From User to Human

USINET ALUMNI EVENT 2024

Jun Hu
Javed Khan (Eds.)

From User to Human

USINET Alumni Event 2024

Jun Hu
Javed Khan
(Eds.)

June, 2024

This event is organized by

- USI Alumni Stichting USINET
- EngD Designing Human-System Interaction
  Department of Industrial Design, Eindhoven University of Technology

and sponsored by

- Ambianti B.V.
- Department of Industrial Design, Eindhoven University of Technology

A catalogue record is available from the Eindhoven University of Technology library.

ISBN: 978-90-386-6016-5

Cover design: Sichen Guo

Contents

I

From User to Human

USINET Alumni Organization

It is almost a quarter of a century (24 years to be exact) since the first trainees graduated from TU/e's User System Interaction (USI as it was known at the time) programme. In 2004, after trainees from six generations had graduated, the need for an alumni network became apparent, and that is how the USINET alumni was born.

This network has grown to count over 300 members strong. That number is silent of the quality of our USINET network. We have members from all continents who are highly esteemed professionals in a wide-range of organizations, which include Fortune 500 companies, innovative startups, top higher-education institutes, government, just to name a few. Some of our entrepreneurial members have started themselves companies which have become very successful.

Completing the USI programme has equipped our members to improve the interaction between people and systems and make technology even more accessible and beneficial for people in all walks of life. Given the relative recent developments of artificial intelligence, our alumni is well-positioned to steer the proper use of such technologies for the future digital society.

The USINET alumni organization (stichting) is based in Eindhoven, the Netherlands and aims to:

- Support the exchange of information between alumni's and current trainees to keep a balance between academia and industry.
- Keep the members updated about professional developments in the field of Human-Computer Interaction (HCI) and User Experience (UX) former USI students.
- Provide an outlet for keeping in touch with other alumni members and interface with other UX and HCI professionals.
- Organize social and professional activities to enrich the discipline as a whole.

Javed Khan, president USINET

EngD Program

Designing Human-System Interaction

The USI (User System Interaction) Program

The User System Interaction (USI) program started in 1998, founded by the Stan Ackermans Institute at the Technical University Eindhoven (TU/e), the Netherlands. It was a two-year post-master program leading to the degree of Master of Technological Design (MTD), for students who wanted to become user system interaction designers in R&D organizations. The overall goal of the program is to provide students with the skills and capabilities for conceptualizing, designing, implementing, and evaluating new products, services, and applications that exploit possibilities of new technologies for the benefit of users in the domain of information and communication technology. The content of the curriculum is regulated by staff of the TU/e from different faculties working in the field of USI and Human-Computer Interaction (HCI), and is carried out by nationally and internationally well-known experts from academia and industry [1, 2]. The program accepts about 20 applicants per year. Each USI cohort is composed of a balanced group of students of multidisciplinary and multicultural background. The target composition is 50% engineering, 50% behavioral sciences, 50% Dutch citizenship and 50% nonDutch citizenship [3].

The last cohort of the USI program graduated in 2018, and the last individual USI graduated in 2020. During the later years of USI, there have been technological changes as well as changes in the group of companies that regularly provide relevant design projects. We decided that the program needed a substantial revision to keep aligned with technological developments and to respond to changes in and the needs of the industry.

HSI: EngD Designing Human-System Interaction

Recognizing the need of the industry for highly trained professionals and for transferring research results and specialized expertise through people, the department started the Engineering Doctorate (EngD) program in Designing Human-System Interaction in June 2022 [4]. We recruit Master's graduates with a degree in a discipline related to interaction design. The training program includes an individualized curricular component that covers design and professional skills to prepare them for work in industry, specialized training based on the research of the department, and an industrial project which takes in total about one year. The projects are funded by industry and from research projects. EngD candidates are in excellent positions to put the knowledge acquired from research into industrial practice in these projects. Currently, there are 13 trainees engaged in this new program, and three of them will be graduating in Fall 2024.

The new EngD program, Designing Human-System Interaction (HSI), aims to train professionals to develop competencies in designing and evaluating interactive, intelligent, and innovative systems, services, and products, ensuring positive user experiences that support their values and needs. It pays attention to the frontier of the complex systems enabled by artificial intelligence, data science, and other emerging technologies in high-tech systems, health applications, and smart mobility, as well as their impact on individuals, organizations, and society. The new program is based on three pillars (Figure I):

- Design thinking, systems thinking, and business innovation.
- Scientifically founded and methodologically sound user-centered and data-driven approaches.
- Understanding of artificial intelligence and machine learning.

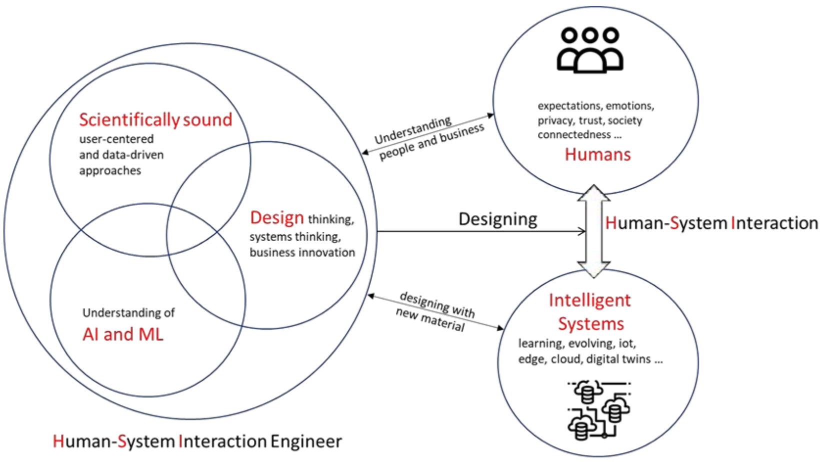


Figure 1: Human-System Interaction Engineer

Next to its new focus, the new HSI program strengthens itself in the following aspects in order to better its position in relation to industry, research and education.

- *Individualized program instead of cohort-based training.* Like PhD programs, the HSI EngDs can start at any time in the year, working on a project over two years, while taking the courses according to their personal development plans (Figure I). This allows flexibility in training and in collaboration.

- *Structural Involvement of Eindhoven Engine.* The program is strongly supported by Eindhoven Engine to reach out to the core interests of the AI and Data related industry and business innovations. The connection to Eindhoven Engine is helpful in project acquisition and provides the context for training the trainees with real-world challenges.

- *Valorizing results in research projects.* Most of the research projects at ID involve industrial partners, and EngD candidates are in excellent positions to put the knowledge acquired from research into the industrial practice in these projects. ID has been strategically promoting the EngD positions among the researchers when acquiring research funding.

- *Strengthening the link between education and research.* EngD candidates are encouraged to define their individual design cases by cooperating with Master students in courses and squad projects and further the student work for more solid research purposes and results.

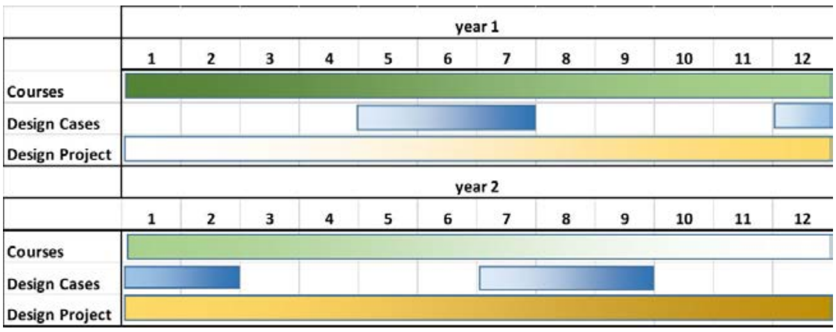


Figure 2: HSI curriculum structure

MTD ⇒ PDEng ⇒ EngD

(Below is edited based on [5])

In 2022, the Executive Boards of 4TU (3TU + Wageningen University & Research) decided to change the degree from Professional Doctorate in Engineering (PDEng) to Engineering Doctorate (EngD) for graduates from the Technological Design (MTD) or PDEng programs as of 1 September 2022.

The reason for the most recent name change is that the universities of applied sciences in the Netherlands are starting with Professional Doctorate (PD) trajectories. Therefore, we want to avoid confusion between those trajectories and our programs at an academic level.

EngD is already used as a degree in the United Kingdom for programs resembling Dutch programs. The national associations for both programs (AEngD and 4TU.SAI) have been affiliated members since 2012. It has been agreed that EngD and PDEng are on the same level.

Only the name of the degree changes; the contents of the programs are not affected. Alumni with an MTD or PDEng diploma are free to use the EngD degree as of 1 September 2022.

References

[1] M. Rauterberg, M. Janse, and P. Vinken, "User-system-interaction design program: an overview," in *Human-Computer Interaction – INTERACT 2003*, pp. 1069–1070, IOS Press, 06 2003.

[2] M. Janse, P. Markopoulos, and P. Vinken, "Eindhoven's user-system-interaction design program: an overview," *Interactions*, vol. 12, p. 33–34, sep 2005.

[3] P. Markopoulos, M. Janse, S. Pajo, P. Deisz, A. van Ruiten, V. Sawirjo, and A. Visser, "User system interaction program," in *Human-Computer Interaction – INTERACT 2007* (C. Baranauskas, P. Palanque, J. Abascal, and S. D. J. Barbosa, eds.), (Berlin, Heidelberg), pp. 688–689, Springer Berlin Heidelberg, 2007.

[4] Department of Industrial Design, TU/e, *Self-evaluation Report 2017-2023*. 2024.

[5] "From PDEng to EngD." `https://www.tue.nl/en/education/graduate-school/engd-at-tue/from-pdeng-to-engd`. Accessed: 2024-06-05.

II

USINET Alumni Event 2024

Event Program

June 14, 2024
Blauwe Zaal, Auditorium, Eindhoven University of Technology

- 11:30 - 12:00 Welcome - registration
- 12:00 - 13:00 Lunch
- 13:00 - 15:00 Talks
  - General introduction, Javed Khan, president USINET
  - Welcome by Paul Koenraad, Dean of Graduate School, TU/e
  - Welcome by Lin-Lin Chen, Dean of Industrial Design, TU/e
  - Update by Jun Hu, Scientific Director, EngD HSI
  - Panos Markopoulos, former director USI
  - Jacques Terken, former director USI
  - Alumni: Aga Szóstek,
  - Alumni: Serena Dorrestijn
  - Alumni: Suleman Shahid
  - Alumni: Beant Kaur Dhillon
- 15:00-16:00
  - Coffee
  - Posters by EngD candidates
- 16:00 - 17:00 Panel discussion hosted by Javed Khan and Amalia Mantziari Zafeiri. Panellists:
  - Alumni: Denys Tserkovnyi,
  - Alumni: Ingrid Halters
  - Alumni: Koen van Turnhout
  - Alumni: Tjeerd de Boer
  - Alumni: Olha Bondarenko
- 17:00 - 18:00 Borrel
- 18:00 - Dinner

Design process for life

Aga Szóstek

Over twenty years ago I was conditionally accepted to the USI program. That day was one of those days that influenced the rest of my life.

I would like to share with you the gist of how USI has shaped my professional career and personal adventures, and how it helped me have a truly fulfilling life.

Aga Szóstek

Aga Szóstek, PhD (and an USI alumna:) is a strategic designer and the author of “The Umami Strategy" and "Leadership by Design". She has worked at the forefront of combining design, technology, and business for the past two decades. She is mentoring leaders and executives in their strategic approach to change. Next to consulting, Aga designs tools aiming to deliver unique experiences and undergo personal and professional transformations and co-hosts a podcast about leadership in volatile times: Catching The Next Wave. She is also a founding partner of the World Experience Organization.

Ensuring Safe Digital Journeys with

Value Driven Design

Serena Dorrestijn

In this talk, I will explore how service design principles have been instrumental in shaping the services and portfolio within the domain of cyber security. The complexity of this domain often renders 'traditional' methodologies insufficient. To address this, we investigated and combined methodologies from service design, aerospace engineering, product management, user experience, and customer experience. This multidisciplinary approach has led to the development of what we now call "Value Driven Design." This method emphasises designing for value across the entire ecosystem rather than focusing solely on the end user.

You will hear me advocate that we should not only focus on the end user with our design methodologies, but on the entire ecosystem that the solution or service resides in. This shift also requires an update to the tools and techniques that we use to analyse and design our experiences and services.

Serena Dorrestijn

Service Portfolio Manager
Northwave Cyber Security

Serena holds an MSc in Psychology and a PDEng in User System Interaction. Her areas of expertise are service design, portfolio management and (b2b) customer experience. She leverages the knowledge gained through the USI program on a daily basis to innovate at Northwave. In her role, Serena is responsible for overseeing and developing Northwave's service portfolio and product development to ensure that clients are protected against constantly changing emerging threats.

Before joining Northwave, Serena served as the Manager of DDEX (Data Driven Experiences) at Bright Cape, where she started experimenting with new data driven design methodologies.

The ‘Us’ in Users
- Practicing Equity-Centred Design in the Global South

Suleman Shahid

In this talk, Suleman Shahid will share a few highlights of his USI and post-USI journey and how the experiences and learnings during the USI Program shaped his career. He will then present his university-based research and applied work on designing immersive and interactive technologies for people on the margins. He will discuss his approach of combining elements of playfulness, serious storytelling, gamification and social interaction for designing digital interventions and systems for diverse user groups in Pakistan.

He will discuss why the concept of “Us” in “Users" is critical in the context of the Global South and how the cultural background, religion, literacy level, socioeconomic status and previous exposure to technology shape the way different users, mainly at the margins, interact with and accept or reject new technologies. He will also present the pros and cons of using the “quick and dirty approach” for technology development and discuss how people on margins associate short- and long-term expectations with half-cooked interventions and digital solutions.

Suleman Shahid

Suleman Shahid is an Associate Professor of Computer Science at the Lahore University of Management Sciences (LUMS) and the founding director of the LUMSx - Center for Online Learning and Professional Development at LUMS. At LUMS, he directs the HCI+Design Lab (CHISEL), manages the university’s Usability Lab and is the Faculty Lead for the Virtual Reality Lab at LUMS. He is also a faculty member at the LUMS Center for Entrepreneurship (LCE) and LUMS’ Executive Education Center (REDC). Suleman is also a Senior Design and Strategy Consultant. He works with private and public sectors on design-driven innovation, people-centred change management, and digital transformation projects. He also founded the Design for Pakistan organisation, the annual UX Pakistan design conference, the Lahore Design Festival, and the National UX Camps program.

Suleman's primary area of interest is designing learning and healthcare technologies for/with the margins in the global south with an emphasis on using participatory and inclusive design to drive innovation. His research focuses on the intersection of design, technology, and inclusion with applications to (1) assistive technologies (mobile apps and VR/AR systems) to enhance the quality of life of persons with disabilities (e.g. autism, dyslexia, dementia), and persons with mental health conditions (e.g. anxiety, depression), and (2) learning and playful technologies for children.

How do we build a heart-aligned life and career?

Beant Kaur Dhillon

It is human to desire a fulfilled and meaningful life and career. I have been on the journey to design a heart-aligned life and career for over a decade while navigating health conditions. Being an embodiment coach is my 3rd career in the last 14ish years. And with every transition, I meet fear, doubt, more fear, joy, possibility, spaciousness, softness, and vulnerability. And I keep (re)learning that our relationship with our bodies and mind is a big part of living and working with pleasure and fulfillment.

Join me in this talk, where I share what I have learned about designing a purposeful and joyful life and career.

Beant Kaur Dhillon

Beant is an Embodiment life and career coach & Lead user researcher with 14+ years of international experience, and author & artist of Amazon NL Bestseller - A Delicious Summer.

As an embodiment coach, she helps people move from their heads to their bodies, reveal and dissolve blocks through the mind-body connection, and move forward with clarity and confidence.

Her previous work has been at Philips, adidas, Google, Schiphol Airport, Ultimaker, and several start-ups. She regularly speaks at events like Rosenfeld Media Conference, UXCopenhagen, ResearchOps Berlin, etc. on topics like embodiment, creative thinking, listening, and learning from failure.

You can find more of her writing at www.beantdhillon.com

Panel Discussion

Hosts: Javed Khan and Amalia Mantziari Zafeiri
Panelists: Denys Tserkovnyi, Ingrid Halters, Koen van Turnhout, Tjeerd de Boer, and Olha Bondarenko

Denys Tserkovnyi

UX Architect/UX Lead at ASML

Denys has worked in diverse industries like Automotive navigation, Web, Education, and Semiconductors. He holds an educational background in Multimedia Technologies, Interaction Design, and User-System Interaction. Denys began his career as a freelancer, developing his web design, development, and graphic design skills through various projects and collaborations. At Carte Blanche Ukraine, he furthered his expertise in interaction design, 3D modelling, and web development. After relocating to the Netherlands for the USI program, he settled in sunny Eindhoven. At Studyportals, Denys worked as a UX Designer and Scrum master, pioneering Agile and UX practices while redesigning platforms to help millions of students find their dream education. As the UX Design Team Lead, he expanded and mentored a diverse UXD team, driving UX and strategy with continuous research and occasional redesign initiatives. At ASML, he focuses on information architecture, user research, and experience design for the semiconductor industry.

Ingrid Halters

Product UX Design Director at IKEA - Growth & Marketing (Global)

Ingrid is currently leading Experience Design teams within Growth & Marketing focusing on Personalisation, Customer Engagement and Loyalty across IKEA. Prior to this role she was leading and growing Experience Design teams focused on the upper funnel of IKEA.com, as well as the IKEA App team.

At Elsevier, Ingrid worked on building the Elsevier digital identity from the ground up, and it is used across the company. She also lead multiple user experience design teams supporting the re-platforming and redesign of ScienceDirect and Scopus - two Elsevier flagship products. This included growing new teams in London and Amsterdam. The team established a more prominent and attractive way of presenting scholarly content and data across platforms.

Her successes at TomTom include building up a multi-disciplinary User Experience team from scratch. This unit was responsible for the delivery of user experience design of all navigation products and services for different embedded platforms as well as web-applications.

Specialties: User Experience Design, Experience Design Management, Design Strategy, and Design Organisation team building.

Koen van Turnhout

Lector
Utrecht University of Applied Sciences

Dr. Koen van Turnhout is Professor of Applied Science at Utrecht University of Applied Science. His research group explores the design of interactions between digital media and humans, with a particular focus on interactions with AI based systems and digital media in sensitive settings. His group does this work in close cooperation with industry (in particular UX professionals) and other societal players. Koen is also expert in the field of design research methodology and author of a Dutch handbook on the topic: "Handboek Wetenschappelijk Ontwerpgericht Onderzoek", as well as editor of the Applied Design Research book series.

Tjeerd de Boer

Managing Director - User Intelligence

Tjeerd de Boer is managing director of User Intelligence. He started his career at Razorfish and started his own company in 2023. This company, User Intelligence, is specialised in UX research of interactive products and services, with special attention to global research and inclusiveness research. User Intelligence is part of the UXalliance, a global network of UX companies.

Olha Bondarenko

Strategic Advisor on Citizen Engagement
@ Eindhoven

I am a former journalist who got fascinated with innovation, new technologies and their impact on people lives. Throughout my career I worked in the media, academia, private & public sectors, the last few years focusing on strategic advise at the city of Eindhoven - the city that took my heart but gave me a job. At the moment I am working on the subject of Citizen & Government Engagement at the Strategy department of the municipality. In my free time I enjoy living in the vibrant De Bergen neighborhood and getting engaged in the city life from a citizen perspective.

III

Projects and Papers

Transfiguring Soil Health Monitoring: AI-Integrated Digital Twins with Edge Computing in Regenerative Agriculture

YuvaSri Borra[0000−0003−3641−1541]

Industrial Design, Eindhoven University of Technology, The Netherlands
y.s.borra@tue.nl

Abstract. Confronted with escalating challenges in global food security, the imperatives of amplifying agricultural yield while preserving crop integrity have galvanized innovations in regenerative agriculture. This exposition introduces an avant-garde Digital Twin framework poised to revolutionize soil health monitoring in the agronomic domain. With its exceptional ability to manage the salient real-time processing contingencies of quenching network bottlenecks through a symbiotic integration of Edge computing, this framework paves the pathway for sustainable farming approaches in the future. Such an ingenious trajectory denotes a noteworthy augmentation in resource stewardship and agricultural harvest, empowering cultivators with strategic insight. The presented architecture will contribute to a notable incursion into agricultural technologization and a salient discourse.

Keywords: ReGen Agriculture · Digital Twin· Edge Computing · Real-Time Data Monitoring · Data Visualization · Soil Health · AI/ML · Data Integration · Resource Optimization.

1 Introduction

1.1 Background

Regenerative agriculture has been acknowledged as an essential panacea to the global food security challenge, emphasizing its vital role in improving crop productivity and encouraging sustainable agricultural methods [10]. This paper examines how regenerative agriculture can effectively manage the volatile balance between higher yields and ecological responsibility and explores the persistent difficulties encountered by modern agronomy: maintaining the authenticity of agricultural quality while meeting the increasing needs of a growing global population. This use case study explores the complex interaction between these two aims, highlighting the importance of creative solutions in agriculture. Integrating Artificial Intelligence into Soil Digital Twin technology is an innovative solution at the intersection of technical advancement and agricultural research. This work is presented as a noteworthy scholarly contribution to the discussion on combining AI and Digital Twins, explicitly focusing on their use in enhancing soil health analysis. This use-case study aims to outline the precise research goals, situating the study within the framework of advancements in smart farming. The following conversation is designed to create a smooth and coherent narrative that explores these key themes, marking the beginning of a new era in regenerative agriculture.

1.1.1 Regenerative Agriculture: A Key to Global Food Security

Amidst escalating food demands and shrinking cultivable land, regenerative agriculture emerges as a quintessential instrument in the quest for global food security. Leveraging the confluence of advanced agronomic technologies, that propounds a paradigm shift towards resource optimization, yield augmentation, and minimized ecological footprints. This lexicon of modern farming amplifies production and is instrumental in navigating the vicissitudes of climatic aberrations.

1.1.2 The Synergy of AI and Digital Twin Paradigms in Agronomy

At the intersection of agritech innovation lies the synergistic fusion of Artificial Intelligence and Digital Twins - a tandem that heralds a new era of farming intelligence systems. Digital Twins, as intricate cyberspace fabrications, mirror agrarian phenomena, creating a platform for nuanced monitoring and prognostications of soil vitality and crop maturation [2]. Concurrently, by distilling voluminous and multifaceted agronomic data streams into actionable wisdom, AI algorithms refine the integrity of predictive endeavors and operative precision, forming the theoretical foundation of regenerative agriculture.

1.1.3 Research Focus: Amalgamation of AI-forged Digital Twins with Edge Computing for Soil Health Surveillance

The overarching intent of this scholarly treatise is to dissect and elucidate the amalgamation of AI-forged Digital Twins with the nuanced Edge computing infrastructures for enhancing soil health surveillance [1]. The discourse examines the efficacy of this concatenated framework in orchestrating real-time data stewardship and executive decision-making. It further endeavors to deconstruct complexities associated mainly with optimizing data conveyance networks [4]. This inquiry seeks to furnish a granular blueprint potent in fostering resilient and fecund agrarian regimes within the agricultural Commodus.

1.2 Motivation of Research

1.2.1 Addressing Real-Time Processing Challenges in Regenerative Agriculture

Regenerative agriculture's reliance on advanced technologies predicates a slew of real-time processing challenges, which become the focal point of this study. It dissects the implications of intensive energy requisites, the augmentation of greenhouse gas emissions, and the burgeoning strain on data transmission networks [11] [6]. Exploring these issues underscores the urgency of refining and innovating computational processes to bolster sustainable agricultural practices and enhance system scalability and agility.

1.2.2 Refining Decision-Making Capabilities in Agricultural Management

As the agricultural industry wades through the intricacies of data-intensive practices an unyielding need emerges for the evolution of decision-making tools.

This section is devoted to probing the necessity for heightened analytical competence that equips stakeholders with the proficiency to parse complex datasets expediently, thereby influencing pivotal operational outcomes [6]. It brings to light the transformative potential of Artificial Intelligence and Digital Twins in surmounting these challenges, advocating for their amalgamation to cultivate a more precise and forward-thinking agricultural epoch [9].

1.3 Objectives

The development of advanced, solar-powered data collection systems that integrate seamlessly with sensor technology and Edge-computing is fundamental to the evolution and success of the digitalization of Regenerative agriculture. Despite numerous studies examining the utility of solar-powered devices in agricultural data collection, a significant void exists in operational technology capable of enduring extreme weather conditions and conveying data wirelessly with high efficiency, mainly through networks like LoRaWAN [5]. The present study seeks to bridge this knowledge gap by conceiving, designing, and constructing a prototype that is not only robust in the face of water and wind exposure but also capable of capturing real-time data from a suite of sensors and wirelessly channeling it to the cloud using LoRaWAN technology [12]. This research is driven by the rationale that a precise, sturdy, and reliable compilation of soil health data is indispensable to optimizing regenerative agriculture practices.

This study aims to design and prototype a solar-powered, water—and wind-resistant, and robust AI-IoT system that can operate efficiently in challenging resource-constrained environments, ensuring seamless data collection, transmission, and analysis to support sustainable agricultural practices. It also facilitates real-time soil health monitoring and decision-making in regenerative agriculture.

2 Literature Review

2.1 Overview of Related work

The advent of regenerative and innovative farming methods has intensified the exploration of sustainable, data-driven agricultural practices. In light of this progression, technologies like the Internet of Things, wireless sensor networks, and edge computing have been focal points of recent scholarly discourse [7]. They have been recognized for their ability to yield high-precision data critical for informed agricultural decision-making [8] [9]. However, application-focused research has highlighted a substantial gap in developing prototypes that gather data efficiently and endure the vicissitudes of agricultural environments, particularly those powered by renewable energy sources like solar power. Studying the current status of nanotechnology in agriculture reveals an exciting intersection of innovative materials and sensor design, with future research pinpointed towards the sustainable development of such technologies [5]. Concurrently, the emergence of soil sensors and plant wearables presents a novel opportunity to navigate the intricacies of Regenerative agriculture. However, challenges related to their deployment and effective data communication persist. The transformation of farming equipment, guided by the principles of "going green" and satellite navigation, provides an instructive parallel for IT advancements and demonstrates a convergence of ecological and technological perspectives [8]. Literature underscores the potential benefits of cross-pollinating ideas between farming and IT, advocating for adapting lessons from each field to bolster the other. Recent studies have recognized the need for sophisticated, automated agricultural frameworks to manage incoming data and corresponding processes(Lytos et al., 2020). These frameworks capitalize on recent ICT advances, aiming to enhance agricultural intelligence through extensive data utilization, alluding to the potential of improving agricultural productivity.

2.2 Research Gaps in Design

Despite strides in technology and conceptual frameworks, deploying a solar-powered, water- and wind-resistant device capable of transmitting data via LoRaWAN remains a relatively uncharted domain [3] [10]. This gap underscores the need for comprehensive research into designing and developing robust prototypes tailored to the harsh realities of the agricultural environment. The proposed study endeavors to build upon this foundation by delivering a prototype that meets the technological demands and withstands environmental challenges, thereby contributing to a significant advancement in smart and precision agriculture for regenerative farming practices.

2.3 Research Question

Pursuing this goal raises the research question: How can we engineer a solar-powered prototype capable of assembling real-time data from multiple sensors fused to Raspberry Pi Pico W and transmitting this information wirelessly to the cloud through LoRaWAN, all the while maintaining water and wind resistance?

3 Methodology

3.1 System Design

This elucidates the prototype's conceptual framework, addressing the architectural blueprint and its accordance with agrarian imperatives. It includes a critical dissection of each component's role within the system and reflects upon design considerations crucial for durability and environmental resilience.

3.2 Prototype Components and their Functions

The prototype is an AI-IoT system for real-time soil health monitoring for regenerative agriculture. The system is built around the Raspberry Pi Pico W microcontroller, selected for its low power consumption, compact size, and built-in wireless connectivity. The prototype integrates various sensors to collect detailed soil and environmental data, which are then processed and transmitted using LoRaWAN technology.

1. Raspberry Pi Pico W: Central data collection, processing, and transmission unit.
2. Capacitive Soil Moisture Sensor (V1.0): Detects soil moisture content, vital for irrigation management.
3. DS18B20 Soil Temperature Sensor: Measures soil temperature, critical for understanding soil conditions.
4. Soil NPK Sensor: Measures concentrations of nitrogen (N), phosphorus (P), and potassium (K), key indicators of soil fertility.
5. Analog Electrical Conductivity Meter V2: Assesses soil salinity by measuring electrical conductivity.
6. Analog pH Sensor Kit V2: Monitors soil pH levels, influencing nutrient availability and microbial activity.
7. LoRaWAN RFM95W Transceiver Module: Enables long-range, low-power data transmission to a central server.
8. Satellite Weather Forecast API: Provides real-time weather data to complement sensor data.
9. BME280 Sensor: This sensor measures environmental conditions above the ground, including temperature, humidity, and atmospheric pressure.

3.3 Design Considerations

- Weatherproofing: The prototype is housed in a durable enclosure to protect the electronic components from harsh environmental conditions. The enclosure includes rubber gaskets and seals to prevent water ingress and ensure durability.
- Robustness: The system is designed to be sturdy and reliable, capable of withstanding the rigors of deployment in agricultural fields. Components are securely mounted to minimize the risk of damage from physical shocks or vibrations.
- Power Management: The system is powered by a solar power bank, ensuring continuous operation without reliance on external power sources. This setup includes a solar panel and a rechargeable battery, providing a sustainable power solution.

3.4 Breadboard Schematic Diagram of Prototype design

see Fig. 1.

3.5 Implementation

An overview of how the whole system will be prototyped, from sourcing components to assembly and eventual deployment in the field, including quality control checkpoints to ensure system reliability.

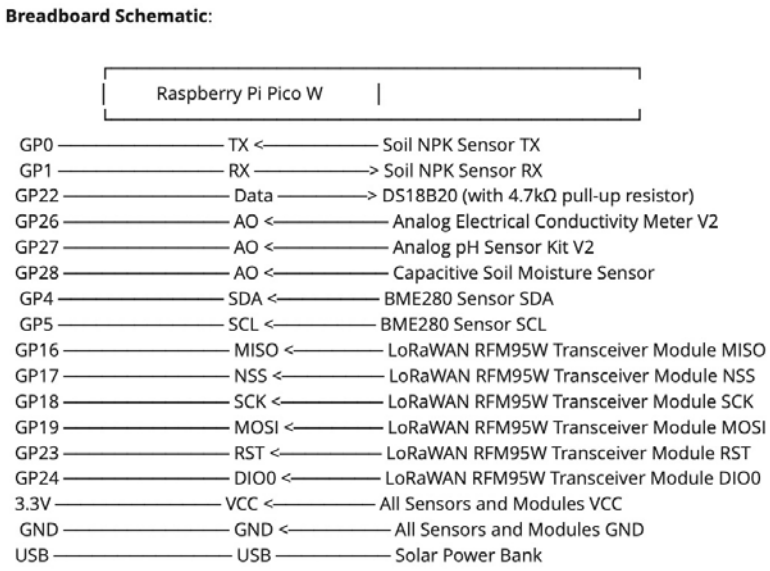


Fig. 1. How the sensors are fused around Raspberry Pi Pico W

- A robust, weatherproof enclosure was selected to provide ample space for all components and allow sensor probes to extend outside.
- The Raspberry Pi Pico W was securely mounted inside the enclosure using brackets and adhesive pads for stability.
- The DS18B20 Soil Temperature Sensor's VCC, GND, and data pins were connected to the respective power, ground, and GPIO pins on the Raspberry Pi Pico W, with a 4.7k ohm pull-up resistor between the VCC and data line to maintain a reliable signal and prevent floating. The Soil NPK Sensor's TX and RX pins were connected to the UART pins (GP0 and GP1), enabling serial communication for accurate nutrient measurement data transmission.
- The Analog Electrical Conductivity Meter V2's analog output was connected to an ADC pin (GP26) for precise analog-to-digital conversion. In contrast, the Analog pH Sensor Kit V2's analog output was connected to another ADC pin (GP27) for accurate pH level readings.
- The LoRaWAN RFM95W Transceiver Module's SPI interface (MOSI MISO SCK NSS) was linked to the corresponding SPI pins (GP19 GP16 GP18 GP17) on the Raspberry Pi Pico W, ensuring efficient data communication for long-range transmission.
- The BME280 Sensor's I2C interface (SDA, SCL) was connected to the I2C pins (GP4 and GP5) on the Raspberry Pi Pico W to collect environmental data, including temperature, humidity, and atmospheric pressure.

- The Capacitive Soil Moisture Sensor's analog output was connected to an ADC pin (GP28) for accurate soil moisture readings.
- The Raspberry Pi Pico W was then connected to the solar power bank using a USB cable to ensure a continuous power supply.

Finally, all components were meticulously arranged inside the enclosure, securing connections to minimize cable clutter, and the enclosure was sealed with rubber gaskets to protect against water and dust ingress, ensuring the system's robustness and reliability in the field.

4 Conclusions and Future steps

To conclude, this paper aims to create a solar-powered prototype using a Raspberry Pi Pico W equipped with wireless connectivity through LoRaWAN. The design will ensure robustness, weatherproofing, and visual appeal, incorporating components such as a solar power bank, Raspberry Pi Pico W, LoRaWAN module, and a waterproof casing.

As the project is ongoing, the following steps will be the prototype engineered to be water- and wind-resistant, sturdy, and rigid using durable casing materials, proper sealing of electronic components, and robust connectors. Real-time sensor readings and weather forecast data will be collected by interfacing multiple soil sensors with the Raspberry Pi Pico W, writing code to gather data, and integrating a weather API for satellite data. The prototypes will be deployed in various controlled locations to test data collection and transmission, ensuring reliability and validating system functionality.

Edge computing will be implemented on the Raspberry Pi Pico W to preprocess and summarize data locally before transferring it to the cloud, such as TU/e DATA FOUNDRY. High-resolution images or videos will be gathered using drones for spatial interpolation, integrating visual field data with sensor data. Data from sampled areas will be compared with visual data. Kriging will be applied to estimate values for unsampled regions, combining both datasets to create a detailed data map for the entire field. AI/ML techniques will be used to identify patterns in the data, optimizing prototype deployment zones. Data accuracy will be enhanced by fusing data from sensors, drones, and satellites using AI algorithms to create a comprehensive dataset. The combined data will be cleaned and standardized, preparing it for machine learning model training. The ML model will be trained to predict future soil health using historical and current data, including weather forecasts.

Weather forecast API data will be integrated with real-time sensor data, identifying commonalities and training the ML model. The accuracy of ML model predictions will be validated by comparing API forecast data with sensor readings, fixing errors, and improving predictions. Errors will be fed back into the model to regenerate real-time predictions, refining accuracy. All data will be integrated into a real-time AI/ML-driven system, combining datasets into a single fused dataset for real-time monitoring and prediction. An interactive, user-friendly web application dashboard will be developed to visualize real-time soil health data, providing an accessible tool for farmers. The dashboard will be designed to be accessible on tablets and mobile devices and tailored to meet the needs of farmers.

About YuvaSri Borra I am an EngD Trainee in the Human System Interaction Group of the Industrial Design Department at Eindhoven University of Technology Netherlands. My Research project focuses on Designing, developing, and integrating AI and ML models for soil analysis in ecological-based farming. I am working on my final project in the direction of an integrated IoT Prototype for Real-time Data collection of respective soils for Soil health Monitoring on an Interactive Data Visualization Dashboard, i.e., titled "An AI-driven IoT system for Real-time Soil Health Monitoring in Ecology farming."

