLM-based reasoner.[4] A response generator based on an LLM can be added as a new module by creating an API endpoint `/process` that accepts data from the reasoner and passes it into the new response generator. Creating this module is a matter of implementing the code in Listing 1, which contains only code to generate an LLM-based response, along with a minimal amount of boilerplate code to link to the correct internal API calls following the microservices design philosophy. Conducting an experiment with the updated module now only requires updating one of the values in the above-mentioned `modules.yaml` configuration file.

4. Impact

CHIP's main impact lies in enabling and facilitating the study of various hybrid intelligence research questions. Broadly speaking, CHIP helps in answering two types of research questions: technical and empirical research questions.

4.1. Technical research

Under the heading of technical research we subsume algorithmic innovations involving user preferences, dialogue, knowledge representation, explanations, and reasoning. A major enabler of this type of research is the platform's modular microservice-based design, allowing for components to be exchanged to study their effects.

CHIP has been used to develop an AI-based diabetes lifestyle management support prototype [24], and to show how to integrate dialogue, information extraction and reasoning in hybrid intelligence systems [42]. A recent study on dialogue-based knowledge acquisition has used the CHIP knowledge base to show that an agent can learn effective policies for acquiring knowledge about a user's preferences by assessing its current knowledge in tandem with adaptively determining the effectiveness

[2] Gemini requires access to a proprietary API. All default implementations come with open source licenses; for a full list, see https://github.com/hybrid-intelligence/CHIP/blob/main/DEPENDENCIES.md.

[3] By default, the web interface can be reached at http://localhost:9000.

[4] A note of caution with this example, as the use of LLMs in healthcare is still under debate [41]. The example illustrates that CHIP offers the infrastructure to study how to use LLMs responsibly.

```python
from flask import Blueprint, current_app, request
import some_llm
import requests

bp = Blueprint('main', __name__)

def generate(advice):
    prompt=f"""
    You act the role of a medical chat bot, that is able to link facts about the patient and about medical science in order to give advice. You do not need to do this linking yourself as this will be given to you if available. I will give you personalized advice for this patient, and you will attempt to formulate this into an appropriate response to the user. Your replies should be succinct, to the point, and strictly in line with the advice given.
    Advice:
    {advice}
    """
    response = some_llm.generate_content(model="my_selected_model", contents=[prompt])
    return response.text

@bp.route('/process', methods=['POST'])
def submit_reasoner_response():
    # obtain advice from reasoner message
    reasoner_resp = request.json["advice"]
    current_app.logger.info(f"From reasoner: {reasoner_resp}")
    # generate natural language response
    language_response = generate(reasoner_resp)

    # send natural language response to the front-end UI module
    address = current_app.config.get("FRONT_END_ADDRESS", None)
    requests.post(f"http://{address}/process",
                  json={"message": language_response}
                  )
    return 'OK'
```

Listing 1. Example module in CHIP: an LLM-based response generator.

of several generic knowledge-acquisition capabilities during interactions with users [37].

Detecting potential instances of deception and self-deception is another challenge that has already benefited from CHIP. This line of research is concerned with flagging when someone is not reporting truthfully about their physical activity or diet. In this area, Koot and van Paridon [43,44] used CHIP to study the detection of deception in diabetic patients, while Madaras [45] compared empathetic and affirming interventions to address the sources of deception.

CHIP has furthermore been deployed to compare the efficacy and flexibility of various knowledge representation, engineering and explanation approaches [46]. Specifically, CHIP has been used to compare approaches to extract knowledge structured as RDF triples from natural language text by comparing various Knowledge Extractors [47].

CHIP is envisioned to support studies into dynamically updating preference rankings from a mix of behavioral data and natural language interactions using techniques from machine learning [48], particularly reinforcement learning [26]. For example, can mobility and weather data be combined to generate hypotheses about a person's activity preferences, e.g., "[user] likes cycling, except when it is pouring." Such hypotheses could be subtly tested in conversation, e.g., "It's sure coming down at the moment. I wouldn't want to be caught in this kind of weather, would you?".

Another line of inquiry pertains to the integration of heterogeneous viewpoints from multiple stakeholders. In hybrid intelligence settings, conflicting viewpoints and advice at various levels of detail are inevitable: a dietitian can recommend a high-level dietary strategy such as 'portion control,' while a general practitioner may emphasize limiting the intake of refined sugar. At the same time, daily caregivers might hold the opinion that neither process will be successful unless a daily routine is established and awareness is created about how food intake affects blood sugar levels. CHIP can help study how to weigh and resolve different viewpoints, and how to select the right stakeholders to acquire certain information.

