VA-PEPR
VOICE ASSISTANTS – PEOPLE, EXPERIENCES, PRACTICES AND ROUTINES

WORKBOOK TWO

Acknowledgment
The VA-PEPR (Voice Assistants - People, Experiences, Practices and Routines) project was funded by the Swiss National Science Foundation Sinergia Program with a grant of 2.23 million CHF (Grant Nr. CRSII5_189955). It is co-led by Dr. Sabine Junginger (lead) (HSLU School of Design, Film and Art), Dr. Ulrich Reimer (OST), Dr. Jens O. Meissner (HSLU School of Business) and Dr. Jon Rogers (Northumbria University). We would like to express our gratitude for the support received from the wider project team from the Lucerne University of Applied Sciences and Arts (HSLU), Ostschweizer Fachhochschule (OST), and Northumbria University. We are grateful for the financial support from The Interdisciplinary Clusters HSLU. Additionally, we extend our appreciation to all the study participants for their valuable contribution to the project and for allowing us to include their photos in this report. These photos were taken during the studies and are published here with our study participants' consent. Permission to use all images in this publication has been granted by the owners or are under the Creative Commons License.

Ethical Approval
The project, including the diary study phase, was pre-approved by the HSLU Ethics Committee.

DOI
https://doi.org/10.5281/zenodo.10259121

SITES.HSLU.CH/VA-PEPR

Contents

What is a Workbook?

Design workbooks can be considered as a design method tool. They recognise that ideas may develop slowly over time. Important issues and perspectives may emerge from multiple studies, ideas or speculations within the workbook.

On a more practical not they can also:

- Document a project, or part of a project
- Be a finished document that help show a process or be used to gain input into certain elements of a project
- Give a good overview of the work from a top level down to very granular details
- Act as an engaging Interim report
- Be a helpful way to show all project stakeholders how everyone works on a project
- Contain any or all of the following: literature review, contextual Review, research, studies, experiments, insight generation, ideation, idea development, testing/reactions

The following workbook has contributions from all the VA-PEPR project members. There are many different writing styles, perspectives and personalities included within. The end result is a fairly un-edited or un-styled publication. We feel this captures the messy yet exciting nature of working in such a large interdisciplinary team.

A selection of other project workbooks

The VA-PEPR Team

Aurelio Todisco - Project Contact Person

Aurelio Todisco recently completed his Master in Design at the Lucerne University of Applied Sciences and Arts (HSLU). He is interested in human-centered and participatory design processes. In the VA-PEPR research project, he is the contact person for study participants.

Aysun Aytac - Product Design

Aysun Aytac, PhD, design researcher at the Competence Centre Design and Management at Lucerne University of Applied Sciences and Arts (HSLU). She combines theories and approaches from material culture, visual culture, visual ethnography, everyday life, everyday life creativity and design thinking to explore interactions between people and artefacts.

Beat Tödtli, - Data Science and Data Analysis

PhD, data scientist, researcher and lecturer at the Eastern Switzerland University of Applied Sciences (OST). He develops and teaches machine learning courses and has a background in industrial algorithm development for fraud detection.

Bettina Minder - Innovation Processes

Bettina Minder holds a PhD in organisational design and innovation processes from Aalborg University. She focuses on participatory and collaborative research approaches that translates research insights into proposals and design briefs for actionable design.

Edith Maier - Social Anthropology and Applied Linguistics

Edith Maier, PhD, professor at the Eastern Switzerland University of Applied Sciences (OST) and has a background in applied linguistics, social anthropology and information science. More recently, her research focuses on cultural aspects of human-computer interaction and behavioural change support.

Inesa Halilovic - Information Systems

Inesa Halilovic is a research assistant and Master student in Information systems at the Eastern Switzerland University of Applied Sciences (OST). She works on various projects in application development.

Jens O. Meissner - Management and Organizational Studies

Jens O. Meissner, PhD, professor of organizational design and head of the interdisciplinary future laboratory CreaLab at Lucerne University of Applied Sciences and Arts (HSLU). His research focuses on questions around organizational resilience in the digital age.

Jon Rogers - Human Computer Interaction

Professor of design at Northumbria University with years of experience working at the interface between design and emerging technologies. Following a recent three year fellowship based in Mozilla Berlin, he is co-directing OpenDoTT, an EU funded doctoral training programme with Mozilla in trusted IoT.

Magalie Jost Naranjo - Research Coordination

Magalie Jost Naranjo is a research coordinator in the School of Art and Design at Lucerne University of Applied Sciences and Arts (HSLU). In VA-PEPR research project, she is responsible for coordinating the reporting to SNF and is responsible for the financial controlling and the financial reporting.

Matthias Baldauf - Information Systems

Matthias Baldauf, PhD, professor of information systems at the Eastern Switzerland University of Applied Sciences (OST). His research centres on pervasive computing and human-computer interaction with a recent focus on voice-based user interfaces.

Melanie Rickenmann - Mathematics

Melanie Rickenmann is a research assistant at the Eastern Switzerland University of Applied Sciences (OST). She works in various projects in data science and in application development.

Michael Doerk - Psychology

Michael Doerk, PhD, professor at the Lucerne University of Applied Sciences and Arts, a psychologist, expert in health promotion, social computing, risk and quality management and innovation management. He developed the award-winning HSLU business application 'relax-concentrate-create' .

Michelle Murri - Law

Michelle Murri, research assistant in the School of Business at Lucerne University of Applied Sciences and Arts (HSLU). She received her Master's degree in law and is particularly interested in data protection law. She supports the project management in VA-PEPR research project.

Mihn-Nguyet Le - Design Research & Intersectionality

Minh-Nguyet Le is a Researcher at Lucerne University of Applied Sciences and Arts. She is interested in more inclusive, egalitarian, and just ways of applying designing and therefore brings an intersectional lens in her approach to research and design. She is currently pursuing doctorate at University of Potsdam in Germany in the field of public administration.

Mike Shorter - Creative Technologist

Dr. Mike Shorter is a Research Fellow at Northumbria University. In his time, he has worked as a Researcher, Creative Technologist, Product Designer, Craftsperson and Innovation Strategist. Across all these roles Mike consistently explores new technology with reflective and playful processes in order to make meaningful objects and experiences.

Paola Pierri - Anthropology of Design & Technologies

Dr. Paola Pierri has a doctorate in Design Anthropology and a background in political theory. Her research explores the impact of digital technologies on societies and its implications for policy-making and for democracy more broadly. She has been researching on issues of digital inequalities as Research Fellow at the Weizenbaum Institute.

Patricia Wolf - Innovation Management

Patricia Wolf a professor of innovation management and future research at Lucerne University of Applied Sciences and Arts. She teaches at ETH Zurich and she is a professor of integrative innovation management at the University of Southern Denmark since 2018. Her research focuses on the link between innovation, emergent technologies and social change.

Sabine Junginger - Design Research

Sabine Junginger, PhD, professor in the Lucerne School of Art and Design and head of the Competence Centre Design and Management at Lucerne University of Applied Sciences and Arts. She contributes her expertise on human-centred design and studies its contribution to social change and economic innovation in the context of digital transformation.

Tom Ulmer - Data Analysis and IT Architecture

Lecturer at the Eastern Switzerland University of Applied Sciences (OST) and has a background in media design, computer sciences and human computer interaction. His main research interests include digital health, active assisted living as well as smart living.

Ulrich Reimer - Computer Science

Ulrich Reimer, PhD, professor at the Eastern Switzerland University of Applied Sciences and computer scientist with a background in AI, semantic technologies and knowledge management. His current research focuses on behavioural change support systems for digital health, with an emphasis on sensor data mining and self-learning to adapt to individual users.

Ute Klotz - Imformatics

Ute Klotz, PhD, co-leader of the focus group 'Technologies for Tomorrow's Digital Working World' in the interdisciplinary thematic cluster 'Digital Transformation of the Working World' at Lucerne University of Applied Sciences and Arts. She deals with future studies and the interfaces between humans, work and technology.

Uwe Riss - Information Systems

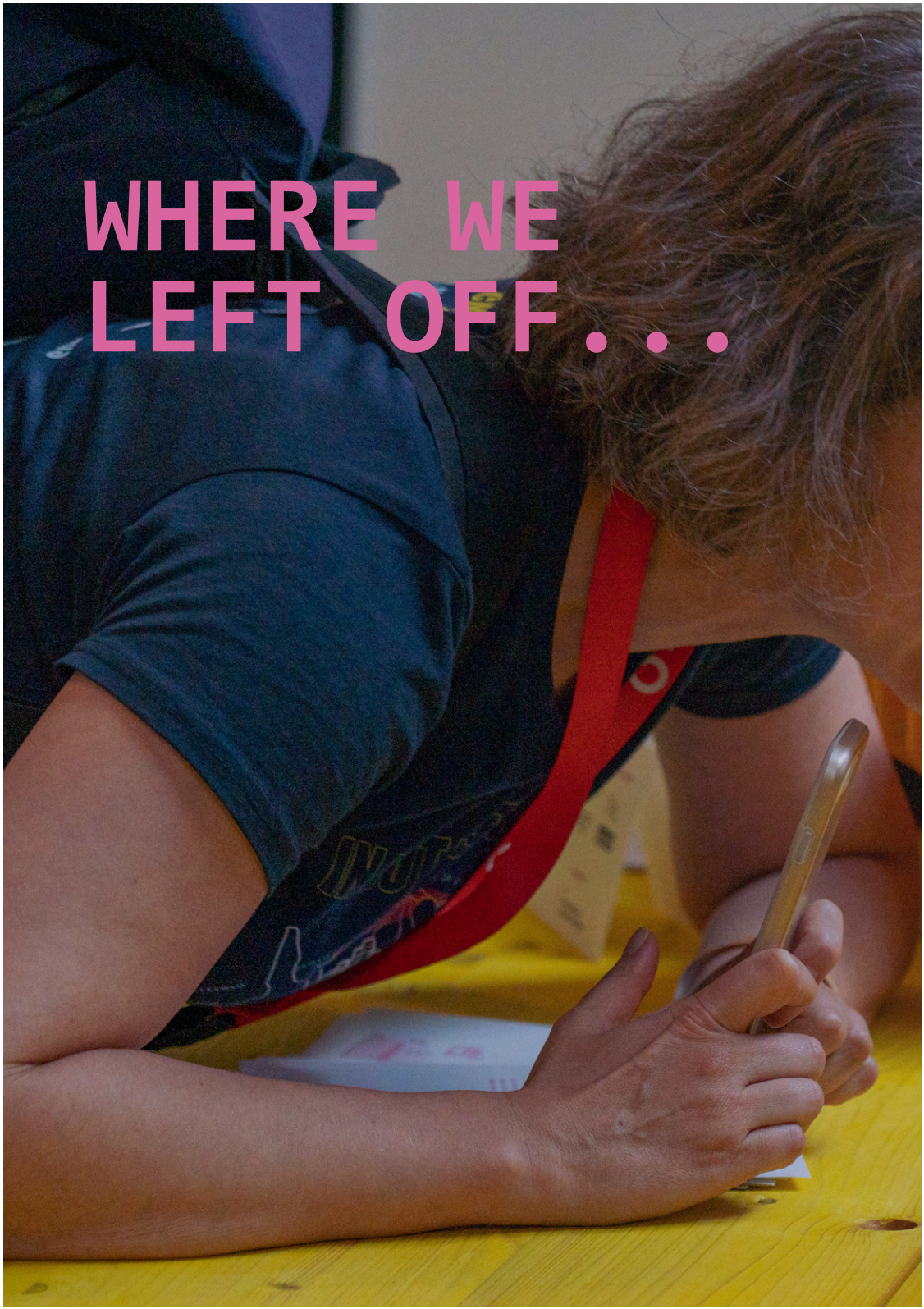
Uwe Riss, Ph.D., lecturer and senior researcher at the Eastern Switzerland University of Applied Sciences (OST), has conducted research in areas such as business information systems, natural sciences, and philosophy.

Currently, his main interest is in the various aspects of digitalization.

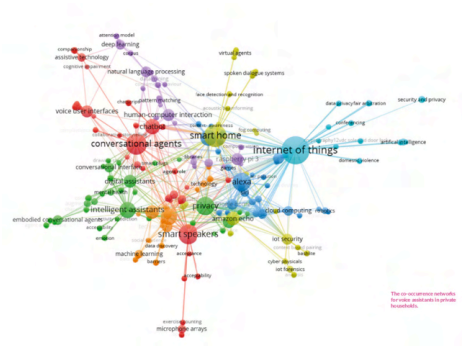
Vivien Luong - Research Coordination

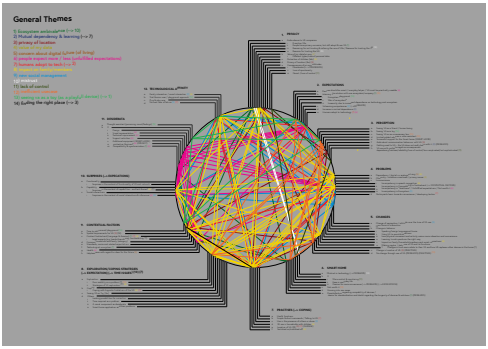
Vivien Luong is a research coordinator in the School of Art and Design at Lucerne University of Applied Sciences and Arts (HSLU). In VA-PEPR research project, she is responsible for coordinating the reporting to SNF and is responsible for the financial controlling and the financial reporting.

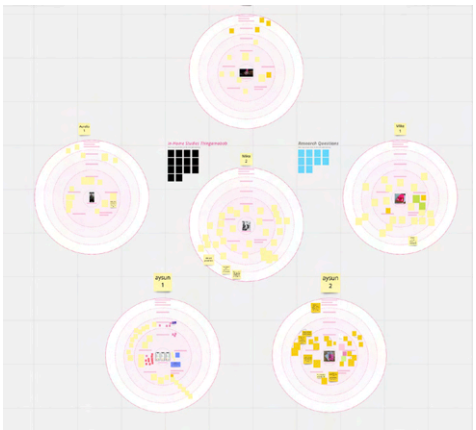
WHERE WE LEFT OFF...




actually
things
people







1 - Exploring Literature review themes
2 - Emerging topics from the In-Home Study
3 - The Network Analysis 'sniffers'
4 - A selection of the Provotypes
5 - Initial rcc Data
6 - A Miro board during one of the Speculative Design Workshops

Where We Left Off

At the end of VA-PEPR Workbook One we had completed six initial research studies exploring how people experience voice assistants in their homes and private lives and how they develop new practices and routines around their use of VAs. These studies focussed on the home environment, user experience and ethical issues. This research allowed the team to contribute to a deeper understanding of this new technology.