YuvaSri holds an International Master's in Biomass and Waste to Energy and Materials from IMT Mines Albi, France, specializing in thermochemical conversion technologies and biochemical conversion processes. Before joining Eindhoven University of Technology, I worked as a Research Engineer at CIRAD RD. I led and successfully achieved a comprehensive LCA of wine-growing systems in Bordeaux, France. I also worked as a Research Analyst at L'Institut Agro Rennes-Angers, where I participated in developing future scenarios for food and water systems, employing spatially explicit methods to study land use patterns and implications.

References

1. Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., Zomaya, A.Y.: Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence. IEEE Internet of Things Journal **7**(8), 7457–7469 (Aug 2020). https://doi.org/10.1109/JIOT.2020.2984887, https://ieeexplore.ieee.org/document/9052677/
2. Diaz, F.J., Ahmad, A., Parra, L., Sendra, S., Lloret, J.: Low-Cost Optical Sensors for Soil Composition Monitoring. Sensors **24**(4), 1140 (Feb 2024). https://doi.org/10.3390/s24041140, https://www.mdpi.com/1424-8220/24/4/1140
3. Hareesh, S.: The latest applications of remote sensing technologies for soil management in precision agriculture practices. In: Remote Sensing in Precision Agriculture, pp. 105–135. Elsevier (2024). https://doi.org/10.1016/B978-0-323-91068-2.00008-4, https://linkinghub.elsevier.com/retrieve/pii/B9780323910682000084

4. Javeed, D., Gao, T., Saeed, M.S., Kumar, P.: An Intrusion Detection System for Edge-Envisioned Smart Agriculture in Extreme Environment. IEEE Internet of Things Journal pp. 1–1 (2023). https://doi.org/10.1109/JIOT.2023.3288544, https://ieeexplore.ieee.org/document/10159378/
5. Muthumalai, K., Gokila, N., Haldorai, Y., Rajendra Kumar, R.T.: Advanced Wearable Sensing Technologies for Sustainable Precision Agriculture – a Review on Chemical Sensors. Advanced Sensor Research **3**(3), 2300107 (Mar 2024). https://doi.org/10.1002/adsr.202300107, https://onlinelibrary.wiley.com/doi/10.1002/adsr.202300107
6. Ngu, N.H., Thanh, N.N., Duc, T.T., Non, D.Q., Thuy An, N.T., Chotpantarat, S.: Active learning-based random forest algorithm used for soil texture classification mapping in Central Vietnam. CATENA **234**, 107629 (Jan 2024). https://doi.org/10.1016/j.catena.2023.107629, https://linkinghub.elsevier.com/retrieve/pii/S0341816223007208
7. Nguyen, H.H., Shin, D.Y., Jung, W.S., Kim, T.Y., Lee, D.H.: An Integrated IoT Sensor-Camera System toward Leveraging Edge Computing for Smart Greenhouse Mushroom Cultivation. Agriculture **14**(3), 489 (Mar 2024). https://doi.org/10.3390/agriculture14030489, https://www.mdpi.com/2077-0472/14/3/489
8. Raja Gopal, S., Prabhakar, V.S.V.: Intelligent edge based smart farming with LoRa and IoT. International Journal of System Assurance Engineering and Management **15**(1), 21–27 (Jan 2024). https://doi.org/10.1007/s13198-021-01576-z, https://link.springer.com/10.1007/s13198-021-01576-z
9. Selvanarayanan, R., Rajendran, S., Algburi, S., Ibrahim Khalaf, O., Hamam, H.: Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system. Scientific Reports **14**(1), 6269 (Mar 2024). https://doi.org/10.1038/s41598-024-56954-x, https://www.nature.com/articles/s41598-024-56954-x
10. Sophocleous, M., Karkotis, A., Papasavva, A., Goldberger, M., Vassiliou, L., Ros-Lis, J.V., Shacham-Diamand, Y., Georgiou, J.: A Stand-Alone, In Situ, Soil Quality Sensing System for Precision Agriculture. IEEE Transactions on AgriFood Electronics **2**(1), 43–50 (Mar 2024). https://doi.org/10.1109/TAFE.2024.3351953, https://ieeexplore.ieee.org/document/10414425/
11. Vibhute, A.D., Kale, K.V., Gaikwad, S.V.: Machine learning-enabled soil classification for precision agriculture: a study on spectral analysis and soil property determination. Applied Geomatics **16**(1), 181–190 (Mar 2024). https://doi.org/10.1007/s12518-023-00546-3, https://link.springer.com/10.1007/s12518-023-00546-3
12. Xu, Y., Smirnov, M., Kohler, M.C., Dong, Z., Amineh, R.K., Li, F., Rojas-Cessa, R.: A Survey of Wireless Soil Sensing Technologies. IEEE Access **12**, 12010–12038 (2024). https://doi.org/10.1109/ACCESS.2024.3352006, https://ieeexplore.ieee.org/document/10387447/

Human-AI Collaboration for Advanced Cancer Care

Victoria Bruno[0009−0003−8227−0855]

Industrial Design Department, Eindhoven University of Technology
v.bruno@tue.nl

The EngD project entails the design and prototyping of a professional software package for clinicians to analyse medical images and accurately detect pancreatic tumours, leveraging artificial intelligence (AI) to support clinical decision-making processes. With some of the lowest survival rates among all cancers, pancreatic cancer often remains undetected until it has reached an advanced stage also due to its intricate anatomical location. The solution in development aims to advance research in early detection and therapeutic strategies.

It is part of a collaborative effort between researchers from the Industrial Design and Electrical Engineering departments at Eindhoven University of Technology, the healthcare professionals and researchers from the Catharina Hospital, and Philips. As a multidisciplinary team they work together in the Eindhoven MedTech Innovation Center (e/MTIC).

The software being developed supports visualization of CT scans and 3D models. It also facilitates manipulation and annotation of observations. It will enable radiologists and surgeons to review a set of AI-generated findings produced by algorithms under development by the Electrical Engineering PhD candidates. The findings consist of segmentations of the anatomical structures, namely pancreas, surrounding veins and arteries, and tumour, as well as the involvement between the tumour and the vessels. The Industrial Design researchers are focusing on aspects such as improving user experience, visualization methods of the AI findings, clinician's trust in AI, and design guidelines for AI systems in healthcare.

At the beginning of the project, the prototype existed as a user interface design concept and mockup with limited functionality that had been developed during previous student projects at Philips. Images and segmentation data were hardcoded and could not be dynamically imported, and many features in the user interface existed solely as visual representation.

The initial phase of the EngD project consisted of further developing the prototype, implementing functionalities according to the exisiting design concepts, and further improving them working with the designers in the team. The Industrial Design department was leading this effort supported by Bureau Moeilijke Dingen (BMD), a software and design agency based in Eindhoven, for a period of three months.

One of the focuses in the frontend development was on modularization of the prototype. Given the software's intended use in hospital and research settings, prioritizing modular design was crucial to facilitate updates and design adjustments anticipatedmfrom upcoming studies conducted by the design researchers, without disrupting other components within the user interface. This involved extracting state management from the user interface components. To support this modular software design, state management was handled by defining workflows and reimplementing application state. Moreover, given the frequent handovers inherent to EngD, PhD, and internship programs, a dedicated component library in Storybook was developed for centralized documentation. This library enables the creation and visualization of individual components in isolation, and its documentation, streamlining the transition and collaboration processes.

The codebase underwent refactoring where necessary. The most notable change involved transitioning from a mix of CSS and styled components to Tailwind CSS. This shift was essential because applying style changes became cumbersome, given the conflicting properties stemming from the diverse styling methods previously employed.The EngD project involved acting as a product manager in this phase, as several team members were contributing to the development. This included streamlining communication, data and algorithm exchange between backend developer and the Electrical Engineering PhDs, working together with the designers and research team in general.

The next phase of the project will require maintaining the software, incorporating design changes derived from ongoing clinical study executions, and integrating AI algorithms developed by Electrical Engineering PhD candidates through the Algorithm Execution Platform developed for the backend.

About Victoria Bruno Victoria Bruno is pursuing her Engineering Doctorate in the Human System Interaction program at the Department of Industrial Design, Eindhoven University of Technology (TU/e) in the Netherlands.

She holds a Bachelor of Science degree in Mechanical Engineering and a Master of Science in Biomedical Engineering with a specialization in Biomaterials and Tissue Biomechanics. Following her academic experience, she worked for over three years in the medical device sector in Product Management, Clinical Evaluation and Medical Writing roles at one of the leading companies of the sector.

Her curiosity and passion for technology led her to pursue a career switch through the EngD program towards software engineering, transforming her web development hobby into her profession, and focusing on health technology. For her doctorate, she is currently collaborating with designers, AI engineers and clinicians of the Catharina hospital in the Eindhoven MedTech Innovation Center (e/MTIC) as a software engineer.

‘Met Mij’: The Theoretical Model Behind AI-Enhanced Identification and Connecting Tool

Jéssica Goss[1][0000−0003−3809−7293] and Walter Baets[2][0000−0003−1860−2483]

[1] Department of Industrial Design, Eindhoven University of Technology
[2] Eindhoven Engine
j.messias.goss.dos.santos@tue.nl

Abstract. In an era of rapid technological advancements, bridging the gap between emergent technologies and societal needs is paramount. This paper discusses the 'Met Mij' project, which addresses the prevalence of low literacy issues in the Netherlands. With over 2.5 million adults lacking basic skills in the country, ‘Met Mij’ leverages Artificial Intelligence and Large Language Models to democratize information access through personalized, adaptive interfaces that enhance user-system interactions for low-literate individuals. One significant issue is that there are no homogeneous groups of low-literate individuals, and many existing solutions fail to reach those in need correctly. In this sense, the core of this research explores the creation of a functional ecosystem that assists adults with limited basic skills by effectively connecting them with available help through innovative technological interventions. Employing a narrative literature review, the study integrates diverse insights into the wicked problem of low literacy, focusing specifically on the Dutch context. It critically examines the role of innovation in crafting impactful social solutions, particularly highlighting how AI and LLMs contribute to making the interaction experience more personal and individualized, leading to a meaningful match between human problems due to low literacy and services from social organizations. This paper underscores the theoretical foundations and practical implications of the ‘Met Mij’ project, positioning it as a significant step towards transforming human interaction with systems through accessible and effective technology. It aligns with broader global and regional goals to enhance societal engagement through technology, making a strong case for the role of innovation in solving wicked problems.

Keywords: Low Literacy · Human-System Interaction · Artificial Intelligence · Innovation.

1 Introduction

In today’s world, technological innovation is seen not only as a driver of economic progress but also as a vital response to urgent societal demands. Eindhoven Engine, created through a collaboration between TU/e, Fontys, and TNO, exemplifies this approach by accelerating innovation in the Brainport region with a special focus on addressing societal challenges through technology-based solutions. In particular, the Emergence Lab seeks to catalyze the convergence of technology and design to create more human and meaningful interactions between systems and their users [13]. Embedded in this context, the 'Met Mij' project that has been developed as part of the involvement with the TU/e's Human-System Interaction program proposes an innovative approach to facilitate interaction between individuals and systems, with a special focus on the needs of humans lacking basic skills in the Eindhoven region.

Low literacy, a prevalent issue in the Netherlands and many other countries, prevents a significant portion of the population from fully benefiting from the advantages offered by emerging technologies [35]. Only in the Netherlands, more than 2.5 million of people aged 16 years or older are considered to be low literate, which means they don't master the MBO level of the Dutch language, creating a lack of basic skills: literacy, numeracy, and digital skills [27, 18]. The absence of these skills poses challenges for individuals to actively engage in society, resulting in various detrimental impacts on both personal and societal fronts. In this context, 'Met Mij' aims to overcome these barriers by making the interaction individualized, recognizing that the group of people with low literacy is heterogeneous with diverse challenges and needs [10]. This approach provides a simplified interface that democratizes access to technology and information, thus aligning with the mission of radically transforming interaction experience with systems through accessible and effective technological solutions [37].

This paper is part of the ongoing Engineering Doctorate research aimed to answer the following research question: "How a functional ecosystem can be created by helping adults who lack basic skills in the Netherlands to find the right help when they interact with the system to facilitate the connection between organizations and citizens?". It discusses the theoretical foundation of the proposed solution called 'Met Mij', exploring the intersection of low literacy, innovation to solve complex problems, and using technologies such as Artificial Intelligence and Large Language Models to create practical and impactful solutions.

An overview of studies conducted about those main domains related to 'Met Mij' development is presented in this paper. A narrative literature revision was made to get a holistic and integrative view to understand better the wicked problem of low literacy, focused on The Netherlands [25, 36].

2 Methodology

The nature of this research is applicable since it is focused on developing a solution addressing the low literacy problem by using technology [11, 16]. Its goal is descriptive since it aims to study the characteristics of the main topics related to the present research [16]. As it requires an in-depth analysis of the object of study involving subjectivity and interpretation of results for the construction of knowledge, the approach is qualitative [31], and its process is inductive since "starting from specific, sufficiently established data it infers a general or universal truth that is not contained within the examined parts" [24].

A narrative literature review is particularly suited for addressing wicked problems due to its flexibility in exploring multiple definitions and perspectives, which aligns with the inherent ambiguity of wicked problems where no definitive problem statement exists. This review type accommodates the subjective nature of solutions—acknowledging that solutions are not right or wrong but rather better or worse—allowing for a nuanced assessment of various proposed solutions. Moreover, narrative reviews can effectively map the interconnected and symptomatic nature of wicked problems, highlighting how one issue may be a symptom of another. By synthesizing a wide range of possible solutions and interpretations, narrative reviews provide a comprehensive overview that is crucial for understanding and addressing the complex and multifaceted nature of wicked problems [25, 29, 36].

This paper aims to answer the following question: how can emerging technologies be utilized to create innovative, personalized, and effective solutions for adults lacking basic skills? The databases consulted were Google Scholar, the Online Library from Eindhoven University of Technology, Emergency Lab research files, and Dutch Government websites. The results of the narrative review are presented below.

3 Literature review

3.1 Wicked problem of low literacy

Rittel and Webber [29] introduced the concept of "wicked problems," which are complex and unique issues in planning and policy that are essentially different from "tame" or "benign" problems, which are more common in exact sciences or engineering. These problems lack clear and definitive definitions, and each time a solution is presented, the understanding of the problem changes. Also, it has solutions evaluated subjectively and involves high-risk experiments with real impacts. Solutions are not universally applicable, and each problem is interconnected with others. Moreover, solutions must be implemented carefully, as there is limited room for error.

Low literacy is a social problem that can be classified as wicked. The reasons someone lacks basic skills depend on several factors and equally impact several areas of people's lives. In the world, 773 million youth and adults have difficulty reading, writing, calculating, and using a computer or smartphone skills [35]. This issue is also a reality in the Netherlands, where approximately 2.5 million people aged 16 and above face this challenge in their daily lives [28]. Out of this, 65% are natives (NT1), and 35% are immigrants (NT2) and have Dutch as a second language.

Other issues can be found when a close look is taken into this scenario. Many initiatives were created to tackle this problem; however, they are usually discontinued over time, or the participants don't feel engaged enough to continue in the support programs. This wicked problem impacts various aspects of daily life, including health, finances, and overall well-being, which contributes to a widening economic gap, highlighting its complexity. Approaches to tackling it should focus on raising awareness about the problem and lowering the threshold for seeking help for those lacking basic skills. This calls for solutions that are sensitive to the diverse challenges faced by each person in this group, thereby advocating for individualized support mechanisms [18, 19].

As a result of a co-creation session between Emergence Lab and stakeholders [12], involving experts and practitioners in tackling this literacy challenge, three primary areas of action were identified to address the challenge effectively. These areas emphasize a multifaceted approach to understanding and mitigating low literacy issues:

1. Finding Cluster: Focus on identifying individuals struggling with low literacy. This involves creating strategies to reach out to those who might not readily seek help due to stigma or the hidden nature of their challenges.

2. Motivating Cluster: Concentrate on motivating individuals to recognize and address their literacy issues. It includes developing interventions that inspire and encourage people to engage with learning resources and support systems.

3. Helping Cluster: Aim at providing direct assistance through innovative educational tools and programs that go beyond conventional methods, addressing the unique needs of each individual.

The study of Simonis and Derks [32] tried to address the wicked problem of low literacy in the Eindhoven region by finding and familiarizing low-literate individuals with existing learning initiatives and using innovative technological solutions. One of the key insights from their study was the recognition of the significant barriers that low-literate individuals face in seeking help, including the stigma and shame associated with low literacy. This stigma often leads to reluctance to participate in literacy programs or other initiatives that could expose their literacy challenges.

Their project initially focused on creating a minimum viable product (MVP) that utilized eye-tracking software to identify low-literate individuals in various public settings like supermarkets and hospitals. This approach aimed to make the identification process both anonymous and accessible, thereby reducing the practical barriers to seeking help. However, they encountered substantial difficulties in locating low-literate individuals willing to participate in such initiatives, highlighting the deep-seated challenges related to social stigma.

Another key insight from their study was the importance of designing user-friendly and non-intrusive interventions. The students experimented with several innovative concepts, such as a karaoke booth and a recipe hub, to measure literacy levels through interactive means. They later pivoted towards a virtual reality (VR) solution intended to raise awareness and empathy by simulating the experiences of low-literate individuals. Despite its potential, they found significant challenges in applying VR for direct intervention, particularly the risk of users feeling confronted by their limitations. This led to the development of the "Help-O-Matic" concept, a mobile snack machine that discreetly provides learning resources and login details for further education. This iterative process, informed by continuous stakeholder and user feedback, underscored the necessity for sensitive, user-friendly interventions that can effectively connect low-literate individuals with the support they need. Another important aspect to consider is self-reliance. Usually, people lacking basic skills rely on others to resolve their issues. This, in turn, compromises their independence and autonomy [18].

The empathize mode in human-centered design is foundational, aiming to understand the human for whom you are designing deeply. It involves observing them within their environment, engaging through interviews and interactions, and immersing oneself in their experiences to gain a holistic view of their needs, thoughts, and emotions. By observing users, designers can gather valuable insights into what users think and feel, which helps uncover their needs and informs the development of innovative solutions. Engaging with users reveals their values and beliefs, providing a deeper understanding that goes beyond surface-level observations. This understanding is crucial for designing solutions that are truly empathetic and human-centered [7].

In the Eindhoven region alone, over 200 social organizations provide services to address citizens’ problems [34]. A social map contains a list of organizations that provide services in six domains: health, finance, work, education, social life, and society. Hootsmans [18] emphasizes that despite the availability of various interventions, a significant disconnect remains due to the lack of an integrated approach that considers the individual’s unique challenges and needs. The author also highlights that one crucial action step for Eindhoven to tackle low literacy effectively is to lower the threshold for people to find the help they need, especially for the group between 18 and 40 years old, since they usually are not willing to seek for help, since the stigma is bigger about their literacy challenges. Also, another study has shown that around 75% of the citizens seeking guidance in one of those organizations do not find what they need on their initial attempt and must be redirected to another institution. This can increase the feeling of shame and stigma around low literacy [2].

Building trust and reducing feelings of shame and insecurity through discreet and effective assistance could be a pathway to enhance the autonomy and engagement of people lacking basic skills [15]. Providing clear, understandable information tailored to the person’s context and needs will ensure that the support offered is both empathetic and effective. This complexity of the wicked problem can be transcended by leveraging innovative approaches that navigate the intricate web of interconnected issues.

3.2 Innovation

When dealing with wicked problems, innovation can be understood as a new way of thinking and doing business. Baets and Oldenboom [4] argue that innovation requires a shift in perspective—a new ontology at the base of business thinking—which affects not only the process of innovation but also the outcomes. This approach underscores the importance of changing how we perceive and approach problems, which is essential in addressing complex or ‘wicked’ problems. Solutions created to tackle wicked problems should lead to true innovation, that is, moving beyond traditional methods and challenging conventional boundaries and norms to embrace new ways of understanding and acting.

Buchanan, Ney et al. [8, 26] argue that innovation, especially through design thinking, is crucial for tackling wicked problems. They suggest that design thinking's iterative and human-centered approach enables practitioners to explore multiple perspectives and develop more comprehensive solutions to complex problems. This framework is based on 5 phases process that can be used to develop innovative solutions and ensure the continuous improvement of the system. It allows the designer to iterate through the phases non-linearly, making it possible to reshape and refine the solution [21]. A sixth phase, called implementation, can be added after the whole iteration process through the other phases, and it is related to launching the solution into the market. Those 6 phases follow a flow based on 3 steps: understand, explore, and materialize [14] (see Fig. 1).

During the co-creation session with the Eindhoven Engine team and stakeholders [12], participants highlighted the necessity of viewing the problem from different angles and conducting studies to understand the complexities of low literacy. This innovative perspective is crucial for developing meaningful solutions.

Besides looking at it from a different angle, it is important to seek innovative approaches when designing for such wicked problems. Biomimicry, which considers nature a source of inspiration for solving human challenges, has been used as a groundbreaking method. By observing and emulating nature's time-tested strategies, we can develop sustainable solutions that have evolved and persisted over millions of years. As Benyus [5] highlights, nature offers a wide range of successful "designs" that have been refined through the process of evolution, providing resilient and adaptable solutions that can inform and enhance human innovation. Biomimicry encourages us to learn from the natural world's vast knowledge repository, thereby fostering a deeper understanding of the issues at hand and promoting the development of more effective and sustainable interventions. It operates under three core principles:

- Nature as Model: this approach uses nature's designs and processes as inspiration to solve human challenges.

- Nature as Measure: it utilizes an ecological standard to judge the effectiveness of innovations. After 3.8 billion years of evolution, nature has learned what works, what is appropriate, and what lasts.

- Nature as Mentor: It promotes a new way of viewing and valuing nature, emphasizing learning from nature rather than merely extracting from it.

The author also points out a set of "Life Principles" that guide the application of biomimicry. These principles serve as a nature-inspired strategy framework for designing more resilient solutions that cater to problem complexity [5, 6] (Fig. 2).

Integrating biomimicry with design thinking, as explored in various studies, provides a robust framework for developing innovative solutions [22, 23]. While these studies primarily discuss sustainability, they demonstrate how biomimetic and design thinking approaches can effectively tackle other wicked problems, such as climate change. These insights underscore the potential of biomimetic approaches in redefining problem-solving and fostering sustainable innovation,

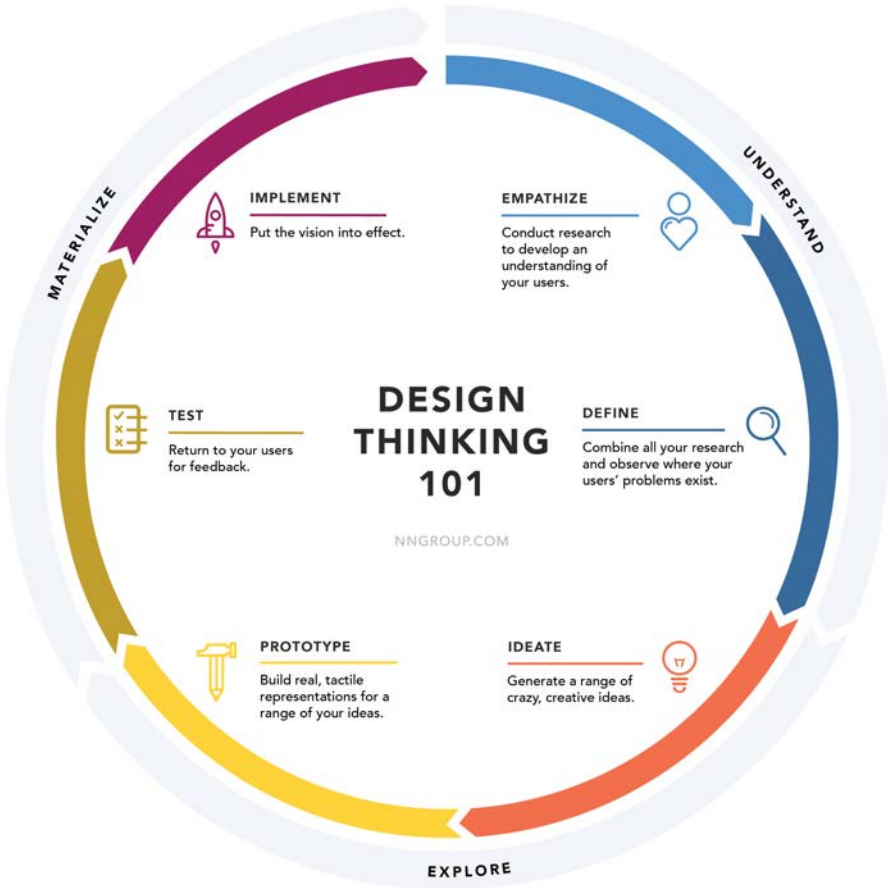


Fig. 1. Design Thinking phases and steps [14].

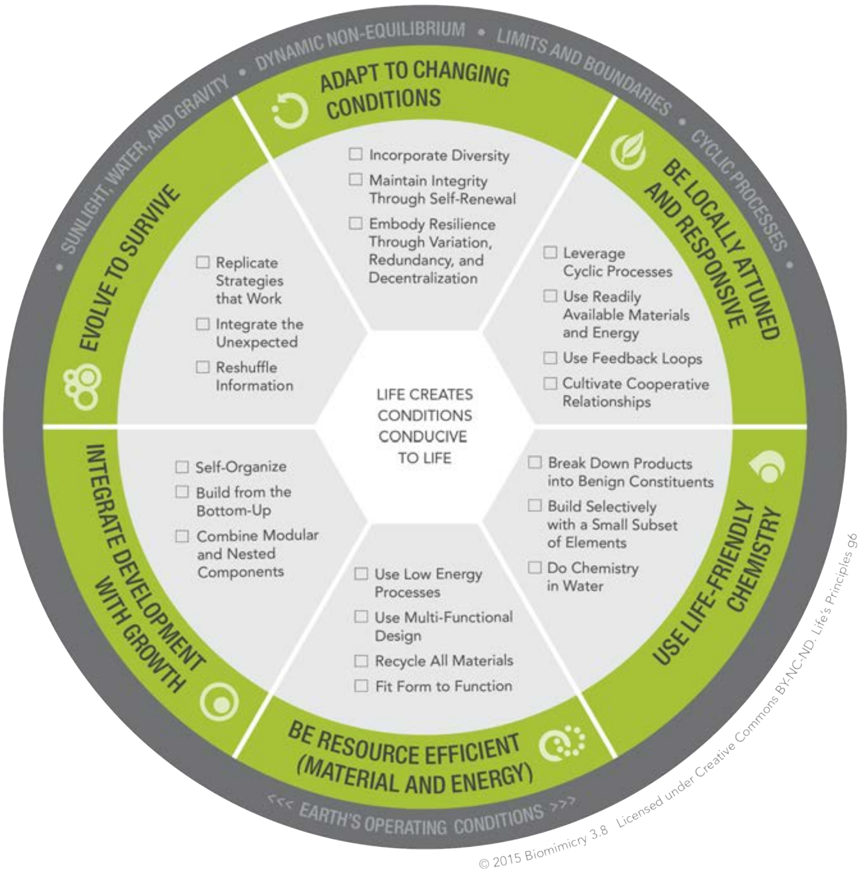


Fig. 2. Biomimicry: Life Principles [5, 6].

not only by minimizing environmental impact but also by creating resilient and adaptable solutions for complex societal issues.

Positioned within what has been termed the fourth industrial revolution, AI's role extends beyond commercial applications to address complex societal challenges. Ruysenaar [30] discussed that AI is increasingly seen as a crucial ally in tackling wicked problems through innovative solutions that reshape traditional problem-solving paradigms.

3.3 Technology

According to Ruysenaar [30], technologies such as AI significantly redefine the complexity of some wicked problems. The author emphasizes the need to rethink traditional technology uses since the fourth industrial revolution encompasses both technological and social dimensions. AI can both complicate and potentially solve aspects of wicked problems; the complexity of AI systems, their capabilities, and their unforeseen impacts position them as both contributors to and potential solutions for these problems. As Rittel and Webber [29] noted, solutions to wicked problems are typically judged as good or bad rather than right or wrong, underscoring the need to carefully evaluate proposed solutions before their implementation in the real world. By employing design thinking—a non-linear method that brings humans to the center of the design process—it is possible to enable new forms of problem-solving that are adaptive, integrative, and capable of handling the intrinsic uncertainty of wicked problems.

Machine learning can be integrated into the user experience to promote design innovations [37]. This field of study within Artificial Intelligence consists of developing algorithms that allow a computer to learn from examples or past data to perform tasks, improve, adapt over time, and learn from inputs. These algorithms enable a computer to gain insights, recognize patterns, and generalize knowledge to make decisions based on the available information [20].

Yang et al. [37] highlight usability and accessibility, which can be enhanced when system interfaces are designed to incorporate machine learning elements. This new design perspective unlocks new ways of how humans interact with systems, promoting a "more intuitive and effective" interaction (p.9). The usage of machine learning in user experience can be seen from 3 lenses. It was adapted to the present project context:

1. User's lens: focus on how machine learning affects users directly. This can be done by adapting content based on individual needs.

2. Provider-centric perspective: focus on engaging the user by iteratively interacting with the system to improve the algorithm.

3. Humanistic Perspective: focus on the ethical issues. It is related to possible biases that the tool might have due to the dataset used to train the model or the type of humans that interact with it.

Large Language Models (LLMs) represent a significant advancement in the field of artificial intelligence, particularly in the domain of natural language processing. According to Hadi et al. [17], LLMs are sophisticated AI systems designed to understand and generate human-like text by leveraging deep learning techniques. These models are trained on extensive datasets comprising a wide array of textual information, enabling them to generate coherent, contextually relevant responses across various topics. LLMs' capabilities include generating text and the ability to comprehend complex language nuances, making them invaluable tools for applications ranging from automated customer service to content creation and beyond.

AI-driven chatbots have enhanced the user experience in customer service interactions by employing anthropomorphic (human-like) cues, like using a conversational tone or exhibiting empathy, making them appear more relatable. These strategies make interactions more engaging, comfortable, and meaningful. These well-designed AI chatbots can significantly improve the user experience by making interactions more personable and increasing the effectiveness of service encounters in digital platforms [1].

ChatGPT, a conversational Large Language Model (LLM) developed by OpenAI, is one standout application of AI that creates unique human-system interactions. ChatGPT has revolutionized how businesses engage with customers, making the interaction more natural and engaging. Skjuve et al.'s [33] study shows that the user experience with ChatGPT has been predominantly positive, with many users appreciating its ability to provide detailed and useful responses, which significantly aided in both personal and professional tasks. Users found ChatGPT particularly helpful in enhancing creative work, such as writing stories, and many reported that the tool exceeded their expectations by delivering entertaining and engaging interactions. The simple language and clear structure of ChatGPT's responses were also well-received, contributing to a pleasant communication style. Some users also noted a reduction in anxiety and a feeling of companionship during interactions, highlighting the human-like communication present in the tool. However, there were some negative experiences, mainly involving technical issues or bugs, irrelevant or useless responses (sometimes referred to as hallucinations), and biases or inaccuracies in certain answers.

Digital games represent a distinct domain where AI-powered tools have been used to explore this novel opportunity. It is the case of the Akinator game that tries to guess the character, object, or animal the player is thinking of by asking a series of questions [3]. It leverages a sophisticated AI algorithm called a "decision tree" to narrow down the possibilities based on the player's answers. By continuously updating its database and learning from previous interactions, Akinator improves its accuracy and efficiency over time. The game showcases the power of AI in providing an interactive and engaging user experience by simulating a seemingly intelligent and intuitive guessing process.

Regarding low literacy from a technological perspective, Hootsmans [18] highlights the necessity of digital inclusion for people who lack basic skills. Although the digital world is often seen as a barrier, this group actively engages with digital platforms like social media, WhatsApp, and YouTube [18, 32]. It underscores the importance of considering designing digital solutions to tackle this problem, considering that it should be easily understandable and accessible.

4 'Met Mij' - Design Directions

Based on the insights previously discussed regarding wicked problems, low literacy, innovation, and technology, an empathy map was developed following the first phase of the design thinking design' method. This map synthesizes what people lacking basic skills say, do, think, and feel, highlighting their experiences and challenges. By capturing these elements, the empathy map serves as a tool to guide the design of personalized and effective interventions tailored to the diverse needs of individuals with low literacy (see Fig. 3).

Fig. 3. Empathy map: what does the person say, think, do, and feel?

Firstly, the statements captured in the "Says" quadrant highlight the frequent need for assistance and clarification among people from this group. Also, it represents the necessity of a trusty person that they usually rely on to receive assistance in what they need. This indicates a strong demand for accessible, supportive interfaces that provide clear guidance and self-reliable tools. Additionally, other statements reflect a tendency towards denial, the shame and stigma around low literacy, and the significant time constraints in participating in support programs that can prevent individuals from seeking help.

In the "Thinks" quadrant, questions reveal the confusion and uncertainty that many individuals feel about their situation and the available help resources. This underscores the importance of designing solutions that are easy to understand and access. Additionally, other questions show a reliance on personal networks, emphasizing the need for trustful solutions that support autonomy.

The behaviors noted in the "Does" quadrant, such as seeking help from trusted individuals and engaging in distracting activities like social media, indicate coping mechanisms used to avoid exposing their literacy issues. Avoiding situations that highlight their condition suggests a need for discreet, supportive tools to assist without drawing attention to their literacy challenges.

Finally, the "Feels" quadrant reveals a range of emotions, including confusion, insecurity, shame, and isolation. These feelings highlight the significant emotional impact of low literacy and the necessity for empathetic and non-stigmatizing support systems. Addressing these emotional needs is crucial for encouraging individuals to seek and accept help.

Regarding the definition phase of the design process, it is important to think of a solution that embodies an innovative approach by focusing on personalized interactions and interventions. By viewing each human's challenges through a unique lens, aiming to provide customized support based on specific needs and personal circumstances. This approach might ensure that interventions are more effective and more engaging for the system's users, facilitating a better understanding of each human and connecting them with the right organization. This strategy aligns with the session's emphasis on rethinking the problem with fresh eyes, showcasing how a nuanced approach to tackling low literacy can lead to more successful outcomes.

AI-driven chatbots have been shown to be an ally in enhancing user experience, making interactions personalized, and providing human-like conversations. At the same time, User Interfaces similar to those from social media and conversational apps seem to be promising solutions. Especially because the use of social media is extremely high among young people and adults in the Netherlands. Specifically, 96.8% of individuals aged 12 to 24 and 96% of those aged 25 to 44 use social media regularly. Based on these data, it can be inferred that a significant proportion of people aged 18 to 40 are also active on social media [9]. Analyzing the main features and simulating the interaction on those apps could be a design strategy to develop the solution.

In reference to the ideation phase, some biomimetic principles were chosen for this project. By being flexible enough to serve a wide range of needs and conditions, addressing different people's problems to different help organizations, the solution reflects the adapt to changing conditions principle. On the other hand, the principle of evolve to survive can be understood as offering customization based on the user's specific language or cultural context, simulating human-like conversations, adapting to the needs of its users, learning from their interactions to serve their needs over time better, and using feedback loops to train the machine learning model and to improve the match between person and organization. Also, the solution embodies this principle by replicating strategies that work, such as social media interfaces and AI-driven chatbots.

5 Conclusion

'Met Mij' emerges from the insights gathered from the above theoretical basis, especially by seeing the problem with a fresh lens. The project aims to shift the focus from merely identifying individuals with low literacy to understanding and addressing the specific problems caused by their literacy challenges. This approach recognizes that low literacy is not merely an educational shortfall but a multifaceted issue affecting various life aspects.

The 'Met Mij' application is designed with two primary functionalities:

1. Problem-Tracking Tool: This tool helps users with low literacy to identify specific areas where they require assistance due to their lack of basic skills, focusing on the six main domains presented before. The tool aims to detect these challenges non-invasively, using individualized interactions to engage users effectively without making them feel stigmatized.