The last set of technical questions we highlight relates to reasoning, for example, about dealing with inconsistencies in a user's medical or behavioral history [49]. Hybrid human-AI reasoning strategies might help to effectively use the breadth of knowledge available in a patient's wider support network, but designing such strategies remains largely an open research question [50,51]. To study such strategies, the default reasoner can be replaced by a reasoner using, e.g., default logic to draw tentative conclusions based on imperfect information about the users' preferences and context, enabling comparative research on automated decision-making logic.

4.2. Empirical research

CHIP also supports empirical research, particularly HCI studies involving human subjects. First, the technical innovations discussed in the previous section might be evaluated with actual end users, e.g., to see whether novel reasoning techniques actually lead to better recommendations. Moreover, CHIP can be used to study how trust in the system and treatment adherence are affected by factors such as the actual or perceived involvement of other stakeholders [52] or the availability of explanations of recommendations [53]. Likewise, the incorporation of, e.g., personas and response styles can be studied by replacing the default implementation of a response generator with a consistent persona or, in contrast, one that tailors the style of its outputs to the user based on feedback signals.

There are also empirical questions around the effectiveness of dynamically generated explanations that are adapted to a user's level of understanding and style of engagement [54]. Likewise, CHIP lends itself well to conducting user studies on how particular types of recommendations affect people's self-efficacy [55,56] and self-determination [57,58]. Such studies are particularly relevant as empirical input to discussions around the ethics of (automated) health coaching, for example with respect to how such systems might affect personal autonomy [59] or (self-)stigmatization [60]. These ethical dimensions will become increasingly important as the public interest in virtual coaching grows [61].

4.3. Application scope

The initial idea for CHIP was conceived in the context of lifestyle support for people dealing with chronic illness, specifically Type II diabetes [24,42]. However, CHIP has been designed to be reusable for different domains from the start of development. This design choice ensures its potential impact is maximized, as its modular architecture allows researchers to replace the default domain knowledge in the Knowledge Store with knowledge about smoking cessation, stress reduction, or other domains, making it a flexible platform for various other hybrid intelligence research questions.

5. Conclusions

CHIP offers a modular platform for hybrid intelligence research into behavior change. Through its modular architecture and focus on symbolic reasoning strategies, the software aims to facilitate human-AI co-reasoning for complex, personalized health decisions that lead to better user adherence and overall more effective and sustainable behavior change interventions.

Going forward, the software itself will undergo continued improvement. The primary focus will lie in growing the suite of available components to expand CHIP's capabilities. Planned upgrades include improved interoperability with external platforms, for example by consuming a user's data streams about a user's behavior, including data from the users' calendar, fitness trackers, nutritional apps and sleep monitoring devices. Future development will also prioritize support for different kinds of user interactions through existing smartphone apps such as Signal by utilizing their APIs.

CHIP is backed by a national consortium consisting of a translational research institute and several universities committed to advancing open science and responsible innovation. This backing ensures the software's ongoing development and use in scientific projects. By releasing the software to the wider public, the project explicitly aims to create and foster an open community of researchers, developers and practitioners beyond the core consortium. Such a community-driven model is expected to not only stimulate collaborative innovation and cross-domain knowledge exchange, but also enhance the long-term viability and maintenance of the project. Building the software as open source is further intended facilitate the reproducibility of research results and to accelerate the refinement of system components. The hope is for the platform to become a shared research infrastructure that advances the scientific study and implementation of hybrid intelligence behavior change technologies.

CRediT authorship contribution statement

Floris den Hengst: Writing – original draft, Supervision, Software, Project administration, Conceptualization. **Shaad Alaka:** Writing – review & editing, Software, Conceptualization. **Bart A. Kamphorst:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support provided by the members of the Hybrid Intelligence Center and the Netherlands Organisation for Applied Scientific Research (TNO). Special thanks are extended to Selene Báez Santamaria, Pei-Yu Chen, Maaike H.T. de Boer, Stergios Ntanavaras, Quirine Smit, Shihan Wang, and Johanna Wolff. This research was funded by the Hybrid Intelligence Center, a 10-year programme supported by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research (NWO), under grant number 024.004.022.