Literature/contextual review - An overview of existing literature and work in relevant sectors to the VA-PEPR project.

In-Home Study - An ethnographic study exploring how Swiss German speaking people use VAs in homes and incorporate them into their everyday lives, develop new practices and routines, affect the interactions and relations around the house as well as explore notions of privacy.

Network Analysis Study - Exploring network traffic in the domestic setting.

The rcc study - Exploring the impact VAs had on the routines and behaviours of students using the HSLU rcc web-based application.

Provotyping - Unpacking the complexities of Voice Assistants through a series of twelve provocative prototypes.

Speculative Design - Exploring voice assistant futures through a series of workshops engaging with the general public and industry experts.

If you are interested in learning more about these studies please seek out the VA-PEPR Workbook One (https://doi.org/10.5281/zenodo.10259119), alternatively please request a physical copy by sending an email to michael.shorter@northumbria.ac.uk

THE VA-PEPR CHALLENGES

The VA-PEPR Challenges

We have generated many insights and discoveries through the Literature and Contextual review, the In-Home study, the Network Analysis Study, The rcc study and the Provotype Workshops. There are many commonalities across these studies. There is a point where we need to step back from the research as a group and decide where we want to focus our attention. It is our jobs as researchers to find the common, and interesting, threads linking insights across these studies, but also to spot where there are gaps in the research. As a group we generated challenges that we felt were pertinent to the research in hand.

Each VA-PEPR Challenge had a theme, a problem statement, an insight (a-ha moment), a brief and a reference to anchor the challenge in the real world. We asked that each VA-PEPR Challenge:

- **Framed the challenge as an opportunity (How might we...)**
- **Had a focus on a 'Who' - Swiss people**
- **Had a focus on a 'What' - Voice assistants**
- **Had a focus on a 'Where' - In the home**
- **Must tie into the social side of VAs (behaviours and routines)**
- **Must be routed in a problem we've discovered**
- **Must relate to an insight or opportunity we've generated**
- **Must anchor the brief in some key literature or reference.**

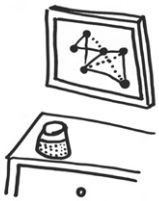
The next few pages show all the VA-PEPR Challenges.

Challenge Theme: Making the cloud tangible

VA-PEPR

Problem Identified
Voice technology is an example of how the end-user has limited visibility into the operations of the cloud. (Limited overview of data flows, privacy, setting options, etc.) They felt exposed to technology.

VA-PEPR Insight
When we make our network data visible it may not give us any agency, but it can allow for education and behavioural changes.



VA-PEPR Challenge
How might we make the cloud processes of VA technology more tangible/understandable to participants via a visual channel/artefact? This would explore how VA users would change behaviours and routines with this more informative device.

Key References:
Microsoft Physical Charts
Mozilla LightBeam

Challenge Theme: Trust

VA-PEPR

Problem Identified
Users were frustrated by a VAs lack of feedback - virtually and visually - this typically led to a lack of trust. Users would double check requests on a different device or enquire of the content source.

VA-PEPR Insight
A small visual element could improve trust between a user and their VA. This would allow users to be more involved in a conversation with a VA.

VA-PEPR Challenge
How might we help users trust their VA more.
1. How can we make users trust their requests have been met.
2. How can we show the source of content being provided from the VA is trustworthy.
3. Consider how the research will investigate participants' behaviours and routines in relation to ethical issues such as privacy and trust in their homes.

Key References:
Mozilla LightBeam

Challenge Theme: Privacy/Smartness

VA-PEPR

Problem Identified
Participants were unsure if a VA could ever become a threat to them if it knew too much information.

VA-PEPR Insight
Are our VAs user modelling or modelling users?

VA-PEPR Challenge
How might we allow VA users to have control or even transparency over device preferences and GDPR. Can we put something in place that gives users reassurance that VAs are well intentioned?

Key References:

Challenge Theme: Smartness

VA-PEPR

Problem Identified
Societal control of engineering strategies.

VA-PEPR Insight
"We want VAs to be smart" But smart for whom?

VA-PEPR Challenge
How might we provide regulations/processes that industry need to comply to or follow?

Key References:
Trustable Technology Mark

Challenge Theme: Smartness

VA PEPR

Problem Identified
Users are disappointed by the contextual awareness of a VA (In-Home study)

VA-PEPR Insight
VAs have the technology and intelligence to make sense of audio to build better contextual models of their environment.

VA-PEPR Challenge
How might we make a VA better understand its environment and context. Consider using audio context as well as adding new sensors or devices to the VA ecosystem.

Key References:
Chatter Baby

Challenge Theme: Application/Form/Affordances

VA PEPR

Problem Identified
The applications for a VA are not clear due to its 'black box' nature (In-Home, NTA, Provotypes).

VA-PEPR Insight
iPhones used games to help users understand what an app could do - making use of the sensors and tech built into a device. We had to be taught the affordances of this new technology.

VA-PEPR Challenge
How might we hero the technological capabilities of a VA in the context of a toy or game in order to educate users on the devices capabilities?

Key References:
How Apple Taught the World to Smartphone - WIRED

Challenge Theme: Data/Privacy

VA PEPR

Problem Identified
Today's data generators (users) are giving away their data in return for services without having any control on how their data is used. Data value generation happens far down the data analysis pipeline, where home owners no longer have any control. Also, data producers are not well coordinated and give away their data before any substantial value is apparent from aggregated data.

VA-PEPR Insight
The dominant business model today is the one of big tech giants owning access rights, and data providers interact with them directly, having little power to demand any return value beyond the services offered. In particular, value generated by data users is not remunerated to the data providers in any form.

VA-PEPR Challenge
How might me develop stories of what a privacy-protecting intermediary would mean, exploring situations for all involved parties. Explore the interconnections of data owners, data intermediaries and data users. Explore flexible options/products on how users can control their privacy and on how to invest their data.

Key References:
Unlocking the value of data: Exploring the role of data intermediaries.

Challenge Theme: Data

VA PEPR

Problem Identified
Participants in the In-Home study and NTA study were unsure of to what extent their voice data was being captured by a VA and also the implications of surrendering their voice data.

VA-PEPR Insight
The Privacy Paradox only exists with VAs because users due to a lack of understanding around our data collection and flow, as well as a feeling of disconnect from the implications as the data is merely a drop in the ocean far far away.

VA-PEPR Challenge
How might we make our voice data visible and understandable, allowing us to have more control over it?
1. Consider how the research (object) will investigate participants' behaviours and routines in relation to VA data in their homes.
2. Embed the research into a speculative narrative (design fiction) that locates this object into the wider findings and objectives of the VA-PEPR project.

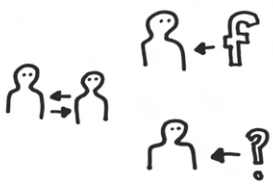
Key References:
Mozilla Common Voice
BBC R&D - The Future of the Living Room

Challenge Theme: Privacy

VA PEPR

Problem Identified
Users understanding of VAs and privacy can be murky at best. There is a lack of understanding of what users need to keep private, and how to achieve this.

VA-PEPR Insight
There are three kinds of privacy - against other users, against service providers, against intrusion.



VA-PEPR Challenge
How might we educate VA users on the different stakeholders that they may want to enforce privacy blockers against.

Key References:

Challenge Theme: Cultural context

VA PEPR

Problem Identified
VA is considered for daily use only in some contexts.

VA-PEPR Insight
IT-Students were emphasising that VA does only play a marginal role for the live in shared flats where – mostly temporarily. It does not make sense to invest money and time. Indian students where emphasising that VA is too expensive for most people.

VA-PEPR Challenge
How might we demonstrate the dependencies between usefulness of VA and cultural contexts?

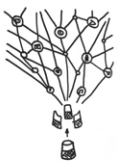
Key References:
Exploring Cultural Awareness When Designing Voice UX. Erica Pang

Challenge Theme: Form

VA PEPR

Problem Identified
Voice assistants are like a T.A.R.D.I.S*. The form of the VA does not accurately give the affordances of its functionality or capabilities (Provotypes).

VA-PEPR Insight
Users lack a mental model of the relationship between the form of a VA and its relationship to the cloud.



VA-PEPR Challenge
How might we communicate the functionality and capability of a voice assistant through design?
1. Consider how the research will investigate participants' behaviours and routines in relation to ethical issues such as privacy and control in their homes.
2. Embed the research into a speculative narrative (design fiction) that locates this research into the wider findings and objectives of the VA-PEPR project.

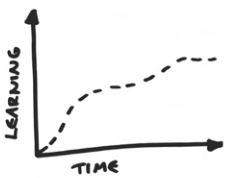
Key References:
Crawford, K. and Joler, V. (2018) An Anatomy of an Ai System.
*Wallace, J., Rogers, J., Shorter, M., Thomas, P.,Cook, R. (2019). The SelfReflector: Design, IoT and the High Street. CHI 2019.
https://thedoctorwhosite.co.uk/tardis/
Mozilla Common Voice

Challenge Theme: Smartness and Learning

VA PEPR

Problem Identified
Participants did on various occasions mention that they needed to learn - e.g. how to address VA; what the right key words were; how to link up VA with other devices etc. Conversely, some participants also had the feeling that VA learned - e.g. to understand Swiss German better - or it learned about the address and closest shops etc.

VA-PEPR Insight
In the discussion in St. Gallen we understood that "there is no learning taking place." Instead it might just be something that users experience?



VA-PEPR Challenge
How might we explore this perceived learning curve? An assessment/ test we need to take (or pass) after a certain period of VA use? Can we make VA pass a test after an initial phase of 'serving' in a household?

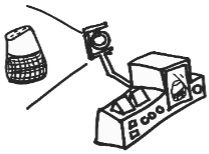
Key References:

Challenge Theme: Smartness

VA PEPR

Problem Identified
In home participants expectations of the smartness of VAs were never met.

VA-PEPR Insight
Users understanding of the terms smartness and intelligence are blurred. This coupled with a lack of language to determine the smartness or intelligence on a VA left users a lost.



VA-PEPR Challenge
How might we demonstrates how a VA intelligence scale and test would work in Switzerland. Consider how the research (object) will investigate participants behaviours and routines with a range of possible intelligences of VAs in their homes. This should include VAs that have very limited (current day) capability to VAs that have beyond-human general intelligence. Embed the research object into a speculative narrative (design fiction) that locates this object into the wider findings and objectives of the VA-PEPR project.

Key References:
Herczeg, M. (2011) 'The smart, the intelligent and the wise.
The Reith Lectures, Stuart Russell - Living With Artificial Intelligence,
The Biggest Event in Human History
Philip K Dick - The Voight Kampf Machine

Challenge Theme: Smartness

VA PEPR

Problem Identified
In-home study participants complained that the VA is not smart enough, e.g. doesn't understand a question, doesn't allow for a dialog (one-shot question and answer only). Expectations haven't been met, many people do not find a VA useful enough and soon lose interest, except for limited usage scenarios such as turning lights on and off.

VA-PEPR Insight
Smartness can mean a lot of things and it is mostly unclear what the users actually wish. Smartness is multidimensional, rather than linear.

VA-PEPR Challenge
Identify the various dimensions of smartness (depth of understanding, dialog capabilities, user modelling, etc.), and for each dimension degrees of capabilities be identified. Explore which smartness capabilities (over the various dimensions) are most relevant to users and have the biggest impact on user satisfaction - allowing for recommendations for the future design of VAs.

Key References:
Moussawi, S., Koufaris, M. & Benbunan-Fich, R. (2020). How perceptions of intelligence and an-thropomorphism affect adoption of personal intelligent agents.
Sun, N. & Botev, J. (2021). Intelligent autonomous agents and trust in virtual reality.

Selecting the Challenges

All the proposed challenges were on brief for the VA-PEPR project. They all explore how people experience voice assistants in their homes and private lives and how they develop new practices and routines around their use of VAs. They focus on the Swiss/German speaking home environment, user experience and ethical issues.

To help us narrow down these challenges to the ones that the team wanted to progress with, we decided on some methods and outputs that we wanted to see from the project, and understand how these could bring value to the various VA-PEPR Challenges.

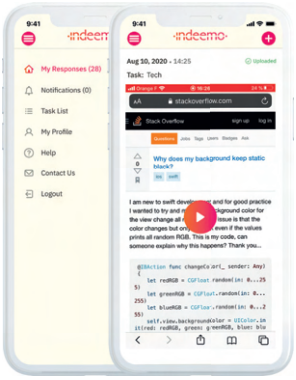
After this, we displayed each Challenge in an 'art gallery' style environment and spent time as a group reading them and writing notes - exploring if any of the challenges could be merged together to strengthen the challenge. This left us with five Challenges. For each Challenge we decided on a lead and team to drive the work associated with this Challenge. The outputs for each challenge were left very open in terms of number of responses and form of responses. They could be anything from prototypes to short stories, from studies to workshops. The five Challenges developed in parallel at different speeds.

METHOD

Diary

HOW: A method which allows us to capture the lives/experiences of people from their eyes as they happen. Participants are asked to keep a diary for a certain period about a product/service/experience which can include texts, photos, videos, drawings, etc..

WHY: Recording events/experiences as they happen in their natural settings. No presence of the researchers meaning participants feel better to express their private reflections.



METHOD

Qualitative Video Analysis

HOW: Research videos are interpreted and analysed by researchers.

WHY: Video can capture a subtle, multi-layered record of situations and environments we observe and study. It captures not only what is said but how it is said, actions and reactions, the appearance of the environment and individuals and how they interact, and so much more.

METHOD

Fictional Dialogues

HOW: Speculating and creating fictional dialogues between people and/or things. These fictional dialogues can look like written stories, videos or roleplay.

WHY: Fictional dialogues can aid researchers in exploring future scenarios without the boundaries of what is possible. They can be quick to produce without much resource. This method is closely related to Speculative Design.

METHOD

Documentary Method

HOW: Qualitative analysis of written documents or imagery.

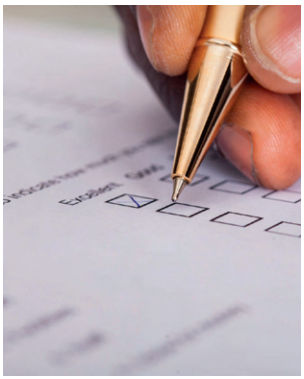
WHY: Documentary method of interpretation is used to categorise, investigate, interpret and identify the limitations of physical sources, most commonly written documents but also imagery.

METHOD

Questionnaires

HOW: A research instrument that consists of a set of questions to collect information from a respondent.

WHY: Offers a fast, efficient, and inexpensive means of gathering large amounts of information from sizeable sample volumes. These tools are particularly effective for measuring subject behaviour, preferences, intentions, attitudes and opinions.