2. Guidance System: Once the specific challenges are identified, the 'Met Mij' tool guides users toward appropriate specialized help. It connects them with existing social organizations. This part of the application is crucial for providing targeted, personalized support that goes beyond general literacy improvement, aiming to empower users to overcome the specific obstacles posed by low literacy and give them self-reliance and autonomy.

By utilizing advanced technologies like LLM, 'Met Mij' intends to enhance the user experience by making the interactions as human-like and intuitive as possible. This technology enables the tool to understand and respond to the unique needs of each user, making the assistance highly personalized and directly relevant to each individual's situation. Properly identifying individual needs can significantly enhance the efficiency of connecting people with the precise support they require and strengthen their autonomy and self-reliance. By focusing on tailored problem identification and subsequent guidance to appropriate services, 'Met Mij' aims to directly address these issues, leveraging technology to enhance accessibility and user experience for those with limited basic skills.

As the 'Met Mij' project progresses, the next steps are clearly outlined: developing a prototype and a rigorous evaluation of its acceptance among users since it is crucial for its successful implementation. This phase involves real-world feedback from users to refinement of the tool, ensuring that it is both effective and user-friendly. Initial testing with the target group has shown promising results, which will be presented in a further publication.

Given the complexity of low literacy as a wicked problem, as discussed by Rittel and Webber [29] and further explored by Hootsmans [18], the necessity for such a tool is evident. By continuously adapting to user needs through iterative development and user feedback, 'Met Mij' will help connect individuals to the right services and foster a more inclusive society. This aligns with the broader goals of technological innovations discussed throughout this paper, highlighting the pivotal role of AI and LLMs in addressing societal challenges.

About Jéssica Goss Jéssica Goss is an Engineering Doctorate trainee from the Human-System Interaction program at the Department of Industrial Design at the Eindhoven University of Technology (TU/e) in The Netherlands. She started the program in 2023 and has worked at the Emergence Lab at Eindhoven Engine since then. She received her master's degree in Design (Information Systems) centered on visual information and rhetoric on digital games from the Federal University of Paraná in 2020 and has published papers in digital entertainment, Human-System Interaction, visual communication, and game design. Her current research at TU/e is based on creating an innovative solution to tackle the low literacy wicked problem in The Netherlands. By borrowing concepts from game design and interactive design, her research seeks to develop an innovative, creative, and meaningful solution for those lacking basic skills in the city of Eindhoven.

References

1. Adam, M., Wessel, M., Benlian, A.: AI-based chatbots in customer service and their effects on user compliance. Electron Markets **31**, pp. 427—445 (2021). https://doi.org/10.1007/s12525-020-00414-7
2. Addokwei, S. A.: Using digital technologies to build a more equal and inclusive society, especially for people affected by low literacy. Master's degree thesis. Fontys Pulsed Academy, Eindhoven (2024)
3. Akinator Homepage, http://www.akinator.com, last accessed 2024/04/27
4. Baets, W., Oldenboom, E.: Values Based Leadership in Business Innovation. 1st edn. Bookboon, London (2013)
5. Benyus, J. M.: Biomimicry: Innovation Inspired by Nature. Harper-Collins e-books, New York (2009)

6. Biomimicry 3.8 Design Lens, https://www.biomimicry.net/the-buzz/resources/designlens-download-2/, (2015)
7. Both, T., Baggereor, D.: Design Thinking Bootcamp Bootleg. Institute of Design at Stanford. Retrieved from https://dschool.stanford.edu/resources/the-bootcamp-bootleg, (n.d)
8. Buchanan, R.: Wicked Problems in Design Thinking. Design Issues **8**(2), 5–21 (1992). https://doi.org/10.2307/1511637
9. Central Bureau voor de Statistiek. Nederland in Cijfers. (2020). Retrieved from https://www.longreads.cbs.nl/nederland-in-cijfers-2020/
10. Cubiss. Ken uw doelgroep: Onvoldoende basisvaardigheden Eindhoven. Retrieved from https://www.cubiss.nl/ken-uw-doelgroep-analyse-limburg-en-brabant, (2016)
11. Dresch, A., Lacerda, D. P., Júnior, J. A. V. A.: Design Science Research: método de pesquisa para avanço da ciência e tecnologia. Bookman, Porto Alegre (2015)
12. Eindhoven Engine. Emergence Lab - Summary of Co-creation Session, company file, (2022)
13. Emergence Lab Homepage, https://eindhovenengine.nl/emergence-lab/, last accessed 2024/04/27
14. Gibbons, S.: Design Thinking 101, NN/g Nilsen Norman Group. Retrieved from https://www.nngroup.com/articles/design-thinking/, (2016)
15. Gijsbers, N.: Onderbouwingsrapport Laaggeletterdheid en Gezondheidszorg. Avans University of Applied Sciences, Eindhoven Engine - Emergence Lab, Eindhoven (2024)
16. Gil, A. C.: Como Elaborar Projeto de Pesquisa. 2nd edn. Atlas, São Paulo (2002)
17. Hadi, M.U., Al-Tashi, Q., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M.B., Akhtar, N., Al-Garadi, M.A., Wu, J., Mirjalili, S.: Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects. In: TechRxiv (2023). https://doi.org/10.36227/techrxiv.23589741.v4
18. Hootsmans, I.: Understanding Low Literacy: Defining Design Opportunities. Eindhoven University of Technology, Eindhoven Engine - Emergence Lab, Eindhoven, (2022)
19. Hootsmans, I.: Navigating Through Low Literacy. Master’s degree thesis, Eindhoven University of Technology, Eindhoven, (2024)
20. Huang, J.: AI/ML, Data and Intelligence. In: Creativity and Aesthetics of Data and AI, Eindhoven University of Technology, Industrial Design Master’s degree course, class notes, (2023)
21. Interaction Design Foundation - IxDF, What is Design Thinking (DT)?, https://www.interaction-design.org/literature/topics/design-thinking, last accessed 2024/03/20
22. Kadar, T., Kadar, M.: Sustainability Is Not Enough: Towards AI Supported Regenerative Design. IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), pp. 1–6. Cardiff, United Kingdom (2020). https://doi.org/10.1109/ICE/ITMC49519.2020.9198554
23. Kennedy, E. B., Marting, T. A.: Biomimicry: Streamlining the Front End of Innovation for Environmentally Sustainable Products. Research-Technology Management, **59**(4), pp. 40—48 (2016). https://doi.org/10.1080/08956308.2016.1185342
24. Lakatos, E. M., Marconi, M. A.: Fundamentos de metodologia científica. 5th edn. Atlas, São Paulo (2003)
25. Lönngrena, J., Van Poeck, K.: Wicked problems: a mapping review of the literature. International Journal of Sustainable Development and World Ecology **28**(6), pp. 481–502 (2021). https://doi.org/10.1080/13504509.2020.1859415

26. Ney, S., Meinel, C.: Innovation, Wicked Problems and Design Thinking. In: Putting Design Thinking to Work: Understanding Innovation, pp. 21–42. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19609-7_2
27. Rijksoverheid: Anpaak Laaggeletterdheid, http://www.rijksoverheid.nl/onderwerpen/laaggeletterdheid/aanpak-laaggeletterdheid, last accessed 2023/10/25
28. Rijksoverheid: Laaggeletterdheid, (2022) http://www.rijksoverheid.nl/onderwerpen/laaggeletterdheid/aanpak-laaggeletterdheid, last accessed 2023/08/10
29. Rittel, H.W.J., Webber, M.M.: Dilemmas in a general theory of planning. Policy Sci **4**, pp. 155–169 (1973). https://doi.org/10.1007/BF01405730
30. Ruysenaar, S.: Thinking Critically About the Fourth Industrial Revolution as a Wicked Problem. In: Fields, Z. (Ed.) Handbook of Research on Using Global Collective Intelligence and Creativity to Solve Wicked Problems 2021, pp. 1–26. IGI Global (2021). https://doi.org/10.4018/978-1-7998-2385-8.ch001
31. Silva, E. L., Menezes, E. M.: Metodologia da pesquisa e elaboração de dissertação. 4th edn. UFSC, Florianópolis (2005)
32. Simonis, W., Derks, M.: Challenge report: Finding and familiarizing low-literate people with existing learning initiatives in Eindhoven region. Fontys Pulsed Academy, Eindhoven Engine - Emergence Lab, Eindhoven (2023)
33. Skjuve, M., Følstad, A., Brandtzaeg, P.B.: The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In: Proceedings of the 5th International Conference on Conversational User Interfaces (CUI '23) **2**, 1–10 (2023). https://doi.org/10.1145/3571884.3597144
34. Steunwijzer. Sociale kaart Eindhoven, https://www.eindhoven.steunwijzer.nl/, last accessed 2024/02/02
35. UNESCO. Guidelines on Open and Distance Learning for Youth and Adult Literacy (2021), https://www.unesdoc.unesco.org/ark:/48223/pf0000379397, last accessed 2024/04/27
36. Varpio, L., Aschenbrener, C., Bates, J.: Tackling wicked problems: how theories of agency can provide new insights. Medical Education **51**, pp. 353–365 (2017) https://doi.org/10.1111/medu.13160
37. Yang, Q., Banovic, N., Zimmerman, J.: Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), vol.130, pp. 1–11. Association for Computing Machinery, New York (2018) https://doi.org/10.1145/3173574.3173704

Can AI Bridge the Literacy Gap? Developing a GPT-4 Summarization Tool for Low Literacy

Sichen Guo[1[0009-0005-2163-4218]], François Leborgne[1[0009-0000-9639-9519]], Jun Hu[1[0000-0003-2714-6264]], and Walter Baets[2[0000-0003-1860-2483]]

[1] Department of Industrial Design, Eindhoven University of Technology
[2] Eindhoven Engine
s.guo3@tue.nl

Abstract. In the Netherlands, 2.5 million adults grapple with low literacy—a gap between integration and an increasingly information-dense society. To address this, we have developed an innovative intervention using GPT-4, integrated within the user interface built via Streamlit, to summarize and simplify the complexity of official documents and everyday text. This paper outlines the design and anticipated functionality of our tool, which uses the summarization capabilities of GPT-4 to transform complicated language into concise, accessible content with clear action points and connect with related organizations. Six participants were invited to this study. We evaluate the tool's impact on enhancing the societal participation of adults lacking basic skills, proposing that such an application can empower low-literacy individuals and bridge the gap to an informed and equitable society.

Keywords: LLM · AI · Human AI Interaction · low literacy

1 Introduction

On a typical quiet afternoon, you exit your room to find a letter waiting in your mailbox. The envelope's light blue color and prominent logo indicate its importance, originating from the government. Nevertheless, feelings of anxiety and irritation prompt you to delay retrieving the letter. Fears of unfavorable news, such as a fine or additional payment, weigh heavily on your mind. These concerns cause you to leave the letter unopened for six months. Ultimately, you muster the courage to open it. Initially, you experience uncertainty. Soon, this uncertainty gives way to discouragement. The reason lies in the letter's content, which features complex numbers and multicolored hyperlinks, rendering its message unclear.

This vignette illustrates a harsh reality. Lacking basic reading, math, and writing skills leads to many daily challenges. It makes managing money, getting information, dealing with health issues, and relationships difficult [1, 2]. The Survey of Adult Skills (PIAAC) shows adults' ability in three key skills: literacy, numeracy, and problem solving in technology-rich environments. Literacy is the ability to understand and respond to written texts. Numeracy is the ability to use numbers and math. Problem solving in technology-rich environments is the ability to access, interpret, and analyze information in digital environments [3]. Low literacy refers to individuals who struggle with reading and writing without being completely illiterate. It is known to be a significant issue in developed countries, including the Netherlands and the consequences are usually severe, leading to feelings of shame, insecurity, and dependence [4].

In the Netherlands, about 2.5 million people struggle with reading, math, and using digital devices due to low literacy. This includes about 1.3 million residents aged 16-65. Research shows that many adults are hesitant to admit their struggles with literacy and, as a result, rarely seek help [5, 6]. These people often struggle to find and keep jobs. Nearly half of those with poor basic skills are unemployed. Additionally, people with limited reading and writing abilities often face health and communication challenges [4, 7, 8]. Therefore, the Dutch government aims for every municipality to assist individuals with low literacy by 2024 [9].

Nowadays, many researchers focus on early intervention. It targets children's and teenagers' education [10]. However, the Survey of Adult Skills (PIAAC) reveals that, like most participating countries, a large minority of adults in the Netherlands are poor at literacy, numeracy, and problem-solving. Moreover, foreign-language immigrants in the Netherlands exhibit significantly lower Dutch proficiency levels than native-born adults with Dutch as their first language [3]. The National Library's Library and Basic Skills program supports efforts to address this issue by establishing a nationwide library infrastructure to improve basic skills for 45,000 vulnerable adults, including those who are low literate, refugees, migrants, unemployed, elderly, and computer illiterate [5].

In this context, the Emergency Lab from Eindhoven Engine[3] and Eindhoven University of Technology[4] aim to explore a more innovative solution to this problem. This work is part of the Met Mij [11] project from Eindhoven Engine. It investigates how AI can improve interaction and user experience for individuals with low literacy in Dutch using a well-known large language model (GPT-4) designed in an online interface (Streamlit) to tailor to low-literacy people. This study focuses on the development, design, and evaluation of a document summarization tool based on GPT-4. With a user-centered approach, qualitative and quantitative data are collected and analyzed. The goal of this study is to contribute to exploring how Artificial Intelligence (AI) can address the low literacy problem through more accessible design. The findings provide key insights into user interaction, covering information on informed effectiveness and efficiency, personal needs, and trust. These insights give a clearer picture of how low-literacy individuals interact with and obtain information from official documents despite their lack of proficiency in the Dutch language.

[3] https://eindhovenengine.nl/
[4] https://www.tue.nl/en/

2 Related work

2.1 Wicked Problem

In 1973, design theorists Horst Rittel and Melvin Webber introduced the term "wicked problem". They did this to show the complexities and challenges of addressing planning and social policy problems. Wicked problems lack clear aims and solutions. They face real-world constraints that prevent many risk-free attempts to solve them [12]. For these kinds of problems, there are no clear rules to stop complex problems, and there is no standard "right" or "wrong" criterion for solving them, only "good" and "bad" solutions [13].

Mari Suoheimo et al. [14] observed a significant shift in design practice over the last fifteen years, influenced by increased interaction with social and political matters. They suggest that designers should improve at navigating complex contexts. They should also get better at aligning designs with challenging situations. This will help them to effectively address tricky issues. Therefore, designers face complex, unpredictable, wicked problems. They must evolve their understanding and proposals while being open and sensitive.

Low literacy is a wicked problem in Dutch society, affecting a diverse group with various needs. Researcher Inge Hootsmans has identified a big constraint. It's the hidden group of young adults aged 18-40 who struggle with literacy issues [15]. This problem leads to feelings of shame, lack of well-being, and marginalization. As a result, it is crucial to address this long-standing issue in a more creative and innovative way.

2.2 Target Audience

In Eindhoven, 7% of the working population does not have the necessary basic skills, which is approximately 19,000 inhabitants [16]. The number is even higher in the report of the National Audit Office [17]. In recent years, efforts have been made to assist residents with limited basic skills. Although many hope for a decline in these numbers, the expectation is that they will actually rise. This increase is expected because more young people are leaving education without enough language skills. There are also more non-native speakers, and society is becoming more digital [18].

Low literacy individuals are split into different levels, as shown in Table 1. The "Everyone Basic Skills City Plan 2024-2028" outlines the levels in Eindhoven. It categorizes people based on their Dutch proficiency as a first (NT1) or second (NT2) language [18]. The levels are aligned with the Common European Framework of Reference for Languages (CEFR) and other educational standards. Below is a summary of the different literacy levels and the corresponding number of individuals in each category [19, 18]. The NT1 group represents about 54.2 percent of the low literacy population.

The city plan of Eindhoven aims to ensure that every adult in Eindhoven possesses basic language, arithmetic, and digital skills. A significant challenge lies in reaching and motivating NT1 individuals mentioned by other research [11,

Table 1. Language Level Reference Framework for Language and Arithmetic

Dutch as First Language (NT1)	Illiterate	1F	2F	3F	4F
Dutch as Second Language (NT2)	A0	A1-A2	B1	B2	C1-C2
Comparable Education Level		End of primary school	End of VMBO[5] and MBO[6]1,2,3	End of MBO-4 or HAVO[7]	End of pre-university education (vwo), higher professional education (hbo), or university (wo)
	Low literate		Not low literate		
Legal regulations	Integration, Language Requirements, WEB				

source https://www.eindhoven.nl/

[5] Preparatory secondary vocational education
[6] Secondary vocational education
[7] General secondary education

15]. To gain more insights into their behaviors and feedback for improvement, we studied NT1 together with the NT2 group.

According to the city plan, this study will focus on both NT1 and NT2 groups who belong to 1F (Table 1). These individuals, at rudimentary and basic literacy levels, are the target audience for this study.

2.3 Large Language Models(LLMs) and GPT-4

Large language models (LLMs), like GPT-3 and GPT4, have greatly improved natural language processing (NLP). They were trained on huge amounts of text data to generate human-like text and do language tasks accurately [20–22]. Recent advancements in LLMs have demonstrated their potential to achieve a level of intelligence comparable to humans [23, 24].

OpenAI developed the Generative Pre-Trained Transformer (GPT), a language model capable of producing human-like text [25]. The development of GPT follows a two-step process: generative, unsupervised pretraining using unlabeled data, and discriminative, supervised fine-tuning [26, 27]. GPT stands out due to the scale of its training program and the extensive amount of data used. With access to the entire internet, the algorithm is trained on billions of data sources [28]. As a result, GPT can perform a wide range of language-based tasks, such as translation, question answering, and text generation.

The latest large-scale, multimodal model, GPT-4, developed by OpenAI, is a significant area of study due to its ability to process both image and text input and generate text output. This model holds great potential for various applications, such as dialogue systems, text summarization, and machine translation [29].

Currently, there is much research related to how the GPT model can solve or improve financial literacy and health literacy [30–34]. Moreover, there is also research on GPT-3 on simplified mathematics problems in education and people with cognitive impairments, which all contribute to better understandable results for the target audience [35]. However, research also indicates that the current GPT interface is not suitable and accessible for people with low literacy, especially those who have difficulties in reading and writing [36].

Additionally, XuanXin Wu et al. found that GPT-4 makes simpler outputs with fewer errors. It also keeps the original meaning better than the best current model, Control-T5. Their research highlights GPT-4's superior performance in text simplification tasks [37].

In conclusion, the GPT-4 model demonstrates an effective ability to comprehend context, summarize, and simplify text, thereby generating more understandable and accessible content. Due to its superior performance, we have chosen the GPT-4 model for text summarization and simplification for our work.

3 Method

3.1 Interactive Prototype

The interface prototype prioritizes logic and practicality. It has an interactive interface tailored to our research. It also takes into account the significant cultural disparity between users and developers. Dutch experts conducted initial testing and interviews during the first iteration. Figure 1 illustrates the logic underlying the prototype. This prototype includes a scenario and interactive interface built with Streamlit[8]. They help users form immersive during the evaluation study and are easy to interact with.

The prototype uses the prompt test to give clear context and visuals through icons. It categorizes the summarized content into four sections. The comprehensive user summary page is illustrated in Figure 2.

- *Sender Information* - Identification of the party responsible for sending the correspondence
- *Action Points* - Specific tasks the recipient is expected to undertake upon receiving the letter
- *Contact Details* - Means of communication
- *Direct Action* - A direct avenue for recipients to establish contact

[8] https://streamlit.io/

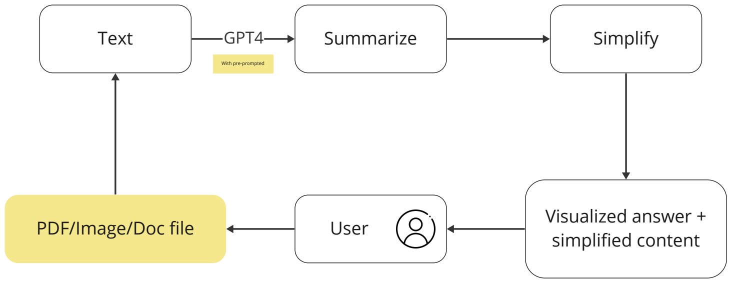


Fig. 1. user information flow.

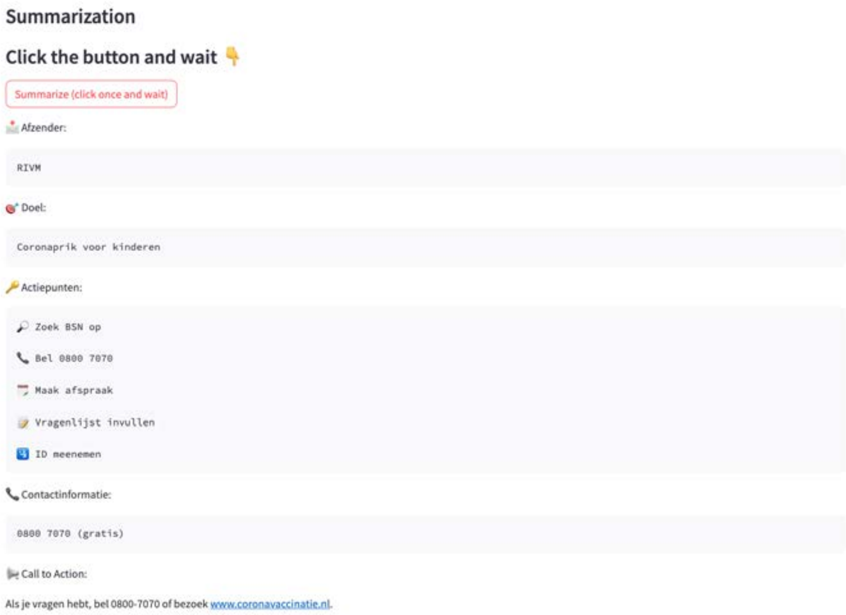


Fig. 2. Summarized content and interface

3.2 Evaluation Study

In the Evaluation study, we first did the whole process test with baseline and the prototype with the NT1 group (P1, P2). An oral assistant and a text content Dutch assistant together with researcher work on the study. Then, we tested with four NT2 group members (P3-P6) with a link they can test on their digital devices. The qualitative and quantitative data are collected. It was based on the six questions about the four categories from the letter. The User Experience Questionnaire (UEQ) combined with Cognitive Load Questionnaire(CLQ) are used in this study [38, 39]. It focuses on effectiveness, efficiency, and cognitive load. Additionally, we did semi-structured interviews for the NT1 and NT2 groups. We asked open-ended questions to learn more about their user experience.

3.3 Participants

In total, 6 users were evaluated and interacted with the tool and did the interview. Hereby, there are 2 people from the NT1 group and four people from the NT2 group. (n=6) We found five themes in the data. They are: (1) user interaction, (2) informing effectively, (3) reading efficiently, (4) understanding context, and (5) personalization and customization needs

Set up The evaluation consists of two sessions, as shown in Figure 3. In the first session, the user chooses the letter and reads the original text. There are four designated topics. However, only two are currently accessible: health and finance. After thoroughly reviewing the correspondence, participants will be prompted to use the evaluation questionnaire. Subsequently, respondents will address six inquiries about content across three distinct categories. A secondary questionnaire focusing on user experience and cognitive load will ensue, followed by an invitation for participants to partake in an interactive session involving our prototype. Upon completing the text overview, participants will once again complete the evaluation questionnaire and the content-related inquiry. Lastly, an in-depth semi-structured interview is reserved for individuals categorized under NT1 to glean additional insights.

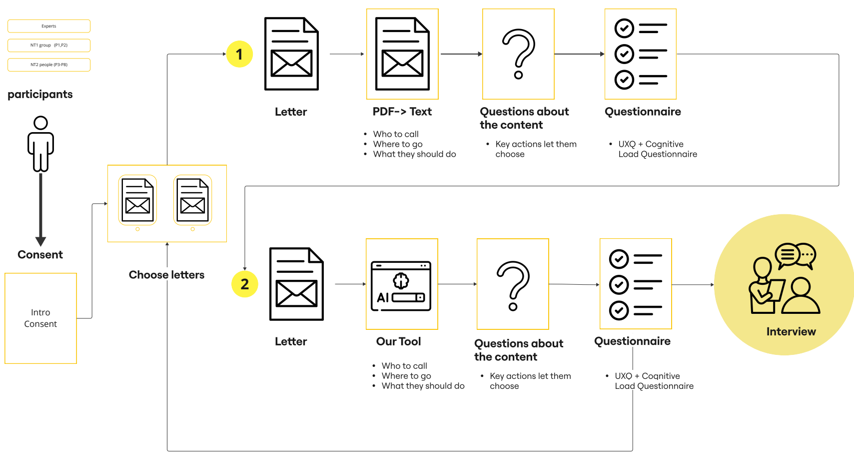


Fig. 3. Evaluation set up.

4 Findings

4.1 Quantitative Data Analysis

The data analysis shows participant performance across two groups (NT1 and NT2) and two letter types (health and financial). It covers before and after an intervention. The metrics evaluated include sum scores (SS) and accuracy ratings (AR). The data is summarized in Table 2

Table 2. Participant Performance Before and After Intervention

Group	**Participant**	**Letter Type**	**Sum (Before)**	**Accuracy (Before)**	**Sum (After)**	**Accuracy (After)**
NT1	P1	health	11	5	44	5
NT1	P2	health	25	2	34	1
NT2	P3	health	29	5	42	5
NT2	P3	financial	24	5	40	5
NT2	P4	health	32	6	45	6
NT2	P4	financial	45	6	45	6
NT2	P5	health	31	5	31	6
NT2	P5	financial	31	6	44	6
NT2	P6	health	34	6	44	6
NT2	P6	financial	34	6	43	5
Average			29.6	5.2	41.2	5.1

The data in Table 2 demonstrate a significant increase in average sum scores (SS) from 29.6 before the intervention to 41.2 after, indicating enhanced participant performance across both health and financial letter types. Despite this improvement, the average accuracy rating (AR) slightly decreased from 5.2 to 5.1, suggesting that overall task performance improved while accuracy did not uniformly benefit. However it is notable for P2 from NT1 group, even though P2 claimed to get a higher score, the accuracy after the tool is actually less.

Figure 4 and Figure 5 show a substantial increase in cognitive load post-intervention, with median and mean values rising significantly. Participants handling financial letters generally maintained or improved their SS and AR, as exemplified by P4. Besides, the bar chart in Figure 5 highlights big increases in cognitive load related to effectiveness and efficiency. Post-intervention scores increased to 4.53 and 4.55, respectively. This suggests that participants found tasks more effective and efficient.

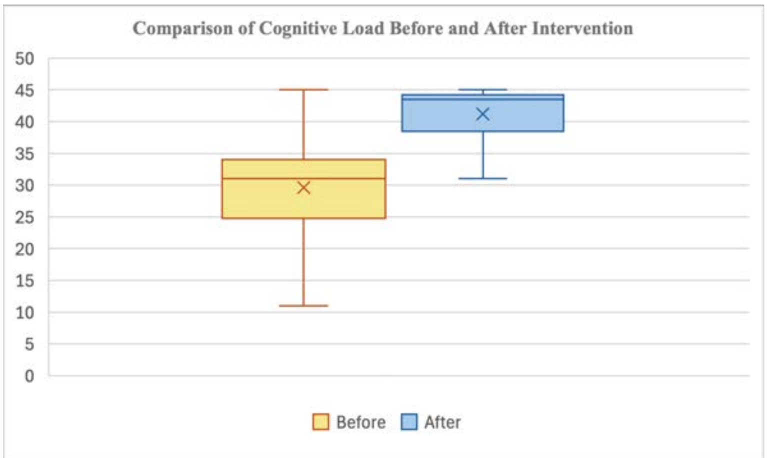


Fig. 4. Comparison of Cognitive Load Before and After Intervention.

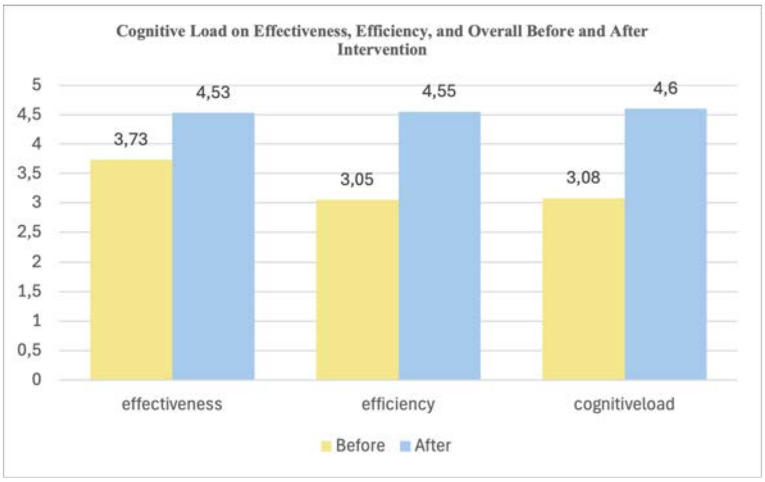


Fig. 5. Cognitive Load on Effectiveness, Efficiency, and Overall Before and After Intervention.

4.2 Qualitative Data Analysis

A significant difference in attitude between NT1 and NT2 groups Aiming to explore the whole user experience further, each participant got a semi-structured interview and provided qualitative data. Researchers found that there were significant behaviors in the NT1 group and the NT2 group. The NT1 and NT2 groups can quickly search for answers when they get the letter. However, NT2 has more confidence in answering due to their better searching skills, as shown in Table 2. When P1 from the NT1 group got the letter, it was noticeable that the participant mentioned, "If it's blue or thick... I don't open it right away... I just put it aside." This reaction highlights the emotional barrier and procrastination associated with dense, official-looking mail." The NT2 group has a learning attitude toward reading letters. P3 (NT2) mentioned, " I can understand most of the content, about 50%." Generally, the NT2 group showed more confidence than the NT1 group when they received official documents. Also, P1 mentioned that after finishing the letters, "I found it took a lot of effort because I found these list things very unclear." This underscores the need for clear and concise communication in official documents. Most of the NT2 group mentioned that they have memorable letter-related to the daily topics .

Preference for Summarized Content There was a clear preference for summarized content over full text. One participant mentioned, "But if something just says like... Immediately what you have to do. Then there is very little room for noise and panic." This preference highlights the importance of direct and clear communication. Furthermore, an NT1 participant mentioned twice that the text was perfect. They said they wanted to use it for other letters. "I always find it very nice because reading a (summarized) letter is exciting. It helps me avoid misunderstanding the information and prevents me from creating fake news in my panic." From the NT2 group, there are three participants who mentioned that it would be good to use it for financial problems or to help them select them of different letters. One participant mentioned specifically that "it would be nice if I could use it for reporting taxes, and I don't need to search for pieces of information online."

Improved Confidence and Efficiency Using the tool especially made NT1 participants feel more confident and efficient in handling official documents. One participant stated, "This app with summarizing is always faster than waiting two weeks to open that letter." This reflects the tool's ability to streamline the process and reduce the cognitive load. Additionally, the NT2 group mentioned that this tool can enhance the time spent reviewing official documents. One mentioned that if the tool can help them know the topic quicker, it can also help them tell if this letter is for them or not.

Accessible user experience design elements. Participants expressed a need for customization, such as larger font sizes and clear visual indicators. One from the NT1 group noted, "I would have made it a bit bigger personally." This feedback points to the importance of adaptable design elements in enhancing user experience. Meanwhile, participants mentioned that it would be nice if there were an audio option or if some people could read the content. NT2 people want more layers; for example, they can adjust how detailed information they can see. All participants mentioned that it would be nice if the tool could scan the document on the mobile.

Trust between technology and users. Trust is the biggest issue mentioned by NT2 participants mostly; one participant noted, "I want to know more explanation on how the tool summarized the key action points. For example, if I click on the summarization, it can turn to the original file automatically". There is a trust gap between LLMs and users; they want to know more about why and how the information is given. Besides, it depends on the topic. If they know the documents are important, they prefer to use a summarization tool first and then double-check with the original files.

5 Discussion

The digital summarization tool enhances the reading of official documents and streamlines the process of responding to them. NT1 participants, who had previously delayed engaging with such documents due to anxiety and confusion, reported significant improvements in their ability to handle these tasks promptly. The tool effectively presents key information, enabling users to make decisions and act quickly without prolonged hesitation. This efficiency reduces both the time spent and the cognitive burden associated with understanding complex official documents.

Accessibility emerged as a critical factor in the tool's effectiveness. Participants emphasized the need for customization options, such as larger font sizes, clear visual indicators, and audio support, to cater to their diverse needs. These features are needed for an inclusive user experience. They are especially important for low-literacy people who may struggle with standard text. By adding accessible design elements, the tool can serve a wider range of users, ensuring that individuals with different literacy levels can manage official documents effectively.

The qualitative feedback underscored the emotional challenges faced by NT1 participants when dealing with official documents. The tool's design addressed these challenges by reducing anxiety and fostering a sense of control over the information. Participants prefer summarized content and they mention it is important to communicate directly and clearly. This helps reduce overwhelm and panic. Researchers analyze the users' feelings and aims to reduce negative ones. This approach improves user engagement and satisfaction.

The study revealed distinct behavioral differences between NT1 and NT2 groups. NT1 participants often feel anxiety and delay dealing with official documents. In contrast, NT2 participants show more confidence and a proactive attitude, likely due to their better-developed basic skills. This difference shows the importance of tailored interventions that address the specific needs and challenges of each group. The tool enhances NT1 participants. on getting information simpler and clearer. This is especially helpful for overcoming emotional barriers and promoting timely action.

Trust is crucial for adopting digital summarization tools. This is especially true for NT2 participants. Since they want more transparency in how the tool works. Users need explainable AI features to understand how the summarization works. The functions let users compare the summary with the original documents, which will be useful to solve this problem. These features can boost trust and reliability. Building this trust is essential. Users need it to feel confident in using the tool for critical tasks. This trust will increase the tool's overall effectiveness and user satisfaction.

6 Conclusion

In conclusion, the digital summarization tool with GPT-4 has shown great potential. It can improve the efficiency, accessibility, and emotional comfort of low-literacy individuals with official documents. While the tool enhances performance and reduces cognitive load, further refinements in accessible design, emotional considerations, and explainability are necessary to fully meet the needs of both NT1 and NT2 groups. By addressing these areas, the tool can play a more transformative role in empowering low-literacy individuals and bridging the gap to an informed and equitable society.

About Sichen Guo

Sichen Guo is an Engineering Doctorate (EngD) working in Eindhoven University's Technology, Netherlands, Human System Interaction program. She is also a member of the Emergency Lab at Eindhoven Engine, which focuses on designing to bridge the gap between people who lack basic skills and society. She has a bachelor's and master's degree in Industrial Design. During her master's study, she focused on socially inclusive design, user interaction, VR, and gamification. As a designer, engineer, and researcher, she takes a user-centered approach to real-life problems, combining innovative technology and design methodologies.

References

1. B Suresh Lal. The economic and social cost of illiteracy: an overview. *International Journal of Advance Research and Innovative Ideas in Education*, 1(5):663–670, 2015.
2. Nancy D Berkman, Darren A DeWalt, Michael P Pignone, Stacey L Sheridan, Kathleen N Lohr, Linda Lux, Sharon F Sutton, Tameka Swinson, and Arthur J Bonito. Literacy and health outcomes: summary. *AHRQ evidence report summaries*, 2004.
3. Education gps, oecd. http://gpseducation.oecd.org. last accessed: 2024/05/21.