References

[1] Chaker L, Falla A, van der Lee SJ, Muka T, Imo D, Jaspers L, Colpani V, Mendis S, Chowdhury R, Bramer WM, et al. The global impact of non-communicable diseases on macro-economic productivity: a systematic review. Eur J Epidemiol 2015;30(5):357–95.
[2] Chen S, Kuhn M, Prettner K, Bloom DE. The macroeconomic burden of non-communicable diseases in the united states: estimates and projections. PLoS One 2018;13(11):e0206702.
[3] Stubbs B, Vancampfort D, Hallgren M, Firth J, Veronese N, Solmi M, Brand S, Cordes J, Malchow B, Gerber M, et al. Epa guidance on physical activity as a treatment for severe mental illness: a meta-review of the evidence and position statement from the european psychiatric association (epa), supported by the international organization of physical therapists in mental health (ioptmh). Eur Psychiatry 2018;54:124–44.
[4] Lichtenstein AH, Appel LJ, Vadiveloo M, Hu FB, Kris-Etherton PM, Rebholz CM, Sacks FM, Thorndike AN, Van Horn L, Wylie-Rosett J, et al. 2021 dietary guidance to improve cardiovascular health: a scientific statement from the American heart association. Circulation 2021;144(23):472–87.
[5] Brown A, McArdle P, Taplin J, Unwin D, Unwin J, Deakin T, Wheatley S, Murdoch C, Malhotra A, Mellor D. Dietary strategies for remission of type 2 diabetes: a narrative review. J Hum Nutr Diet 2022;35(1):165–78.
[6] Mensink M, Blaak EE, Corpeleijn E, Saris WH, De Bruin TW, Feskens EJ. Lifestyle intervention according to general recommendations improves glucose tolerance. Obes Res 2003;11(12):1588–96.
[7] Roumen C, Blaak EE, Corpeleijn E. Lifestyle intervention for prevention of diabetes: determinants of success for future implementation. Nutr Rev 2009;67(3):132–46. https://doi.org/10.1111/j.1753-4887.2009.00181.x
[8] Tuomilehto J, Schwarz P, Lindström J. Long-term benefits from lifestyle interventions for type 2 diabetes prevention: time to expand the efforts. Diabetes Care 2011;34(Suppl 2):S210.
[9] Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Medicine 2017;15(1):1–11.
[10] Galaviz KI, Narayan KMV, Lobelo F, Weber MB. Lifestyle and the prevention of type 2 diabetes: a status report. Am J Lifestyle Med 2018;12(1):4–20.
[11] Koenigsberg MR, Bartlett D, Cramer JS. Facilitating treatment adherence with lifestyle changes in diabetes. Am Fam Physician 2004;69(2):309–16.
[12] Mumu SJ, Saleh F, Ara F, Afnan F, Ali L. Non-adherence to life-style modification and its factors among type 2 diabetic patients. Indian J Public Health 2014;58(1). https://journals.lww.com/ijph/fulltext/2014/58010/non_adherence_to_life_style_modification_and_its.9.aspx
[13] Kavookjian J, Berger BA, Grimley DM, Villaume WA, Anderson HM, Barker KN. Patient decision making: strategies for diabetes diet adherence intervention. Res Soc Adm Pharm 2005;1(3):389–407. https://doi.org/10.1016/j.sapharm.2005.06.006. https://www.sciencedirect.com/science/article/pii/S1551741105000707
[14] Hayes E, McCahon C, Panahi MR, Hamre T, Pohlman K. Alliance not compliance: coaching strategies to improve type 2 diabetes outcomes. J Am Acad Nurse Pract 2008;20(3):155–62. https://doi.org/10.1111/j.1745-7599.2007.00297.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-7599.2007.00297.x