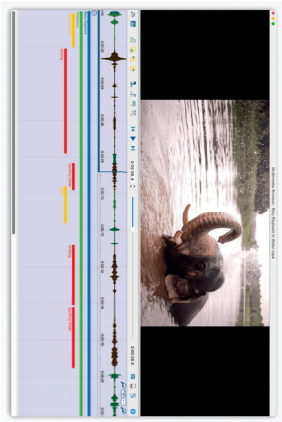


METHOD

Wizard of Oz

HOW: Wizard of Oz (WoZ) is a method where participants interact with a system that they believe to be autonomous, but in reality, is controlled by an unseen human operator in the next room.

WHY: To convey and test the experience of hard to create or impossible scenarios rapidly.

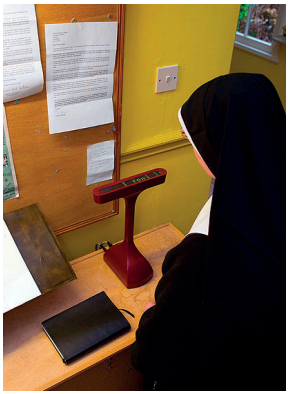


METHOD

Deployment

HOW: Deploying research objects with participants to live with for a period of time.

WHY: Allowing participants to live with something for a period of time gives space for views and use to change. Deployments are typically accompanied with ongoing check-ins such as interviews.

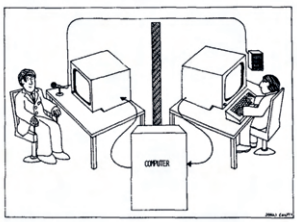


METHOD

Digital Probes

HOW: A digital object that is given to participants to record specific events, feelings or interactions.

WHY: Digital probes serve as a novel method for collecting information on participants lives, values and thoughts.

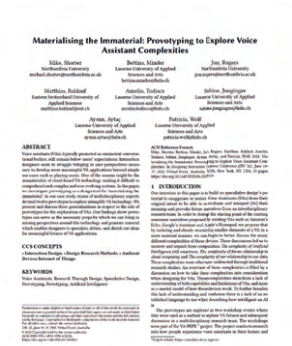


The methods selected by the team to work into the VA-PEPR Challenges.

OUTPUT

Academic Paper

EXAMPLE: A paper/article submitted to a conference to gain peer review and discuss original contributions to knowledge.



Materialising the Immaterial: Provotyping to Explore Voice Assistant Complexities

Mike Skarter, Northumbria University, michael.skarter@northumbria.ac.uk
Bettina Nissler, Lucerne University of Applied Sciences and Arts, bettina.nissler@hslu.ch
Jon Rogers, Northumbria University, jon.rogers@northumbria.ac.uk
Matthias Rolland, Eastern Switzerland University of Applied Sciences, matthias.holland@ost.ch
Annelin Todisco, Lucerne University of Applied Sciences and Arts, annelin.todisco@hslu.ch
Sabine Jungjurger, Lucerne University of Applied Sciences and Arts, sabine.jungjurger@hslu.ch
Ayman. Ayhar, Lucerne University of Applied Sciences and Arts, ayman.ayhar@hslu.ch
Patricia. Wolf, Lucerne University of Applied Sciences and Arts, patricia.wolf@hslu.ch

ABSTRACT

CCS CONCEPTS

KEYWORDS

1 INTRODUCTION

/A PEPR

OUTPUT

Video

EXAMPLE: A way of capturing and documenting a project beyond the life of the funding/prototypes. The video can present itself in many ways from documentary style content to adverts.

/A PEPR

OUTPUT

Prototypes

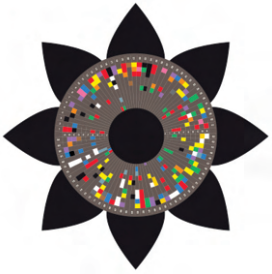
EXAMPLE: An object/service that is built to test a concept or process. Prototyping enables the designer to communicate ideas to a wide audience in response to a defined problem.

/A PEPR

OUTPUT

Infographics

EXAMPLE: Infographics are graphic visual representations of information. Because people process visual data better than any other type of data.



VA-PEPR

OUTPUT

Service Robots

EXAMPLE: Service robots assist human beings, typically by performing a job that is dirty, dull, distant, dangerous or repetitive. They typically are autonomous and/or operated by a built-in control system, with manual override options

VA-PEPR

The outputs selected by the team to work into the VA-PEPR Challenges.

1. What if your VA could have added inputs and outputs?
2. What if you could dial something up and down?
3. What if your VA was not associated with a big tech provider?
4. What if your VA looked like a Speaker?
5. What if you could instantly block your VA
6. What if your VA had eyes?
7. What if the VA looked like a microphone rather than a speaker?
8. What if a VA had only one function?
9. What if a VA could print off everything it heard that day?

(Opposite page)
Development pictures of some of the provotypes.

Challenge Theme: Smartness

VA PEPR

Problem Identified

In-home study participants complained that the VA is not smart enough, e.g. doesn't understand a question, doesn't allow for a dialog (one-shot question and answer only). Expectations haven't been met, many people do not find a VA useful enough and soon lose interest, except for limited usage scenarios such as turning lights on and off.

VA-PEPR Insight

Smartness can mean a lot of things and it is mostly unclear what the users actually wish. Smartness is multidimensional, rather than linear.

VA-PEPR Challenge

Identify the various dimensions of smartness (depth of understanding, dialog capabilities, user modelling, etc.), and for each dimension degrees of capabilities be identified. Explore which smartness capabilities (over the various dimensions) are most relevant to users and have the biggest impact on user satisfaction - allowing for recommendations for the future design of VAs.

Key References:

Moussawi, S., Koufaris, M. & Benbunan-Fich, R. (2020). How perceptions of intelligence and an-thropomorphism affect adoption of personal intelligent agents.

Sun, N. & Botev, J. (2021). Intelligent autonomous agents and trust in virtual reality.

Challenge Theme: Data

VA PEPR

Problem Identified

Participants in the In-Home study and NTA study were unsure of to what extent their voice data was being captured by a VA and also the implications of surrendering their voice data.

VA-PEPR Insight

The Privacy Paradox only exists with VAs because users due to a lack of understanding around our data collection and flow, as well as a feeling of disconnect from the implications as the data is merely a drop in the ocean far far away.

VA-PEPR Challenge

How might we make our voice data visible and understandable, allowing us to have more control over it?

1. Consider how the research (object) will investigate participants' behaviours and routines in relation to VA data in their homes.
2. Embed the research into a speculative narrative (design fiction) that locates this object into the wider findings and objectives of the VA-PEPR project.

Key References:

Mozilla Common Voice

BBC R&D - The Future of the Living Room

Challenge Theme: Smartness and Learning

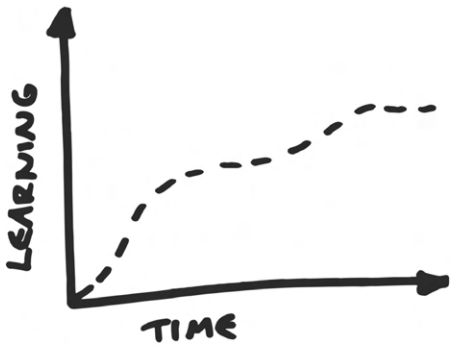
VA PEPR

Problem Identified

Participants did on various occasions mention that they needed to learn - e.g. how to address VA; what the right key words were; how to link up VA with other devices etc. Conversely, some participants also had the feeling that VA learned - e.g. to understand Swiss German better - or it learned about the address and closest shops etc.

VA-PEPR Insight

In the discussion in St. Gallen we understood that "there is no learning taking place." Instead it might just be something that users experience?



VA-PEPR Challenge

How might we explore this perceived learning curve? An assessment/ test we need to take (or pass) after a certain period of VA use? Can we make VA pass a test after an initial phase of 'serving' in a household?

Key References:

Challenge Theme: Trust

VA PEPR

Problem Identified
Users were frustrated by a VAs lack of feedback - virtually and visually - this typically led to a lack of trust. Users would double check requests on a different device or enquire of the content source.

VA-PEPR Insight
A small visual element could improve trust between a user and their VA. This would allow users to be more involved in a conversation with a VA.

VA-PEPR Challenge
How might we help users trust their VA more.
1. How can we make users trust their requests have been met.
2. How can we show the source of content being provided from the VA is trustworthy.
3. Consider how the research will investigate participants' behaviours and routines in relation to ethical issues such as privacy and trust in their homes.

Key References:
Mozilla LightBeam

Challenge Theme: Data/Privacy

VA PEPR

Problem Identified

Today's data generators (users) are giving away their data in return for services without having any control on how their data is used. Data value generation happens far down the data analysis pipeline, where home owners no longer have any control. Also, data producers are not well coordinated and give away their data before any substantial value is apparent from aggregated data.

VA-PEPR Insight

The dominant business model today is the one of big tech giants owning access rights, and data providers interact with them directly, having little power to demand any return value beyond the services offered. In particular, value generated by data users is not remunerated to the data providers in any form.

VA-PEPR Challenge

How might me develop stories of what a privacy-protecting intermediary would mean, exploring situations for all involved parties. Explore the interconnections of data owners, data intermediaries and data users. Explore flexible options/products on how users can control their privacy and on how to invest their data.

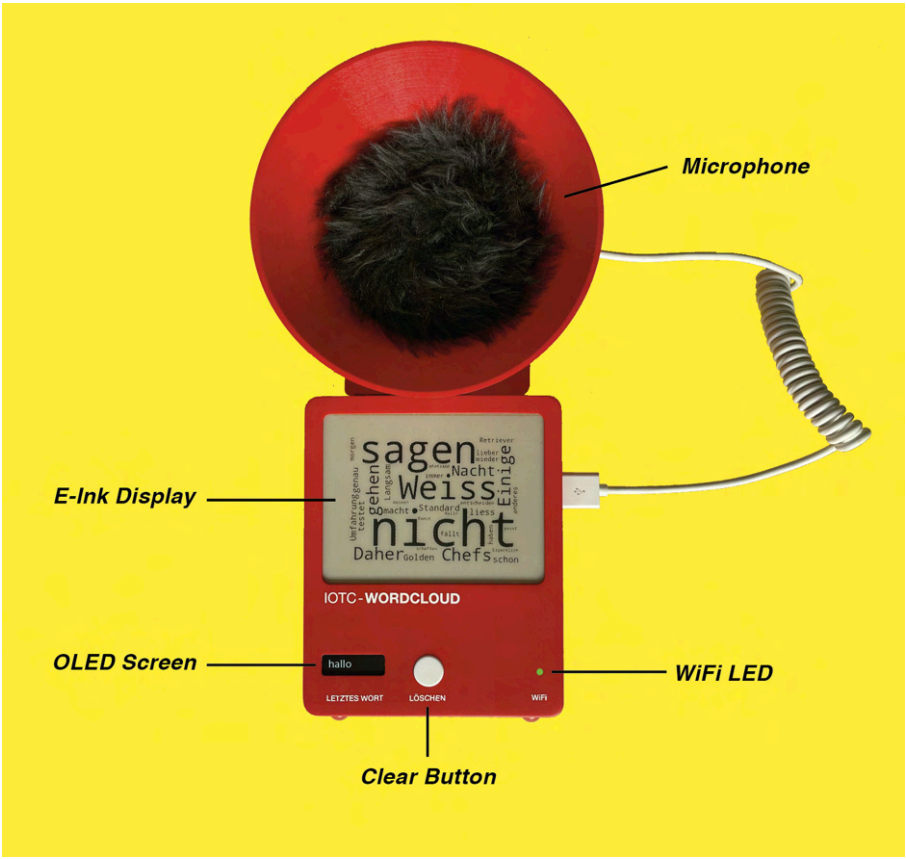
Key References:

Unlocking the value of data: Exploring the role of data intermediaries,

CHALLENGE 01: MATERIALISING THE IMMATERIAL

The WordCloud

The WordCloud is a research product that allows people to understand how VAs hear the human voice. This research product is a response to the research question of 'Can Designing in friction to VAs change domestic behaviours and routines?'The WordCloud consists of a microphone that captures audio, an E-Ink display and an OLED display. Listening to the environment it has been placed in, the device presents spoken words it recognises into visible text and produces a WordCloud on the E-Ink display. The E-Ink display is not entirely dynamic and refreshes at a set time delay. The small OLED display shows users the last word the device has heard to provide instant feedback of its functionality to the user.



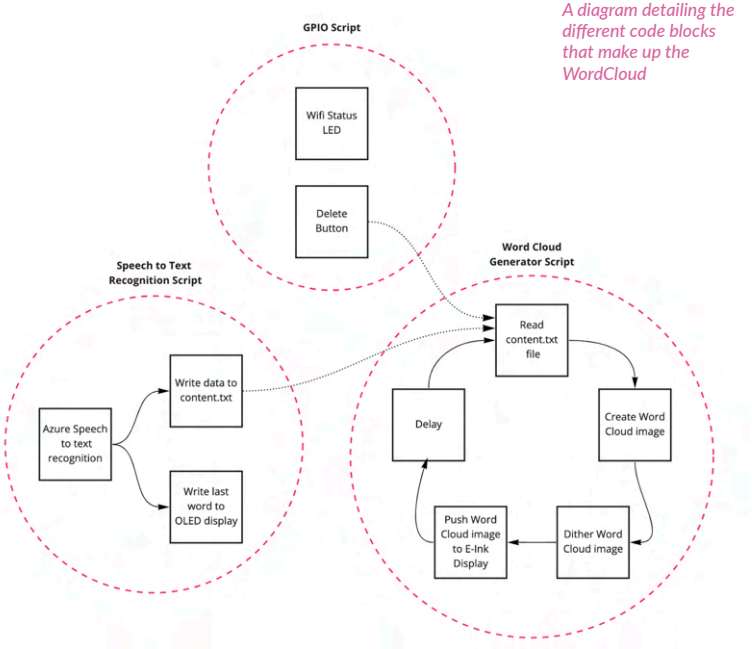
The WordCLoud

The WordCloud is predominantly designed to be a glanceable display with limited controls. A single button allows users to clear the memory of the WordCloud, i.e., delete all the words that the device had heard.

The core hardware of the WordCloud consists of a Raspberry Pi 4, an array of four high performance digital microphones (ReSpeaker Mic-Array V.2.0), an E-ink Display (Inky WHat), and an OLED display (PiOLED). The Raspberry Pi was running Ubuntu and the real-time speech to text recognition was achieved using Microsoft Azure Cognitive Services. During the development of the WordCloud the researchers found only two APIs that had real-time speech to text capabilities that would run on a Raspberry Pi with any sort of competence - Microsoft's Azure Cognitive Services and Assembly.ai. The research team selected the Microsoft API as it could translate many languages, including Swiss German, whereas Assembly.ai was only effective in English.