4. Laaggeletterdheid-slo. https://www.slo.nl/thema/vakspecifieke-thema/nederlands/laaggeletterdheid/. last accessed: 2024-05-21.
5. Adriaan Langendonk and Maaike Toonen. Dutch approach to prevent and curate low literacy. 2017.
6. World Literacy Foundation. The economic and social cost of illiteracy: A white paper by the world literacy foundation. *The World Literacy Summit*, 2018.
7. Lezenenschrijven. https://www.lezenenschrijven.nl/. last accessed: 2024-05-21.
8. Basisvaardigheden. https://basisvaardigheden.nl/themas/thema-basis-over-basisvaardigheden. last accessed: 2024-05-21.
9. Rijksoverheid. https://www.rijksoverheid.nl/onderwerpen/laaggeletterdheid/aanpak-laaggeletterdheid. last accessed: 2024-05-21.
10. Sanne L Tamboer, Inge Molenaar, Tibor Bosse, and Mariska Kleemans. Testing an intervention to stimulate early adolescents' news literacy application in the netherlands: A classroom experiment. *Journal of Children and Media*, 18(1):60–79, 2024.
11. Jéssica Messias Goss Dos Santos. 'met mij': Ai-enhanced tool for identifying and connecting people and services, 2024.
12. Randall Bass. What's the problem now? *To Improve the Academy: A Journal of Educational Development*, 39(1), 2020.
13. Horst WJ Rittel and Melvin M Webber. Dilemmas in a general theory of planning. *Policy sciences*, 4(2):155–169, 1973.
14. Bo Westerlund and Katarina Wetter-Edman. Dealing with wicked problems, in messy contexts, through prototyping. *The Design Journal*, 20(sup1):S886–S899, 2017.
15. Inge Hootsmans. Navigating through low literacy, 2024.
16. Geletterdheid in zicht. https://geletterdheidinzicht.nl/. last accessed: 2024-05-21.
17. Rekenkamer. https://www.rekenkamer.nl/. last accessed: 2024-05-21.
18. eindhoven.nl.
19. CONSIGLIO D'EUROPA. Common european framework of reference for languages (cefr). *Learning, teaching*, 2018.
20. Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and individual differences*, 103:102274, 2023.
21. Elizabeth D Liddy. Document retrieval, automatic. 2005.
22. Mark Burry. A new agenda for ai-based urban design and planning. In *Artificial Intelligence in Urban Planning and Design*, pages 3–20. Elsevier, 2022.
23. Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
24. Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.
25. Robert Dale. Gpt-3: What's it good for? *Natural Language Engineering*, 27(1):113–118, 2021.
26. Dumitru Erhan, Aaron Courville, Yoshua Bengio, and Pascal Vincent. Why does unsupervised pre-training help deep learning? In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 201–208. JMLR Workshop and Conference Proceedings, 2010.

27. Paweł Budzianowski and Ivan Vulić. Hello, it's gpt-2–how can i help you? towards the use of pretrained language models for task-oriented dialogue systems. *arXiv preprint arXiv:1907.05774*, 2019.
28. Luciano Floridi and Massimo Chiriatti. Gpt-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30:681–694, 2020.
29. Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
30. Pragnya Ramjee, Bhuvan Sachdeva, Satvik Golechha, Shreyas Kulkarni, Geeta Fulari, Kaushik Murali, and Mohit Jain. Cataractbot: An llm-powered expert-in-the-loop chatbot for cataract patients. *arXiv preprint arXiv:2402.04620*, 2024.
31. Ziqi Yang, Xuhai Xu, Bingsheng Yao, Ethan Rogers, Shao Zhang, Stephen Intille, Nawar Shara, Guodong Gordon Gao, and Dakuo Wang. Talk2care: An llm-based voice assistant for communication between healthcare providers and older adults. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(2):1–35, 2024.
32. Md Mushfiqur Rahman, Mohammad Sabik Irbaz, Kai North, Michelle S Williams, Marcos Zampieri, and Kevin Lybarger. Health text simplification: An annotated corpus for digestive cancer education and novel strategies for reinforcement learning. *arXiv preprint arXiv:2401.15043*, 2024.
33. Hana L Haver, Anuj K Gupta, Emily B Ambinder, Manisha Bahl, Eniola T Oluyemi, Jean Jeudy, and Paul H Yi. Evaluating the use of chatgpt to accurately simplify patient-centered information about breast cancer prevention and screening. *Radiology: Imaging Cancer*, 6(2):e230086, 2024.
34. Julie Ayre, Olivia Mac, Kirsten McCaffery, Brad R McKay, Mingyi Liu, Yi Shi, Atria Rezwan, and Adam G Dunn. New frontiers in health literacy: using chatgpt to simplify health information for people in the community. *Journal of General Internal Medicine*, 39(4):573–577, 2024.
35. Nirmal Patel, Pooja Nagpal, Tirth Shah, Aditya Sharma, Shrey Malvi, and Derek Lomas. Improving mathematics assessment readability: Do large language models help? *Journal of Computer Assisted Learning*, 39(3):804–822, 2023.
36. A Fiora, F Piferi, P Crovari, and F Garzotto. Exploring large language models for the education of individuals with cognitive impairments. In *INTED2024 Proceedings*, pages 4479–4487. IATED, 2024.
37. Xuanxin Wu and Yuki Arase. An in-depth evaluation of gpt-4 in sentence simplification with error-based human assessment. *arXiv preprint arXiv:2403.04963*, 2024.
38. M Schrepp. User experience questionnaire handbook version 8. 2019. *URL: https://wWww. ueq-online. org/Material/Handbook. pdf*, 2021.
39. Gordon Matthew. The effect of adding same-language subtitles to recorded lectures for non-native, english speakers in e-learning environments. *Research in Learning Technology*, 28, 2020.

Empowering AI Education by Enhancing Scalability and Accessibility of the Mix & Match Machine Learning Toolkit

Anniek Jansen[1[0000−0003−3474−457X]] and Sara Colombo[2[0000−0003−3902−1581]]

[1] Eindhoven University of Technology, The Netherlands
[2] Delft University of Technology, The Netherlands
a.jansen@tue.nl

Abstract. With the integration of Artificial Intelligence (AI) into our daily lives, fostering AI literacy among users is essential. Recognizing this need, AI education is increasingly finding its place in curricula, with a growing array of tools catering to diverse age groups. A selection of these tools is based on tangible user interfaces, as they offer advantages for (collaborative) learning. However, the scalability of these tools is often very limited as they are resource intensive to make. In this work, we address this scalability challenge for the Mix & Match Machine Learning toolkit by replacing the sensing board and sensorized tokens with a mobile web application that can scan printed tokens. Moreover, some updates to the toolkit are presented and evaluated that can engage and support students in their learning. With these updates, we aim to make the Mix & Match ML toolkit more accessible and scalable, thereby facilitating AI literacy education on a broader scale.

Keywords: Machine learning · Tangible user interface · Education · Mobile application

1 Introduction

In a short period, the role of Artificial Intelligence (AI) in everyday life has undergone a significant transformation. Previously, AI was integrated into some products and services, often unbeknownst to users. For instance, in the form of music or video recommendations, or suggestions for products similar to ones previously viewed. However, this dynamic shifted with the increasing accessibility and emergence of generative AI. Now, users have the power to harness generative AI for their own needs and objectives. Students utilize it for crafting essays, programmers employ it to co-create code, and even artists utilize it for generating imagery.

As AI becomes more deeply ingrained into our lives, it becomes paramount for users to attain AI literacy [13]. This encompasses not only the ability to utilize AI, but also the capacity to think critically and comprehend associated risks [13]. Acquiring these competencies necessitates integrating AI education into curricula and developing methods and tools suitable for educational settings. Prior research has demonstrated how, for example, ethical considerations of AI can be taught to high school students [9], or how fundamental principles of the learning and evaluation phases can be explained to children [6, 10], as well as design students [7]. In this work, special attention is paid to tangible user interfaces (TUIs) used in AI education [6, 10, 7, 2]. While TUIs have several advantages when it comes to learning [12, 15, 26, 18], they suffer from limited accessibility and scalability due to their reliance on custom parts and electronics. In this work, we therefore address this scalability challenge by enhancing the Mix & Match Machine Learning toolkit through the utilization of printed tokens and a mobile web application for token scanning. Furthermore, the web interface of the toolkit was updated to include interactive examples featuring real ML models, facilitate the creation of sequential tokens and enable direct comparison of two tokens of the same category (e.g. two data tokens). A small evaluation study was conducted with students to evaluate the first updates that were done to the physical toolkit, and identify other pain points. This study highlighted once more that the toolkit had to become more scalable if to be used in education. After this study, the mobile web application was designed and implemented.

With these enhancements, we aim to make the Mix & Match ML toolkit more scalable and accessible for educational purposes. All the materials and instructions are open source and can be found online [3].

2 Related Work

2.1 AI education

With the increase of AI in every day (professional;) life, to illustrate a 2023 report from McKinsey report that one third of the companies started using Generative AI within a year after its release [16], AI literacy becomes a skill for everyone and not just computer science students. AI literacy can be defined as *"a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace"*[13]. From this definition, it also becomes clear that AI literacy mainly focus on equipping lay people with the right knowledge and skills to responsibly use AI, something that will only become more important in our society [17].

Users across different age groups will need to learn these skills. Students can learn it at primary, high school and at university, while adult learners will need to equip themselves with these skills at their work or during their free time. The toolkit highlighted in this work, the Mix & Match ML toolkit, is targeted at students at the end of their high school and older but can also be used within companies to familiarize teams with the possibilities of AI.

[3] https://github.com/MixMatchMLtoolkit

2.2 Educational resources

To teach AI literacy, a wide range of demos, activities, tools, and methods have been developed [3] that explain different AI concepts. For instance, Teachable Machine [5] explains the training phase to users and the AI+ethics curriculum from the MIT media lab [4] focuses on the ethical issues. In literature, several studies have also investigated this topic of teaching AI to different age groups [25, 24, 19, 6, 22, 23, 9, 10]. Sulmont et al. [20] identified challenging topics when teaching AI while for instance Hitron et al. and Kaspersen et al. demonstrated that children can understand the training and evaluation phase of AI using a tangible user interface (TUI) [6, 10]. While in literature, there are several studies using TUIs to teach AI to mainly children, the majority of resources identified by Druga et al. is digital or might make use of some off-screen activities but hardly ever of a TUI [3].

2.3 Tangible User Interfaces

This difference might be explained by the limited scalability of TUIs [1]. They are often used in research as they offer a multitude of potential advantages for learning and collaboration [12, 15, 26, 18]. TUIs for example have the benefit of natural affordances which can reduce the cognitive effort and lower the threshold for participating [15, 26]. Moreover, due to their increased visibility, the possibility to observe the current state of the work and to interact with multiple people, TUIs work well for collaborative learning [26]. However, most TUIs presented in research cannot easily be made by others to use in education. The maker files are not always available, they might require specific materials and equipment or are simply to resource intensive to create for a whole class [1]. In this work, we therefore focus on how we can improve the scalability of the Mix & Match ML toolkit [7], a tangible user interface which uses a token+constraint approach to communicate information through tangible tokens and their constraints [21].

3 The Mix & Match Machine Learning Toolkit

The Mix & Match ML toolkit was originally introduced in [7] as an ideation toolkit to design ML-enabled solutions. However, the potential of the toolkit reaches beyond ideation for designers. It could also serve as an educational toolkit potentially for a broader audience. The study conducted in [7] showed that the toolkit allowed users to achieve learning goals associated to the "remember", "understand", and "apply" levels of the Bloom's Taxonomy [11].

The Mix & Match ML toolkit is based on a framework of ML knowledge that consists of data types (*labeled and unlabeled audio, image, table, text, time series, and video data*), ML capabilities (see Table 1) and exemplars of ML applications. The toolkit as presented in [7] itself consists of a set of tangible tokens, a sensing board and a web interface. We will refer to this version as

[4] https://www.media.mit.edu/projects/ai-ethics-for-middle-school/overview/

Table 1: The twelve ML capabilities included in the toolkit with their respective learning approaches and definitions.

Machine learning capability	Learning approach	Definition
Categorize	Supervised	Match an item with a category
Foresee	Supervised	Predict an action, event, state, behavior, intention, preferences, etc.
Identify	Supervised	Recognize the identity of a specific individual/item from a trait
Translate	Supervised	Transform contents from one domain to another
Understand	Supervised	Comprehend topics, themes, or sentiments; interpret language
Communicate	Supervised	Convey messages/content in understandable languages
Cluster	Unsupervised	Group items based on their similarities
Distinguish	Unsupervised	Differentiate certain items from a group or average (find outliers, anomalies, etc.)
Recommend	Unsupervised	Provide suggestions or guidance; propose contents, activities, etc.
Generate	Unsupervised	Create content (e.g. videos, images, music, text) from scratch
Optimize	Reinforcement	Improve or perfect a certain task/route/process
Navigate	Reinforcement	Steer autonomously through a physical or virtual environment

v1.0. The tokens represent the different data types and ML capabilities and have an RFID label embedded. These sensorized tokens can be placed on the sensing board, which will detect the ID of the tokens and communicate this using BLE to the web interface on a laptop. In the web interface, users can find explanations about individual tokens, as well as explanations and combinations for valid combinations between two tokens.

Fig. 1: The original version of the toolkit with the sensing board, sensorized tokens and web interface

3.1 Adding interactivity

Based on the findings of the first studies with the Mix & Match ML toolkit [7], several changes were implemented to better support users during learning and ideation. These changes together resulted in v 1.1.

Compare mode To support learning, we implemented the option to compare two tokens of the same category. Users can use this to for example compare the unlabeled and labeled image data token to learn what it means to have labeled data. It could also support in sensemaking of the different ML capabilities as they can be directly compared.

Previously in the toolkit v1.0, this would result in an error message on the website explaining you made an invalid combination. For version v1.1, we redesigned the top plate of the sensing board and made it two-sided (see Fig. 2a). By turning the plate, the user can switch from *normal* mode to *compare* mode. The orientation of the top plate is detected using an analog hall sensor and a magnet in the top plate.

(a) The toolkit v1.1 in compare mode. This mode is accessed by flipping the top plate and allows users to compare two tokens of the same category, here two data type tokens

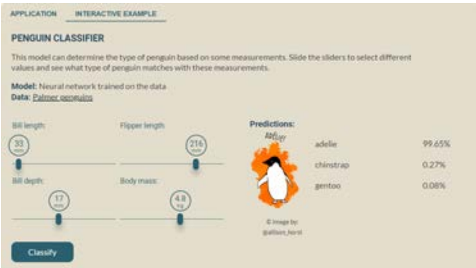
(b) The compare page in the web interface, comparing labeled and unlabeled image data

Fig. 2: Comparing tokens

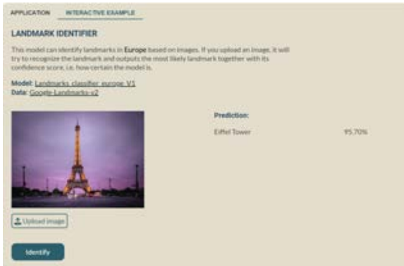
Sequential models During the user studies, participants often created concepts that required sequential models. For instance, they would first recognize from image data what someone was eating and next recommend healthier alternatives. To accommodate this, a *connector token* was created and added to the tokens. With this connector tokens, it becomes possible to link an ML capability token with a next ML capability token (see Fig. 3). The connector token has a whiteboard surface allowing users to write down the in-between output. An extra piece can be added to the connector token to make it labeled data. At this moment, it is not possible to place this chain of tokens on the sensing board.

Fig. 3: The connector piece can be used to create sequential models.

Interactive examples The exemplars featured in the web interface were highly valued for both learning and ideation purposes. To enhance these resources, we introduce *interactive examples* for several combinations. These interactive examples comprise trained ML models that operate within the browser, enabling users to experiment with them by adjusting the inputs and observing the resulting outputs. The inclusion of these interactive examples serves multiple purposes. Firstly, it aids in conveying the nature of the model's output. In the previous study, this was sometimes unclear for users [7]. Secondly, it can help with understanding the capabilities and limitations, as users will be able to experiment with an ML model and experience that the ML model is non-deterministic and can make strange mistakes. Finally, it can increase the engagement as users are able to interact with the web interface while before the example were static.



(a) For the combination of labeled tabular data and categorize, the interactive example is one of a penguin classifier.



(b) For the combination of labeled image data and identify, the interactive is the landmarks classifier for recognizing landmarks in Europe.

Fig. 4: Two instances of the interactive examples included in the Mix & Match ML toolkit

4 Evaluation

To evaluate how the toolkit could be used for educational purposes, a small study was conducted with students of a University of Applied Sciences. These students all were part of a minor about big data and design. During this minor, they followed different modules, where each module focussed on a different aspect of big data and design. The toolkit was evaluated during the first lecture and workgroup (one afternoon) of the module about ML.

4.1 Participants

Participants were informed by email that they would be testing the toolkit during class and that if they would like, they could participate in the study by completing two short surveys during the class. In total 9 students completed both of the surveys (female =4, male=5, average age = 23,4). Four of these participants had no experience with AI and five had read some articles or books. Participants had different study backgrounds including law, business, mechatronics and chemistry, IT and Industrial Engineering.

4.2 Method

The toolkit was evaluated within an educational setting, this would provide more ecological validity compared to a lab study as the students now used it within a real-world environment. Using a before and after survey we asked questions about their previous experience with AI and afterward what they liked and disliked about the toolkit in terms of i) learning about AI/ML, ii) how it supported them during the assignment to reverse engineer a solution, and iii) collaboration.

4.3 Procedure

The class started with a lecture where students were introduced to ML using different examples and explaining the different types of learning (supervised, unsupervised and reinforcement learning) and the importance of training. After this, students shortly experimented with Google's Teachable Machine [5]. Next, the students received their assignment for that specific module - recreating and/or extending the algorithm of an existing platform. With this assignment in mind and the task to come up with questions for a quiz, students could start exploring the Mix & Match ML toolkit. However, before doing so they could fill in the first survey. In total there were three groups and two physical toolkits, therefore the physical toolkit rotated and the remaining group used just the web interface. The total interaction time with the toolkit was 45 minutes. Surveys were administered just before starting with the toolkit and at the end of the class.

4.4 Findings

Ease of use and learning As in the previous study with novices [7], participants appreciated the ease of use. Having received preliminary explanations about basic ML concepts, participants demonstrated a quicker grasp of the toolkit's functionalities. The presence of this basic knowledge could also be observed from their responses when asked about their experience with the toolkit in terms of learning about Machine Learning (ML). Rather than mentioning how the toolkit helped them to understand the ML concepts (*understand* level of the Bloom's taxonomy [11]), they would directly start using the toolkit to *apply* their knowledge by matching it with concepts and their assignment. For instance, Participant 5 noted, *"It was handy to puzzle in between things, think of your subject and try to find what fits your subject."*

Engaging Interactivity Additionally, participants found the toolkit's interactive features to be both enjoyable and stimulating. Some remarked on the tactile nature of the experience, expressing a preference for physical interaction over traditional screen-based interfaces. Participant 2 stated, *"[I liked] The interactive side of it. I like the physical side of it. The NFC stickers blocks are way more engaging than just buttons on a screen."* Moreover, the inclusion of interactive examples, such as image classification, was highlighted by some participants as a feature they liked: *"I liked the image recognition where could upload a picture to try it out."* (P8).

Facilitated Collaboration Furthermore, as in the findings of the earlier study, the toolkit continued to foster group exploration and collaboration [7]. Participants appreciated how it encouraged teamwork and facilitated discussions among team members. Participant 4 remarked, *"You explore it together and find out what works and what does not which helps with the collaboration."* Similarly, Participant 9 highlighted the inclusivity of the toolkit, stating, *"Everybody from the team could interact."*

Limited number of toolkits During the study, a larger-than-anticipated number of students presented scalability challenges. Due to the time-intensive nature of crafting additional sensing boards and tokens, it was not feasible to create more physical toolkits for this study. This necessitated the rotation of physical toolkits among groups. In educational settings, fluctuations in student numbers are common, highlighting the importance to accommodate for this within the toolkit, something that is still challenging with toolkit v1.1.

While the tokens and sensing board support collaboration in groups, having only one small laptop screen for bigger groups does make it more challenging to read the information when groups consists of three or more members. This view was also echoed by Participant 2: *"I would've liked to use the kit in pairs. Then you can control the process more and read everything in a slow pace. In groups of 4 there were 2 people just staring at the kit basically which leads to less concentration"*.

5 The Mix & Match Machine Learning Toolkit 2.0

5.1 Scalability challenge

As the evaluation made clear, the toolkit should improve its scalability before it can be widely adopted in education. Currently, all files and instructions for creating the toolkit are published open source[5]. However, to create one's own Mix & Match ML toolkit one still needs materials, equipment and certain skills. For instance, a laser cutter is required to make the tokens and sensing board and the toolkit requires soldering a PCB. While one might find laser cutters and

[5] https://github.com/MixMatchMLtoolkit

soldering irons at their work facilities or in a local maker space, not everyone will have the access or skills to operate these tools. Moreover, the process of making the toolkit is time-consuming as it requires a great deal of manual effort. As a result, the barrier to developing a single toolkit is already substantial, not to mention the challenges associated with producing sufficient toolkits for an entire cohort of students. To improve the scalability and accessibility of the toolkit, we set out to replace the sensing board and sensorized tokens with alternatives that would be more scalable and accessible to create. As the tangible element of the toolkit is one of its key features, we preserved this and did not opt for a full digital version.

5.2 Design process

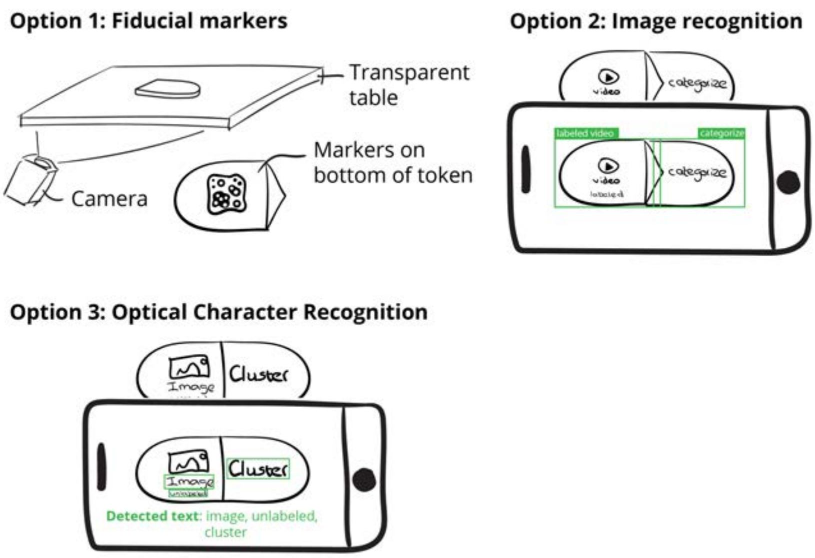


Fig. 5: Three possible techniques for detecting the tangible tokens

Starting off from the existing toolkit, we explored how we could improve the scalability. The main bottleneck is the soldering of the electronics for the sensing board and creating and initializing the sensorized tokens. The creation of these is resource intensive (in terms of time and money) when producing multiple toolkit. To eliminate the soldering of the electronics, it is an option to develop and manufacture a custom PCB that could be sold. While this would help certain educators, the time-saving is minimal as it still requires the creation of the sensing board and sensorized tokens. Moreover, the costs that are paired with this solution make the toolkit less accessible.

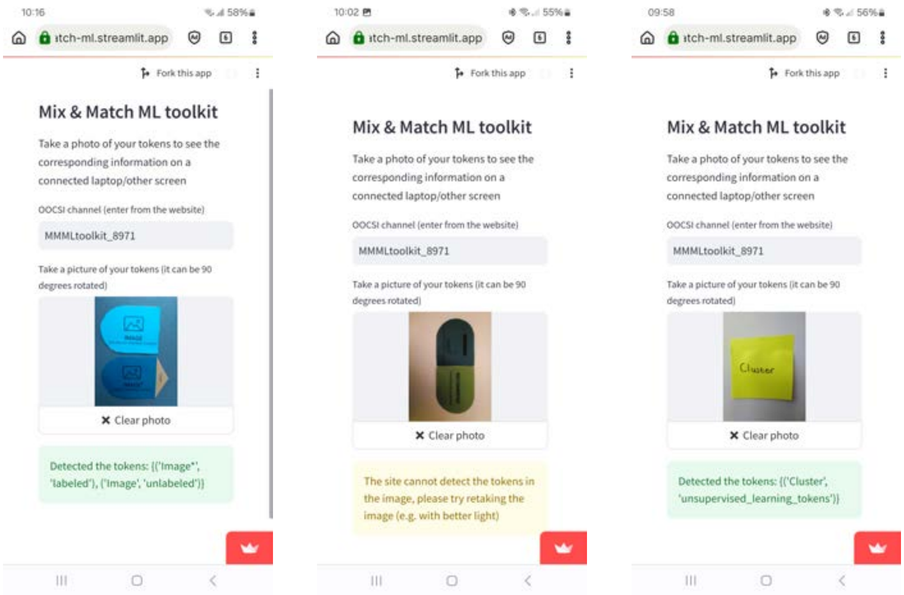
Instead, we considered how we could completely eliminate the sensing board and replace it with a different sensing technology. Several technologies were considered (see Fig. 5). The first uses computer vision to track fiducial markers, for instance using the reacTIVision framework[8]. However, this would still require special equipment such as camera's and software for computer vision. For the best results it is moreover advised to attach the markers to the bottoms of the tokens and place the camera below a transparent table. This is preferred as it leaves the top of the token visible to the user, and users cannot obscure the marker with their hands. Considering these requirements, it becomes clear that this is not a feasible setup within classrooms, and neither is more scalable with low resources.

Looking at what is already available typically in a classroom, we identified the potential of using a mobile application that could scan tokens and send this information to the web interface. Next to eliminating the sensing board, this option also allows the use of printed tokens. Two methods for scanning the tokens were explored. Firstly, by training an image classification model to recognize the different types of tokens. While technical feasible, this would require a large set of training images of all tokens on different backgrounds, in different position and in all possible combinations. With 24 different tokens, this would result in more than 250 combinations. We therefore opted for a second approach, using optical character recognition (OCR) to recognize the text on the tokens. Multiple pre-trained models are available to use and perform well with different fonts and backgrounds. Moreover, OCR models are also able to recognize handwriting (if relatively neat) making it even possible to quickly make your own tokens from post-its.

5.3 Mobile web application

For the mobile application, we tested several pretrained OCR models to determine which one performed best. Several models had trouble with recognizing the text in the tokens and the best and most reliable results were obtained using the python EasyOCR library. EasyOCR uses a multistep framework to recognize the text in the images [14]. First, it extracts useful features form the image using deep learning models. Next, it uses Long Short-Term Memory (LSTM) networks for sequence labeling, enabling the interpretation of the sequential context of the extracted features. The final step is decoding, where the labeled sequences are decoded and transcribed into recognizable text. Using streamlit a website that could run on a mobile phone was created (see Fig. 6). Streamlit allows developers to build websites using Python and Python libraries making it a suitable platform for our application.

Next, the recognized text should be checked against the possible tokens, and if one or more tokens are recognized, the ID of this token should be sent to the web interface. To check for available tokens, the mobile application runs a cross-check to match the found words with the full list of existing tokens. If there is a



(a) The mobile web application recognizing printed tokens of unlabeled and labeled image data

(b) When the text is not recognized, the application will give a warning

(c) Clear handwritten text can also be recognized

Fig. 6: The mobile web application

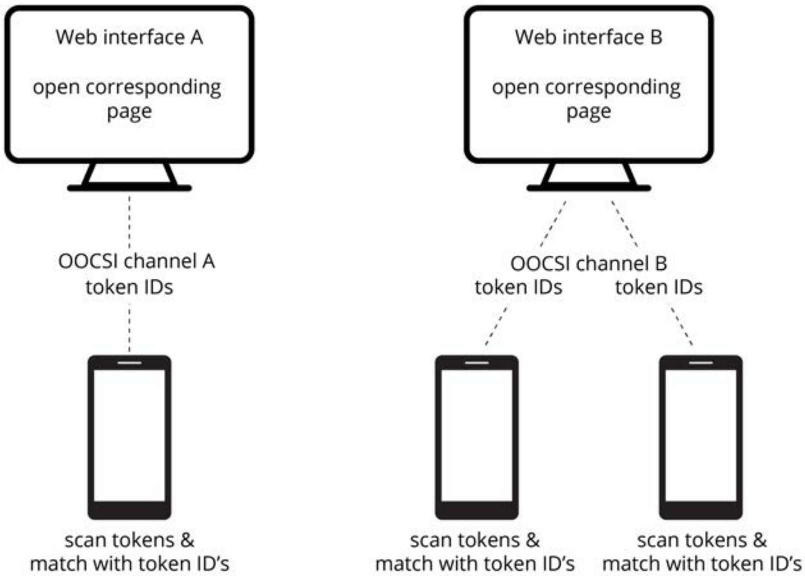


Fig. 7: Diagram of how the mobile web applications are connected to the web interface

match, the ID of that token is added to an array and the search will continue to match more tokens if possible. Once all words are checked, the list of token ID's is ready to be sent. A small change to the tokens was required to ensure proper recognition of the data tokens. Since these tokens exist in two types, labeled and unlabeled data, it was hard for the OCR model to detect if it was a labeled or unlabeled token. The bounding boxes are available for each word which makes it in theory possible to find the word closest to the type of data (e.g. image), this proved to not be reliable. Instead, a small asterisk (*) was added to the end of the labeled data tokens. Using this technique, the model will detect *Image** for the labeled image data tokens and *Image* for the unlabeled image tokens, making it possible to distinguish them easily.

Sending the data from the sensing board to the web interface was done in the v1.0 version using BLE, for the mobile application we chose to send this data over internet using OOCSI [4]. OOCSI is a message-based communication framework enabling cross-platform and cross-language communication among clients. The web interface of the Mix & Match ML toolkit dynamically generates a distinct channel name upon page load - preventing interferences between different users. The interface communicates this channel name to the user when they click on "connect using phone" and starts listening to this channel. Users can scan the QR code displayed on the web interface to launch the mobile app and are prompted to input the OOCSI channel After having completed these steps, the user can take photos in the mobile application and the recognized tokens are matched to their IDs and these are sent to the OOCSI channel. The web interface will receive these IDs and show the corresponding pages.

In scenarios where multiple users collaborate and view the same web interface (remote or in person), each user can participate by connecting to the identical channel and capturing token images. This setup fosters remote collaboration.

5.4 Tokens

By shifting from the sensing board with sensorized tokens to scanning the text on the tokens, new fabrication methods for the tokens are possible. Depending on the fidelity required, tokens can be made using different materials and equipment. We will present four levels of fidelity, but more options are possible (see Fig. 8). Instructions and files to make these can be found at our Github page [6].

Fig. 8: The four different options for creating the tokens.

Handwritten tokens Starting with the lowest level of fidelity, tokens can be made by simply writing the name of the tokens on post-its with a neat handwriting. This option is ideal when the toolkit is used without planning and tokens need to be created on the spot. The disadvantage is that the tokens lose their shape and this shape plays an important role in communicating the constraints of the different types of ML, e.g. supervised learning tokens only fit with labeled data tokens. The interface will stay report that an invalid combination is made, but this cannot be observed directly from the tokens. Moreover, this method depends on the OCR model recognizing handwriting and this is less reliable compared to printed tokens.

[6] https://github.com/MixMatchMLtoolkit/DIY-toolkit

Printed tokens One step more advanced from the handwritten tokens is printing the tokens. These tokens have the same colors and shape as the original tokens. However, since they are flat, the hidden fit for unsupervised learning with labeled data is less clearly communicated. To still show that this is an option, a small dashed triangle is added to the unsupervised tokens to suggest that this combination is possible. With printable templates for tokens, it is easy to scale up and create a large number of tokens for students. Students can be given the task of cutting out the tokens before starting, which would reduce the workload for the educator and students already have a chance to familiarize themselves with the tokens. Depending on how careful the printed tokens are treated, they are easier to lose, tear or get creased.

Basic MDF laser cut tokens If tokens have to be reused frequently, laser cutting tokens out of MDF might be the best option. With the basic version, the entire token is made out of MDF and the names are engraved. The unlabeled data tokens, the supervised and reinforcement learning tokens can all be made in one piece out of 6 mm MDF. The tokens for labeled data and unsupervised learning need to be made in two pieces and glued together, this to ensure that they will fit together. Optionally, these tokens can be painted or cut out of materials with different colors.

Felt + MDF laser cut tokens Finally, for the high-fidelity version, the original tokens can be made minus the RFID tags. This version is made using the base of the basic MDF laser cut tokens, but an additional layer is added of engraved felt. By using different color felt, the tokens will have different colors.

6 Limitations

The evaluation study presented in this study was limited in terms of sample size and data collected on the different aspects of the toolkit. Moreover, it would be even more interesting to evaluate if the toolkit still holds its value for collaboration and learning when replacing the sensing board with the mobile web application and the tokens with printed tokens. The affordance of the tokens decrease and an extra step is required to access the information. A future study could compare the plain web interface with the version with printed tokens and the original toolkit on both learning and ideation goals. In addition, conducting an evaluation with educators to assess the ease of creation and the usability of the toolkit could further enhance its development.

7 Conclusion

In this paper, we set out to update the Mix & Match Machine Learning toolkit to include the changes suggested after previous studies and most importantly, to increase its scalability by designing and implementing an alternative for the sensorized tokens and sensing board. A first update to the toolkit facilitated the comparison of two tokens, the creation of sequential models, and introduced interactive examples into the web interface. This version was evaluated in a small study with students. The results were in line with the previous findings, but highlighted the need for a more scalable toolkit if to be used in an educational context. This scalability was achieved by replacing the sensing board with a mobile web application that can scan different kinds of tokens (handwritten, printed, laser cut) using optical character recognition. The materials and mobile web interface are available for educators and others who are interested under a creative common license at https://github.com/MixMatchMLtoolkit. We hope that by eliminating the resource-intensive step of making the sensing board and sensorized tokens, the Mix & Match ML toolkit becomes more accessible for educators to use in their classrooms.

About Anniek Jansen Ir. Anniek Jansen is a Postgraduate Design Engineer at the Designing Human System Interaction program at Industrial Design (ID), Eindhoven University of Technology (TU/e). She obtained her B.Sc. and M.Sc. degree at ID, TU/e. Anniek Jansen conducts design research at the cross-section of design, AI, and education. During her M.Sc. and EngD program, she acquired technical knowledge and skills to work with AI and combines this with her design and prototyping skills to create interactive prototypes for research. She is particularly interested in developing tangible user interfaces that aid users in understanding data and AI. The work presented in this paper is a follow-up on her Master thesis. After her Engineering Doctorate degree, Anniek will start a PhD on the Responsible democratization of AI at Tilburg University in October 2024.