[15] Kamphorst BA. E-coaching systems: what they are, and what they aren't. Pers Ubiquitous Comput 2017;21(4):625–32.
[16] den Hengst F, Grua EM, el Hassouni A, Hoogendoorn M. Reinforcement learning for personalization: a systematic literature review. Data Sci 2020;3(2):107–47.
[17] Rapp A, Boldi A. Exploring the lived experience of behavior change technologies: towards an existential model of behavior change for HCI. ACM Trans Comput Hum Interact 2023;30(6):1–50.
[18] Grua EM, Hoogendoorn M. Exploring clustering techniques for effective reinforcement learning based personalization for health and wellbeing. In: 2018 IEEE symposium series on computational intelligence (SSCI). IEEE; 2018. p. 813–20.
[19] Albers N, Neerincx MA, Brinkman WP. Reinforcement learning-based persuasion by a conversational agent for behavior change. In: 33rd benelux conference on artificial intelligence and30th Belgian-Dutch conference on machine learning; 2021. p. 729–32.
[20] Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Hochberg I. Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. J Med Internet Res 2017;19(10):e338.
[21] Van Ommen B, Wopereis S, Van Empelen P, Van Keulen HM, Otten W, Kasteleyn M, Molema JJW, De Hoogh IM, Chavannes NH, Numans ME, et al. From diabetes care to diabetes cure—the integration of systems biology, ehealth, and behavioral change. Frontiers in endocrinology 2018;8:381.
[22] European Commission. Healthier together – EU non-communicable diseases initiative. Public Health 2022. https://doi.org/10.2875/195572
[23] Hogan A, Blomqvist E, Cochez M, d'Amato C, Melo de G, Gutierrez C, Kirrane S, Gayo JEL, Navigli R, Neumaier S, et al. Knowledge graphs. ACM Comput Surv 2021;54(4):1–37.
[24] Dudzik BJW, van der Waa JS, Chen P-Y, Dobbe R, de Troya ÍM, Bakker RM, de Boer MHT, Smit QTS, Dell'Anna D, Erdogan E, Yolum P, Wang, S, Santamaria SB Krause L Kamphorst BA. Hybrid intelligence supports application development for diabetes lifestyle management. J Artif Intell Res 2024;80:919–29.
[25] Miller A, Feng W, Batra D, Bordes A, Fisch A, Lu J, Parikh D, Weston J. Parlai: a dialog research software platform. In: Proceedings of the 2017 conference on empirical methods in natural language processing: system demonstrations; 2017. p. 79–84.
[26] Den Hengst F, Hoogendoorn M, Van Harmelen F, Bosman J. Reinforcement learning for personalized dialogue management. In: IEEE/WIC/ACM international conference on web intelligence; 2019. p. 59–67.
[27] Beinema T, op den Akker H, Hofs D, van Schooten B. The wool dialogue platform: enabling interdisciplinary user-friendly development of dialogue for conversational agents. Open Res Eur 2022;2:7.
[28] Castro O, Mair JL, Salamanca-Sanabria A, Alattas A, Keller R, Zheng S, Jabir A, Lin X, Frese BF, Lim CS, et al. Development of "lvl up 1.0": a smartphone-based, conversational agent-delivered holistic lifestyle intervention for the prevention of non-communicable diseases and common mental disorders. Frontiers in digital health 2023;5:1039171.
[29] Carroll M, Shah R, Ho MK, Griffiths T, Seshia S, Abbeel P, Dragan A. On the utility of learning about humans for human-ai coordination. Adv Neural Inf Process Syst 2019;32.
[30] Aydın H, Godin-Dubois K, Goncalves Braz L, den Hengst F, Baraka K, Çelikok MM, Sauter A, Wang S, Oliehoek FA. SHARPIE: a modular framework for reinforcement learning and human-ai interaction experiments. In: AAAI bridge program workshop on collaborative AI and modeling of humans. Philadelphia, Pennsylvania, USA; 2025. p. 1–7. https://doi.org/10.48550/arXiv.2501.19245
[31] Dragoni N, Giallorenzo S, Lafuente AL, Mazzara M, Montesi F, Mustafin R, Safina L. Microservices: yesterday, today, and tomorrow. Present and ulterior software engineering 2017:195–216.
[32] Stoenescu R, Contributors Q. Quasar framework. https://quasar.dev [accessed: 7 August 2025].
[33] Cyganiak R, Wood D, Lanthaler M. RDF 1.1 concepts and abstract syntax, W3c recommendation, W3C. Feb 2014. https://www.w3.org/TR/rdf11-concepts/.
[34] Ontotext. GraphDB, semantic graph database software. 2026. https://www.ontotext.com/products/graphdb/ version 11.2.1.
[35] Harris S, Seaborne A. Sparql 1.1 query language, W3c recommendation, W3C. Mar 2013. https://www.w3.org/TR/sparql11-query/.
[36] Horrocks I, Patel-Schneider PF, Boley H, Tabet S, Grosof B, Dean M. Swrl: a semantic web rule language combining OWL and ruleml, W3c member submission, W3C. May 2004. https://www.w3.org/Submission/SWRL/.