All the WordClouds used the remote terminal access service Dataplicity, allowing the researchers to not only debug but also update and tweak software whilst the devices were on deployment.

A diagram detailing the different code blocks that make up the WordCloud



Developing the WordCLoud

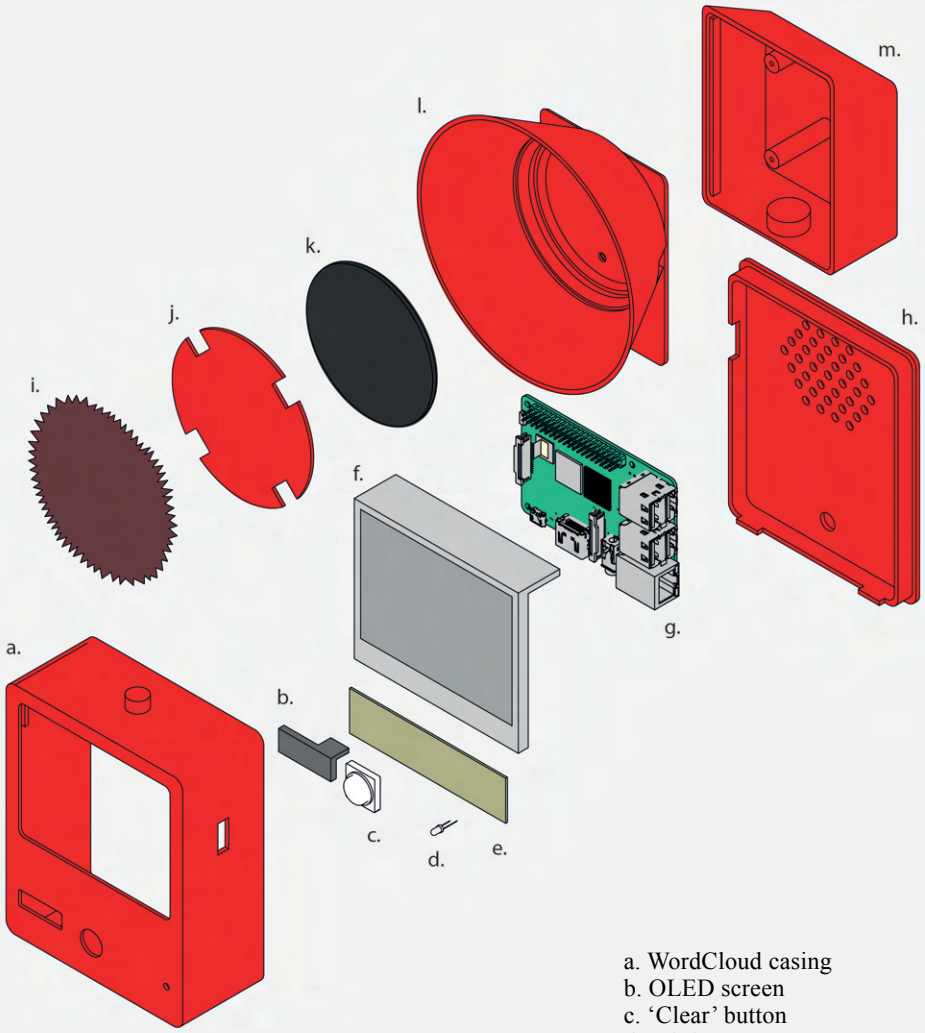
Device Settings

The WordCloud research product has three parameters that had to be decided on to make the device as engaging and meaningful as possible: minimum word length displayed on the word cloud, quantity of words displayed on the word cloud and refresh rate of the E-Ink display. Initial user testing revealed that the WordCloud was most interesting to view with words of six letters or more, a “cloud” of 40 words and a refresh rate of two minutes. These settings appeared to display interesting words with a feedback rate that meant users could observe changes with no long waits.

Component Selection

We decided to display a word cloud on an E-Ink display. The words heard more frequently were displayed larger than those mentioned less frequently. After hours of testing, we set three parameters on the WordCloud. Firstly the refresh rate; the screen would refresh the data displayed to the user every two minutes. Secondly the amount of words displayed on the screen; this was set to 30 allowing for as much content to be displayed as possible, while remaining readable. Thirdly the length of the word displayed; the WordCloud only displayed words of five letters or more to cut out smaller, more common and less notable words. This format of screen could exist in people’s homes without the glaring properties of a standard digital screen.

The OLED was not part of the initial design of the WordCloud. The output on the original design was purely the E-Ink display. This initial design was critiqued during workshops. It was quickly discovered during these workshops that users wanted immediate feedback. Workshop participants perceived the device to be not functioning properly as the words they were saying weren’t showing up on the E-ink display. It became clear that the participants needed feedback that the WordCloud was turned on and listening to what they were saying. And most importantly, users needed to know how the WordCloud was interpreting their voice. A user may say “are you stupid” but the WordCloud hears “are you stew pit”. To overcome this, a small OLED screen was used to display the last word the device had heard, or more to the point, thought it had heard.



An exploded diagram of the physical components that make up the WordCloud

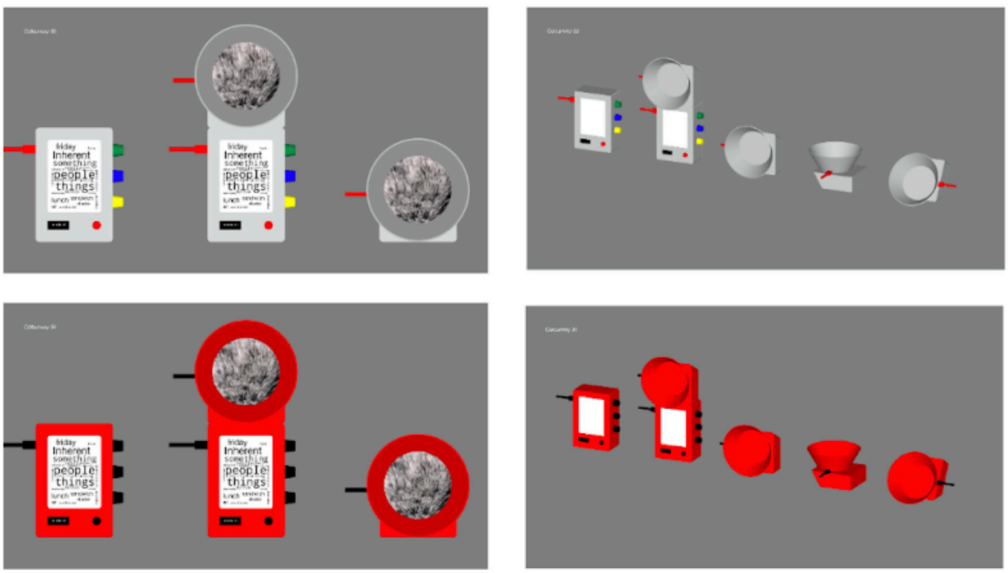
a. WordCloud casing
b. OLED screen
c. ‘Clear’ button
d. WiFi LED
e. Circuit board
f. E-Ink screen
g. Raspberry Pi 4
h. WordCloud casing
i. Microphone muff
j. Microphone muff plate
k. ReSpeaker microphone array
l. Microphone casing
m. Microphone casing

The ‘Löschen’ (German for ‘clear’) button was added to the WordCloud as we noticed that very little change would happen on the screen over longer periods of time (more than an hour of capturing speech). There would be a select few commonly said words dominating the word cloud. The ‘clear’ button simply deletes the content of the text file that all the heard words are saved to. The ‘clear’ button was used to erase all previous memory. It clears the history and is the starting point of taking a reading from the WordCloud, no matter how long the user wanted to run the device for before taking a reading by looking at the display.

Physical Form

It is in the nature of instruments to stand out and be clear. The WordCloud was designed to behave like an instrument. Therefore the WordCloud was designed not to blend into the background, but rather stand out, and make clear its intentions through design affordances. The red and white colour-way achieve this, and is also a cultural nod to the Swiss/German speaking people the WordCloud was designed for.

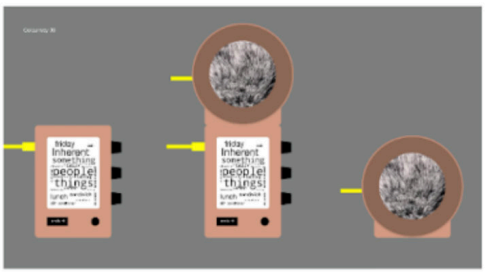
A common issue we had with existing voice assistants was that their form is akin to a smart speaker, which potentially misguides users of the actual

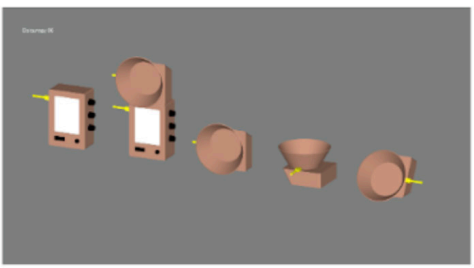


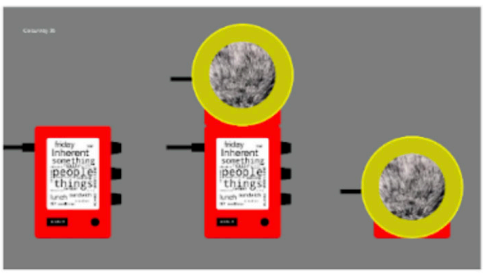
Initial designs and colour-ways for the WordCloud

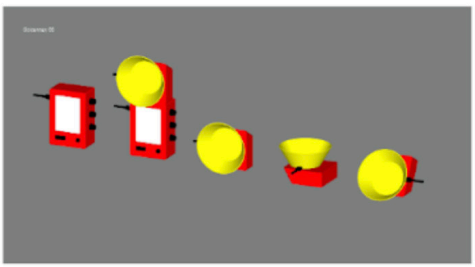
functionality of these devices. Through the design of the WordClouds form it was intended to make it clear that the Instrument was capturing audio from its environment. This led to the conical microphone design with the recognisable microphone muff material covering the technology.

The microphone was designed to be angled in different positions, depending where the user wanted to focus the listening power of the WordCloud in response to different household situations. The microphone could even be turned upside down to dampen its abilities to listen to its environment. Another difference the designers wanted to create to standard voice assistants was for the 'turn off mic' function to be definitive. On standard voice assistants a button indicates the microphone has been disabled by simply illuminating a LED. This LED has a disconnect from the microphone - the microphone could still well be active, the voice assistant just doesn't respond to the user. With the WordCloud, we wanted there to be no doubt that the microphone was turned off if users desired privacy. This was achieved by making it possible to unplug the microphone hardware from the main unit entirely. This also kept the device modular, similar to other scientific instruments, where different sensors and probes can be connected to base units.



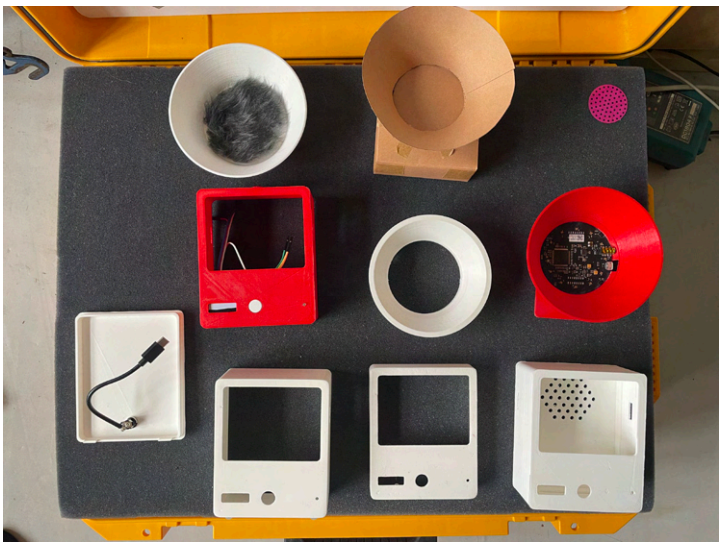






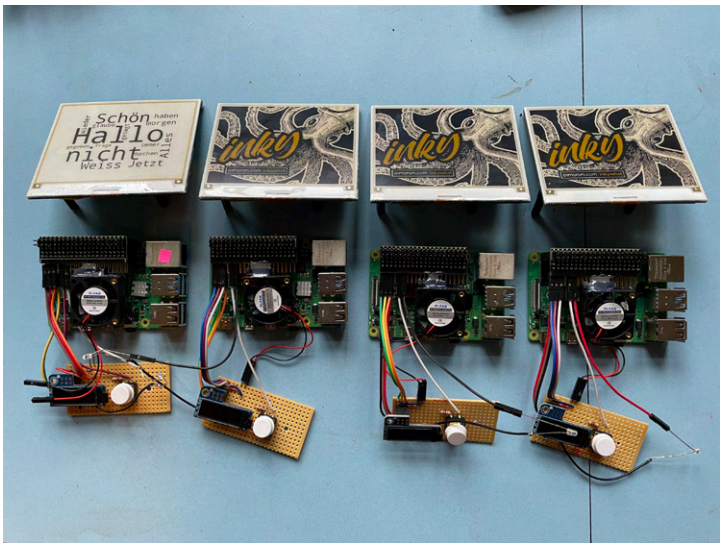
Test 3D print of the WordCloud

Preparing the WordClouds for the journey from Liverpool to Luzern



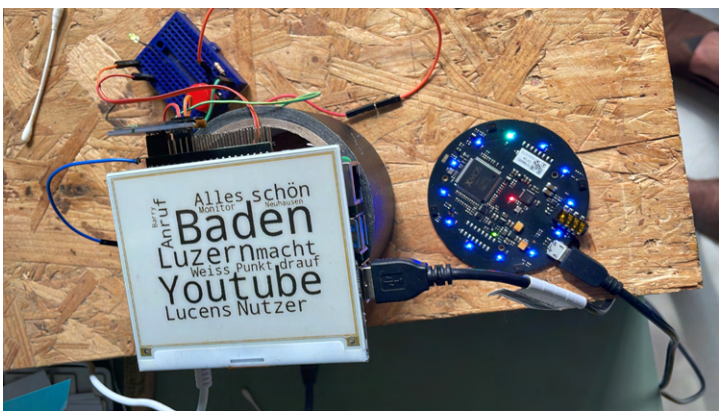
Iterations of the design of the WordCloud.