References

1. Bossy, C.: Scaling up a tangible user interface for education (2013)
2. De Raffaele, C., Smith, S., Gemikonakli, O.: An active tangible user interface framework for teaching and learning artificial intelligence. In: 23rd International Conference on Intelligent User Interfaces. p. 535–546. IUI '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3172944.3172976, https://doi.org/10.1145/3172944.3172976
3. Druga, S., Otero, N., Ko, A.J.: The landscape of teaching resources for ai education. Annual Conference on Innovation and Technology in Computer Science

Education, ITiCSE **1**, 96–102 (7 2022). https://doi.org/10.1145/3502718.3524782, https://dl.acm.org/doi/10.1145/3502718.3524782
4. Funk, M.: Oocsi (May 2019). https://doi.org/10.5281/zenodo.1321220, accessed on 26 May 2021 from: https://doi.org/10.5281/zenodo.1321220
5. Google: Teachable machine, https://teachablemachine.withgoogle.com/
6. Hitron, T., Orlev, Y., Wald, I., Shamir, A., Erel, H., Zuckerman, O.: Can children understand machine learning concepts? the effect of uncovering black boxes. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. p. 1–11. CHI '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300645, https://doi.org/10.1145/3290605.3300645
7. Jansen, A., Colombo, S.: Mix & match machine learning: An ideation toolkit to design machine learning-enabled solutions. In: Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction. TEI '23, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3569009.3572739, https://doi.org/10.1145/3569009.3572739
8. Kaltenbrunner, M., Bencina, R.: reactivision: a computer-vision framework for table-based tangible interaction. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction. p. 69–74. TEI '07, Association for Computing Machinery, New York, NY, USA (2007). https://doi.org/10.1145/1226969.1226983, https://doi.org/10.1145/1226969.1226983
9. Kaspersen, M.H., Bilstrup, K.E.K., Van Mechelen, M., Hjorth, A., Bouvin, N.O., Petersen, M.G.: Votestratesml: A high school learning tool for exploring machine learning and its societal implications. In: FabLearn Europe / MakeEd 2021 - An International Conference on Computing, Design and Making in Education. FabLearn Europe / MakeEd 2021, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3466725.3466728, https://doi.org/10.1145/3466725.3466728
10. Kaspersen, M.H., Bilstrup, K.E.K., Petersen, M.G.: The machine learning machine: A tangible user interface for teaching machine learning. In: Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction. TEI '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3430524.3440638, https://doi.org/10.1145/3430524.3440638
11. Krathwohl, D.R.: A revision of bloom's taxonomy: An overview. Theory into practice **41**(4), 212–218 (2002)
12. Li, Y., Liang, M., Preissing, J., Bachl, N., Dutoit, M.M., Weber, T., Mayer, S., Hussmann, H.: A meta-analysis of tangible learning studies from the tei conference. In: Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction. TEI '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3490149.3501313, https://doi.org/10.1145/3490149.3501313
13. Long, D., Magerko, B.: What is AI Literacy? Competencies and Design Considerations. In: Conference on Human Factors in Computing Systems - Proceedings. pp. 1–16 (2020). https://doi.org/10.1145/3313831.3376727
14. Mahajan, A.: Easyocr: A comprehensive guide (2023), accessed on April 29, 2024 from https://medium.com/@adityamahajan.work/easyocr-a-comprehensive-guide-5ff1cb850168

15. Marshall, P.: Do tangible interfaces enhance learning? In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction. p. 163–170. TEI '07, Association for Computing Machinery, New York, NY, USA (2007). https://doi.org/10.1145/1226969.1227004, https://doi.org/10.1145/1226969.1227004
16. McKinsey: The state of ai in 2023: Generative ai's breakout year (8 2023), https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year#talent
17. Ng, D.T.K., Leung, J.K.L., Chu, S.K.W., Qiao, M.S.: Conceptualizing ai literacy: An exploratory review. Computers and Education: Artificial Intelligence **2**, 100041 (1 2021). https://doi.org/10.1016/J.CAEAI.2021.100041
18. Schneider, B., Jermann, P., Zufferey, G., Dillenbourg, P.: Benefits of a tangible interface for collaborative learning and interaction. IEEE Transactions on Learning Technologies **4**(3), 222–232 (2010)
19. Sulmont, E., Patitsas, E., Cooperstock, J.R.: Can you teach me to machine learn? In: Proceedings of the 50th ACM Technical Symposium on Computer Science Education. p. 948–954. SIGCSE '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3287324.3287392, https://doi.org/10.1145/3287324.3287392
20. Sulmont, E., Patitsas, E., Cooperstock, J.R.: What is hard about teaching machine learning to non-majors? insights from classifying instructors' learning goals. ACM Trans. Comput. Educ. **19**(4) (jul 2019). https://doi.org/10.1145/3336124, https://doi.org/10.1145/3336124
21. Ullmer, B., Ishii, H., Jacob, R.J.K.: Token+constraint systems for tangible interaction with digital information. ACM Trans. Comput.-Hum. Interact. **12**(1), 81–118 (mar 2005). https://doi.org/10.1145/1057237.1057242, https://doi.org/10.1145/1057237.1057242
22. Van Brummelen, J., Heng, T., Tabunshchyk, V.: Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools (2020), http://arxiv.org/abs/2009.05653
23. Van Der Vlist, B., Van De Westelaken, R., Bartneck, C., Hu, J., Ahn, R., Barakova, E., Delbressine, F., Feijs, L.: Teaching machine learning to design students. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **5093 LNCS**, 206–217 (2008). https://doi.org/10.1007/978-3-540-69736-7_23, https://link.springer.com/chapter/10.1007/978-3-540-69736-7_23
24. Zimmermann-Niefield, A.: Machine learning education for young people without programming experience. ICER 2019 - Proceedings of the 2019 ACM Conference on International Computing Education Research pp. 359–360 (2019). https://doi.org/10.1145/3291279.3339446
25. Zimmermann-Niefield, A., Polson, S., Moreno, C., Shapiro, R.B.: Youth making machine learning models for gesture-controlled interactive media. In: Proceedings of the Interaction Design and Children Conference. p. 63–74. IDC '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3392063.3394438, https://doi.org/10.1145/3392063.3394438
26. Zuckerman, O., Gal-Oz, A.: To TUI or not to TUI: Evaluating performance and preference in tangible vs. graphical user interfaces. International Journal of Human Computer Studies **71**(7-8), 803–820 (2013). https://doi.org/10.1016/j.ijhcs.2013.04.003, http://dx.doi.org/10.1016/j.ijhcs.2013.04.003

Qualifying a signal: using AI to improve model prediction

François Leborgne[0009-0000-9639-9519] and
Pierre Cluitmans[0000-0002-3099-4611]

Technical University of Eindhoven, 5600 MB Eindhoven, Netherlands
f.m.g.leborgne@tue.nl, p.j.m.cluitmans@tue.nl

Abstract. Stress prediction using physiological sensors is crucial for effective caregiving, yet real-time monitoring often faces challenges due to data noise and artefacts, impacting model accuracy due to poor data quality. This study utilizes a wearable sensor developed by a Dutch company (Mentech Innovation, Eindhoven) to monitor skin conductance and electrodermal activity (EDA) and detect stress in care settings for individuals with intellectual disabilities and people with dementia. The dataset includes EDA data from participants performing stress-induced and motion tasks. Annotation was conducted by multiple annotators following specific guidelines, with Cohen-kappa assessing inter-rater agreement. Feature selection techniques grounded in physiological knowledge and statistics were employed to identify relevant predictors, such as frequency band contributions and peaks associated with physiological responses. An XGBoost model achieved a balanced accuracy of 94% in distinguishing clean and noisy signal segments of 5 seconds using 5-fold cross-validation. Practical implications involve enhancing stress prediction accuracy by utilizing only clean data for predictive modelling and implementing targeted cleaning methods using wavelet decomposition for noisy segments to preserve valuable information. However, the need for further validation and exploration of model generalizability across diverse populations is still relevant. Future research will focus on refining noise detection algorithms and incorporating additional physiological signals for comprehensive stress assessment.

Keywords: Human Computer Interaction · Stress detection · Electrodermal Activity · Machine learning model · Data validation

1 Introduction

People with intellectual disabilities (ID) and dementia are usually more sensitive to stress, which often leads to problematic behavior due to their difficulty in fully expressing their needs and emotions [1]. The potential implications of these stressful situations can include aggression or even self-injury. A study by Bowring et al. [2] showed that 20% of people with an intellectual disability exhibited challenging behaviors, which ultimately led to a lower quality of life for them, as well as increased stress for caregivers and family members.

Knapp et al. [3] demonstrated that early detection and reporting of stress allow caregivers to respond promptly and appropriately, thereby reducing stress in both caregivers and clients and the risk of escalation. Developing solutions and methods for preventing and recognizing these behaviors, and providing efficient responses in a timely manner, would significantly improve the well-being of residents, their families, and caregivers, and reduce the costs associated with interventions in cases of violent responses [4]. Stress can be described as a state in the autonomic nervous system (ANS), which can be monitored through various physiological signals such as skin conductance and heart rate variability [5] [6]. Monitoring these physiological signals and detecting potential stress moments is challenging due to the ambulatory nature of the monitoring. In laboratory conditions, signals are relatively free from noise, making it easier to gather clean data and develop accurate stress prediction models. However, in real-world conditions, signals often contain noise from sources such as motion artefacts, poor sensor connection, and electrical drift [7]. Traditional time series cleaning techniques may remove valuable information within the signal due to the absence of ground truth, making it difficult to ascertain the desired shape of cleaned data. Extensive research and models have been developed for stress prediction using wearables, as reviewed by Gideon et al. [8].

In collaboration with Mentech Innovation, a novel wearable garment was developed, featuring two electrodes printed onto a sock positioned optimally on the sole of the foot. The sensor's location results in significant motion noise, necessitating a method to separate physiological responses from noise. This study presents the development of a machine learning model - XGBoost - to distinguish clean from noisy signal epochs, achieving a balanced accuracy of 94%. Finally, the study will explore wavelet decomposition to improve signal quality without excessively removing physiological information and discuss methods to quantify this loss.

2 Related Work

2.1 Physiological Signals

The body's stress response is typically triggered by the Peripheral Nervous System, which transmits sensory and motor impulses between the central nervous system and the body. The Autonomic Nervous System (ANS), a part of the Peripheral Nervous System, consists of two main components:

1. The **Sympathetic Nervous System**, often referred to as the 'fight or flight' response, regulates the body's reaction to external stimuli.
2. The **Parasympathetic Nervous System**, known as the 'rest and digest' response, maintains the body's internal balance during rest, feeding, and digestion.

These two branches are crucial in controlling cardiovascular functions like blood pressure, heart rate, and respiration through the activity of noradrenergic and adrenergic neurons. Responsible for producing norepinephrine that regulates arousal, stress responses, mood, attention, and focus, these neurons are crucial for mental and physical well-being [9]. For instance, changes in blood pressure prompt the ANS to adjust the level back to normal (homeostasis principle), which involves the release of hormones (epinephrine and norepinephrine) activating the sweat glands, resulting in sweating. The measure of the response's amplitude- skin conductance response (SRC)- will provide information about sympathetic activity. This provides a non-invasive method to quantify sympathetic nervous system activation [10]. The signal is then recorded by applying a low voltage between two electrodes located in specific places on the body. Boucsein's book 'Electrodermal Activity' [11] identifies these locations as areas with the highest density of sweat glands, such as the tip of the finger, the palm of the hand, or the sole of the foot.

A study by Alberti et al. [12] showed that two main physiological signals are used for real-time stress detection: skin conductance response - called electrodermal activity - (EDA), and heart rate variability (HRV). HRV measures the variation in time between heartbeats, reflecting autonomic nervous system balance. High HRV indicates adaptability, while low HRV suggests stress. On the other hand, EDA response can be separated into two components: the phasic and tonic components [13]. The tonic level, or skin conductance level (SCL), is associated with low frequencies in the EDA spectrum and is supposed to reflect the person's level of attention [11]. The phasic component, associated with high frequency, is linked to specific emotionally arousing stimulus events and occurs one to five seconds after the onset of the emotional stimulus [14]. Both components may vary significantly from person to person depending on hydration, skin dryness, or autonomic regulation [15]. Hence, changes in these physiological parameters provide reliable input for methods that predict emotions in various conditions, including seizure risk and sleep patterns [16].

2.2 Stress Detection in Real-Time Applications

EDA can be measured in different locations on the body, such as the head, shoulders, hand, and foot. However, the signal intensity varies depending on the density of the sweat glands in each area [11] [17]. Current commercially available devices for EDA recording on an ambulatory or long-term basis are typically wrist sensors [18]. However, for people with ID and dementia, this location might not be ideal because of potential discomfort from the target group. Therefore, Mentech Innovation developed a novel wearable garment with two electrodes printed onto a sock on the sole of the foot, resembling a normal commercialized sock. This garment aims to combine user acceptance with good signal quality. In care settings, the detection of stress has to be realised in (pseudo) real-time so that the caregivers can act in a timely manner to prevent any risk of escalation. Developing a high computational model, such as an unsupervised model with many features, might not be suitable due to the computational time for each prediction. Therefore, a balance between time and accuracy needs to be found. Leveraging accessible knowledge about physiological signals (EDA, HRV, ECG, etc.) can help build a model using only suitable features, hence semi-supervised learning. A multi-task learning model was trained and validated in Mentech's lab, achieving a balanced accuracy of 80% in detecting stress.

2.3 Signal Quality

Data quality is another crucial element for increasing the model's stress prediction accuracy in real-time monitoring. Poor data quality inevitably leads to poor machine learning model predictions [19] [20]. Therefore, improving signal quality could enhance model prediction accuracy. Unlike ECG signals, which show stationary and repetitive patterns and are more straightforward to analyze, electrodermal activity is known to be non- or pseudo-stationary and presents high inter-subject differences. Consequently, there is no fully described ground truth regarding what constitutes a perfect response from the signal. Some rules have been described empirically, but no clear golden rule exists. Gaining insight into what constitutes clean and noisy signals (e.g., Gaussian noise, motion noise shape) is valuable for improving overall model prediction. The general physiological rules of a clean signal are listed below.

Good Signal Quality - Common Rules As presented in Boucsein [11], electrodermal activity directly measures sympathetic nervous system activation, reflecting reactions to stress, emotions, and engagement with activities. The main characteristics of EDA are as follows:

- A response to stress typically lasts between 1 to 5 seconds. [11]
- A skin conductance response usually occurs within 1 to 3 seconds with a half recovery period of 2 to 10 seconds. [11]
- The decay after a peak is always exponential in form [11].

In contrast, a noisy signal will show several peaks within a short period (high-frequency term) and sharp drops or increases that are not physiologically acceptable within a fixed amount of time. Those main rules were used for the labelling of the data between clean and noisy windows by internal annotators using a personalized labelling tool presented below.

Labelling Tool To annotate the sessions, an EDA labelling tool was developed for this study inspired by a similar tool developed in [21]. This tool allows users to hover over the session and label every 5-second window of a session as clean, noisy, or unsure. A representation of the tool is presented in Figure 1.

2.4 Signal Processing

Signal processing plays a vital role in analyzing Electrodermal Activity (EDA) signals, particularly in detecting artefacts. EDA signals can be quite noisy, hiding valuable insights into sweat gland activity. To analyse the data and extract

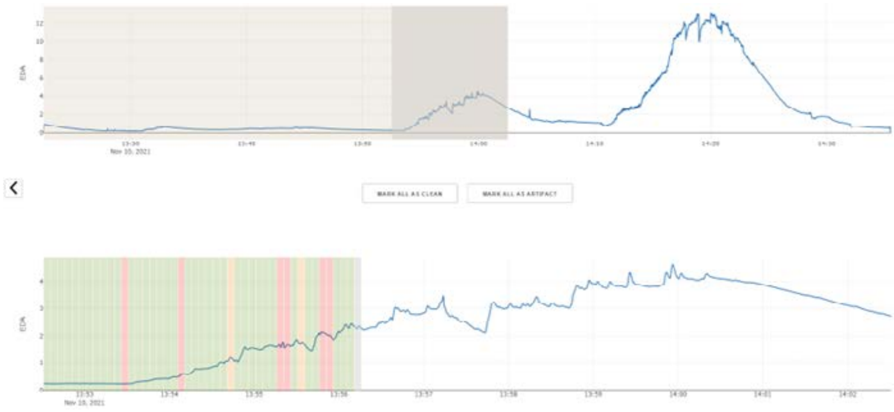


Fig. 1. Representation of the EDA labelling tool.

all information possible, a combination of time-domain and frequency-domain features is utilized to enhance the reliability of the data.

Time-domain analysis examines the signal over time, focusing on key features such as mean, standard deviation, and the number of peaks. This approach aids in identifying sudden changes that may indicate the presence of artefacts or physiological activation.

The frequency-domain analysis involves transforming the signal into the frequency domain using methods like Fourier Transform to examine spectral components. Features such as power spectral density (PSD) provide information on the amount of information located within a specific frequency band (low frequency, for example) and are employed to differentiate between physiological signals and noise. Despite its effectiveness in identifying frequency components, the Fourier Transform lacks time-localized information, making it challenging to detect transient artefacts common in EDA signals.

Wavelet Decomposition Wavelet decomposition seamlessly combines time and frequency analyses. By decomposing the signal into wavelet coefficients at different scales, it captures both temporal and spectral characteristics. This process involves the careful selection of an appropriate wavelet, decomposition of the signal into detail and approximation coefficients by splitting the data sequentially into high and low-frequency bands, analysis of these coefficients to identify and reduce noise, and reconstruction of the signal with minimized artefacts. Wavelet decomposition allows for localized analysis, effectively isolating transient artefacts without impacting the entire signal, thus enhancing the accuracy of EDA-based stress detection models. The choice of mother wavelet is usually made according to the shape of the EDA signal. indeed, it was shown that the mother wavelet that resembles the signal (peaks and small slopes) was giving the best results. However, for this selection, the mother wavelet chosen was producing the minimum value of the mean square error defined as follows:

$$\text{MSE} = \frac{\int |x(t) - \hat{x}(t)|^2 \, dt}{n-1} \tag{1}$$

where x represent the original signal and $\hat{x}$ the reconstructed signal. 15 mother wavelets were compared in total to minimise the MSE. This list is the following:

'haar', 'db1', 'db2', 'db3', 'sym2', 'sym3', 'sym4', 'coif1', 'coif2', 'coif3', 'bior1.1', 'bior1.3', 'bior1.5', 'rbio1.1', 'rbio1.3', 'rbio1.5'.

Integrating time-domain features, frequency-domain features, and wavelet decomposition significantly improves the quality of EDA data, leading to more precise stress detection and advanced psychophysiological research. Furthermore, one critical step in the wavelet decomposition process is thresholding. After decomposing a signal into its wavelet coefficients, thresholding involves setting small coefficients, typically associated with noise, to zero while retaining larger coefficients that represent the true signal features. By applying an appropriate threshold, we can effectively filter out noise and enhance the signal quality. Hard thresholding sets all coefficients below a certain threshold to zero, while soft thresholding also shrinks the remaining coefficients towards zero, further reducing noise. This selective removal of noise components preserves important signal characteristics, making thresholding a crucial step in the wavelet-based signal-cleaning process.

3 Model Development

3.1 Dataset presentation

To train the model, 10 participants were asked to execute various tasks wearing two sock sensors (right and left foot). The experiment aimed to obtain a labelled dataset of physiological signals to distinguish between high and low arousal responses during various conditions. Besides, several annotators labelled the data to analyse further whether a data point could be considered clean or noisy using the labelling tool, as presented in the above-mentioned part. The decomposition of the experiment was as follows. The participant was asked to answer math questions while sitting, standing and walking. Besides, every part was preceded by a baseline condition, where the participants were asked to focus solely on their breathing, and a neutral condition, where the participants were asked to look at coloured bars moving across the screen, which was proved to elicit a neutral nervous state. In total, 6 annotators labelled the 10 sessions. Cohen's kappa inter-agreement coefficient was used to assess the reliability of agreement between pairs of annotators [22] when assigning the categorical ratings. Overall, a value of $\kappa = 91\%$ was found for the training and validating part of the machine learning model. The data with no clear agreement were removed from the training/testing dataset of the model.

3.2 Model creation

Pre-processing & feature extraction Pre-processing is a crucial step in cleaning the data. For time series, every decision should be based on physical or physiological insights. The study of Posada-Quintero et al. showed that the electrodermal activity range was located within 0.005 to 0.4 Hz. Therefore, for this study, a Butterworth band-pass with a cutoff window of [0.005, 0.65] Hz was applied to the signal. As presented in the introduction, the signal was further separated into phasic and tonic components using a median filter window as presented in this method [23]. Because the objective mainly focuses on early stress detection to prevent violent bursts, the phasic component gives valuable detail regarding the nervous system's activity by looking at the amplitude and the amount of peak within the signal. Besides, the signal was decomposed using the wavelet transform using a common wavelet family (Haar wavelet). This wavelet is very efficient for detecting the peak and slope changes, which are related to the nervous system's activation and, therefore, the number of peaks in the signal. Finally, as a typical response lasts between 1 to five (cf. common rules), the data were split into 5-second windows, increasing the data points without losing any information regarding unique activation, potentially improving the quality of the model. The traditional statistical features were extracted (min, max, mean, variance, median, std, value range, means and standard deviations of the first and second derivatives). Number of peaks and amplitude of the peak as well. In total, 57 features were extracted.

Machine learning model Due to potential disconnection with the sensor resulting in incorrect numbers (NaNs) sent into the system, an XGboost model was chosen for the artefact classification because of its capacity to handle them. Besides, previous studies have already shown the ability to use the extreme gradient boosting (XGBoost) model to analyse and detect stressful conditions with physiological data with high accuracy [24].

4 Result

4.1 Artefact detection model

The confusion matrix was computed to analyse the accuracy of the model. This matrix is a compact representation of the performance of a classification model, displaying the counts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions made by the model. It offers a detailed breakdown of the model's accuracy by comparing actual class labels with predicted ones. This matrix allows understanding the model's strengths and weaknesses, identifying common errors (such as false positives and negatives), and calculating performance metrics like accuracy, precision, recall, and F1-score, which are essential for evaluating the model's effectiveness in classification tasks. The accuracy metric measures the proportion of correctly classified instances out of the total instances. The precision measures the proportion of correctly classified positive instances out of all instances classified as positive. The recall metric measures the proportion of correctly classified positive instances out of all actual positive instances. Finally, the F1-score is the harmonic mean of precision and recall, providing a balance between the two. The results are presented in the table below:

Table 1. Model metrics

Metric	**Value**
Accuracy	0.861
Precision	0.826
Recall	0.877
F1 Score	0.836

For the detection of noise in the signal, correct prediction and the detection of positive prediction seem to be the most important. Therefore, improving accuracy and precision should be the most valuable option. To do so, both the feature selection and the Bayesian optimisation were realised using the built-in function of the XGboost library. The new metrics of the model were:

Table 2. Optimised model metrics

Metric	**Value**
Accuracy	0.90
Precision	0.833
Recall	0.823
F1 Score	0.851

After the feature selection and the Bayesian optimisation, the artefact detection model showed a balanced accuracy of 94%. Thus, to qualify the model, a single metric on the overall quality of the signal is given. The ratio of noisy epochs over a specific period of time (1-minute, 2-minute or the whole session) will be chosen. This ratio will be called MR (motion response) in the following steps. Hence, the quality metric EDA_Q is thus:

$$EDA_q = 100(1 - MR) \tag{2}$$

Visual validation In addition to analysing the model's metrics, a visual validation was conducted on new sessions from a different dataset with varying conditions and participants. One of the visualized sessions is depicted in Figure 2 :

Fig. 2. Visual representation of the model predictions. The noisy parts are highlighted in red.

By following the common rules for good signal quality, it is clear to see that the high-frequency parts of the signals were identified as noise, as well as the sharp peaks and drops within the signal.

4.2 Wavelet decomposition

As presented in the section above, the Mean Square error was minimised to determine which wavelet family would fit the data distribution best. Besides the family selection, the initial threshold value was also computed following the research from Donoho et al. [25]. After analysis of the evolution of the EDA_Q with the threshold value - x -, a specific trend appeared in the data. The general equation of the EDA_Q is:

$$EDA_Q = a - (a - b)e^{-cx^d} \tag{3}$$

To simplify this equation, and considering that the data quality cannot be higher than 100, a was set to this value. Besides, the threshold is null, meaning no cleaning has been done, and the parameter b equals EDA_{Q0} the quality of the raw signal. Equation 3 becomes now:

$$EDA_Q = 100 - (100 - EDA_{Q0})e^{-cx^d} \tag{4}$$

Therefore, a Python function was created to compute the quality of the signal two times with random threshold values to determine the correct value c and d for obtaining a sufficient signal quality. The value of $EDA_{Q\infty} = 87\%$ was chosen arbitrarily to balance clean signal and non-modified data.

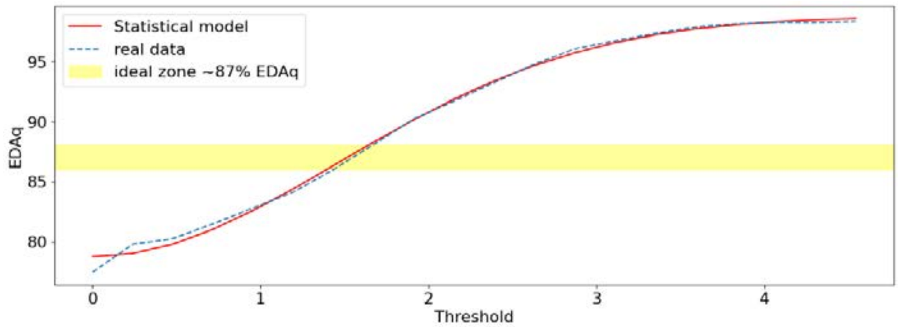


Fig. 3. Example of $f(EDA_Q) = Threshold$

As shown in Fig 3, the fitted model closely follows the real data, showing the robustness of the used distribution. More tests regarding the validity of this distribution will be done in the future.

Wavelet cleaning - Visual validation Using the method presented above, the correct threshold value was selected to get a cleaned data version. The new session is represented in the graph 4 below:

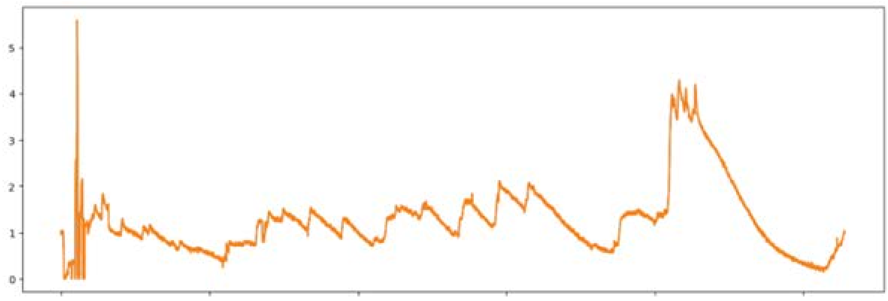


Fig. 4. Cleaned signal - $EDA_Q = 85\%$

5 Discussion

The primary objective of this study was to develop a reliable method for detecting and cleaning artefacts in electrodermal activity (EDA) signals, which are crucial for real-time stress detection. The results demonstrate that the combination of XGBoost for artefact detection and wavelet decomposition for signal cleaning can be considered as a solution for improving the quality of EDA data in a pseudo-real-time condition. The optimized XGBoost model achieved a balanced accuracy of 94%, indicating high precision in distinguishing clean from noisy signals. Additionally, applying wavelet decomposition effectively enhanced signal quality, as evidenced by the substantial reduction in noise. The XGBoost model's precision and recall values suggest that the model can accurately identify artefacts without significantly missing true positives or introducing false positives. This balance is critical in ensuring that subsequent signal analysis is based on reliable data, thereby improving the overall efficacy of stress detection algorithms. Wavelet decomposition proved to be an effective method for cleaning EDA signals. We retained essential signal characteristics while removing noise by minimizing the Mean Square Error (MSE) during wavelet selection and optimizing the threshold value to obtain a good quality in the signal systematically. The empirical equation for $f(EDA) = threshold$ provided a systematic approach to determine the optimal threshold, ensuring a balance between signal cleanliness and data integrity. Visual validation further confirmed the effectiveness of this method, showing clear distinctions between clean and noisy segments of the signal.

The improved artefact detection and signal cleaning methods developed in this study have significant implications for real-time stress detection applications. Accurate and clean EDA signals are essential for reliable stress monitoring, particularly in sensitive populations such as individuals with intellectual disabilities or dementia. The wearable sock sensors designed for this study provided a non-intrusive means of data collection, which is crucial for user acceptance and compliance. The high-quality data obtained through these methods can enhance the performance of stress detection models, potentially leading to better interventions and management strategies for stress-related conditions.

Despite the promising results, this study has several limitations. The sample size was relatively small, with only 10 participants, which may affect the generalizability of the findings. Future studies should include a larger and more diverse participant pool to validate the robustness of the models. Additionally, while the XGBoost and wavelet decomposition methods performed well, other machine learning and signal processing techniques could be explored to enhance artefact detection and signal cleaning further. Finally, the real-time implementation of these methods in wearable devices also requires further investigation to ensure computational efficiency and practicality.

6 Conclusion

In this study, 10 hours of sessions were labelled by 6 different annotators between clean and noisy windows of 5 seconds. Using this dataset, an XGboost model composed of 20 features was trained and validated to detect the different types of epochs. This semi-supervised model showed a balanced accuracy of 94%, showing great capacity in quickly qualifying the signal. Therefore, with the research copay, if the signal quality is already below a certain level (20%), the signal is immediately disregarded because no valid prediction will be given with such poor quality. Besides the data qualification, a novel method using the wavelet decomposition was developed to obtain with 5% inaccuracy a pre-selected final EDA_Q. The company selected a value of 87% to balance having good signal quality and removing a lot of information during the cleaning process. In the future, a series of tests comparing the internal sensor with a medical device will be realised to validate the device's validity and, therefore, validate the EDA artefact model.

About François Leborgne François Leborgne is currently pursuing his Engineering Doctorate in the Human System Interaction program at the Department of Industrial Design, Eindhoven University of Technology (TU/e) in the Netherlands. He is an active member of the MASQUE consortium, focusing on advancing artificial intelligence techniques for stress detection for people with intellectual disabilities and dementia. François holds a Master's degree in Engineering from Les Mines de Nancy in France, specializing in Fluid dynamics. He did a specialization in Biomedical at the University of Alberta, Edmonton. With a background in engineering and industrial design, his research interests lie at the intersection of human-data interaction and XAI, aiming to enhance user understanding and trust in complex health data systems.

References

1. C.G.C. Janssen, C. Schuengel, and J. Stolk. Understanding challenging behaviour in people with severe and profound intellectual disability: a stress-attachment model. *Journal of Intellectual Disability Research*, 46(6):445–453, Sep 2002.
2. D.L. Bowring, V. Totsika, R.P. Hastings, S. Toogood, and G.M. Griffith. Challenging behaviours in adults with an intellectual disability: A total population study and exploration of risk indices. *British Journal of Clinical Psychology*, 56(1):16–32, Mar 2017.
3. M. Knapp et al. Intellectual disability, challenging behaviour and cost in care accommodation: what are the links? *Health and Social Care in the Community*, 13(4):297–306, 2005.
4. R. Lang et al. *Early Signs and Early Behavioral Intervention of Challenging Behavior*, volume 44, pages 1–35. Academic Press, London, 2013.
5. Bradley M. Appelhans and Linda J. Luecken. Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, 10(3):229–240, 2006.
6. I. Ladakis and I. Chouvarda. Overview of biosignal analysis methods for the assessment of stress. *Emerging Science Journal*, 5:233–244, 2021.
7. Andrea Scarciglia, Vincenzo Catrambone, Claudio Bonanno, and Gaetano Valenza. Physiological noise: Definition, estimation, and characterization in complex

biomedical signals. *IEEE Transactions on Biomedical Engineering*, 71(1):45–55, 2024.
8. Gideon Vos, Kelly Trinh, Zoltan Sarnyai, and Mostafa Rahimi Azghadi. Machine learning for stress monitoring from wearable devices: A systematic literature review. 09 2022.
9. Marianne Fillenz. *Noradrenergic neurons*. Cambridge University Press, 1990.
10. M. M. Hossain and P. D. Bradford. *Durability of Smart Electronic Textiles*, pages 27–53. Elsevier Applied Science, 2020.
11. W. Boucsein. *Electrodermal Activity*. Springer US, Boston, MA, 2nd edition, 2012.
12. A. Alberdi, A. Aztiria, and A. Basarab. Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics*, 59:49–75, 2016.
13. H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon. Power spectral density analysis of electrodermal activity for sympathetic function assessment. *Annals of Biomedical Engineering*, 44(10):3124–3135, 2016.
14. F. Hernando-Gallego, D. Luengo, and A. Artes-Rodriguez. Feature extraction of galvanic skin responses by nonnegative sparse deconvolution. *IEEE Journal of Biomedical and Health Informatics*, 22(5):1385–1394, 2018.
15. A. Saeed et al. Model adaptation and personalization for physiological stress detection. In *IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 209–216, 2018.
16. L. Shu et al. A review of emotion recognition using physiological signals. *Sensors (Switzerland)*, 18(7), 2018.
17. Marieke van Dooren, J.J.G. (Gert-Jan) de Vries, and Joris H. Janssen. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology Behavior*, 106(2):298–304, 2012.
18. Jerry Chen, Maysam Abbod, and Jiann-Shing Shieh. Pain and stress detection using wearable sensors and devices—a review. *Sensors*, 21(4), 2021.
19. Lukas Budach, Moritz Feuerpfeil, Nina Ihde, Andrea Nathansen, Nele Noack, Hendrik Patzlaff, Felix Naumann, and Hazar Harmouch. The effects of data quality on machine learning performance, 2022.
20. Oliver Faust, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U Rajendra Acharya. Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, 161:1–13, 2018.
21. Shkurta Gashi, Elena Di Lascio, Bianca Stancu, Vedant Das Swain, Varun Mishra, Martin Gjoreski, and Silvia Santini. Detection of artifacts in ambulatory electrodermal activity data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(2), jun 2020.
22. M. L. McHugh. Interrater reliability: The kappa statistic. *Biochemia medica*, 22(3):276–282, 2012.
23. iMotions Lab - EDA/GSR (Electrodermal Activity). https://imotions.com/products/imotions-lab/modules/eda-gsr-electrodermal-activity/. Accessed: May 24, 2024.
24. Cheng-Ping Hsieh, Yi-Ta Chen, Win-Ken Beh, and An-Yeu Andy Wu. Feature selection framework for xgboost based on electrodermal activity in stress detection. In *2019 IEEE International Workshop on Signal Processing Systems (SiPS)*, pages 330–335, 2019.
25. David L Donoho and Iain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 09 1994.

Improving User Engagement in Sleep Monitoring through Community Involvement

Qingyuan Lin[0000−0003−1262−6214] and Jun Hu[0000−0003−2714−6264]

Department of Industrial Design
Eindhvoen University of Technology
q.lin@tue.nl

Abstract. The Slaap Lekker sleep monitoring device by Bobo Technology is an advanced non-intrusive system designed to monitor and improve sleep quality. This paper outlines a new initiative to enhance user engagement through community involvement. By deploying these devices in community health service centers across China, we aim to improve community health management, elevate health awareness among the elderly, and assist in their social reintegration. The project focuses on building relationships, enhancing health awareness, improving healthcare management, promoting social engagement, and increasing service utilization. This initiative leverages advanced technology to address critical social and health challenges faced by the elderly population. Ultimately, our goal is to reduce healthcare costs and enhance the overall well-being and quality of life for elderly individuals through better sleep health management.

Keywords: Sleep Monitoring · Community Involvement · Elderly Care · Health Awareness · User Engagement · Healthcare Management.

1 Introduction

The sleep monitoring device developed by Bobo Technology, under the brand name Slaap Lekker, is an advanced non-intrusive system designed to monitor and improve sleep quality. This device focuses on remote and continuous monitoring of patients' sleep patterns and physiological data. Key features include heart rate monitoring, respiratory rate tracking, and movement detection during sleep. Utilizing piezoelectric and capacitive sensors embedded in the mattress, the device collects detailed biometric data without disturbing the patient. Our latest initiative aims to enhance user engagement through community involvement by deploying these devices in community health service centers across China.