[37] Baez Santamaria S. Knowledge-centered conversational agents with a drive to learn. In: Cao YT, Papadimitriou I, Ovalle A, Zampieri M, Ferraro F, Swayamdipta S, editors. Proceedings of the 2024 conference of the north American chapter of the association for computational linguistics: human language technologies (volume 4: student research workshop). Mexico City, Mexico: Association for Computational Linguistics; 2024. p. 83–92. https://doi.org/10.18653/v1/2024.naacl-srw.10. https://aclanthology.org/2024.naacl-srw.10/
[38] Wolff J, de Boer V, Heylen D, van Riemsdijk MB. Defining an adaptable framework for behaviour support agents in default logic. In: CEUR workshop proceedings, vol. 3835. CEUR; 2024. p. 72–82.
[39] Team G, Anil R, Borgeaud S, Alayrac J-B, Yu J, Soricut R, Schalkwyk J, Dai AM, Hauth A, Millican K, et al. Gemini: a family of highly capable multimodal models. [arXiv preprint] arXiv:2312.11805. 2023.
[40] Han T, Adams LC, Papaioannou J-M, Grundmann P, Oberhauser T, Löser A, Truhn D, Bressem KK. Medalpaca–an open-source collection of medical conversational ai models and training data. [arXiv preprint] arXiv:2304.08247. 2023.
[41] Bélisle-Pipon J-C. Why we need to be careful with llms in medicine. Front Med 2024;11:1495582.
[42] Chen P-Y, Baez Santamaria S, De Boer MHT, Den Hengst F, Kamphorst BA, Smit Q, Wang S, Wolff J. Intelligent support systems for lifestyle change: integrating dialogue, information extraction, and reasoning. In: HHAI 2024: hybrid human AI systems for the social good. IOS Press; 2024. pp. 457–9.
[43] Koot H. Detecting patient deception and adherence in diabetes support using ai-generated conversation summaries, [Master's thesis], Delft University of Technology; 2025.
[44] van Westerlaak R. Enhancing diabetes care through ai-driven lie detection in a diabetes support system, [Master's thesis], Delft University of Technology; 2025.
[45] Madaras M. To deceive or self-deceive? [Master's thesis], Delft University of Technology; 2025.
[46] Alrabbaa C, Borgwardt S, Friese T, Hirsch A, Knieriemen N, Koopmann P, Kovtunova A, Krüger A, Popovič A, Siahaan I. Explaining reasoning results for owl ontologies with evee. In: Proceedings of the international conference on principles of knowledge representation and reasoning, vol. 21. 2024. p. 709–19.
[47] Ntanavaras S. Conversational triple extraction for diabetes healthcare management using synthetic data, [Master's thesis], Vrije Universiteit Amsterdam; 2024.
[48] Brons A, Wang S, Visser B, Kröse B, Bakkes S, Veltkamp R. Machine learning methods to personalize persuasive strategies in mhealth interventions that promote physical activity: scoping review and categorization overview. J Med Internet Res 2024;26:e47774.
[49] van Paridon J. Detecting patient information conflicts through conflict reasoning in knowledge graphs, [Master's thesis], Delft University of Technology; 2025.
[50] Benferhat S, Sossai C. Reasoning with multiple-source information in a possibilistic logic framework. Information Fusion 2006;7(1):80–96.
[51] Hohenecker P, Lukasiewicz T. Ontology reasoning with deep neural networks. J Artif Intell Res 2020;68:503–40.
[52] Kamphorst BA, Klein MCA, Van Wissen A. Human involvement in e-coaching: effects on effectiveness, perceived influence and trust. In: International workshop on human behavior understanding. Springer; 2014. p. 16–29.
[53] Papenmeier A, Kern D, Englebienne G, Seifert C. It's complicated: the relationship between user trust, model accuracy and explanations in AI. ACM Trans Comput Hum Interact 2022;29(4):1–33.
[54] Li X, Zheng H, Chen J, Zong Y, Yu L. User interaction interface design and innovation based on artificial intelligence technology. J Theory Pract Eng Sci 2024;4(3):1–8.
[55] Warner LM, Schwarzer R. Self-efficacy and health. In: Handbook of concepts in health, health behavior and environmental health. Springer; 2024. pp. 1–26.
[56] Tang MY, Smith DM, Mc Sharry J, Hann M, French DP. Behavior change techniques associated with changes in postintervention and maintained changes in self-efficacy for physical activity: a systematic review with meta-analysis. Ann Behav Med 2019;53(9):801–15.
[57] Alberts L, Lyngs U, Lukoff K. Designing for sustained motivation: a review of self-determination theory in behaviour change technologies. Interact Comput 2024:iwae040.
[58] Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. Health Psychol Rev 2019;13(1):110–30.
[59] Kamphorst BA, Kalis A. Why option generation matters for the design of autonomous e-coaching systems. AI & society 2015;30(1):77–88.
[60] Kamphorst BA, Anderson JH. E-coaching systems and social justice: ethical concerns about inequality, coercion, and stigmatization. AI and Ethics 2024:1–10.
[61] Korn F, Karger E, Ahlemann F. Mapping the research landscape of virtual coaches: what is to come? IEEE Access 2025;13:107093–107111.