Developing the WordCloud hardware

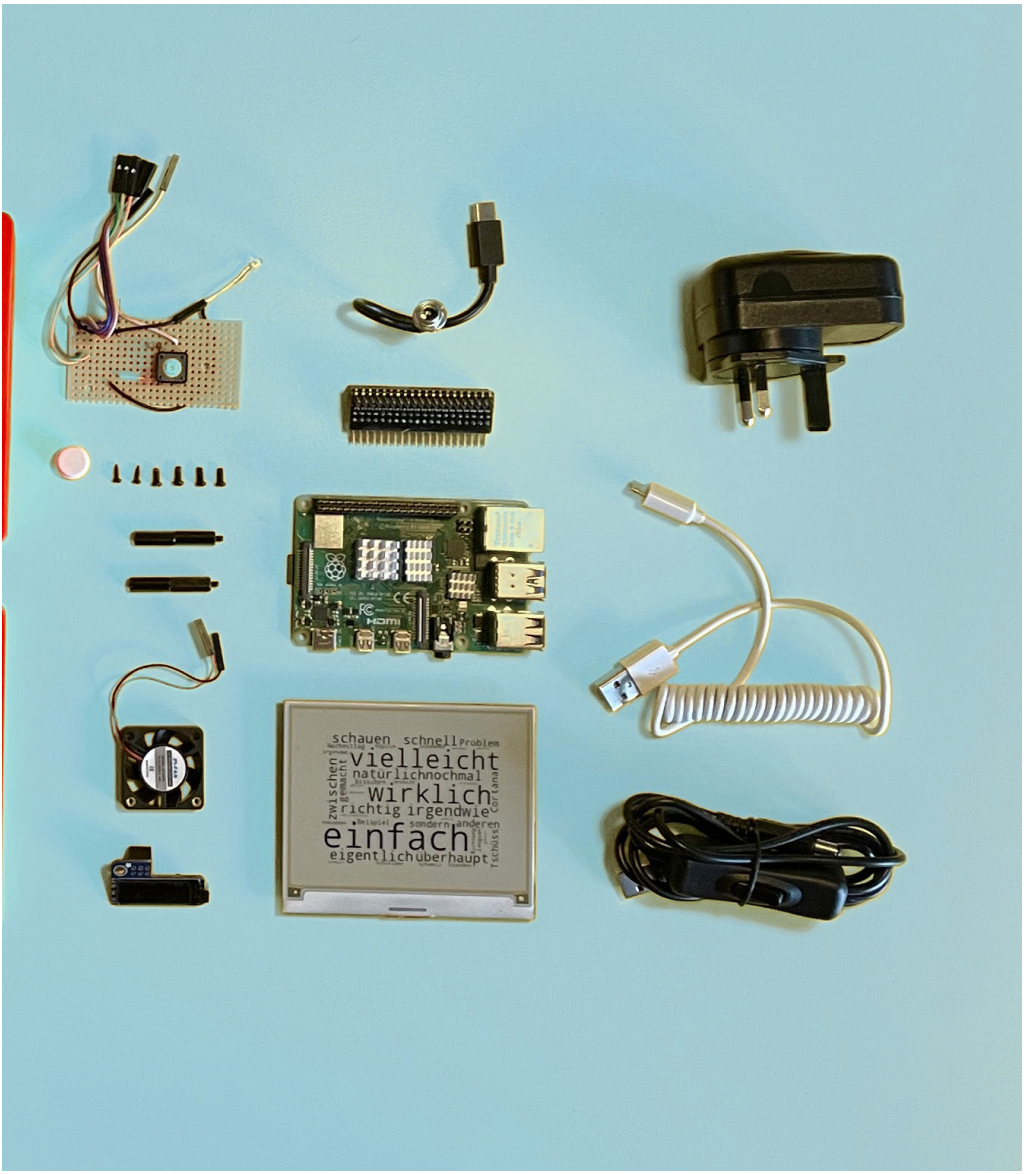


Developing the WordCloud hardware

Developing the WordCloud hardware



All components of the WordCloud laid out and ready to be assembled



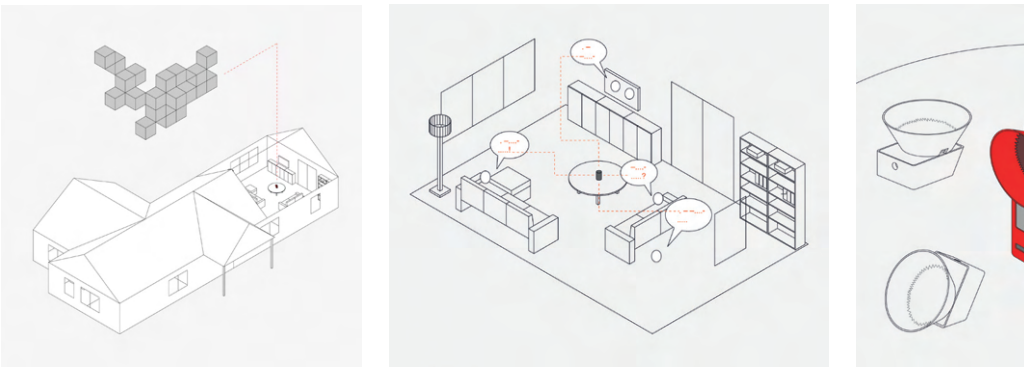
An example of how the WordCloud develops over a one hour conversation.

Observations from building the WordCloud

It became quickly apparent that working at a code level of APIs is similar to lifting the bonnet of an (old) car. You start to see the inner workings of how something functions. Here, that means the way the data is flowing around the software elements and the logic at the core of the service. Working at this level starts to allow the coder to tweak, manipulate and craft the way the code behaves. It also allows the coder to see the strengths and weaknesses of what lies within these black box technologies. In systems, a black box is typically something where the input and output can be measured, but the process is unknown.

Lifting the bonnet on Microsoft's Cognitive Services API revealed that the API does not actually recognise as many words as it presents to the user. For example, the API may recognise "hello john .... you" but presents it to the user via machine learning "Hello John, how are you?". Here the API has not only filled in the gaps, but also added capitalisation and punctuation. Along with this comes a confidence score, i.e. how confident the machine is (0-100%) in turning the speech it heard into individual words. By inspecting the code, and printing its process into a terminal window, we can start to explore the flow of information through the Azure Cognitive Services black box of AI. Speech is the readable input and text is the readable output with this particular black box. It felt like we were opening up the bonnet of the Raspberry Pi Wifi Chip and seeing what was happening within. We found that as we went deeper into the code, we kept finding more black boxes within the initial black box. It was as if the black boxes were nested, and each box was a section of the nested code. Like the initial larger black box, it was still possible to take a data 'reading' before entering and after exiting these smaller black boxes by printing out data at certain moments from within the code. This allowed us to form better mental models of how speech-to-text technology works. Users create mental models to help them understand how devices work. This internal mental representation helps a user create relationships between features of a system

Diagram detailing our connection to cloud connected black box technologies

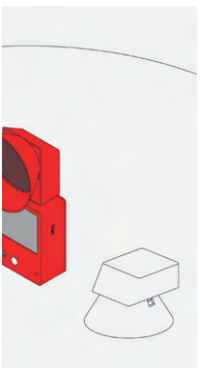


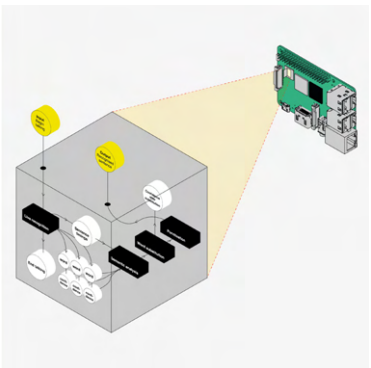
and helps them to both operate and understand a system [Hanisch et al., 1991]. These mental models can be imprecise in many ways, including contradictory, erroneous and incomplete concepts [6]. By making elements of the voice recognition process visible, we make those mental models take smaller leaps, in turn improving their accuracy.

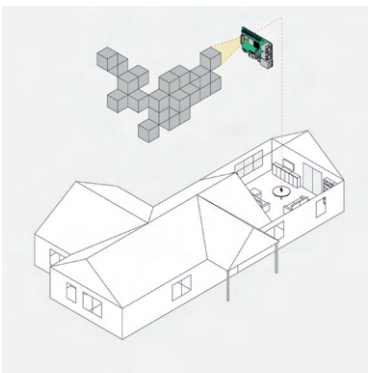
These mental models, along with the use of metaphors (like black boxes or car bonnets) create abstractions of complex processes and services. This is a technique that has been used historically to simplify complexity, particularly in computing [3]. Think servers, clients, daemons, pipes, scripts, stacks and builds. These mental models and metaphors are simply the names we give these complex services to chunk them down into more manageable and human readable concepts.

By stripping down to small enough parts - we were able to decode the meaning of each of the smaller black boxes, and in doing so, reveal a level of computation unseen from the outside. We can make much more informed 'smaller' assumptions of what is happening inside these many black boxes, rather than the larger assumptions and unknowns we may otherwise have without an understanding of this flow of data.

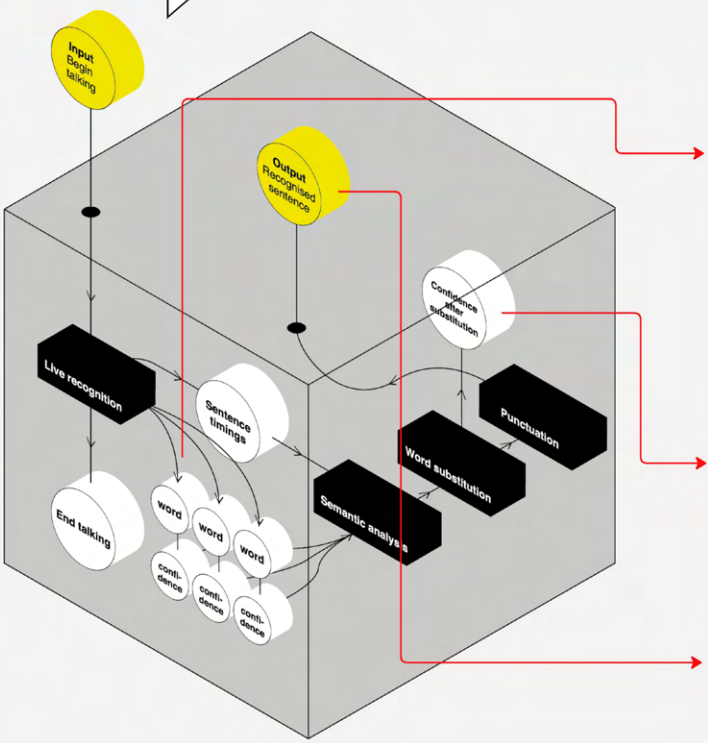
We can now see what words are being heard in the first instance, and what the API transforms them into; i.e. how it fills in the gaps and adds punctuation. We can also see the confidence the machine has in each individual word, even the words the technology did not originally hear.
For the most part, the final output accurately represents what the user has spoken; the API managed to fill in the gaps as it typically was hearing only 50% - 75% of the words spoken. It is clear though, that the API has much less confidence in the words it did not originally hear.







Can you hear lots of words when I say lots of words? I've just said lots of words.



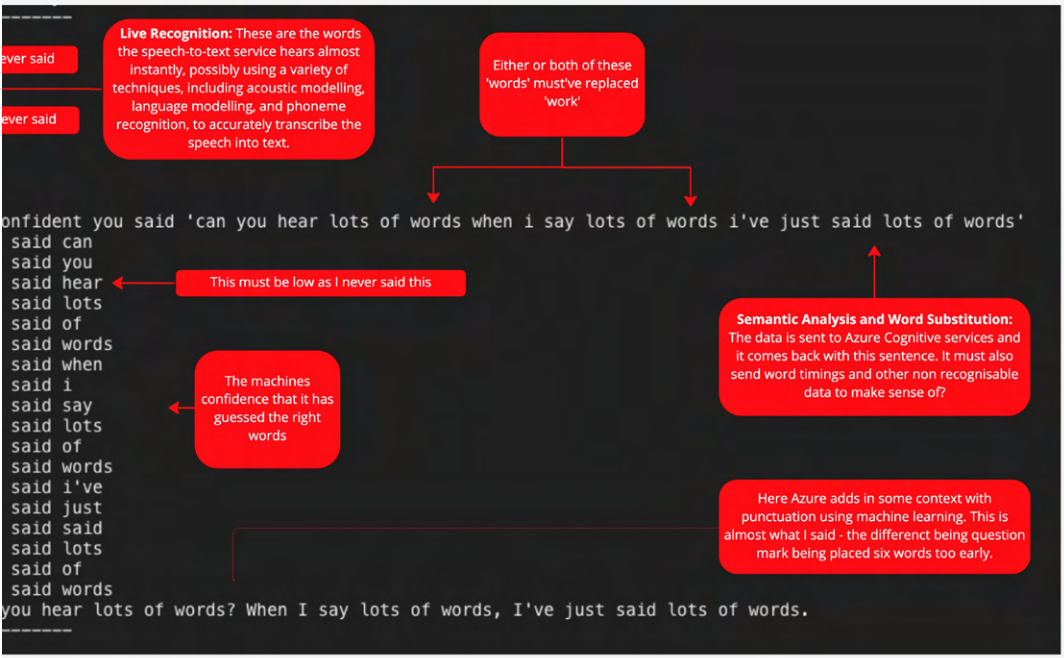

```
I heard:can
I heard:me
I heard:say
I heard:of
I heard:work
I heard:i've
I heard:just
I heard:of
I heard:words
I am 81.32872599999999% c
I am 81.95% confident you
I am 87.51% confident you
I am 49.03% confident you
I am 77.40% confident you
I am 84.78% confident you
I am 53.63% confident you
I am 75.47% confident you
I am 75.88% confident you
I am 87.47% confident you
I am 81.10% confident you
I am 89.93% confident you
I am 88.84% confident you
I am 66.82% confident you
I am 83.72% confident you
I am 91.13% confident you
I am 85.71% confident you
I am 90.20% confident you
I am 97.07% confident you
After AI Processing: Can
```