The Slaap Lekker sleep monitoring device is currently utilized in various hospitals and home care settings. It has garnered positive feedback for its accuracy, non-intrusiveness, and ability to provide valuable health insights. The device's design allows it to be placed under the mattress, where it monitors heart rate, respiratory rate, and movement without disturbing the patient. Next to the initial success in introducing the product to the market, several aspects need further improvements:

- *Low Health Awareness* Many elderly individuals have limited understanding of the importance of sleep health.
- *High Healthcare Costs* Community health management faces challenges due to the high costs associated with healthcare and insufficient service utilization.
- *Social Isolation* Elderly individuals often suffer from social isolation, which negatively impacts their mental and physical well-being.

Our primary objective is to enhance user engagement in sleep monitoring through community involvement. This involves the strategic deployment of Slaap Lekker devices in community health service centers, aiming to improve community health management, elevate health awareness among the elderly, and assist in their social reintegration.

2 Objectives and Approaches

We summarize the objectives and the approaches as follows.

Strengthening Community Involvement

Building Relationships Use the sleep monitoring device to foster stronger connections between the elderly, their families, community health workers, and each other. This approach aims to build a supportive network that encourages active participation in health management.

Community Health Workshops Organize regular workshops and events where community members, including the elderly and their families, can learn about the importance of sleep health and how to use monitoring devices effectively.

Enhancing Health Awareness

Educational Programs Implement educational workshops and seminars to inform the elderly about sleep health and the benefits of using the sleep monitoring device. These programs can also include family members to ensure a supportive environment.

User Training Provide hands-on training sessions to ensure that users are comfortable and proficient in operating the devices. These sessions can be conducted by trained community health workers and volunteers.

Improving Healthcare Management

Integration with Community Centers Equip community health service centers with sleep monitoring devices to offer continuous health monitoring. This integration will enable health workers to track and analyze sleep data, providing timely interventions when necessary.

Data Utilization Utilize the data collected to provide personalized health recommendations and early intervention for sleep-related health issues. Health workers can use this data to create customized care plans for each individual.

Promoting Social Engagement

Community Activities Organize community events and social programs based on data insights to encourage the elderly to participate, thereby reducing social isolation. Activities could include group exercise sessions, social gatherings, and health education classes.

Support Groups Establish support groups where users can share their experiences and support each other in improving their sleep health. These groups can meet regularly and provide a platform for social interaction and mutual support.

Increasing Accessibility and Reducing cost

Accessibility Ensure that the devices and related services are easily accessible to the elderly population. This could involve setting up dedicated support lines and service points within the community centers.

Cost Reduction Implement strategies to reduce healthcare costs through preventive measures and efficient health management facilitated by monitoring devices. This can include subsidizing the cost of devices for low-income families and providing financial incentives for regular health check-ups.

3 Expected Outcomes

The expected outcomes of this project are:

Improved Sleep Health Awareness Educate a larger segment of the elderly population about the significance of sleep health.

Enhanced Healthcare Efficiency Reduce healthcare costs and improve service utilization through proactive health management.

Strengthened Community Ties Foster a sense of community and social support among the elderly, improving their quality of life.

Increased User Engagement Achieve higher levels of engagement and compliance with sleep monitoring practices among users.

4 Conclusion

The Slaap Lekker sleep monitoring device by Bobo Technology is an advanced non-intrusive system designed to monitor and improve sleep quality. This paper outlines a new initiative to enhance user engagement through community involvement. By deploying these devices in community health service centers across China, we aim to improve health management, elevate health awareness among the elderly, and assist in their social reintegration. The project focuses on building relationships, enhancing health awareness, improving healthcare management, promoting social engagement, and increasing service utilization. This initiative leverages advanced technology to address critical social and health challenges faced by the elderly population, ultimately aiming to reduce healthcare costs and improve overall well-being.

About Qingyuan Lin Qingyuan Lin is an EngD trainee specializing in Human-System Interaction at the Department of Industrial Design, Eindhoven University of Technology (TU/e). She holds a Master of Fine Arts (MFA) focused on User Interface and User Experience, and a Bachelor of Finance. Qingyuan has extensive experience as a User Interface Designer, notably contributing to the redesign of a popular weather app. Her skills extend to product management, where she has led efforts in product prototyping and iterative design, enhancing user experiences across digital platforms.

Her academic pursuits include an internship at Zhejiang University's College of Computer Science and Technology, where she contributed to research, resulting in publications in a respected journal and presentations at international conferences. Qingyuan has worked on significant projects, such as the design of sensor systems and the development of emotionally engaging vehicle product designs. She has also led initiatives promoting design exchange and innovation among students. Her achievements have been recognized with awards at Milan Design Week 2022 and the BIEAF2021 Busan International Environmental Art Festival. She also received a First-class scholarship for her academic excellence.

References

1. B. Yu, B. Zhang, L. Xu, P. Fang, and J. Hu, “Automatic Detection of Atrial Fibrillation from Ballistocardiogram (BCG) Using Wavelet Features and Machine Learning,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 4322-4325.
2. R. van Dijk, W. Liang, B. Zhang, and J. Hu, “Development and Evaluation of a Non-obtrusive Patient Monitoring System with Smart Patient Beds,” Distributed, Ambient and Pervasive Interactions, Series, 10291, N. Streitz and P. Markopoulos, eds., pp. 482-490: Springer, 2017.
3. X. Wang, B. Zhang, and J. Hu, Piezoelectric Film Structure, ZL 2017 20 601896.8, 2018.
4. X. Wang, B. Zhang, and J. Hu, Patient Monitoring Bed, ZL 2017 2 0602447.5, 2018.

Exploring Artistic Data Visualization Design for Health Monitoring: A Survey Study

Tianqin Lu[1][0000−0001−8033−4024] and Jun Hu[1][0000−0003−2714−6264]

Department of Industrial Design, Eindhoven University of Technology
{t.lu1,j.hu}@tue.nl

Abstract. As the global population ages, there is an increasing need for innovative approaches to support healthy aging and independent living. Remote measurement technologies (RMTs) offer promising solutions but face challenges in user engagement and sustained utilization. Addressing these challenges requires effective visualization of health data to promote comprehension and sustained engagement. In this study, we proposed an artistic approach to health data visualization and conducted a survey involving 8 participants (3 individuals and 5 informal caregivers) to explore their preferences and needs in this regard. Two preliminary artistic visualization concepts were presented, eliciting feedback to inform design considerations. Findings underscore the importance of prioritizing aesthetics, ensuring comprehensibility, balancing information presentation, and seamlessly integrating visualizations into users' daily routines. Based on these findings, we present four design considerations to guide future research efforts and contribute to the development of artistic visualizations supporting health monitoring.

Keywords: Health monitoring · Data visualization · Art.

1 Introduction

The global population is aging rapidly, with projections indicating that by 2050, the number of individuals aged 65 or older will reach 1.5 billion [1]. This demographic shift poses challenges for the healthcare industry, leading to a rise in healthcare utilization and associated costs [2, 3]. In response, promoting self-management and providing elderly individuals with the necessary tools and support to live independently becomes crucial. Informal caregivers, typically family members or close friends, play a vital role in supporting self-management efforts by providing emotional support, monitoring health data, and assisting with daily activities [4].

In recent years, there has been a growing focus on technology as a means of supporting healthy and independent aging, with remote measurement technologies (RMTs) emerging as a promising solution [5, 6]. RMTs assist in delivering or facilitating support for both individuals and caregivers, offering a range of benefits in healthcare management, including various (mobile) health devices and applications that allow for health data monitoring [6]. Despite their benefits, challenges persist regarding user engagement and long-term utilization. Studies indicate that approximately one-third of devices are abandoned within six to twelve months of use [7]. Specifically, research shows that 50% of Fitbit users abandon the device within the first two weeks of use [8], and around two-thirds of users of RMTs quit using within several weeks to six months of purchase [9].

Moreover, making sense of health monitoring data retrieved from RMTs creates demands for turning data points and trends into actionable insights [10]. Data visualization emerges as a crucial tool in this process, intuitively communicating complex information by harnessing the human visual capacity to perceive differences in size, color, and spatial position [11, ?]. Such visual representation not only enhances understanding but also empowers users to make informed decisions about health and well-being. However, current visualizations of health data in common health-tracking applications often rely on conventional forms such as bar charts, pie charts, scatterplots, and tables [13]. These visualizations mainly focus on triggering a single intended behavior, such as physical activity, rather than promoting review and understanding of the health data itself [14]. Additionally, certain types of conventional visualization, such as pie charts and area-based graphs, can be confusing and require significant time to interpret accurately [15]. This poses challenges, particularly for users lacking in health literacy, who may struggle to understand such visualizations effectively [16]

Recognizing the opportunity to develop visualizations that not only effectively convey information but also foster sustained engagement with health data over time, we propose an artistic approach to visualizing health data. A survey study was conducted involving 8 participants (comprising 3 individuals and 5 informal caregivers), during which we presented two preliminary artistic visualization ideas. The aim of this survey was to investigate the needs and preferences of individuals and informal caregivers regarding health data visualization in the current context of RMTs and to collect feedback on the proposed artistic approach. The findings contribute to identifying four key design considerations for informing subsequent research endeavors and providing insights for the development of artistic visualizations that support health monitoring.

2 Methods

2.1 Inclusion Criteria

This study comprised two participant groups: individuals and informal caregivers. Eligible participants were adults aged 18 years or older. For individuals, inclusion criteria were current or previous use of at least one of RMTs for health data tracking. For informal caregivers, inclusion criteria included: (1) being a caregiver for a “dependent person” or “care recipient,” providing unpaid care, and (2) current or previous use of at least one of the RMTs for tracking the health data of the person under their care.

2.2 Sampling and Recruitment

A purposive convenience-based sampling was used to recruit participants who met the aforementioned criteria. Two approaches were used for recruitment: individuals were recruited from the researcher's network of contacts, while informal caregivers were recruited through Hangzhou Bobo Technology Limited Company (referred to as Hangzhou Bobo). Hangzhou Bobo is a company specializing in intelligent healthcare solutions, utilizing mobile internet technology to address issues such as chronic disease management. Hangzhou Bobo appointed a coordinator who contacted potential participants from customers of Slaap Lekker[1]. Slaap Lekker is a smart sleep and health monitor that monitors health data from the user's vital signs during sleep. Those who provided voluntary informed consent were formally enrolled in the study. This study received ethical approval from the Ethics team of the Eindhoven University of Technology (TU/e). Participation in the study was voluntary.

2.3 Preliminary Design

Two preliminary visualization ideas were created for stimulating discussion and eliciting feedback. Design A utilized pixelated artwork, where the whole image represented the health status through the gradual greying of pixels as the health status deteriorated. The complete pixel artwork indicated optimal health, while the increasing number of gray pixels represented a decline in health (Fig. 1).

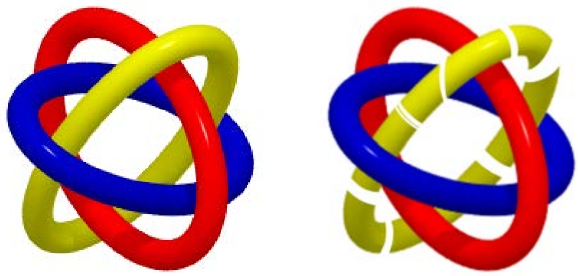
(a) (b)

Fig. 1: Design A: (a) Complete pixel artwork depicting optimal health status; (b) Pixel artwork with added gray pixels indicating deteriorating health. The artwork is based on Van Gogh's "Starry Night".

Design B utilized a more targeted approach by incorporating shapes and colors to visualize specific health indicators. Borromean rings (Fig. 2) were used as a metaphor for overall health. Each ring represents a different health indicator, such as heart rate, blood oxygen level, and breathing rate. When all the

[1] https://www.slaaplekker.cn/

indicators fell within the healthy range, the rings were fully interlocked. If any index exceeded the healthy range, the corresponding ring became disconnected, indicating a suboptimal or bad health status.



(a) (b)

Fig. 2: Design B: (a) Complete Borromean Rings with all health indices within healthy ranges; (b) Broken Borromean Rings with the yellow ring indicating a health index outside the healthy range.

2.4 Survey Study

A survey was developed to collect qualitative data from individuals and informal caregivers. The survey questions were essentially the same for both groups, with minor adjustments to accommodate their different contexts. The data collection methods varied based on the group. For individuals, the survey was administered through face-to-face interviews (in English) conducted by the first author, who served as the interviewer. For informal caregivers, a paper-based survey was provided in Chinese, allowing them to answer in their native language. The coordinator delivered and collected the surveys from informal caregivers and then sent completed surveys back to the researcher as digital scans.

Content of the survey. The survey was structured into three sections: (1) user experience with RMT(s), (2) user experience with health data visualization and suggestions on improvement, and (3) collection of insights into the initial visualization design ideas. Demographic information was collected from participants at the beginning of the survey.

The first section consisted of six questions, covering the duration of use of RMT(s), frequency of examining health data through RMT(s), potential changes in user behavior over time, perception of the health data, specific information sought during the data review, and whether using RMT(s) helped support the health tracking/management.

The second section comprised nine questions that began with asking about participants' experience with health data visualization tools. It further delved into challenges encountered in tracking health data, the impact of visualization in this process, and how it compared to the guidance or support provided by formal healthcare providers. Participants were also asked about their preferences or requirements for health data visualization, including desired features or functionality.

In the final section, participants were shown the two preliminary design ideas, with an explanation of the difference between conventional form data visualization and unconventional form data visualization. They were then asked to share their thoughts on the design ideas. All terminology mentioned in the survey was clearly explained to participants before or during the survey.

3 Results

Sample Characteristics A total of 8 participants attended this study, including 3 individuals (I1 to I3) and 5 informal caregivers (IC1 to IC5). Their characteristics are summarized in Table 1.

Table 1: Demographic characteristics

Characteristic	**Individuals (n=3)**	**Informal caregivers (n=5)**
Age (years), n (%)		
≤ 30	3 (100)	2 (40)
31 - 40		1 (20)
41 - 50		1 (20)
> 50		1 (20)
Gender		
Male: Female	2:1	2:3
Residence, n (%)		
the Netherlands	2 (66.7)	
China	1 (33.3)	5 (100)
Highest education completed, n (%)		
High school		2 (40)
Bachelor's degree		3 (60)
Graduate degree	3 (100)	
Duration of using RMTs (years), n (%)		
≤ 1	1 (33.3)	4 (80)
1 - 3	2 (66.7)	1 (20)
Duration of caregiving (years), n (%)	Not applicable	
≤ 1		4 (80)
1 - 3		1 (20)
Type of caregiving	Not applicable	
Parent		5 (100)

User Experience with RMTs

Impact on health management. There was a general view among the participants that RMTs helped them promote self-awareness and support health management by providing them with convenient and timely health data updating. Moreover, informal caregivers reported that RMTs enabled more effective communication with their care recipients since information can be transmitted at any time, which further improved their relationship and emotional well-being. However, participants also expressed an expectation to see more information beyond the data itself, such as information on the overall health condition, which is currently lacking in existing RMTs. They also suggested that RMTs should be more personalized and interactive, and provide more feedback and guidance on how to improve their health outcomes.

Changes in usage over time. Most participants reported reduced frequency in checking their own health data or that of their care recipients compared to when they initially began using RMTs. This could be attributed to reduced interest, as they did not find the data to be valuable in providing useful insights, as one individual stated: "I don't understand what they mean [...] I lost my interest to viewing them without not getting much valuable information." [I3]

Another factor that contributed to reduced checking frequency related to the informal caregivers was an improved understanding of the data. Some of them who gained a better understanding of the data during the usage period reported a reduced concern for minor fluctuations in the health data over time: "I learned the corresponding explanation of the data later, I would not be too entangled in the small changes in the data." [IC1]

However, for some individuals, the frequency of using RMTs remained relatively stable over time. Those who relied on RMTs to track their fitness data reported that their continued use was driven by the need to track their physical activities.

Emotional impact. Several participants expressed concerns about the reliability of RMTs, leading to a lack of trust in the health data they provide. Apart from the doubt about the technologies, the lack of clarity between data and user performance was also a major contributor to this concern. One individual reported feeling anxious while using RMTs: "Sometimes it even caused me anxiety since some indexes can be abnormal all the time, which makes me very worried [...]" [I2] These negative emotional experiences can have a significant impact on the user's mental health. Therefore, it is necessary to consider the emotional impact of RMTs and design them to deliver information in a more user-friendly manner.

Health Data

Dissatisfaction with raw data and preference for information. While most participants acknowledged the importance of health data, they expressed their dissatisfaction and uncertainty about raw health data. This was especially reported by informal caregiver participants, who lacked professional training and familiarity with basic health terminologies and analysis. Due to the high demand for health-related knowledge in caregiving tasks, some of them have even resorted to seeking interpretation and meaning of the data online. They, therefore, have expressed a desire for a clear picture of how good or bad the health condition was along with some analysis and explanation, which matches the user experience mentioned earlier with RMTs. Moreover, healthy individuals among the participants have expressed strong disinterest in knowing about specific health data. Instead, they have expressed a preference for general information regarding their health and the option to check specific data only when they want to. As one individual noted: " [...] most of the time it can be a general information but if I want to know the specific data I can also find it somewhere." [I2]

Desire for actionable insights. Participants expressed a desire for specific and understandable recommendations on how to use health data to make better health management decisions. They expressed concern about the actions or steps they should take based on the health information they received.

Health Data Visualization

Meaningful visualization of health data. Participants reported that while visualizing health data in a different format (e.g., graphs) seems straightforward, it may not be helpful. Different formats did not add any value if the user did not understand the meaning behind the data. Therefore, participants expressed a desire to see an overview of their overall health status rather than numerous bar charts showing different health data. This opinion was evident in their responses to the question about the missing features in the current health tracking devices and visualization tools.

Simplicity, aesthetics, and balance. Participants emphasized the importance of simplicity in health data visualization as it could contribute to better comprehension. Videos and animations were highly mentioned as potential approaches to visualize health data more engagingly and interactively. While most participants did not provide specific suggestions for making health data more userfriendly, many expressed a desire for visualizations to be beautiful, suggesting that aesthetics play an important role in their preferences for health data visualization. Interestingly, one informal caregiver specifically mentioned the wish for "a combination of digital and abstract, pictorial representations" [IC5], which aligns with the idea of providing a general overview of health status along with specific data as needed. This highlights the importance of striking a balance between the big picture and detailed information.

Insights into the Preliminary Design Ideas Participants had mixed opinions on Design A and Design B. Design A was praised for its novelty and creativity, but some participants expressed concerns about the clarity and effectiveness of using pixel art to represent health data. For instance, one participant stated that said that the pixel artwork might be too blurry or low-resolution to convey health information. Another participant questioned whether the pixel variations would be noticeable enough to draw the user's attention. Design B was appreciated for its simplicity and straightforwardness ins using simple shapes and colors to convey the health data effectively. However, it faced criticism regarding its visual appeal, as it was perceived as lacking in aesthetics. Moreover, one participant was puzzled about how the three health data were selected over other health indexes and presented in three rings.

4 Findings

The results highlight the need to refine data visualization used in RMTs. While acknowledging the positive aspects of RMTs, the findings reveal shortcomings that require further research.

Current data visualizations in RMTs, such as bar and line graphs, are perceived as lacking meaningful insights by participants, leading to a decline in interest in health data over time. This indicates a clear gap between the value of health data and users' perceived usefulness, where conventional visualizations fail to support users' ability to derive meaningful insights from the data.

Informal caregivers lacking professional training and familiarity with basic health terminologies and analysis expressed incomprehension and uncertainty when faced with raw health data. They sought interpretations, analyses, and explanations accompanying the health data. On the other hand, healthy individuals expressed disinterest in specific health data and preferred general health information as it related to their overall health outcomes and well-being. These findings suggest a need for a flexible design approach that can address the diverse requirements and accommodate the varying needs of different user groups.

Additionally, participants emphasized the importance of aesthetics and suggested the use of dynamic presentations, to engage users. This coincides with the growing recognition of the critical role that aesthetics play in engaging users and facilitating meaningful interactions with health data (Sutcliffe, 2009). However, it is important to strike a balance between artistic expression and conveying meaningful information. A well-executed visualization should harmonize visual appeal with the ability to communicate health knowledge in an understandable manner.

Moreover, the emotional impact caused by RMTs was a significant topic among participants, as concerns about the reliability of the provided health data led to a lack of trust, or the overwhelming focus on health data sometimes led to negative emotional experiences, including feelings of anxiety and annoyance. Such negative emotional responses can potentially diminish one's motivation to seek health information and, therefore, lead users to disengage from RMTs during periods of disease relapse or noticeable progression (Lee et al., 2008). Thus, addressing these emotional concerns and designing RMTs to deliver information in a pleasing and user-friendly manner is important to promote positive emotional experiences.

5 Design Considerations

Based on the findings, we have established four design considerations:

C1 Emphasize Aesthetics: The visualization should prioritize the incorporation of visually appealing elements that captivate users and deliver health information in a userfriendly manner to promote positive emotional experiences. Consider using color, shape, and other artistic elements to evoke interest and convey health information effectively.

C2 Comprehensibility: The visualization should strive to convey health information in an understandable manner. The visualization should not sacrifice clarity for aesthetics, ensuring that users can easily interpret the presented visuals. Avoid the use of complex metaphors.

C3 Balance of information: The visualization should balance general information with specific details, ensuring that users can quickly grasp the key insights while also allowing them to dive deeper into more specific details if needed. Provide visualizations that offer a comprehensive overview of users' overall health status, addressing the preference for collective health data rather than relying solely on conventional forms like bar or line graphs.

C4 Seamless integration into daily routines: The visualization should be designed in a way that seamlessly integrates with users' routines. This can be achieved by considering factors such as simplicity, convenience, and compatibility with existing technologies or devices commonly used in daily activities. By integrating visualizations seamlessly, users can effortlessly access and engage with health information without disruptions, ensuring a harmonious coexistence with their regular routines.

6 Conclusion

This study reveals the need for innovative data visualization strategies for health tracking/management purposes. Although precise design opinions for artistic and abstract data visualization were not extracted in this study, common aspects emerged, including providing a comprehensive overview of health information, addressing emotional implications, incorporating aesthetics, and ensuring consideration of the diverse needs of different stakeholders. The proposed design considerations provide valuable insights for the development of future artistic visualizations in the realm of health monitoring and self-management. By emphasizing aesthetics, comprehensibility, information balance, and seamless integration, designers can create visualizations that not only captivate users but also empower them to make informed decisions about their health and well-being.

About Tianqin Lu Tianqin Lu is pursuing her Engineering Doctorate in the Human System Interaction program at the Department of Industrial Design, Eindhoven University of Technology (TU/e) in the Netherlands. She is a member of the STRAP (Self Tracking for Prevention and Diagnosis of Heart Disease) consortium. She holds a Bachelor of Science degree in Mechatronics Engineering with a concentration in Mechatronics, Robotics, and Automation Engineering. Additionally, she obtained a Master of Science in Industrial Design with a specialization in Constructive Design Research. Her recent research interests include Human-Data Interaction and Health Data Visualization.

References

1. United Nations Department of Economic and Social Affairs, P. D. World Population Ageing 2020: Highlights: Living Arrangements of Older Persons. (2021) https://tinyurl.com/4z2j8m5m, last accessed 2023/03/21
2. de Meijer, C., Wouterse, B., Polder, J., & Koopmanschap, M. The effect of population aging on health expenditure growth: a critical review. Eur J Ageing, **10**(4), pp. 353–361. (2013) https://doi.org/10.1007/s10433-013-0280-x
3. Paul, S., Riffat, M., Yasir, A., Mahim, M., Sharnali, B., Naheen, I. T., Rahman, A., & Kulkarni, A. Industry 4.0 Applications for Medical/Healthcare Services. Journal of Sensor and Actuator Networks, 10, 43. (2021) https://doi.org/10.3390/jsan10030043
4. Roth, D. L., Fredman, L., & Haley, W. E. Informal Caregiving and Its Impact on Health: A Reappraisal From Population-Based Studies. The Gerontologist, **55**(2), 309319. (2015) https://doi.org/10.1093/geront/gnu177
5. Chang, F., and Östlund, B. Perspectives of older adults and informal caregivers on information Visualization for Smart home Monitoring Systems: A Critical review. In Advances in intelligent systems and computing pp. 681-–690. (2018)
6. Polhemus, A., Novak, J., Majid, S., Simblett, S., Morris, D., Bruce, S., Burke, P., Dockendorf, M. F., Temesi, G., & Wykes, T. Data Visualization for Chronic Neurological and Mental Health Condition Self-management: Systematic Review of User Perspectives [Review]. JMIR Ment Health, **9**(4), e25249. (2022) https://doi.org/10.2196/25249
7. Ledger, D., and McCaffrey, D. Inside wearables: How the science of human behavior change offers the secret to long-term engagement. Endeavour Partners, LLC 93, 1, pp. 36-–45. (2014)
8. Shih, P. C., Han, K., Poole, E. S., Rosson, M. B., & Carroll, J. M. Use and adoption challenges of wearable activity trackers. IConference 2015 proceedings. (2015)
9. Hendker, A., Jetzke, M., Eils, E., & Voelcker-Rehage, C. The Implication of Wearables and the Factors Affecting Their Usage among Recreationally Active People.

International journal of environmental research and public health, 17(22), 8532. (2020) https://doi.org/10.3390/ijerph17228532
10. Faisal S, Blandford A, Potts HW. Making sense of personal health information: challenges for information visualization. Health Informatics J. textbf19(3), pp. 198–217. (2013)
11. Grossman, L. V., Feiner, S. K., Mitchell, E. G., & Masterson Creber, R. M. Leveraging Patient-Reported Outcomes Using Data Visualization. Applied clinical informatics, **9**(3), pp. 565-–575. (2018) https://doi.org/10.1055/s-0038-1667041
12. Card M. Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann; Burlington, MA, USA (1999).
13. Kim, S. H. A systematic review on visualizations for self-generated health data for daily activities. International Journal of Environmental Research and Public Health, **19**(18), 11166. (2022)
14. Meyer, J., Kazakova, A., Büsing, M., & Boll, S.C. Visualization of Complex Health Data on Mobile Devices. Proceedings of the 2016 ACM Workshop on Multimedia for Personal Health and Health Care. (2016)
15. Laubheimer, P. Dashboards: Making Charts and Graphs Easier to Understand. https://www.nngroup.com/articles/dashboards-preattentive/, last accessed 2024/04/28
16. Visvalingam, A., Dhillon, J. S., Gunasekaran, S. S., & Hoe, A. C. K. Self-reflective visualizations of patient-centered health information systems. In Recent Trends in Data Science and Soft Computing: Proceedings of the 3rd International Conference of Reliable Information and Communication Technology (IRICT 2018), pp. 711–725. Springer International Publishing. (2019)

Towards a Gamified Intervention Platform for Patients with Atrial Fibrillation

Amalia Mantziari Zafeiri

Eindhoven University of Technology, The Netherlands
a.mantziari.zafeiri@tue.nl

Abstract. Lifestyle behaviors including physical activity, nutrition, sleep, and management of mental stress are significant factors for the clinical course and overall treatment results of chronic cardiac diseases such as Coronary Artery Disease and Atrial Fibrillation. Even though there is evidence that structured lifestyle interventions and Cardiac Rehabilitation can improve the quality of life for these patients it often fails to be implemented efficiently. The reasons behind that can be patient-related (i.e., transportation problems, working obligations, not appealing for specific subgroups) or/and healthcare-related (i.e. low referral rates, limited facilities, and resources). Consequently, there is a need for personalized lifestyle interventions for these patients which can be implemented easily, while avoiding burdening the healthcare system. The proposed solution is the development of a structured lifestyle intervention platform that will combine the holistic gamified lifestyle interventions of an existing app integrated with a recently developed continuous lifestyle monitoring system (digital platform and dashboard). The project is ongoing, and the project team has started working on different work packages defining different functional and technical requirements, involving the professional experts and developers of the consortium partners.

Keywords: Cardiac Rehabilitation · Lifestyle factors · Cardiac telerehabilitation.

1 Background

Cardiovascular diseases (CVDs) have the highest mortality rates worldwide, making them the leading cause of death globally. The factors affecting CVDs include genetic predispositions which are non-modifiable and modifiable factors, such as lifestyle behaviors like nutrition, physical activity, sleep quality and smoking [16].

Cardiac rehabilitation (CR) has been proven to be an effective and safe way to reduce the mortality rates and improve the life quality of CVD patients [4, 12]. In order for the CR to be successful, lifestyle behavior change, and monitoring are needed [13].

Lifestyle behaviors and psychological well-being are considered essential elements of cardiac rehabilitation, as dietary habits [19], physical fitness [20], daily physical activity levels, mental stress [19], and sleep quality [8, 22] are closely linked to the development, clinical course and overall treatment outcomes of common cardiovascular diseases (CVDs), such as coronary artery disease (CAD), and cardiac arrhythmias (irregular heartbeat), such as atrial fibrillation (AF) [1].

Furthermore, the quality of life often does not improve after major cardiac interventions [10, 11, 18], which may be due to the negative effects of ongoing unhealthy lifestyle behaviors on the clinical course of coronary artery disease and atrial fibrillation [10]).

Despite the beneficial effects of participation in CR [4, 12, 5, 21] the participation and completion rates in the CR programs are low [9, 12, 17]. The reasons behind this can be patient-related and consist of work-related responsibilities and transportation issues [3, 15], and the healthcare-related reasons can be due to low referral rate, lack of facilities and resources [12, 15].

Consequently, there is a need for personalized lifestyle interventions for these patients which can be implemented easily, while avoiding burdening the healthcare system.

2 The project

The project is a collaboration between two companies Greenhabit and MiBida [7, 14], Eindhoven University of Technology and Máxima Medical Center.

The proposed solution is the development and evaluation of a structured lifestyle intervention that will create awareness and provide accurate and objective feedback in patients with atrial fibrillation (AF) who are selected or underwent elective radio frequent catheter ablation (RFCA). This can be achieved by the integration of the existing holistic gamified lifestyle interventions of the Greenhabit app [7] integrated with the recently developed continuous lifestyle monitoring system of Care-On (digital platform and dashboard) [6].

2.1 Hypothesis

The general hypothesis of this project is that the integration of an existing gamification intervention combined with a newly developed platform for continuous monitoring before or after elective radiofrequency ablation (RFCA) will result in significant improvements in: i) lifestyle behavior (physical activity, nutrition behavior, mental stress and sleeping habits), ii) traditional risk factors (blood pressure, LDL cholesterol, glucose), quality of life, self-management, usability and adherence.

2.2 Study population and clinical evaluation

The study population is patients who are selected or underwent elective RFCA for AF. They need to be older than 18 years old, speak and understand the Dutch language, they own and can use a smartphone or a tablet, they are intrinsically motivated to start their cardiac rehabilitation through the specific study.

For the clinical evaluation 30 patients who are selected or underwent elective RFCA for AF will be recruited at the Maxima Medical Center of Eindhoven and Veldhoven. The results of these patients will be compared with the ones of a matched controlled group consisting of patients who participated in the Care-On study, using the Care-On platform only.

3 The different components of the project

As mentioned, the integration of a holistic gamified lifestyle intervention app and a lifestyle monitoring system has to be achieved to create the new gamified intervention.

Greenhabit Greenhabit works through an app (mobile or tablet) that helps people to make healthy choices and develop new habits through science-based behavior change models. The training program has a 12-week duration, and it supports and educates people about their physical and mental health. Moreover, this is a gamified experience where the user receives challenges and rewards during the "journey". It uses a holistic approach focused on physical, mental and social health. They are divided into five elements, i) exercise (move), ii) healthy diet, iii) social environment, iv) positive thinking, and v) relaxation.

Based on a Quality-of-Life questionnaire which is provided at the baseline assessment, and it is repeated every four weeks the algorithm provides personalized challenges and content [7].

Care-On monitoring system Care-On is a recently designed lifestyle monitoring system for patients with coronary artery disease, arrhythmias, and valvular disease. It consists of i) a smartwatch to track daily activity, ii) a chatbot that patients use to self-report their nutrition, stress, and sleep habits, iii) a personal dashboard where the patients can receive feedback on their self-reporting, iv) a goal tracking module and quarterly assessments.

Care-On aims to assist these patients in monitoring their lifestyle factors which can lead to better self-management and improved self-motivation for a healthy lifestyle [6].

3.1 The newly developed platform

To achieve personalization the patients need to provide data, this will happen through the Care-On platform and the CardiacSense 3 medical watch [2]. Through these two ways, self-reported and objective data respectively will be collected and will feed the algorithm of Greenhabit.

As the project is ongoing the project team is working on different working packages defining different functional and technical requirements, involving the professional experts and developers of the consortium partners. This is happening through co-creation sessions, expert interviews, regular meetings and co-creation workshops.

Acknowledgments. This paper was written based on an approved proposal sponsored by the Hartstichting (Dutch Heart Foundation) and enabled by the four partners; Greenhabit, MiBida, Eindhoven University of Technology, and Máxima Medical Center.

About Amalia Mantziari Zafeiri Amalia is an Engineering Doctorate trainee from the Human-System Interaction program at the department of Industrial Design at the Eindhoven University of Technology (TU/e) in the Netherlands. She holds a Bachelor's Degree in Product and Systems Design Engineering from the University of the Aegean in Greece, and a Master of Science in Industrial Design Engineering following the track of Human-Technology Relations from the University of Twente in the Netherlands. She loves to combine information from different backgrounds and to work on projects related to sports performance and healthcare. Her aim is to assist people in these sectors through her background in product design engineering and design research, while she is developing and acquiring new knowledge and skills in her current position at the TU/e.

References

1. Ambrosetti, M., Abreu, A., Corrà, U., Davos, C.H., Hansen, D., et al.: Secondary prevention through comprehensive cardiovascular rehabilitation: From knowledge to implementation. 2020 update. a position paper from the secondary prevention and rehabilitation section of the european association of preventive cardiology. European journal of preventive cardiology **28**(5), 460–495 (2021)
2. CardiacSense: Cardiacsense. https://www.cardiacsense.com/en/, accessed: 2024-05-26
3. Clark, A.M., King-Shier, K.M., Spaling, M.A., Duncan, A.S., Stone, J.A., Jaglal, S.B., Thompson, D.R., Angus, J.E.: Factors influencing participation in cardiac rehabilitation programmes after referral and initial attendance: qualitative systematic review and meta-synthesis. Clinical rehabilitation **27**(10), 948–959 (2013)
4. Dibben, G.O., Faulkner, J., Oldridge, N., Rees, K., Thompson, D.R., Zwisler, A.D., Taylor, R.S.: Exercise-based cardiac rehabilitation for coronary heart disease: a meta-analysis. European heart journal **44**(6), 452–469 (2023)
5. Eijsvogels, T.M., Maessen, M.F., Bakker, E.A., Meindersma, E.P., et al.: Association of cardiac rehabilitation with all-cause mortality among patients with cardiovascular disease in the netherlands. JAMA Network Open **3**(7), e2011686–e2011686 (2020)
6. Goevaerts, W.F., Tenbült-van Limpt, N., Kop, W.J., Birk, M.V., Liu, Y., Brouwers, R.W.M., Lu, Y., Kemps, H.M.C.: Adherence to a lifestyle monitoring system in patients with heart disease: protocol for the care-on prospective observational trial. BMC Cardiovascular Disorders **23**(1), 196 (2023)
7. Greenhabit: Greenhabit. https://greenhabit.nl/en/, accessed: 2024-05-26

8. Jackson, C.L., Redline, S., Emmons, K.M.: Sleep as a potential fundamental contributor to disparities in cardiovascular health. Annual review of public health **36**, 417–440 (2015)
9. Kotseva, K., Wood, D., De Bacquer, D.: Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in europe: Euroaspire iv survey. European journal of preventive cardiology **25**(12), 1242–1251 (2018)
10. Kotseva, K., Wood, D., De Bacquer, D., De Backer, G., Rydén, L., , et al.: Euroaspire iv: A european society of cardiology survey on the lifestyle, risk factor and therapeutic management of coronary patients from 24 european countries. European journal of preventive cardiology **23**(6), 636–648 (2016)
11. Maisano, F., Viganò, G., Calabrese, C., Taramasso, M., Denti, P., Blasio, A., Guidotti, A., Alfieri, O.: Quality of life of elderly patients following valve surgery for chronic organic mitral regurgitation. European journal of cardio-thoracic surgery **36**(2), 261–266 (2009)
12. Mampuya, W.M.: Cardiac rehabilitation past, present and future: an overview. Cardiovascular diagnosis and therapy **2**(1), 38 (2012)
13. MC, M., Ziekenhuis, C.: Hartrevalidatieprogramma folder. https://webshare.zenya.work/t25004jn9vryq9ss/Document.aspx?websharedocumentid=a6f70893-655c-4ced-90f8-0315c8d869e9, accessed: 2024-05-26
14. MiBida: Mibida. https://mibida.nl/, accessed: 2024-05-26
15. Neubeck, L., Freedman, S.B., Clark, A.M., Briffa, T., Bauman, A., Redfern, J.: Participating in cardiac rehabilitation: a systematic review and meta-synthesis of qualitative data. European journal of preventive cardiology **19**(3), 494–503 (2012)
16. Organization, W.H.: World health statistics 2020: monitoring health for the sdgs, sustainable development goals. https://apps.who.int/iris/handle/10665/332070 (2020), accessed: 2024-05-26
17. Peters, A.E., Keeley, E.C.: Trends and predictors of participation in cardiac rehabilitation following acute myocardial infarction: data from the behavioral risk factor surveillance system. Journal of the American Heart Association **7**(1), e007664 (2017)
18. Risom, S.S., Zwisler, A.D., Thygesen, L.C., Svendsen, J.H., Berg, S.K.: High readmission rates and mental distress 1 yr after ablation for atrial fibrillation or atrial flutter: a nationwide survey. Journal of cardiopulmonary rehabilitation and prevention **39**(1), 33–38 (2019)
19. Sabzmakan, L., Morowatisharifabad, M.A., Mohammadi, E., Mazloomy-Mahmoodabad, S.S., Rabiei, K., Naseri, M.H., Shakibazadeh, E., Mirzaei, M.: Behavioral determinants of cardiovascular diseases risk factors: A qualitative directed content analysis. ARYA atherosclerosis **10**(2), 71 (2014)
20. SALLIS, J.F., PATTERSON, T.L., BUONO, M.J., NADER, P.R.: Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. American Journal of Epidemiology **127**(5), 933–941 (1988)
21. Salzwedel, A., Jensen, K., Rauch, B., Doherty, P., Metzendorf, M.I., Hackbusch, M., Völler, H., Schmid, J.P., Davos, C.H.: Effectiveness of comprehensive cardiac rehabilitation in coronary artery disease patients treated according to contemporary evidence based medicine: Update of the cardiac rehabilitation outcome study (cros-ii). European journal of preventive cardiology **27**(16), 1756–1774 (2020)
22. Shahrbabaki, S.S., Linz, D., Hartmann, S., Redline, S., Baumert, M.: Sleep arousal burden is associated with long-term all-cause and cardiovascular mortality in 8001 community-dwelling older men and women. European heart journal **42**(21), 2088–2099 (2021)

WalkEase Mapper: The Digital Twin of Walkability with Light

Golnoosh Sabahifard[0009-0002-4139-2867],
Hanmeng Wang[0009-0008-4840-7797],
Wout Vissers, and Regina Bernhaupt[0000-0003-3522-5990]

Department of Industrial Design, Eindhoven University of Technology
g.sabahifard@tue.nl

Abstract. This study presents a prototype integrating a digital and physical system designed to enhance urban walkability through real-time data processing. The core technology involves a Raspberry Pi interfaced with an ESP32 module for capturing video feeds, which are then processed using an object detection model trained on the COCO dataset. The collected sensor data is used to evaluate pedestrian traffic and environmental variables, which are logged in CSV format for further analysis. This data is the backbone for a predictive model employing the Support Vector Machine (SVM) algorithm. The SVM model, tested with various kernel functions (linear, polynomial, RBF, and sigmoid), excels at classifying input data into distinct classes, facilitating the accurate prediction of walkability scores based on the environmental inputs. Although the system is operational in real-time and capable of supporting an application or dashboard interface for user interaction, the connection to a user interface remains theoretical. This study offers a novel approach to urban design that prioritizes pedestrian convenience and safety. This project not only showcases the integration of machine learning in urban planning but also highlights the potential for future enhancements in interactive city planning tools.

Keywords: Artificial Intelligence · Digital twin city · Light design · Walkability.

1 Introduction

1.1 Problem definition

One of the crucial indicators of urban design quality is walkability, the measure of how friendly an area is. Currently, many cities continue to face the problem of inadequate or poorly designed street lighting, which further influences the area's walkability. Old lighting infrastructure and lack of maintenance are unable to meet the diversity of pedestrian light requirements. The different lighting needs of pedestrians, especially at night, are not adequately met by outdated infrastructure and inadequate maintenance. This contributes to feelings of insecurity and discourages walking. In addition, inefficient lighting systems not only contribute to the pollution of light but also result in energy waste and rising maintenance costs that pose a long-term sustainability problem [11].

[2] emphasized that walkability is an indicator that integrates different factors in order to evaluate how friendly a region is to walking activities. It should be an excellent shorthand for good urban design [16] and crucial for fostering healthy communities and enhancing city livability [8]. From the previous research, [15] acknowledged four factors influencing the walking experience: useful, safe, comfortable, and interesting. Therefore, ensuring adequate lighting levels is imperative to augment an area's walkability, safeguarding pedestrians against risks such as falls, accidents, and personal security concerns [14]. Well-designed lighting not only enhances visibility but also empowers pedestrians to identify potential hazards and threats in their surroundings effectively.

In addition to safety concerns, lighting has an important influence on pedestrians' environment. The space can be transformed into an inviting place, encouraging leisurely walks and leisure activities by engaging lighting designs. Architectural features, landscapes, public art, and the various elements of city architecture may be effectively highlighted by illumination accents [3]. In addition, it is essential to make lighting arrangements that consider various needs at different times of the day and night. With the flexibility of street lighting, energy efficiency can be optimized, and adequate illumination can be ensured at all times.

Therefore, the main objective of optimizing street lighting is to increase walkability through improvements in pedestrian safety, comfort, and overall aesthetic appeal of the urban environment. Cities may increase the number of people walking, resulting in numerous benefits such as reducing carbon dioxide emissions, improving public health, and increasing social interaction by providing appropriate and well-planned lighting.

A prevalent strategy is to exploit digital twin city (DTC) frameworks for the aggregation of multiple data on environmental needs in urban locations. Four main areas of application are found to be widespread in this approach: a design for monitoring, optimization, simulation, and prediction. This is what we are trying to exploit by drawing a visual map of walkability, which allows the simulation and forecasting of walking ability indices at different lighting conditions on the street. In order to optimize walkability, urban planning is empowered by this comprehensive approach to make sound decisions on street lighting improvements.

A comprehensive walkability index that takes into account not only lighting conditions but also other relevant factors is our focus in order to achieve this objective. A key component of our strategy is to introduce "WalkEase Mapper," a mobile application. This innovative tool provides real-time information on the walkability of specific areas and offers simulations under different lighting scenarios. This solution provides valuable data to guide targeted interventions by examining the complex interplay between lighting and pedestrian experience. With this information, urban planning can launch initiatives to improve lighting in order to enhance the overall accessibility of a city area.

1.2 Related works

Digital twin application in walkability The advent of digital twins has revolutionized the ability to explore diverse hypothetical scenarios concerning physical entities and processes. Extending this concept, [10] did a literature review on the application of Digital Twins in the context of lighting analysis. Their findings delineated two primary categories of applications: 1) Digital twins for urban lighting planning and 2) Digital twins for walkability and wildlife impact assessment. These approaches aim to guide the definition of the most efficient lighting configurations for indoor and outdoor environments to make further lighting intervention planning aimed at urban operations.

Our research aims to assess the effects of different lighting scenarios, such as the influence of high light intensity on walkability, in contrast to popular approaches focused on real-time monitoring and energy consumption analysis. We will use our digital twin model to predict outcomes based on selected lighting scenarios and current environmental conditions, thereby providing invaluable support for the professionals who design effective illumination interventions. In this forward-thinking methodology, a deeper understanding of the relationship between light and city mobility is to be offered in order to inform more informed decision-making processes.

Expanding the scope of digital twins in the realm of walkability extends toward health quantification. One compelling application involves creating a Digital Twin City (DTC) model to quantify the collective distress of older adults and correlate it with environmental conditions [4]. Our study, inspired by this approach, uses the potential of the DTC model to quantify the walkability index in different areas, thereby promoting walking and health benefits, particularly for older people. This innovation has highlighted the crucial role that digital twin technologies play as a key tool to tackle public health challenges in cities and improve the quality of life for different ethnic groups.

Walkability index measurement Many researchers have investigated what could be utilized as components of the walkability index. [8] proposed 6D's framework, and even earlier studies, such as [9] built a parcel-level walkability index. Subsequent studies in measuring the walkability index generally take the components from these previous frameworks and add them together. It is either by simply adding the partial indexes, such as in the case study of the walkability index in Olomouc city, [7], or by adding the scaled z-standardized components when [12] measured the walkability index in the Netherlands.

In addition to Lam's methodology, our study is based on seven key components that make up the total walkability index. These components include population density, street connectivity, retailers' and service densities, land use mix, green space, sidewalks accessibility as well as public transport access. We draw upon detailed descriptions and technical Geographic Information System (GIS) techniques, referencing works such as Wagtendonk and Lakerveld for guidance and implementation.

2 Methods and Materials

2.1 Approach and methods

The key to understanding the impact of street lighting on walkability is to recognize the different needs of pedestrians. It is very clear that pedestrians are the primary users, including:

- Commuters moving to their destination, for example: work, school, or public transport.
- Residents taking part in leisure activities, such as dog walks.
- Fitness fanatics use the streets to exercise, for example, running, walking.

These users have a variety of tasks and requirements, such as navigation from point A to point B and enjoying outdoor recreational activities. Our study focuses on creating a digital twin of the walking experience influenced by lighting for residents in habitual areas where walking is integral. By targeting both young and elderly adults in residential areas, we aim to provide insights into how walkability maps and real-time lighting conditions can inform their decision-making when selecting walking routes.

Furthermore, the concept of Active Living Environments (ALE) proposes enhancing walkability by considering the combined impact of natural, built, and social properties of neighborhoods to encourage physical activity and improve health [1]. In line with this approach, we also incorporate other natural factors, such as real-time weather conditions and crowd density, into our analysis to provide a comprehensive understanding of walkability. These environmental variables could be specified as:

- **Real-time light**: When assessing current light levels, considering both brightness and color temperature is crucial. While natural light predominantly influences outdoor mobility during the day, artificial lighting becomes increasingly significant as daylight fades. Particularly at night, artificial lighting significantly impacts the perceived safety and comfort of pedestrians navigating streets. Thus, evaluating both natural and artificial light conditions is essential for understanding the overall walking experience throughout the day.
- **Real-time weather**: In addition to light conditions, real-time weather parameters such as temperature, humidity, precipitation, wind speed, and direction significantly influence pedestrian walkability [5,18]. These factors are particularly impactful for older adults, whose mobility can be more sensitive to environmental conditions. By considering these weather variables alongside lighting, we gain a more comprehensive understanding of the factors shaping the walking experience for different demographic groups [6].
- **Real-time crowdedness**: In our study, we define a "Crowdedness spot" as an area densely populated with an unusually high number of objects which can impede walkability. The level of crowdedness directly correlates with the walkability of a spot: the more crowded an area, the less walkable it

becomes. To quantify crowdedness, we employ a method that detects the number of people within a specific area at any given time. This approach provides a dynamic understanding of pedestrian density, enabling us to assess walkability in real-time and identify areas that may require interventions to improve pedestrian flow.

By providing real-time information on weather conditions and walkability, the inclusion of this information in the application enhances the user experience. In view of factors such as weather and walkability displayed on the map, pedestrian users can now make more informed decisions about their routes. The digital twin acts as a guide, indicating the best times and routes to walk through its map display and information panel. This data integration will enable pedestrians to plan their journeys more effectively, ensuring a safer and more pleasant experience of walking. Table 1 presents a guide for system designers and developers to ensure that the system's functionality aligns with the users' needs and goals.

Table 1: User Task Description

User	**Task Description**
Pedestrians	Use the walkability index on the heatmap to make walking decisions.
Pedestrians	Check real-time light on the application, weather, and traffic information to assist in decision-making.
Pedestrians	Give feedback on the walking experience of one spot on the map.
Designers	Adjust real-time light, weather, and traffic information to predict walkability on the application logged in as a "domain expert."
Designers	Based on the walkability index to possibly optimize and redesign lighting in the future (Not covered in current application).

3 Design

3.1 Data and system states

Physical sensors of the digital twin We have chosen a simplified method of collecting input data in the interest of accessibility and feasibility for this project, which is limited to 4 weeks. In designing and implementing a digital twin, the use of complex or additional sensors would lead to unnecessary complications. Therefore, the current digital twin is primarily based on real-time data from a specially designed probe test. There are various components in this probe, each of which serves a specific purpose in the collection of data. Below are the components included in the probe, along with brief descriptions of their functions and significance (Fig 1).

- **ESP 32 camera board[ESP32CAMLINK]**: The central hub of our probe is the ESP32CAMLINK camera board, which functions not only as a sensor but also as a processing unit. The ESP32, which organizes the collection and transmission of data, is connected to every other sensor. The board formats this data into JSON and sends it over HTTP in order to retrieve the sensor readings on a regular basis. The camera on the module also operates through HTTP; for the project, the default 'example' [17] is used, where the camera can be configured and set to a certain resolution. The camera's streaming capability was disabled in the code to ensure privacy, so that only images could be captured. In order to determine the density of the crowd, these images are then used by an object detection model. Throughout the project, the probe remained connected to the local network via a mobile hotspot, limiting access to ESP32 HTTP posts to connected devices.
- **Lux sensor**: The lux sensor is an advanced device for measuring the intensity of sunlight in a small dome. It uses I2C to communicate with the probe, so that it can be integrated as smoothly as possible. The operation of the sensor is simplified by using the provided library, which facilitates its implementation in a relatively easy way. The sensor's location on the probe provides optimum exposure to sunlight and light from the surrounding environment, which allows for precise measurement of illumination levels within the surroundings.
- **DHT 11**: The DHT11, well-known for its affordability, serves as a reliable sensor for capturing both humidity and temperature levels. It is easy to integrate this sensor into a project with several available libraries, such as Adafruit's DHT library. Positioned within the probe, the DHT11 accurately measures both humidity and temperature, offering comprehensive environmental data crucial for the digital twin's functionality.
- **Rain sensor**: Mounted near the backside of the probe, the rain sensor mimics a miniature roof, angled slightly to maximize its efficiency. It's based on a simple principle: if the rain contacts its surface, it will be connected. This connection is assessed by a small board to determine whether there is rainfall (1) or not (0).

Essentially, the ESP32 orchestrates all sensors' functions to make it easier for a Raspberry Pi to collect and store sensor data. The probe, designed specifically for this project, has a low-cost design that is easily accessible and easy to replicate in order to streamline the grid integration procedures.

Digital processing A Raspberry Pi, which serves as a gateway and data logging hub, is responsible for managing the digital component of the Digital Twin along with physical sensors. In order to ensure the integrity of the data stream, the Pi is connected to the same local hotspot as ESP32, ensuring the security of the data stream. Data logging is integral to the project's success, feeding into the model responsible for computing the walkability score. The logging process is divided into two segments, with one dedicated to sensor data and the other focusing on object detection for people counting.

Fig. 1: The designed probe outside the case

The first part is the object detection for counting people in frame, this is entirely done on the raspberry pi. The used model is the smallest model of YOLOV8 family [19], which is nano. This model is relatively small, and a good fit for devices with limited computing power. The inference times on the pi are around 1100ms which is okay since a live feed is not needed. Every YOLOV8 model comes pretrained on the COCO dataset [13] which is used as a typical benchmark, however still powerful in detecting everyday objects and things. This is also the case for people, therefore no pretraining is required.

The video streams from the ESP32 are being captured by the Raspberry Pi via HTTP. Then the model processes these frames, which are basically wide arrays with RGB values. The model increments the counter for each detected instance by detecting RGB patterns indicative of human beings. It is important to point out that images were not examined or stored in the course of testing, only accessible and analyzed by object detection models.

There are a lot of potential uses after the Raspberry Pi has read sensor data. For user access through the application, one possibility is to process and save the data to lamp posts. However, only data logging has been carried out in the project. Therefore, the CSV file format has been used for that data stream. The model is then trained on these logged CSV files to calculate the walkability score.

In summary, the prototype for this project consists of both Digital and Physical Components which are completely functional and capable of processing time series data. While the digital twin operates in real time, there is currently no direct connection between the designed user interface and the digital twin itself. For the time being, this aspect continues to be conceptual and future work aims at establishing links among these components.

3.2 Implementation of predictive model

ML Model The SVM algorithm is particularly adept at classifying input data by creating a hyperplane that maximizes the separation between different classes, as illustrated in Figure 2. Support vectors, highlighted in the figure, represent the data points closest to the hyperplane within each class. By maximising the margin between classes, this strategic positioning ensures that classification is efficient.

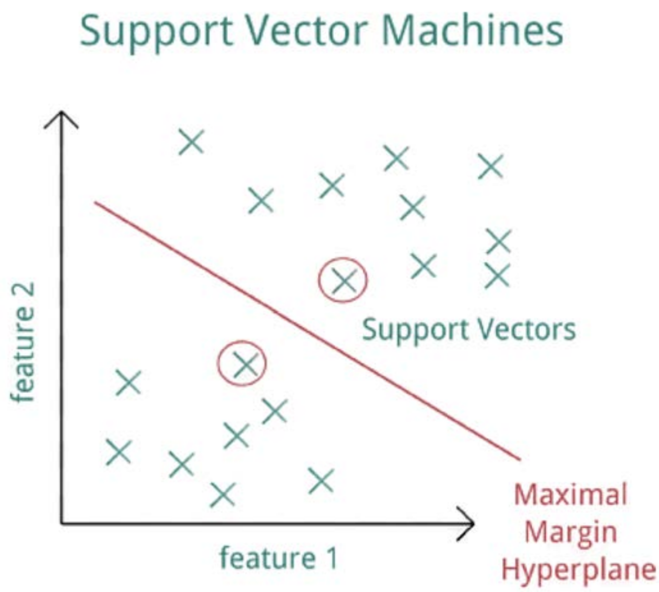


Fig. 2: SVM Model

We explored several kernel functions such as linear, polynomial, RBF (Radial Basis Function), and sigmoid in our quest to optimize the SVM algorithm's performance. The RBF kernel, which has an excellent accuracy rate of 83%, was identified as the most notable performer in this category, as shown in Table 2. The selection of kernels has played an important role in achieving the intended classification results, demonstrating its ability to deal with a number of data distribution issues.

Walkability index & Predictive Model Here we will review the walkability index in relation to the predictive model.

Walkability index: While Lam's seven components model of walkability is available to us, it provides a somewhat stable measurement which remains

Table 2: Accuracy of various kernels

Kernel	**Accuracy**
Linear	0.77
Polynomial	0.74
RBF	0.83
Sigmoid	0.27

consistent as long as the key factors are unchanged. The objective of the project is to establish a quantifiable walking ability index that reflects changes before and following lighting interventions. In order to achieve this flexibility, we use a five point Likert survey in our application to collect perceived walkability data from pedestrians, including self-reported comfort and safety. While the actual user testing is still ongoing, researchers have scored perceptions of walkability twice an hour as a proxy. This approach ensures that the dynamic index is adapted to changes in light conditions.

$$WAI = 0.5 \times (PD + SC + RSD + LUM + GS + SD + PTD) + 0.5 \times (PC + PS)$$

Where:

- WAI: Walkability index
- PD: the standard value of the population density
- SC: the standard value of street connectivity
- RSD: the standard value of retail and service density
- LUM: the standard value of land use mix
- GS: the standard value of green space
- SD: the standard value of sidewalk density
- PTD: the standard value of public transport density
- PC: the standard value of perceived comfort
- PS: the standard value of perceived safety

To obtain the overall walkability index, we average z-score values for each component after calculating them. We minimize the index, ensuring that it is between 0 and 100, to aid interpretation. The scale makes it easier to interpret the analysis results, with higher values suggesting a better walking experience for pedestrians. This method offers a standardized metric to compare walking ability in different areas and scenarios.

Predictive model: The forecasting model is intended to anticipate the level of walkability in a variety of climatic conditions. The flow of the predictive model, as well as the inputs and outputs that form part of the classification process, are presented in Figure 3. This visual representation provides a complete overview of its functions and outputs, explaining how the various factors are contributing to models' predictions.

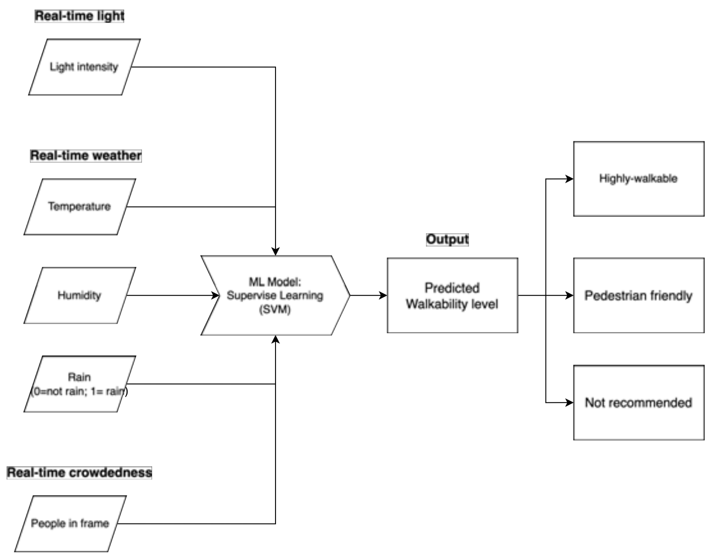


Fig. 3: Flow of predictive model

The other inputs are collected from temperature sensors and so on, as described in the section Data and system states above. Before we trained the model, we organized the walkability index into three labels:

- highly-walkable (walkability index $\in (66, 100]$),
- pedestrian-friendly (walkability index $\in (33, 66]$),
- and not-recommended (walkability index $\in (0, 33]$).

These are also what the outputs are as a result of the training.

3.3 Testing and Analysis

Data description To enrich the dataset and provide a more comprehensive analysis, experiments should encompass diverse times of the day and various locations within Eindhoven. However, due to practical constraints in obtaining real-world data, our model was trained and evaluated specifically in

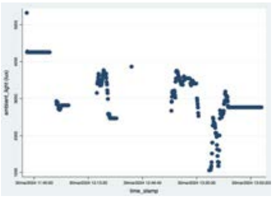
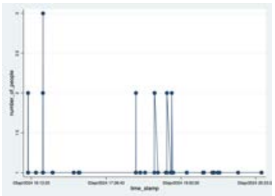
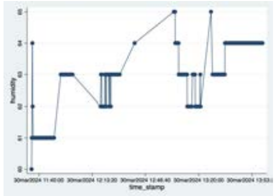
Laboratoriumstraat, Eindhoven. This focused approach involved two sessions: one conducted on 30/03/2024 from 11:30 to 14:00 and another on 03/04/2024 from 15:00 to 21:00. While limited in scope; these sessions provided valuable insights into the model's performance under different lighting and environmental conditions.

Prior to conducting descriptive statistics analysis, rigorous data cleaning procedures were implemented to ensure the integrity and reliability of the dataset. Outlier data points, including those observed at the beginning of each experiment and instances where temperature readings exceeded 100 degrees Celsius, were systematically removed. Additionally, ambient light values that exhibited unusually high or low values during nighttime or daytime conditions, respectively, were filtered out. To maintain consistency and coherence in the dataset, continuous variables were averaged per minute, mitigating any potential variations that could affect the final analysis. These meticulous steps in data preparation contributed to the robustness and accuracy of subsequent statistical analyses.

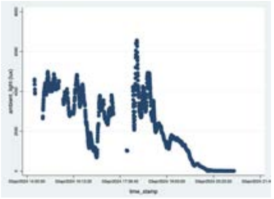
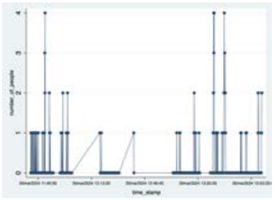
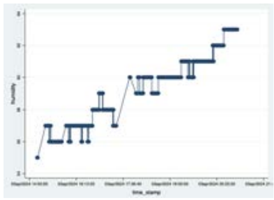
Figure 4 presents the temporal variation of key environmental variables, focusing on ambient light, crowdedness, and humidity. A discernible trend emerges wherein light intensity gradually diminishes over time during both observation periods. Conversely, humidity displays an upward trajectory during the latter phase of data collection. The fluctuation in real-time crowdedness exhibits no discernible pattern, showcasing diverse patterns of pedestrian activity throughout the observation period. Notably, certain instances of missing data were observed, particularly during the initial data collection phase. However, given the precision of timestamps, which are accurate to over 2 digits, the impact of these missing data points on subsequent modeling endeavors is expected to be negligible.

Prior to model training, preprocessing of the dataset was conducted to facilitate accurate analysis. Continuous variables such as light intensity, temperature, and humidity were averaged on an hourly basis to provide a representative snapshot of environmental conditions. Additionally, the total count of individuals within the frame was calculated per hour to capture variations in pedestrian activity. Rainfall levels were dichotomized, with a value of 1 assigned if rain was detected at any point within the preceding hour and 0 otherwise. This preprocessing step ensured that the dataset was appropriately formatted for subsequent modeling tasks.

Evaluation of the ML Model A Python script was developed to automate dataset generation. This script employs predefined thresholds to label data based on real-time light, weather, and crowdedness inputs, facilitating the creation of a structured dataset. Subsequently, the dataset was partitioned into training and testing sets, with 70% allocated for training and 30% for testing. Leveraging this dataset, a clustering algorithm was deployed to delineate three distinct clusters. Evaluation of the learning algorithm revealed an impressive accuracy rate of 83% when tested against the reserved dataset. This indicates robust performance

(a) Ambient light on 30/03/2024 (b) Number of people on 30/03/2024 (c) Humidity on 30/03/2024

(d) Ambient light on 03/04/2024 (e) Number of people on 03/04/2024 (f) Humidity on 03/04/2024

Fig. 4: Several environmental variables variation across different timestamps (30/03/2024, 03/04/2024 from top to bottom) (a) Ambient light; (b) Number of people; (c) Humidity

and suggests the efficacy of the machine learning model in predicting walkability levels under varying environmental conditions.

3.4 Final Design

For pedestrians Figure 5 (a) exemplifies our digital twin system's user interface, prioritizing pedestrian safety and comfort. Central to its design is real-time information delivery and options ensuring an optimal walking experience. Users are empowered to make informed route choices through two primary features: a walkability map showcasing street rankings based on various factors and a mobile application offering detailed insights on weather conditions, lighting quality, infrastructure, proximity, and traffic security. This multifaceted approach aims to enhance walkability by leveraging street lighting optimization.

In the interactive map displayed at the top of the app's homepage, users can instantly assess the walkability ranking of each street within the city. To denote walkability levels, a color-coded system is implemented: green signifies highly walkable areas, yellow indicates pedestrian-friendly zones, and red denotes streets not recommended for walking. This classification mirrors the breakdown of walkability index into distinct levels, as outlined in the walkability index section. Within the map interface, users have the option to select individual streets for more comprehensive details, such as user ratings and reviews, facilitating informed decision-making regarding their walking routes.

On the app's second page, users will have access to a range of customizable features and notifications, enhancing their overall experience. These options encompass free Wi-Fi availability, weather updates, traffic conditions, dynamic lighting control, and traffic signage alerts. Each feature will be accompanied by a dedicated page, offering users detailed information and customization capabilities.

For instance, within the weather conditions section, users can access real-time updates on temperature, precipitation, and wind speed forecasts, aiding in route planning for their walks. Additionally, users can toggle these features on or off based on their preferences, ensuring a personalized experience tailored to their specific needs and preferences.

Upon concluding their walking experience, users will be encouraged to provide feedback and rate the app's effectiveness in facilitating their journey. This feedback mechanism serves a dual purpose: it aids in predicting the walkability level of different areas and identifies potential areas for improvement in the city's walking infrastructure. By soliciting user feedback, the app can continuously refine its features and offerings to better meet the needs and expectations of pedestrians. This iterative process not only enhances the user experience but also contributes to the overall improvement of urban walkability and pedestrian-friendly environments.

For designers In Figure 5 (b), we see how the digital twin serves as a valuable tool for domain experts and designers. Through the operation dashboard, designers can manipulate the scroll bar to gain insights into the walkability of specific areas. By adjusting the light intensity, designers can observe how the walkability classification changes, with more areas transitioning from "not-recommended" to "pedestrian-friendly" as the light intensity increases. Additionally, designers can zoom in to analyze the variation in walkability within a specific area as light intensity varies.

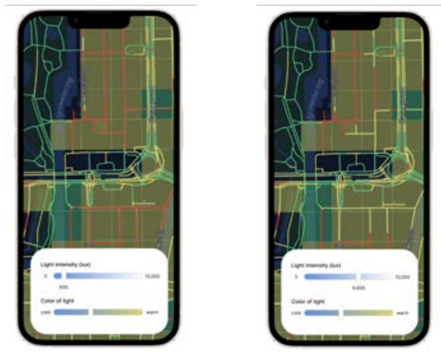
It's important to note that our current model does not incorporate the color of light, but this could be a potential enhancement in the future. By integrating light color into the model, designers could further refine their understanding of how different lighting characteristics impact walkability. The predictive walkability level provided by the digital twin offers valuable insights for designers in optimizing and redesigning lighting infrastructure. These insights pave the way for more informed decision-making processes and ultimately contribute to the creation of safer and more pedestrian-friendly urban environments.

4 Discussion

4.1 Digital twin and real-world

Now that every aspect of the system is explained, the broader system will be explained again. For this project feasibility was a bit factor in the design process and decision making. The flowchart in Figure 6 shows the proposed system in its entirety, where the real-world interaction and the underlying systems in the digital twin are portrayed. This section will go into more depth in this flowchart.

(a) Pedestrians use the application to gain walkability suggestion and environmental information



(b) Domain experts use the application to gain the walkability prediction on their intervention of lighting intensity.

Fig. 5: User interface of *WalkEase Mapper*

Real world As we analyze the real world, it's crucial to understand the motivations behind individuals going for a walk. As mentioned earlier, users may have various reasons for their stroll. In our scenario, when users consult the app, they encounter a stop-light style presentation of the walkability score for a given area. Green signifies high walkability, while red indicates a lower score. If a user decides to venture to that location, they are prompted to provide their own perceived scores, which are then utilized for training and recalibrating purposes. This iterative process ensures that the walkability scores remain accurate and reflect real-world experiences.

Digital twin The digital twin relies on various data sources, collected either through sensor data or calculated by the twin itself. Sensor data is captured by a node, and along with the count of people, it's stored in a central database

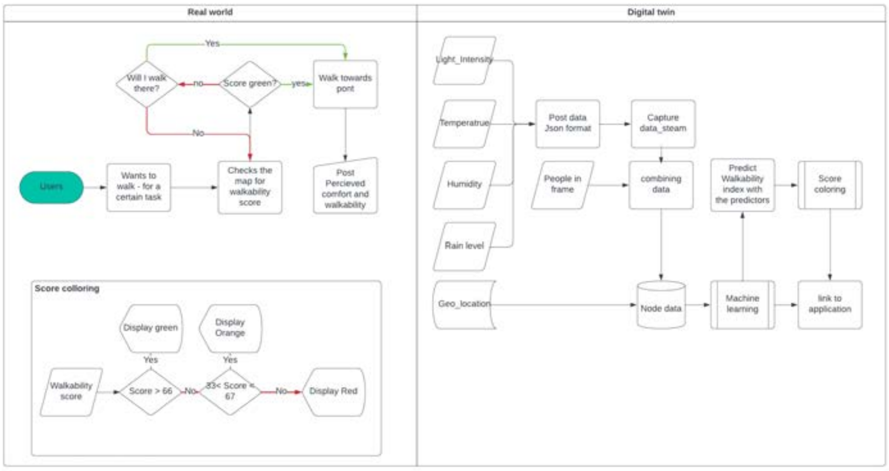


Fig. 6: The digital twin proposed system in its entirety, where the real-world interaction and the underlaying systems are portrayed

for future reference. This database holds values that represent the real-time conditions at each node, forming the basis for displaying values on the map in the user interface. Beyond just coloring areas on the map, users can access real-time information specific to each node. This feature provides users with deeper insights into how scores are calculated and what factors contribute to low or high walkability scores.

4.2 Limitation

While our approach offers valuable insights, a notable limitation is its focus on just one location. Relying solely on time-series data from a single spot restricts our ability to make accurate predictions for the entire residential area of Eindhoven. Similarly, it hampers our capacity to develop reliable walkability indices across various lighting scenarios. To address this constraint, we intend to incorporate self-report data as an additional data source. However, obtaining such data from actual participants remains a challenge due to time constraints.

While our sensor setup provides valuable data, it's essential to acknowledge its limitations. Off-the-shelf components like the DHT sensor and ESP32 camera board are known for their moderate accuracy and resolution. Upgrading to higher-end versions of these sensors could enhance precision, but at a higher cost. Similarly, opting for premium sensors would improve data quality but increase expenses. This trade-off between cost and precision is a crucial consideration for future iterations of our project.

Expanding our data collection beyond ambient light intensity could provide richer insights for domain experts. Adjusting artificial light intensity and color, for instance, could offer a more nuanced understanding of their impact on walkability. Moreover, enhancing the predictive model involves considering more accurate predictors. While we currently use binary rain data, incorporating rainfall intensity could better capture its effect on walkability. Additionally, factors like traffic congestion and emergencies could significantly influence pedestrian experience, warranting inclusion in future model iterations for improved accuracy.

5 Conclusion

In this study, the key role of lighting in improving walking conditions in urban environments and adopting Digital Twin technology to improve street lighting systems is highlighted. This study aims to improve pedestrian safety, comfort, and overall perception of the urban environment to address the problem of insufficient or inadequate street lighting. Future research could focus on integrating sustainability metrics into the walkability index and digital twin framework, engaging communities in the co-design process for more inclusive interventions, leveraging advanced data analytics for real-time adaptability, integrating street lighting optimization with multi-modal transportation networks, evaluating social equity impacts, and testing interventions in real-world settings to inform evidence-based decision-making in urban planning and design.

About Golnoosh Sabahifard Golnoosh Sabahifard is an Engineering Doctorate (EngD) trainee at Eindhoven University of Technology (TU/e) in collaboration with Eindhoven Engine. Her project focuses on developing a multipurpose integrated digital twin for the Brainport region, which is part of a larger initiative to tackle critical urban development challenges such as energy transition, nitrogen reduction, and managing the exponential increase in urban pollution.

Golnoosh's academic journey includes a Master's degree in Sustainable Urban Mobility Transition through the EIT program, where she earned a double degree from Aalto University in Helsinki and Eindhoven University of Technology. She also holds a Master's degree in Urban Planning and Design from Tehran, Iran. Her research interests lie in urban development, sustainability, and Environmental, Social, and Governance (ESG) factors.