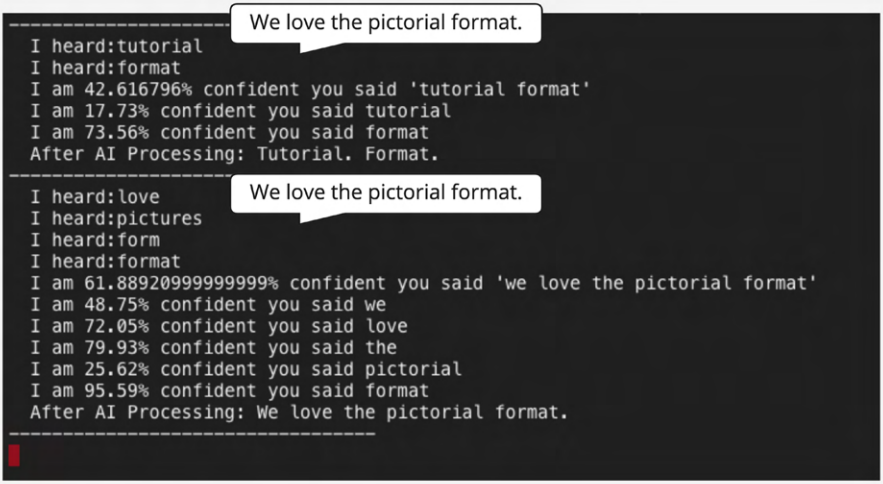
This was n

This was n

Designing Interactive Systems is a great conference.

```
I heard:designing
I heard:interactive
I heard:is
I heard:great
I heard:conference
I am 66.72001% confident you said 'designing interactive systems is a great conference'
I am 47.54% confident you said designing
I am 20.66% confident you said interactive
I am 98.59% confident you said systems
I am 67.29% confident you said is
I am 83.76% confident you said a
I am 86.54% confident you said great
I am 95.07% confident you said conference
After AI Processing: Designing interactive systems is a great conference.
```


ever said
ever said
Live Recognition: These are the words the speech-to-text service hears almost instantly, possibly using a variety of techniques, including acoustic modelling, language modelling, and phoneme recognition, to accurately transcribe the speech into text.
Either or both of these 'words' must've replaced 'work'
onfident you said 'can you hear lots of words when i say lots of words i've just said lots of words'
said can
said you
said hear
said lots
said of
said words
said when
said i
said say
said lots
said of
said words
said i've
said just
said said
said lots
said of
said words
you hear lots of words? When I say lots of words, I've just said lots of words.
This must be low as I never said this
The machines confidence that it has guessed the right words
Semantic Analysis and Word Substitution: The data is sent to Azure Cognitive services and it comes back with this sentence. It must also send word timings and other non recognisable data to make sense of?
Here Azure adds in some context with punctuation using machine learning. This is almost what I said - the differenct being question mark being placed six words too early.


We love the pictorial format.
I heard:tutorial
I heard:format
I am 42.616796% confident you said 'tutorial format'
I am 17.73% confident you said tutorial
I am 73.56% confident you said format
After AI Processing: Tutorial. Format.
We love the pictorial format.
I heard:love
I heard:pictures
I heard:form
I heard:format
I am 61.88920999999999% confident you said 'we love the pictorial format'
I am 48.75% confident you said we
I am 72.05% confident you said love
I am 79.93% confident you said the
I am 25.62% confident you said pictorial
I am 95.59% confident you said format
After AI Processing: We love the pictorial format.

The WordCloud Deployment

The WordCloud was deployed in the households of the three most engaged participants from the earlier In-Home study. Each household spoke Swiss-German as their primary language and were based in the canton of Lucerne, Switzerland.

The WordCloud deployment lasted eight days. At the beginning of each deployment the participants would receive a study handbook, consent forms and a short semi-structured interview. This interview focussed on verifying that the existing data the researchers had from the In-Home study was up to date as there was a year between contact. During the course of the study, the participants were contacted by the researchers using a preferred messaging platform to ask prompt questions and assign bespoke tasks. Depending on the responses from the participants, updates were able to be pushed to each individual WordCloud to make the device more bespoke to them.

After the eight day deployment the participants took part in an exit interview. The researchers encouraged not just the primary participant to these interviews but also the other members of the household. This interview focussed on in-home dynamics, data and privacy, practices and routines, reflections on their own voice and device developments.

WordCloud instruction manuals

WordCloud during a deployment in Switzerland

Doorstep repairs during a deployment

Installing a WordCloud in a participants home

INSTRUCTIONS FOR SAFE USE OF YOUR **WORDCLOUD**

The WordCloud is a prototype device and it is therefore essential that you follow these instructions. HSLU and Northumbria University cannot be held responsible for any loss or damage that arises as a result of your failing to do so.

If any elements of the WordCloud stop working or break, please contact us as soon as possible.

No parts of the WordCloud are repairable by users. Do not attempt to disassemble the WordCloud or power supply.

Only use the power supply we have supplied with the WordCloud. Do not use any other charger.

If the WordCloud or power supply is damaged, becomes hot during use, or if the WordCloud appears to be behaving unusually in any other way, turn it off immediately and contact us as soon as possible.

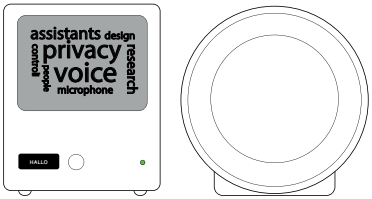
The WordCloud is not waterproof. Do not use in wet environments.

For more information please visit https://sites.hslu.ch/va-pepr/

HSLU Lucerne University of Applied Sciences and Arts · Northumbria University NEWCASTLE · OST Ostschweizer Fachhochschule · Swiss National Science Foundation

YOUR **WORDCLOUD**

The WordCloud is a device that allows you to have a better understanding of how AI machines interpret the human voice.



powered by VA PEPR

HOW TO USE YOUR **WORDCLOUD**

Simply plug the device in and observe what your WordCloud hears.

SD Card Slot
Please do not remove this card, it is not comaptible with other device

E-Ink Display
The E-Ink display will display a wordcloud of the words it can hear. This display will refresh every 30 minutes. The screen contents remain visible even when turned off

Clear Button
The Clear button erases all words stored on the device to allow you to visualise shorter situations

LED Screen
This smaller screen displays the last word(s) that the device has heard

Microphone
The microphone can be attached to the top of the device, or alternatively positioned positioned off the device in different orientations

Microphone Socket
Plug in a microphone here. Otherwise the device cannot hear anything

Power Socket
Only use the power source supplied with your WordCloud

WiFI Status LED
If LED is green you are connected to wifi! If the LED does not light up then please contact the support team

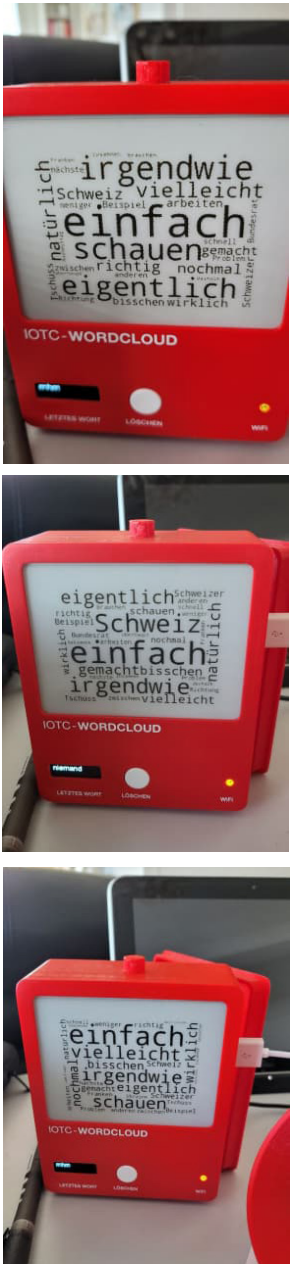
WordCloud Instruction manual

Findings from the WordCloud Deployment

The WordCloud as a Dialogical Object

One of the initial intentions for the WordCloud was to create a dialogical object; an object that could foster a dialogue among people around VA related technologies, in particular speech recognition, AI and cloud computing. Once installed, however, the WordCloud also acted like a mirror for people's own speech patterns and the words they used. While this self-awareness was not described by our participants as particularly stressful, it does echo findings from research into self-views and attitudes to virtual meetings. People simply could not escape seeing how they spoke in a similar way people in a virtual meeting cannot escape noticing how they look and what others see. This might contribute to the sense of the WordCloud being an 'deliberately eye-catching' object as one participant described it. This reflective quality, however, does point to potential opportunities for VAs to be employed in speech training programs.

Seeing in writing what a device hears also made the participants wonder what language is "safe" around the VA or what speaking volume 'protects' them and how they can control what the VA hears. For some participants, this meant to speak in a Swiss dialect as the WordCloud could not recognize these terms as words. Such cultural or linguistic 'escapes' are disappearing fast though. In our wider Swiss study, the VA seems to have an active role in making it acceptable to speak High German in the privacy of the home. The significance of this adjustment goes to the heart of Swissness as a recent Swiss public radio broadcast reveals: Not long ago, children were mobbed on the school yard when speaking High German. This affirms the VA to be a cultural conductor for those engaging with it, demanding what they call 'a code-switch' to interact with it.



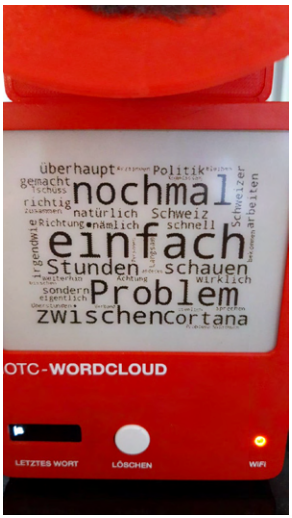
Various examples of content sent back to us from the deployment participants

Various examples of content sent back to us from the deployment participants

As a dialogical object, the WordCloud did bring home that a VA always listens and hears, even though it only responds to one of the wake words. Yet, our participants attached this characteristic and behaviour exclusively to the WordCloud, not to the VA still listening next to it. This suggests that the participants mentally segment, compartmentalise and differentiate the devices in their VA ecosystem. In interviews as part of the first In-Home study, users, too, rarely thought or talked of their smartphones as a smart speaker or VA even when using its voice functions. The WordCloud device, too, was perceived as something distinct from their VA. While the WordCloud with its constant in-your-face visual playback of what users said was described as annoying, irritating or surprising, they stated that they did not extend these perceptions and emotions to their VA.

Though the contents of the WordCloud were still interesting to view and surprised participants by showing them which words they were using most often without ever having noticed themselves. The most common words in English that appeared on the researchers WordClouds were something, people, things, really and actually. Those are in line with the 100 most common words found in the English language. Future developments of the WordCloud may explore new algorithms that go beyond these basics. Exploring the least said word, or the longest word or the rarest word, or even word couplets. It is unclear if this would make the WordCloud more meaningful to participants. It certainly would not have affected behaviour change in removing particularly meaningless words from our own voices.

Technologists and designers need to consider the mirroring of collected data on users to give them awareness and therefore agency to change behaviours. More so this extends towards the physicality of these object. Even without the mirroring of data the WordCloud had a clear



Various examples of content sent back to us from the deployment participants

microphone, as opposed to VAs which typically look like speakers. This alone still gives users awareness of the capabilities of the ubiquitous smart objects in their home.

The Importance of Design Friction

Historically, a meter has no automation, but rather gives humans the autonomy to understand something and make decisions from the data. Participants called the WordCloud a gimmick, but was this purely from unfamiliarity? When home energy meters first arrived in the public domain they were labelled 'gadgets', now they are commonplace and expected as standard from energy providers.

Like the WordCloud, a meter should be human readable, and ideally glanceable - a quick look and the viewer instantly understands their environment better. The WordCloud maybe did not get this right (too many filler words and missed key words), but the researchers now know what could change to make the words displayed more meaningful to the user. By increasing this meaning to the viewer we start to design in more friction. The act of taking a meter reading can be an interruption, slowing down your day, but it creates a talking point. We noticed when the participants and their families took a WordCloud meter reading together this provoked conversations around a technology that they had been living with for years. The designed friction of the WordCloud began to alter domestic behaviours through data transparency, just as it did according to Lovell et al. during their home energy meter study.

At times the participants noticed words coming up on the system that they had not said. It made them wonder about whether it was hearing the radio, their neighbours or people on the street. It is easy to think that a VA is yours and your voice - that the physical walls are the boundary of this. Much like Dunne explored the boundaries of devices beyond the walls. So does the WordCloud. However it is not just the physical boundaries of the microphone it is also the

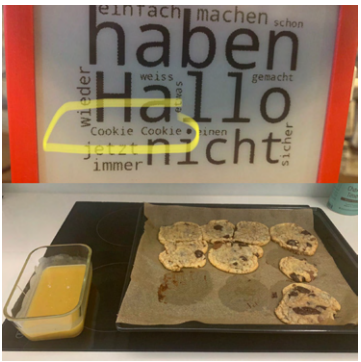
boundaries of the cloud. This is much harder to interpret beyond the speculations Joler and Crawford for example. We argue that by making the words visible back to anyone in line of sight of the VA in a glanceable way, that we are supporting Ahmad et al.'s call for IoT devices that "clearly and unambiguously conveys sensor states to people around them". This of course does not help to identify the physical boundaries of where a VA microphone might be picking up voices from, but it could, we argue, go some way to informing all people in a room, whether resident, guest, visitor, service-provider or perhaps intruder, of the nature and capability of listening power.

Anthropomorphising VAs

For 50 years programmers have been getting computers to say 'Hello World' as a first test - it is a basic program but it shows the computer is working. During the WordCloud study the researchers found that on initial deployment the participants would say Hello or Hallo to the WordCloud to see if the device was heating them correctly. Once a WordCloud had been deployed in a home the researcher would ask the participant to send a picture of their initial word cloud. Every response had 'Hello' or 'Hallo' on their word clouds. In the absence of a wake-word, the humans used the most human introduction "hello" to greet the new device. Moreover, participants would refer to the device with the pronoun he. Stemming from the German masculine pronoun used for Cloud "Der WortWelle" .

While our VA didn't directly talk back (verbally) to the participants, we would like to return here to Haas et al. who argue that making VAs sound as human as possible is not the most efficient way for humans to talk to machines - that shorter (or no) response from VAs is preferable. In addition, they argue that this will also help to break the illusion that a VA is an equal partner. The distinction between borrowing-from and the copying-from is a well-trodden path in design (for example biomorphic design).

Various examples of content sent back to us from the deployment participants



Various examples of content sent back to us from the deployment participants

Copying directly from nature however can make things feel creepy or unnatural - a door handle that resembled a human hand would be incredibly creepy and clearly not be the optimal, and inclusive, option for a universal interface to a door. Making machines as close to humans as possible poses serious ethical challenges but also is very likely to not be the best way for us to interface with computers using our voice. We can be the humans, but we can design interactions to allow for machine intelligence in potentially more universal ways. Of course we do not want our smart speakers to look like people, but there is a wide gulf of design intent and value between hiding technological capability through design and making 'smart' things transparent. What we might want for the aesthetics of things in our home is an entangled set of conflicting agendas for how we might want to make things more open and transparent.