Professionally, Golnoosh has accumulated a diverse range of experiences. At Panteia, she serves as a research assistant in the transport and mobility department, utilizing a variety of qualitative and quantitative research tools, contributing to project management, and drafting proposals for tender applications. Her work with the consultancy firm Goudappel involved a master's thesis on the experiences and evaluation of cyclists regarding bicycle parking facilities in Dutch city centres. In this role, she conducted structured literature reviews and quantitative research using SPSS. Additionally, Golnoosh collaborated with the Danish company Skipit to develop a strategy and 3D model for local bike shops in Helsinki, illustrating how they could benefit from renting bikes through a central app.

Golnoosh is actively engaged in extracurricular activities, serving as the vice-president of the EIT Alumni Board in the Urban Mobility group. In this capacity, she planned activities aimed at enhancing urban mobility for the alumni community, managing both time and budget constraints effectively. Golnoosh has a great personal commitment to sustainable urban development and she has extensive background in research and project management aimed at improving urban environments through innovative and sustainable solutions.

References

1. Bandurist, P.S.: Rethinking walkability and developing a conceptual definition of active living environments to guide research and practice. BMC Public Health **22**(1) (March 2022). https://doi.org/10.1186/s12889-022-12747-3
2. Beiler, M.O., Phillips, B.: Prioritizing pedestrian corridors using walkability performance metrics and decision analysis. Journal of Urban Planning **142**(1) (March 2016). https://doi.org/10.1061/(ASCE)UP.1943-5444.0000290
3. Boučková, V.: Smart city transportation trends: Urban digital twins (2023), good Vision, June 7, 2023
4. Chaspari, T., Kim, J., Lee, C., Ham, Y., Lee, S., Ahn, C.R.: A digital twin city model for capturing walkability distress of elderly adults. IEEE (2022). https://doi.org/10.1109/dtpi55838.2022.9998907
5. Clarke, P., Hirsch, J.A., Melendez, R., Winters, M., Gould, J.S., Ashe, M.C., Furst, S., McKay, H.: Snow and rain modify neighbourhood walkability for older adults. Canadian Journal on Aging **36**(2) (April 2017). https://doi.org/10.1017/S071498081700006X
6. Delclòs-Alió, X., Marquet, O., Vich, G., Schipperijn, J., Zhang, K., Maciejewska, M., Miralles-Guasch, C.: Temperature and rain moderate the effect of neighborhood walkability on walking time for seniors in barcelona. International Journal of Environmental Research and Public Health **17**(1) (2019). https://doi.org/10.3390/ijerph17010014
7. Dobešová, Z., Křivka, T.: Walkability index in urban planning: A case study in olomouc city (2012). https://doi.org/10.5772/36587, inTech eBooks
8. Ewing, R., Cervero, R.: Travel and the built environment. Journal of the American Planning Association **76**(3), 265–294 (2010). https://doi.org/10.1080/01944361003766766
9. Frank, L.D., Schmid, T.L., Sallis, J.F., Chapman, J.E., Saelens, B.E.: Linking objectively measured physical activity with objectively measured urban form: findings from smartraq. American Journal of Preventive Medicine **28**(2) (February 2005). https://doi.org/10.1016/J.AMEPRE.2004.11.001

10. Hassan, M.U., Angelaki, S., de Lima, C.A.V., Major, P., Styve, A., Alaliyat, S., Hameed, I.A., Besenecker, U.S., Torres, R.D.S.: Digital twins for lighting analysis: Literature review, challenges, and research opportunities. Proceedings of the 36th ECMS International Conference on Modelling and Simulation ECMS 2022 (2022). https://doi.org/10.7148/2022-0226
11. Kyba, C.C.M., Kuester, T., de Miguel, A.S., Baugh, K.E., Jechow, A., Hölker, F., Bennie, J., Elvidge, C.D., Gaston, K.J., Guanter, L.: Artificially lit surface of earth at night increasing in radiance and extent. Science Advances **3**(11) (November 2017). https://doi.org/10.1126/SCIADV.1701528
12. Lam, T.M., Wang, Z., Vaartjes, I., et al.: Development of an objectively measured walkability index for the netherlands. International Journal of Behavioral Nutrition and Physical Activity **19**(50) (2022). https://doi.org/10.1186/s12966-022-01270-8
13. Lin, T.Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Dollár, P.: Microsoft coco: Common objects in context (2015), arXiv [Cs.CV], Retrieved from http://arxiv.org/abs/1405.0312
14. Pammer, K., Gauld, C., McKerral, A., Reeves, C.: "they have to be better than human drivers!" motorcyclists' and cyclists' perceptions of autonomous vehicles. Transportation Research Part F: Traffic Psychology and Behaviour **78** (April 2021). https://doi.org/10.1016/J.TRF.2021.02.009
15. Speck, J.: Walkable City Rules: 101 Steps to Making Better Places. Island Press (2018)
16. SPUR: Getting to great places (2013), https://www.spur.org/publications/spur-report/2013-12-12/getting-great-places, retrieved April 12, 2024
17. Systems, E.: Camerawebserver (2024), https://github.com/espressif/arduino-esp32/tree/master/libraries/ESP32/examples/Camera/CameraWebServer, retrieved April 12, 2024
18. Thielman, J., Chiu, M., Rosella, L., Copes, R., Manson, H.: Does climate matter? how temperature and precipitation affect the relationship between neighbourhood walkability and walking for transportation. Environmental Health Perspectives **126**(7) (2018). https://doi.org/10.1289/isesisee.2018.p02.0020
19. Ultralytics: Ultralytics yolo (2024), retrieved April 12, 2024

Enhancing Workplace Well-being: Integrating Physiological Data into Aromatherapy through Calm Technology *

Qiurui Wang[1 [0000−0002−9994−5359]], Caihong He[2 [0000−0002−0791−6308]], and Jun Hu[1 [0000−0003−2714−6264]]

[1] Department of Industrial Design,
Eindhoven University of Technology, The Netherlands
{q.wang1, j.hu}@tue.nl
[2] Guangzhou Wanqu Cooperative Institute of Design, Guangzhou, China
hch01prima@163.com

Abstract. This study aims to develop an intelligent aromatherapy device that harnesses the principles of calm technology to enhance workplace well-being. The device automatically initiates aromatherapy based on physiological data from a wearable wristband to relieve anxiety and enhance office vitality, thereby increasing productivity. The study recruited 90 participants, divided into three groups of 30 people. Anxiety was induced by the Kraepelin test(arithmetic task) and the efficacy of aromatherapy using lavender essential oil, coconut oil and blank group served as control groups. Our experimental findings highlight the effectiveness of lavender essential oil in significantly reducing anxiety levels when compared to the control group. This case study underscores the potential of intelligent aromatherapy device as natural, non-pharmacological intervention for stress management and vitality enhancement in office settings, aligning with the principles of calm technology. It lays the foundation for the further development and improvement of intelligent aromatherapy equipment to improve the quality of work and life.

Keywords: Stress · Aromatherapy · Calm technology · Physiological data.

1 Introduction

With the rapid development of technology, the increase of work tasks and the blurring of the line between work and life, the problem of stress and anxiety in the modern office environment has become more and more prominent[17]. Stress is often associated with tension, time pressure, and increased workload, while

* This is a copy of:
Q. Wang, C. He, and J. Hu, "Enhancing Workplace Well-being: Integrating Physiological Data into Aromatherapy through Calm Technology," in Proceedings of the Eleventh International Symposium of Chinese CHI, Denpasar, Bali, Indonesia, 2023, pp. 455–460.

anxiety can be due to job demands, uncertainty, and a competitive environment. These problems may manifest as physical symptoms, emotional distress, and mental health issues. Stress and anxiety not only negatively impact an individual's health, but also their productivity and work performance[25]. Research shows that chronic exposure to high stress can lead to fatigue, insomnia, anxiety symptoms, and emotional exhaustion, which can ultimately affect productivity and creativity[23]. Additionally, they may also lead to cardiovascular problems, immune system disturbances, and other health risks[8].

In order to alleviate the problems caused by long-term stress, some employees may need to seek medical assistance, including psychological or medication treatment. Psychological therapy can help employees deal with emotional issues such as anxiety and depression, while medication may be necessary in certain situations to stabilize emotions and alleviate symptoms[19]. However, a better approach is to take proactive preventive measures to reduce and manage work stress before the stress problem becomes severe. This includes using stress relief techniques such as deep breathing, physical exercise, meditation, and aromatherapy, as well as establishing a healthy work life balance[7]. Aromatherapy is a promising option as it can be easily integrated into daily work, providing rapid emotional support, and helping employees cope with stress without relying on medication[28]. Aromatherapy is the use of volatile oils to act on the olfactory nerves of the human body, stimulating the brain to synthesize neurotransmitters of joyful emotions, thereby regulating the body and mind, eliminating negative emotions such as anxiety, depression, and anger, while also avoiding gastrointestinal reactions and other drawbacks. Currently, it is also widely used to regulate human emotions and alleviate stress[24]. Neurologists believe that the reason why fragrance can eliminate people's negative emotions is because the flowers or stem and leaf cells of plants can decompose into a volatile aromatic oil after being exposed to sunlight. If this fragrance comes into contact with olfactory cells in the human nasal cavity, it will produce beneficial physiological reactions[16]. For example, the smell of lavender can eliminate tense and melancholic emotions, thus effectively treating insomnia and depression.

In this work we devised a smart diffuser for office people called AromaMate, which revolutionizes the concept of stress management by tapping into the power of scent to alleviate anxiety and promote office vitality. Through harnessing calm technology[9], AromaMate introduces an innovative approach that leverages aromatherapy in an unobtrusive way to effectively reduce stress levels. Embracing this unique utilization of olfactory stimulation, AromaMate offers individuals a discreet and personalized solution to find calm in the midst of their work.

2 Related Work

Aromatherapy is used to treat chronic pain, depression, and anxiety. Essential oils recommended for aromatherapy include tea tree oil, cinnamon oil, clove oil, eucalyptus oil, thyme oil, rosemary oil, lavender oil, and more[15]. Recommended methods of using these essential oils are bathing, inhalation, and massage. Hosseini et al. found that sniffing lavender volatile oil can effectively reduce anxiety levels in patients who are about to undergo cardiopulmonary bypass heart surgery[11], Kasper et al. found that oral lavender volatile oil preparations can effectively reduce anxiety levels in patients with sleep disorders[12], Kianpour et al. found that continuous sniffing of lavender odor for 4 weeks can effectively prevent postpartum women from developing stress, anxiety, and depression[13]. One clinical study investigating anxiety during childbirth[22] reported no psychological benefits and undesirable physiological effects of orange aroma administration. A study that considered changes to the perceived fragrance of essential oils after mental and physical work reported an unfavorable impression of orange aroma[26].

Traditional fragrance mainly use the volatilization of the fragrance itself, or through heating, natural placement, and atomization of essential oils with water. Smart fragrance device produced by combining traditional products with smart hardware. Cyrano odor diffuser[1] and Moodo smart home fragrance diffuser[3] and other products appear, and users can modulate and broadcast fragrance through mobile phone applications. Smell Kingdom[2]realizes the "coding" and "decoding" of smells in real life, and applies them to consumers watching movies, bringing users a new movie watching and smell experience.

There could be adverse effects on both work and personal life if people fail to recognize their state of anxiety and not take preventive measures. The failure to intervene in a timely manner can lead to a decline in productivity. Research conducted by Lohmann-Haislah et al.[14] has demonstrated that anxious employees tend to exhibit relatively lower job performance. Prolonged anxiety is associated with various health issues. These health concerns may translate into increased healthcare expenses and days lost at work[20]. It is well known that aroma inhalation can cause physiological and psychological changes in human[10], and the effects of aroma are believed to be caused by pharmacological and psychological mechanisms. This underscores the importance of developing intervention tools like smart aromatherapy devices, aimed at helping employees manage their emotional states and enhance both their productivity and overall well-being. AromaMate provides a solution by offering smart aromatherapy devices that facilitate the use of aromatherapy as a preventive measure on a regular basis. By incorporating intelligent features and portability, AromaMate ensures individuals can experience the benefits of aromatherapy continuously. With its intelligent, personalized, and portable features, AromaMate can become a valuable tool for emotional management in individuals' everyday lives.

3 Concept Design and Methods

3.1 Concept Design

AromaMate leverages calm technology to provide personalized aromatherapy, aiding in mood enhancement and anxiety management. AromaMate aims at discreetly improving emotional well-being without interfering with the user's daily life. This smart approach ensures users get the right support when they need it.

By harnessing the power of calm technology and aromatherapy, AromaMate is a new way to find stress management and emotional balance in today's fast-paced world.

The smart device relies on real-time data from wearable sensors that seamlessly connect to the user's wristband to monitor their physiological state. By continuously assessing variables such as heart rate and skin behavior, AromaMate can detect signs of increased anxiety or stress levels. Once these indicators are identified, the device automatically initiates aromatherapy, releasing carefully selected scents known for their calming and mood-enhancing properties, such as lavender. What sets AromaMate apart is its ability to adjust the intensity and duration of aromatherapy based on the user's current emotional state. If the system detects heightened anxiety, it can release a stronger scent over a longer period of time, effectively relieving anxiety. Figure 1. shows the design concept of AromaMate.

The AromaMate is equipped with three fragrance modules, which can accommodate up to three fragrance capsules at the same time. Users can mix different fragrance capsules to create personalized fragrance, and adjust the proportion of corresponding fragrance capsules through mobile applications on their phones. They can save and share the fragrance they have prepared with others, with personal customization and social attributes. We use aromatherapy methods to maintain a moderate concentration for 30 minutes and automatically stop, ensuring that users receive professional stress relief without worrying about adverse reactions caused by excessive or prolonged aromatherapy.

3.2 Experiment Design

Participant. A total of 90 participants were selected as participants for the study, aged from 20 to 50 years old(M=35.42, SD=15.98). The snowball sampling[5] technique was employed for participant recruitment, whereby initial participants were identified and then additional participants. All participants were healthy with normal olfactory function. This study was approved by the local Ethical Review Board (ERB). To avoid causing unwanted feelings, users are informed in the consent form that they can stop participating at any time. They can revoke their permission to use their data under any circumstances.

Experiment Preparation. The laboratory temperature is 23±1°C, humidity is 60±5, the environment is quiet, and the light is soft, which meets the research requirements. On the day of the experiment, the diet of the participants should not contain alcohol, caffeine, and other ingredients. Smoking, taking any drugs, and cosmetics and essential oils with aromatic smells are strictly prohibited, and sleep time at night is guaranteed to be 7 hours or above.

As a laboratory stress method that can induce anxiety emotions in the human body and cause changes in physiological indicators, mental arithmetic tasks are widely used in research. The mental arithmetic task of "continuously subtracting 17 from three digits" as a stressor, which is so-called Kraepelin test, this task is frequently used to induce acute stress in various studies[27].

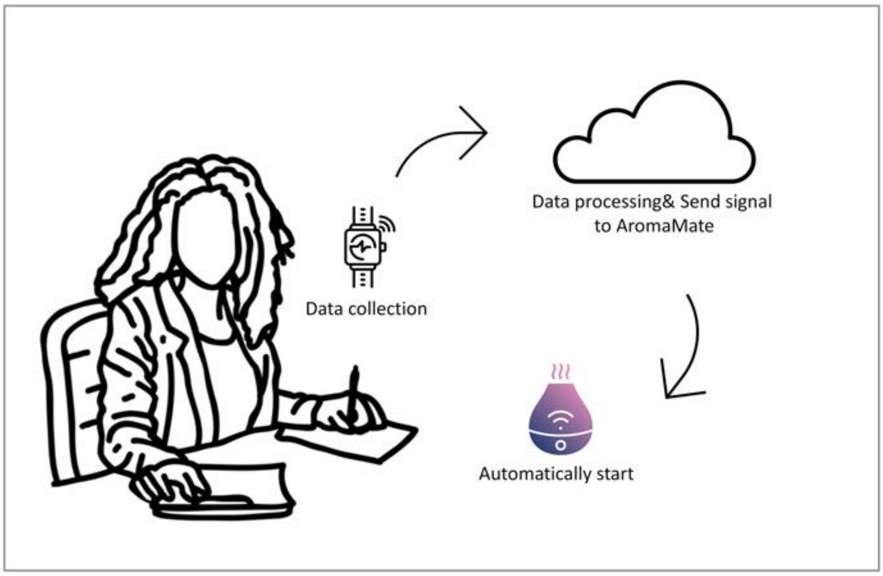


Fig. 1. Design Concept of AromaMate.

By measuring the changes in various physiological indicators, the degree of anxiety in the human body under stress can be objectively reflected, and the low-frequency to high-frequency ratio (LF/HF) of heart rate variability is a sensitive indicator that reflects the balance of sympathetic and parasympathetic nerves[21]. The increase in this value indicates that the subject has successfully induced anxiety under laboratory stress conditions. We use the Empatica E4 wristband to collect the data.

Lavender Essential Oil stands out as a highly researched and popular aromatherapy oil, often considered the most extensively studied within the field. According to aromatherapists, lavender is a cell rejuvenator, antiseptic, and immunostimulant. It is known for its calming effects, and it is said to enhance well-being. Additionally, it offers the advantage of being cost-effective[4]. We opted for a commercially available capsule-based aromatherapy diffuser, with lavender essential oil chosen for its pharmacological properties, coconut oil as a placebo with a pleasant aroma, and a blank group for control. We selected a medium concentration and set the duration to 30 minutes.

In light of AromaMate's current conceptual design stage, which lacked a physical prototype, our experiments necessitated human assistance. This involved continuous monitoring of both the wristband and the aromatherapy machine using the applications. This approach allowed us to simulate the device's intended functionality and gauge its potential efficacy in stress management, even though a physical prototype had not yet been developed.

Experiment Procedure. The experimental process is divided into four stages, each of which records the total number of heartbeats and heart rate variability LF/HF of the participants. The participants were required to fill out the visual analysis scale, and at the end, the participants were required to fill out the Likert scale for the AromaMate design concept.

- Baseline period: After briefly explaining the test requirements and process, the experimenter assisted the participants in adjusting their comfortable sitting posture, and recorded the pre-stress baseline period for 10 minutes. The experimenter dictated the instructions: "Thank you for participating in today's experiment. During the experiment, you will wear equipment that does not cause any harm to your body to test your heart rate under different conditions. At each stage, you will hear my instructions, just focus and follow them."
- Stress period: Each group took the same task by listening to the recording, continuously subtracting 17 from the 3-digit mental calculation program (presenting 1500ms/question), and prompted for the correct answer after an interval of 3 seconds. Participants were required to report their mental calculation results orally as soon as possible within 3 seconds. Regardless of whether the reported results were correct or not, the correct answer was subtracted by 17 from the recorded sound. If the report was made after the next digit sound, it would be counted as an error, and the mental arithmetic task continued for 30 minutes.
- Intervention period: The lavender group and coconut oil group continued to undergo mental arithmetic tasks and essential oil intervention for 30 minutes. The blank group continued to take the arithmetic task for 30 minutes.
- Recovery period: Rest for 10 minutes.

Statistics. We are using IBM SPSS 28 for statistical analysis. The data does not conform to a normal distribution, it is denoted as M (Min, Max), Bonferroni corrections were applied. Kruskal-Wallis test is used for inter-group comparison. Wilcoxon signed-rank test is used for intragroup comparison. The level of statistical significance (p) was set at 0.05.

3.3 Result

Comparison of Baseline Heart Rate Variability LF/HF. Table 1 shows that there were no intergroup differences in baseline heart rate variability LF/HF among the three groups ($P>0.05$), indicating comparability between the groups.

Comparison of Heart Rate Variability LF/HF between Stress and Baseline Period. Intragroup comparison(Table 2) : The LF/HF of heart rate variability during stress period in all three groups of subjects was significantly higher than that of the baseline period. There was a statistical difference ($P<0.01$), indicating that all three groups of subjects successfully induced anxiety under laboratory stress conditions.

Inter-group comparison: There was no inter-group difference in LF/HF during the stress period among the three groups of subjects (P0.05), with comparability.

Table 1. The comparison of heart rate variability LF/HF among participants in the baseline period.

Group	No. of Participants	LF/HF(Min, Max)
Lavender	30	1.28 (0.46, 3.90)
Coconut	30	1.65 (0.41, 3.56)
Control	30	1.40 (0.24, 4.82)

Table 2. Heart rate variability LF/HF among participants in the baseline and stress period.

Group	No. of Participants	Baseline Period	Stress Period
Lavender	30	1.28 (0.46, 3.90)**	2.16 (1.00, 5.65)
Coconut	30	1.65 (0.41, 3.56)**	2.41 (1.06, 5.63)
Control	30	1.40 (0.24, 4.82)**	2.45 (0.81, 6.14)

Comparison of Heart Rate Variability LF/HF in Different Stages of the Experiment. Intragroup comparison (see Table 3): The LF/HF of heart rate variability in the lavender group during stress period was significantly higher than that in the baseline period, intervention period and recovery period ($z=4.78$, $P<0.01$; $z=4.34$, $P<0.01$; $z=3.57$, $P<0.01$); There was no significant difference in heart rate variability LF/HF between baseline period, intervention period and recovery period ($P>0.05$).

The LF/HF of heart rate variability during stress period in the coconut oil group were significantly higher than those in the baseline period, intervention period and recovery period ($z=4.78$, $P<0.01$; $z=2.97$, $P=0.003$; $z=4.10$, $P<0.01$); There was no significant difference in heart rate variability LF/HF between baseline and recovery period ($P>0.05$); The heart rate variability LF/HF of intervention was higher than the baseline and recovery period ($z=2.86$, $P=0.004$; $z=3.10$, $P=0.002$).

The LF/HF of heart rate variability during the stress period in the control group was significantly higher than that during the baseline and recovery periods ($z=4.86$, $P<0.01$; $z=3.44$, $P=0.001$); There was no difference between stress period and intervention period ($P>0.05$).

Inter group comparison (see Table 3): There was no difference in LF/HF of heart rate variability between the baseline and stress periods among the three groups ($P>0.05$). The expected LF/HF of heart rate variability in the lavender group was significantly lower than that of the coconut oil group and the control group ($z=-2.02$, $P=0.04$; $z=-2.86$, $P=0.004$).

Table 3. Heart rate variability LF/HF among participants in different periods.

Group	No.	Baseline Period	Intervention Period	Stress Period	Recovery Period
Lavender	30	1.28 (0.46, 3.90)	2.16 (1.00, 5.65)	1.63 (0.70, 2.98)	1.36 (0.24, 4.66)
Coconut	30	1.65 (0.41, 3.56)	2.41 (1.06, 5.63)	2.01 (1.02, 3.96)	1.74 (0.35, 3.38)
Control	30	1.40 (0.24, 4.82)	2.45 (0.81, 6.14)	2.43 (0.67, 4.66)	1.74 (0.46, 5.21)

4 Discussion and Future Work

The results of all three groups of subjects showed a significant increase in LF/HF heart rate variability during stress, highlighting the effectiveness of arithmetic tasks in inducing stress in subjects. This significant increase indicates the ability of arithmetic tasks to trigger physiological responses typically associated with stress. LF/HF heart rate variability is a reliable marker for assessing the balance of the autonomic nervous system, reflecting the interaction between sympathetic and parasympathetic nervous activities. A high LF/HF ratio usually indicates enhanced sympathetic innervation, which is a typical physiological response in stressful situations. In this study, a significant increase indicates that arithmetic tasks trigger a sympathetic 'fight or escape' response, leading to a tilt in the balance of the autonomic nervous system towards sympathetic activation. This enhanced sympathetic activity may be due to the complexity of the task itself and the nature of induced stress, which often leads to an increase in alertness and arousal levels.

When the human body experiences anxiety induced by laboratory stressors, its subjective psychological experience, behavioral performance, and physiological reactions will undergo corresponding changes[6]. However, subjective emotions and behavioral reactions are to some extent difficult to objectively and truly reflect anxiety levels, and the autonomic nervous system must accompany changes in anxiety emotions. By monitoring changes in heart rate variability LF/HF, the level of anxiety can be quantitatively measured. After successfully inducing anxiety in the participants, the lavender group continued to perform mental arithmetic while inhaling lavender for intervention. The heart rate variability LF/HF was significantly lower than during the stress period, indicating that lavender essential oil may rapidly act on the autonomic nervous system while the participants were facing mental arithmetic stress tasks. There is no difference compared to the baseline and recovery periods, indicating that the effect may be more persistent and stable.

Another interesting finding is that people's expectation of strong aroma effects leading people to relieve stress or divert attention, no matter what aroma they experience, will increase their processing speed. Some aromatherapy effects come from people's expectations. The coconut oil group, as a placebo, showed a slight decrease in expected heart rate variability LF/HF during intervention. Although coconut oil has no pharmacological effect on relieving stress, its friendly aroma may alleviate some anxiety to some extent due to its expected relieving effect. In future research, it would be interesting to test the effectiveness of other aromas in relieving stress. Different aromas have been known to have various effects on mood and cognition. Testing the efficacy of lavender aroma in comparison to other aromas, such as citrus or peppermint, could provide valuable insights into the potential therapeutic benefits of different scents.

In an interesting report, Martin reviewed the current evidence and clinical significance of the "role of olfactory stimulation in the alteration of cognition, mood, and social behavior," concluding there is "a common, if uneasy, relationship with the holistic practice of so-called aromatherapy[18]. The observed effects are the result of a mixture of mechanisms, including aroma-specific pharmacological profiles and user expectation. Thus, the overall outcome of aromatherapy exposure is influenced by different mechanisms, each of which can be tuned to have the most beneficial effect on function. Given that aromatherapy is a rapid intervention with minimal side effects, further research is warranted as it has the potential to enhance cognitive function in response to stress.

Traditional aromatherapy lacked the incorporation of physiological data as signals for applying treatment. The timing of the therapy was not optimal, as individuals had to proactively seek treatment. The design of smart diffuser offers better self-prevention measures for people with low level stress. We should conduct further research to precisely pinpoint the onset of stress and correlate it with the administration of aromatherapy interventions. This precision would enable experiments specifically designed to alleviate stress at the moment it emerges, potentially enhancing the effectiveness of such interventions. Moreover, investigations into optimal dosages and intervention durations for essential oils are crucial. Delving deeper into the relationship between the amount of essential oil used and the duration of exposure could reveal the minimum effective dose and the necessary timeframe for achieving significant stress reduction. Incorporating variables like cortisol levels, EEG scans, or heart rate variability can provide a more holistic understanding of the effects of aromatherapy.

Despite the valuable insights gained from this study, it is imperative to acknowledge its limitations. To bolster the representativeness of results, future research should aim for larger and more diverse participant pools, encompassing a broader spectrum of demographics. Other scents will also be considered to test the effectiveness of alleviating stress. The brief exposure to aromatherapy may not adequately capture the potential long-term effects of such interventions. Extending the duration of interventions would allow for a more comprehensive assessment of their sustained stress-relief potential. Lastly, additional research is imperative to enhance the synergy between the software and hardware components of AromaMate. This pivotal phase is critical in guaranteeing the device's pragmatic utility and effortless operation within authentic stress management contexts. These endeavors aimed at mitigating these constraints will undeniably facilitate the advancement of a more comprehensive comprehension of aromatherapy's viability in stress management.

5 Conclusion

In summary, AromaMate is a creative conceptual design that combines calm technology, and data-driven aromatherapy based on physiological markers. It could be a promising step towards customizing efficient stress management solutions for people in the workplace. Future research should strive to expand the participant population, extend experimental time, explore various essential oil concentrations and durations, and cover a wider range of physiological parameters. AromaMate helps stress management by providing customized, non-invasive, and timely intervention measures. In the dynamic landscape of stress management, AromaMate demonstrates the potential of calm technology to enhance our well-being.

About Qiurui Wang Qiurui Wang started her Engineering Doctorate in the Human System Interaction program at Department of Industrial Design, Eindhoven University of Technology in the Netherlands in 2022. She received her master's degree in Software Engineering (Information Product Design) from Zhejiang University in 2018. Her research spans emotional response, stress relief, emotional resonance, workplace vitality, and the development and application of psychological interventions. She particularly focuses on integrating physiological data and calm technology in the workplace to improve employee well-being and productivity. Through this innovative research, she advances the integration of digital art, workplace vitality, and mental health, aiming to create happier and healthier living and working environments.

References

1. Mashable homepage. https://mashable.com/article/cyrano-scent-playlist (2014), last accessed: 4 April 2014
2. Qiweiwangguo homepage. https://www.qiweiwangguo.com/ (2020)
3. Moodo homepage. https://moodo.co (2022)
4. Ali, B., Al-Wabel, N.A., Shams, S., Ahamad, A., Khan, S.A., Anwar, F.: Essential oils used in aromatherapy: A systemic review. Asian Pacific Journal of Tropical Biomedicine **5**(8), 601–611 (2015)
5. Case, A.: Calm technology: principles and patterns for non-intrusive design. " O'Reilly Media, Inc." (2015)
6. Freeman, R., Chapleau, M.W.: Testing the autonomic nervous system. Handbook of clinical neurology **115**, 115–136 (2013)
7. Geurts, S., Gründemann, R.: Workplace stress and stress prevention in europe. In: Preventing stress, improving productivity, pp. 9–32. Routledge (2012)
8. Heikkilä, K., Nyberg, S.T., Theorell, T., Fransson, E.I., Alfredsson, L., Bjorner, J.B., Bonenfant, S., Borritz, M., Bouillon, K., Burr, H., et al.: Work stress and risk of cancer: meta-analysis of 5700 incident cancer events in 116 000 european men and women. Bmj **346** (2013)
9. Hioki, K., Nozawa, A., Mizuno, T., Ide, H.: Physiological evaluation of mental workload in time pressure. IEEJ Transactions on Electronics, Information and Systems **127**(7), 1000–1006 (2007)
10. Hongratanaworakit, T.: Physiological effects in aromatherapy. Songklanakarin J Sci Technol **26**(1), 117–125 (2004)
11. Hosseini, S., Heydari, A., Vakili, M., Moghadam, S., Tazyky, S.: Effect of lavender essence inhalation on the level of anxiety and blood cortisol in candidates for open-heart surgery. Iranian journal of nursing and midwifery research **21**(4), 397–401 (2016)
12. Kasper, S., Anghelescu, I., Dienel, A.: Efficacy of orally administered silexan in patients with anxiety-related restlessness and disturbed sleep–a randomized, placebo-controlled trial. European neuropsychopharmacology **25**(11), 1960–1967 (2015)
13. Kianpour, M., Mansouri, A., Mehrabi, T., Asghari, G.: Effect of lavender scent inhalation on prevention of stress, anxiety and depression in the postpartum period. Iranian journal of nursing and midwifery research **21**(2), 197–201 (2016)
14. Kuo, N.T., Lin, L.P., Chang, K.C., Cheng, Y.S.: How emotional labor and job stress affect the job performance of tour leaders: Moderating effects of job characteristics and social media use intensity. International Journal of Hospitality & Tourism Administration **25**(1), 30–58 (2024)
15. Lawless, J.: The complete illustrated guide to aromatherapy: A practical approach to the use of essential oils for health and well-being. Element (1999)
16. Lehrner, J., Marwinski, G., Lehr, S., Johren, P., Deecke, L.: Ambient odors of orange and lavender reduce anxiety and improve mood in a dental office. Physiology & Behavior **86**(1-2), 92–95 (2005)
17. Leka, S., Jain, A., Organization, W.H., et al.: Health impact of psychosocial hazards at work: an overview (2010)
18. Martin, G.N.: Olfactory remediation: Current evidence and possible applications. Social Science & Medicine **43**(1), 63–70 (1996)
19. Milner, A., Scovelle, A.J., King, T.L., Madsen, I.: Exposure to work stress and use of psychotropic medications: a systematic review and meta-analysis. J Epidemiol Community Health **73**(6), 569–576 (2019)

20. Nieuwenhuijsen, K., De Boer, A., Verbeek, J., Blonk, R., Van Dijk, F.: The depression anxiety stress scales (dass): detecting anxiety disorder and depression in employees absent from work because of mental health problems. Occupational and Environmental Medicine **60**(suppl 1), i77–i82 (2003)
21. Pomeranz, B., Macaulay, R., Caudill, M.A., Kutz, I., Adam, D., Gordon, D., Kilborn, K.M., Barger, A.C., Shannon, D.C., Cohen, R.J.: Assessment of autonomic function in humans by heart rate spectral analysis. American Journal of Physiology-Heart and Circulatory Physiology **248**(1), H151–H153 (1985)
22. Rashidi-Fakari, F., Tabatabaeichehr, M., Mortazavi, H.: The effect of aromatherapy by essential oil of orange on anxiety during labor: A randomized clinical trial. Iranian Journal of Nursing and Midwifery Research **20**(6), 661–664 (2015)
23. Salovey, P., Mayer, J.D.: Emotional intelligence. Imagination, cognition and personality **9**(3), 185–211 (1990)
24. Seyyed-Rasooli, A., Salehi, F., Mohammadpoorasl, A., Goljaryan, S., Seyyedi, Z., Thomson, B.: Comparing the effects of aromatherapy massage and inhalation aromatherapy on anxiety and pain in burn patients: A single-blind randomized clinical trial. Burns **42**(8), 1774–1780 (2016)
25. Stansfeld, S., Candy, B.: Psychosocial work environment and mental health—a meta-analytic review. Scandinavian journal of work, environment & health pp. 443–462 (2006)
26. Sugawara, Y., Hino, Y., Kawasaki, M., Hara, C., Tamura, K., Sugimoto, N., Yamanishi, Y., Miyauchi, M., Masujima, T., Aoki, T.: Alteration of perceived fragrance of essential oils in relation to type of work: a simple screening test for efficacy of aroma. Chemical Senses **24**(4), 415–421 (1999)
27. Wiesmann, A., Mätzler, C., Weise, T.: Radiometric and structural measurements of snow samples. Radio Science **33**(2), 273–289 (1998)
28. Wu, C.Y., Lee, H.F., Chang, C., Chiang, H.C., Tsai, Y.H., Liu, H.E., et al.: The immediate effects of lavender aromatherapy massage versus massage in work stress, burnout, and hrv parameters: A randomized controlled trial. Evidence-Based Complementary and Alternative Medicine **2020** (2020)

IV

Sponsors

Ambianti B.V.

Ambianti develops high-tech building materials, combining cutting-edge digital intelligence, inventive minimalism and the flexibility of conventional building construction. Ambianti's mission is to reform architecture and industrial design by materializing innovative concepts, which integrate seamlessly into the established processes of the building sector and emancipate the way people interact with indoor environments. A number of proprietary advances in software, hardware and industrial design make it possible to realize the company's mission starting today. Patents pending.

Ambianti was founded in 2012 in Eindhoven, the Netherlands, by an international team of engineers, scientists and experts who share the same vision and passion: to make ambient technologies part of everyday life. Ambianti brings together diverse, high-level academic backgrounds, professional expertise and extensive work experience in engineering, computer science, architecture, building construction and industrial design.

Department of Industrial Design, Eindhoven University of Technology

We offer students and staff an international and academic, in other words intellectually stimulating, study and working environment that will inspire them to broadly-based personal development, social and cultural engagement and an entrepreneurial attitude. We are shaping a new discipline: designing and creating intelligent systems, products and related services. We combine cutting edge research in several domains to educate a new type of engineer.

Our industrial designers create connected solutions, combining products and services to suit individual needs. Both students and researchers can work with all stakeholders, customers and producers alike, who are not only involved, but are also essential contributors to the future sustainability of our world. Their role will be to mediate the developments in technology, the needs of people (culture & society) and the responsibility for properly using the power of ideas and values.

DEPARTMENT OF INDUSTRIAL DESIGN

EINDHOVEN UNIVERSITY OF TECHNOLOGY

PDEng becomes EngD same degree new name

Engineering Doctorate (EngD)
Designing Human-System Interaction