Reflections on the Cost of AI Labour

We selected the Microsoft's Azure speech to text API due to its capabilities of translating multiple languages into text in real time. However, this API was the most expensive of the APIs tried, the cost of the service being $1/hour per device. This seemed significantly more expensive than the close competitors Google (£0.029/h) and Assembly.ai (£0.012/h).

The concept of AIs having 'wages' is an interesting one. Is the Microsoft AI better paid being multilingual and coming from a big tech company? How do these AI wages compare to human wages, and does this seem fair? So what is this cost paying for? Is it the processing power, the data centres, or is it paying for the planetary costs of such a technology? Or are we paying for security and reliability of a service? Or perhaps bank rolling other sides of a business.

Some AI text to speech services are in fact using Artificial Artificial Intelligence (AAI), in that the service pretends to be AI but rather the company employs a team of workers to analyse the data and respond as if they were the machine. This low-wage workforce typically earn less than the legal minimum, and also tend to have no health benefits or job security.
Should we be told what is AI and what is AAI? If AAI is more common than we think, should there be more transparency here? Would our relationship to AI IoT objects change if we suddenly found out a real human with intelligence was sorting our data rather than an emotionless machine? Perhaps the industry needs to show the provenance of where our AI is coming from, what it's using and if it is using AAI is it free range, organic or battery.

Does this higher API cost suggest that we may be getting more of a human touch than expected? There may even be a technological placebo effect in play. Typically in medicine the placebo effect refers to administering a sugar pill to a patient to try and solve their illness using psychology. AI systems have something in common with this sugar pill - their inner workings are hidden and unreadable to the layman. Is our willingness to pay a premium on our AI APIs a placebo effect - do we assign better data protection to a larger price tag? Kosch et al. discuss the idea that AI systems will be more favourable and usable with some intentional placebo theory planned in by the "communication in the system's abilities and functioning, even if highly inaccurate, can potentially increase usability".

Participants watching content develop on the WordCloud

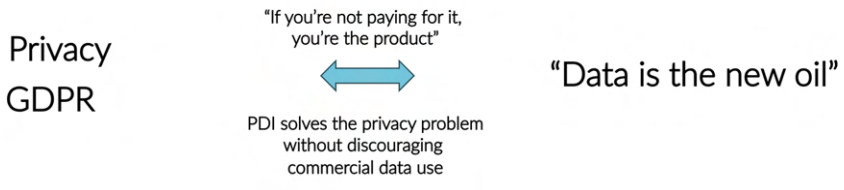
CHALLENGE 02: DATA & PRIVACY

Privacy Data Intermediary

Voice data is very personal data. When a voice assistant or a large language model records and processes a voice command, very private data is transmitted. And once it's uploaded, it is out of the control of the data subject. Who uses it? What for? And who earns money using it? To prevent such a loss of control over one's own private data, it seems like the data must stay with the user, and be processed under the supervision of someone that cares for the data subject, rather than aiming to exploit data for their own benefit.

Hence the idea of a privacy data intermediary was conceived. It's like a bank but for data instead of for money: You get a bank account in which you may store your data, and the bank will protect and invest it. Yes, data needs to be invested: Suitably processed data becomes valuable, so allowing computations to be performed on one's personal data should produce profits. A data intermediary provides the market (and hence establishes prices) for such profits to be made. By "bringing the algorithms to the data" instead of giving up control over one's data, a fair share of the profits might stay both with the data subjects and the digital service providers.

Interestingly, no such privacy data intermediaries exists. In principle it could be a private organisation (like any bank). This study aimed at identifying the key obstacles for a viable business model: Maybe no one wants such a service? There is so much value in data- why is there no institution that tries to provide services to monetize that value on behalf of the data subjects? Is it impossible to build and run?



In our study, we have elaborated convincing arguments that a privacy data intermediary is desirable for data subjects, technically feasible and viable in principle. The Figure shows the outline of the new data economy developed in this project, in which a fair price for evaluation access on data is paid, where data subjects invest their data or limit their use just as banking clients invest their monetary wealth. Such a transparent data market would also have massive benefits for all data hungry businesses except the big tech companies. They currently profit from access to a lot of undervalued data.

It is probable that a significant technical effort would need to be made to build a secure and efficient privacy data intermediary (PDI). Big tech companies will oppose to a PDI to protect their interests, such that it will seem useless to data subjects, as a lot of their private data (such as online activity logs, user profiles and preferences) will not be protected. Hence the positive effect of scale of pooling data in a PDI will not easily materialise. However, we believe that the main reason why PDIs do not exist is that only very few people realise that it is possible to let other entities use one's private data without infringing privacy. A PDI would enable all clients to buy goods in an online store without disclosing what they buy to the online store (via a system of obfuscation by tokenisation, to be implemented by the PDI). Yes, the online store needs to deliver the products and send bills to the correct address. But to this end it is not necessary to provide them with the opportunity to construct and use a user profile. They might need to show personalised recommendations, but, as we show, this does not require them to know any private data about their customers. They only need to obtain the permission from the user (via a PDI) to obtain the harmless results of certain function evaluations.

In the end, the technicalities do not matter much. But the political will to implement a privacy preserving data infrastructure is contingent on a large number of people demanding it. Many do not demand it, because it is not

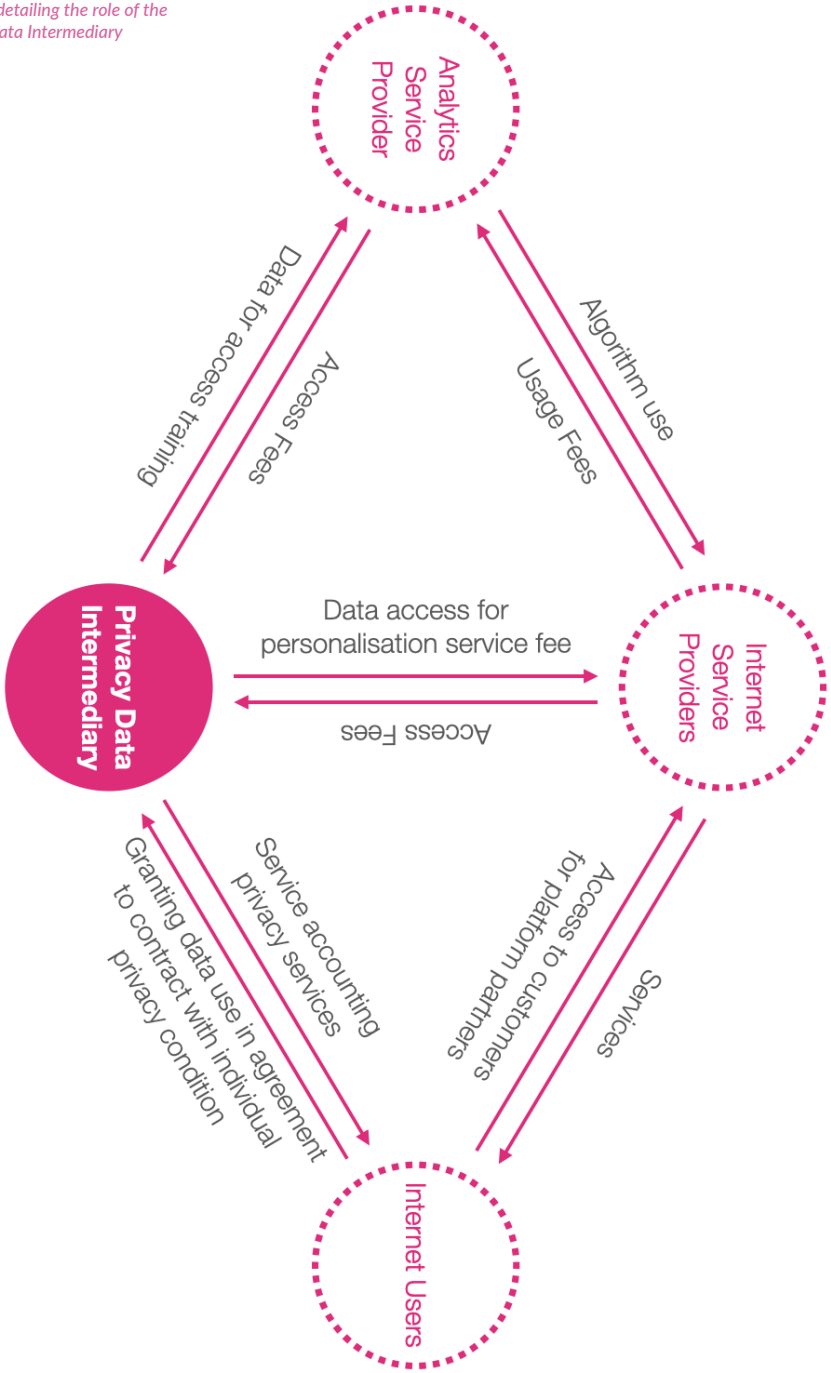
yet present in the public debate, and because their obvious benefits are not perceived as realisable. And because of this, the business model of a private PDI remains non-viable.

Our Proposal

We propose that in a first step, sensitive government data be transferred to a public PDI, and that the upfront investments in building a PDI-based data infrastructure are made using public funds and initiatives. A next challenge is identifying data sources that can suitably demonstrate the value and safety of pooling personally identifiably data to the benefit of a community and to the detriment of no one.

Financial Bank	Privacy Data Intermediary	Category
Personal Finances	Personal Data	Assets
Property Protection, Payments	Privacy Protection, Selective Disclosure	Stewardship
Finance Management	Data Usage Management	Consultancy
Accounting	data use	Transparency

Diagram detailing the role of the Privacy Data Intermediary



CHALLENGE 03: CONFIDENCE

Exploring Confidence & VAs

Problem Identified

Users were frustrated by a VAs lack of feedback - virtually and visually - this typically led to a lack of trust. Users would double check requests on a different device or enquire of the content source.

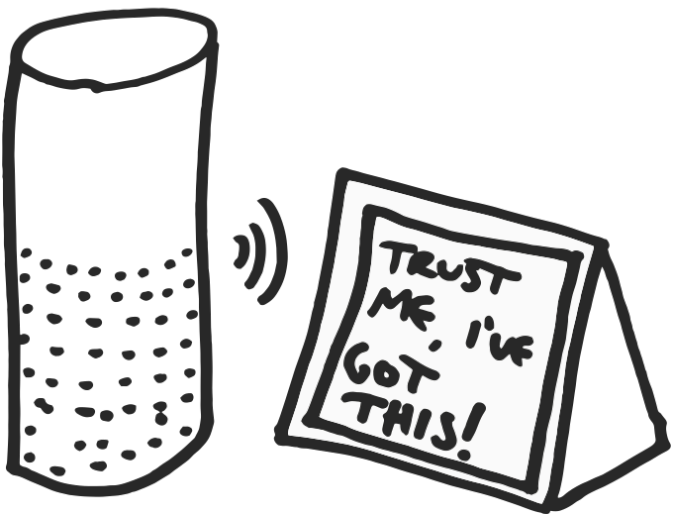
Insight

A small visual element could improve trust between a user and their VA. This would allow users to be more involved in a conversation with a VA.

Challenge

How might we help users trust their VA more.

1. How can we make users trust their requests have been met.
2. How can we show the source of content being provided from the VA is trustworthy.
3. Consider how the research will investigate participants' behaviours and routines in relation to ethical issues such as privacy and trust in their homes.



How does the Confidence challenge relate back to the wider project?

Research Question Seven: What kinds of changes in attitude towards privacy occur?
Privacy and confidence are closely related. How can we be confident in something if it is not looking after our privacy?

Research Question Eight: What does privacy mean to VA users in their homes? Does it differ across the German, French, and Italian cantons?
Research Question Nine: What negative side effects of VAs emerge in the home? And what is needed to mitigate them in terms of services, tools or regulations?
Research Question Ten: What desirable uses of VA in the home emerge? And what would be needed and how could they be realised?

Task one - Read Trust: A Very Short Introduction by Katherine Hawley.

The team member who were interested in this challenge all took time to read this book, and held many discussions around it, and other literature to help the team explore what trust can mean in the realm of Voice Assistants.

Task One: Insights

Aspect of Relation

The description of trust changing based on the parties.

Interpersonal Trust vs. Reliance on Objects

We trust people but we rely on objects.

Trustworthiness requires both skill/ability and good intentions.

Trusting someone to do something has two aspects:
- An expectation of good intentions, (we call this 'social trust')
- and an expectation of skill or ability (we call this 'functional trust')

Aspect of Time

Time will tell to trust or distrust, even maybe re-trust. It is a journey.

Aspect of Collectiveness

Can we talk about Collective trust? How could a user has impact on it?

Aspect of Mutual Benefit - Reciprocal Altruism

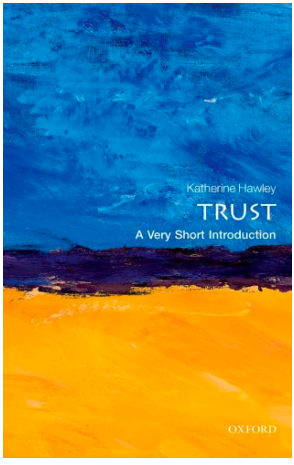
Can we offer trust to the VA? We benefit by giving away our data to receive a more bespoke service. Are there any other benefits for the human side? Does the machine trust us and the data we give it? One-way or two-way relationship? Tit for tat - Bit for byte.

Untrustworthy - unreliable - no confidence

We make back-up plans, we are not surprised when it fails.

Affordances, competences and intentions

Are these the three 'boxes' we want to tick to fully trust a VA?



Trust book cover

Task Two - Critique a non-digital object or service

Aim: To explore how we built confidence in products or services and exchange stories in relation to personal interactions and feelings about confidence.

How?: The exchange can be any form: visual (image, video, drawing), audio (story, talk), text, drawing, emojis, etc.

Questions: Why did you pick that product/service? How much do you have confidence in it? Why? What affected the development of this confidence or the opposite?

Task Two: Insights

Relation/Experience might start out of obligation/necessity, for fun, just random, with/without interest, curiosity, intrigue even with confidence

Satisfaction, Loyalty, Disappointment, Unexpected bonding, Greater confidence

An example of recording thoughts from task two on a Miro board

Aurelio *Task 1 - Confidence in Things - **Water Level Tool***

My thoughts on Trust: I trust many objects, although I do not know or only assume that they tell me the truth and that they reflect reality. Most likely we live in a very complex world and to constantly go to the bottom of everything we simply do not have time - or am I too lazy?

Such a device is the "spirit level". I have never really asked myself how it works exactly and what liquid is inside. Nevertheless, during my work as an architect, I always trusted the spirit level when visiting construction sites.

in German I would say: Ich hatte blindes Vertrauen

Do you first test if it is working properly by seeing if the bubble is in the same place when you turn it over. I was trusting mine... blindly.... Then a friend showed me this trick and it was way off.... I got a new one now and no longer trust them - I test them first! -Jon

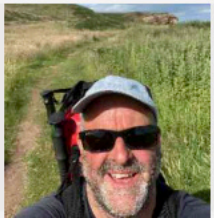
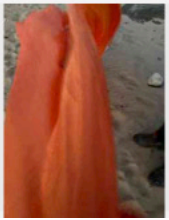
I like the idea of a datum or zero function to tools. Like on scales

An example of recording thoughts from task two on a Miro board

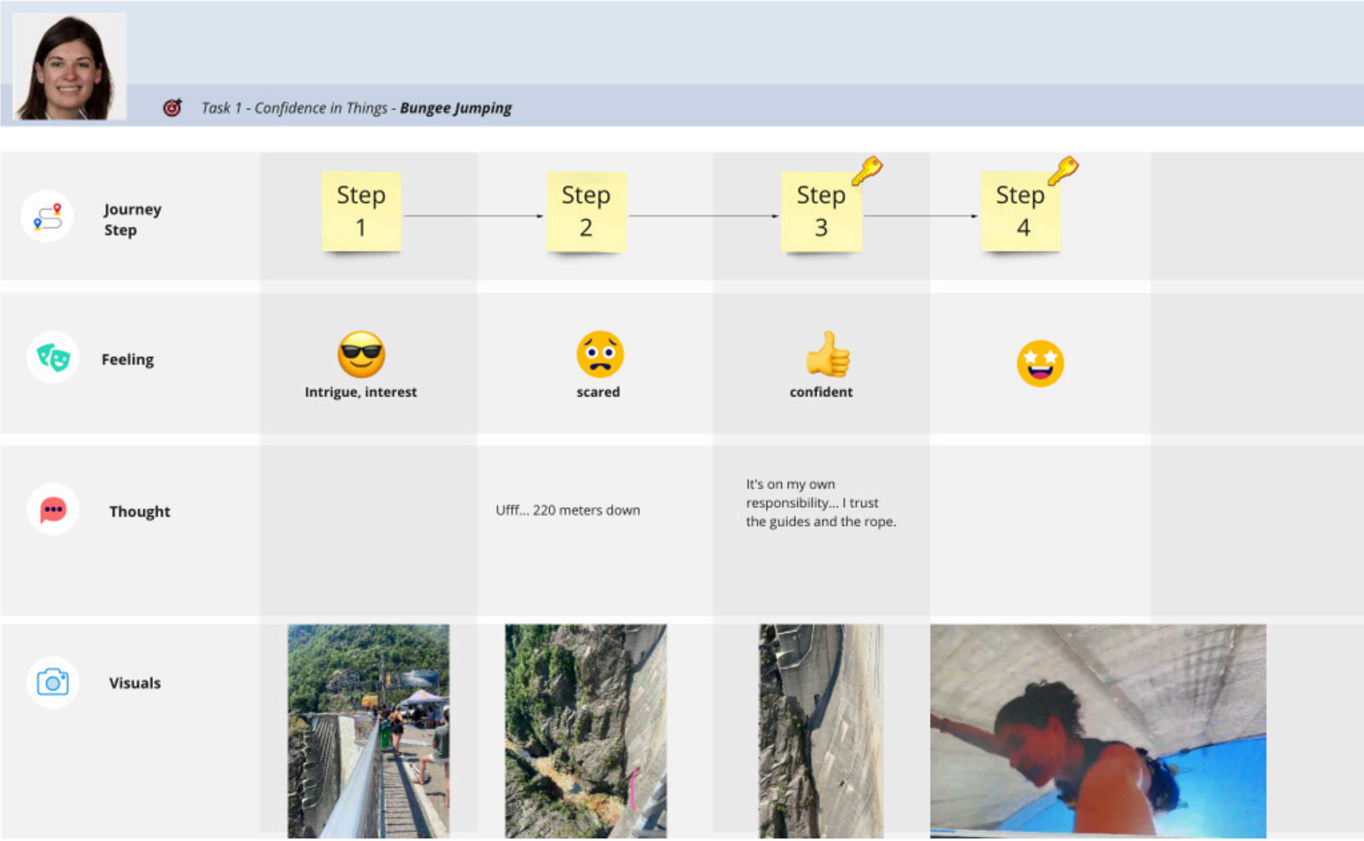
(opposite page) An example of recording thoughts from task two on a Miro board

Michelle

Task 1 - - **Tarp Tent**

	Step 1	Step 2	Step 3	Step 4	Step 5
Journey Step					
Feeling	Intrigue, interest	relaxed and confident	Dissapointment	not happy	OK
Thought	I started my hike full of confidence in my new lightweight adventure tarp.... A night of wild camping on the beach.... Under the tarp.... What could go wrong....	A cheeky beer on arrival.... Confidence rising....	Then set up... it was windier than I thought.... The sand to loose... the pegs too short... confidence fading....	Failing	Diminished.... No tarp that night. Instead just my Bivi bag and a beach... Next time I'll take the tarp to the woods!
Visuals					

An example of recording thoughts from task two on a Miro board



Task Three - Critique a digital object or service

Aim: To explore how we built confidence in products or services and exchange stories in relation to personal interactions and feelings about confidence.

How: The exchange can be any form: visual (image, video, drawing), audio (story, talk), text, drawing, emojis, etc..

Questions: Why did you pick that product/service? How much do you have confidence in it? Why? What affected the development of this confidence or the opposite?

Task Three: Insights

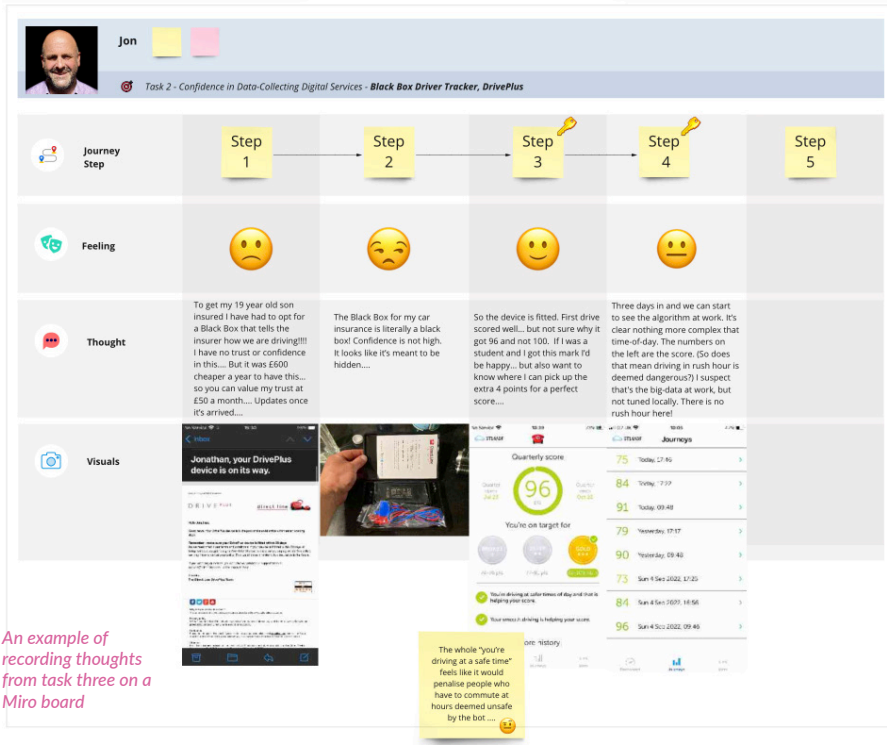
Big Data vs. Local Cases

We somewhat confide in the apps of big companies, their big data to work well, however, we experience that it is not tuned locally very well. This leads to disappointment and a loss of confidence.

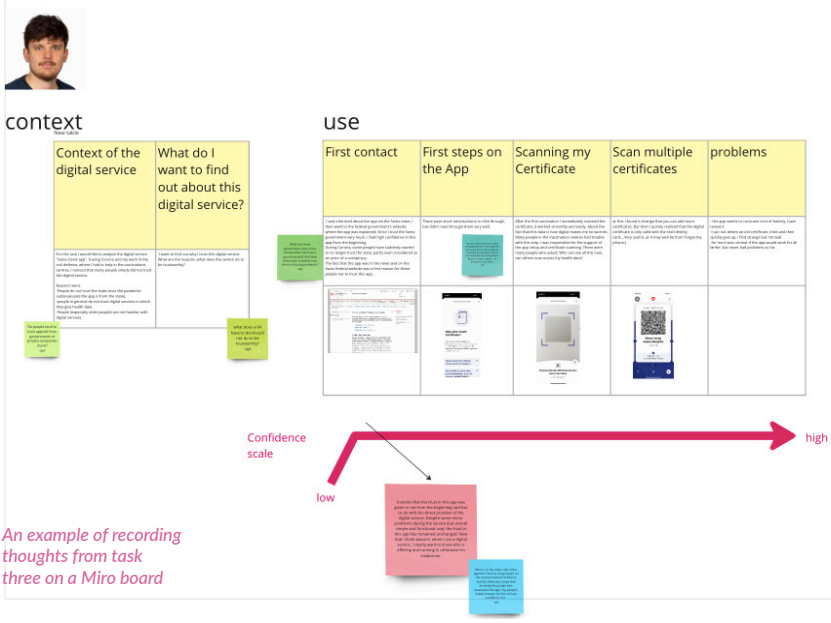
Whom to trust? In what to have confidence?

We seem to trust the services from the trustworthy organisations/institutions/government departments. Also, it is rather easy to decide to trust/distrust the services if you know the community behind it; “my people” / “not my people”

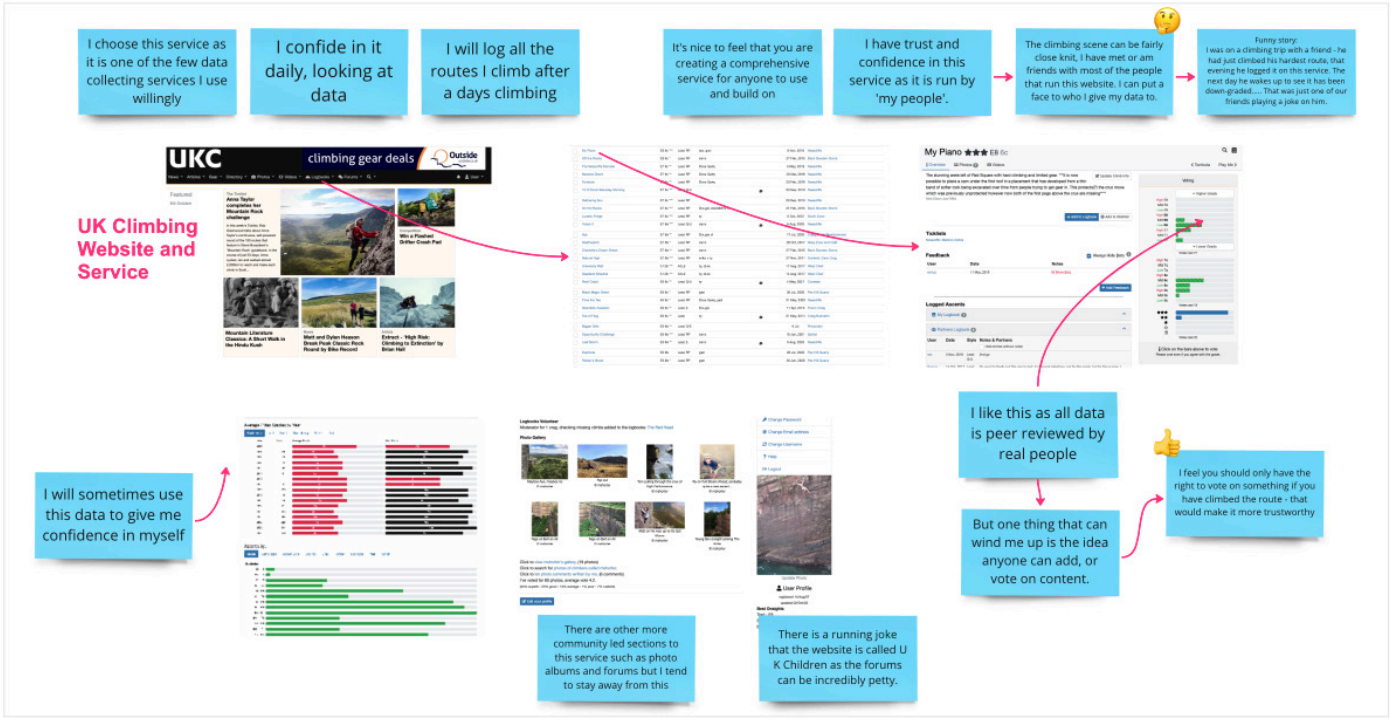
And in some cases we are overwhelmed with all these complexities and find the best solution to be blindly confiding in artefacts or trust services.



An example of recording thoughts from task three on a Miro board



An example of recording thoughts from task three on a Miro board



An example of recording thoughts from task three on a Miro board

An example of recording thoughts from task three on a Miro board

Ute *Task 2 - Confidence in Digital Services - Starbucks Loyality Card = Starbucks Rewards*

What is

It's a loyality card (app and/or card) where cou can collect points whenever you buy something at Starbucks.

You can load money on your loyality card, and use it later in the sense of a debit card, but you don't have to. You can also just collect points and pay cash or by cc.

You can register your loyalty card online (not a must) but then you can check on your purchases, points and vouchers.

Why I picke

It started when we were in the US. We got the loyalty card as a gift and used it and charged it from time to time.

The loyality cards are country-specific. I.e. for Switzerland you need a different card and a different login. Only one card can be assigned per account.

So, I have two cards: US and CH.

Why I use

Because I like it.

And I like their efficient service. At a Boston Starbucks I have seen four cashiers and only two barristas. And there wasn't any long waiting times.

You will get a free beverage at your birthday. Have you ever got a free beverage in your favourite local coffee shop? And you will be invited to test new flavours or new beverages.

I do like the different options you have for a hot beverage.

For many years they had their customer form (MyStarbucksIdea). That was so interesting, all the

Trust -

I don't have any money on my Starbucks card.

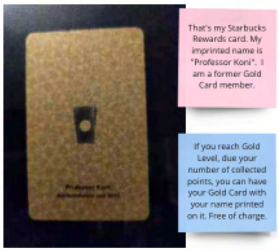
But if you loose it, you have to block your account otherwise the money is gone.

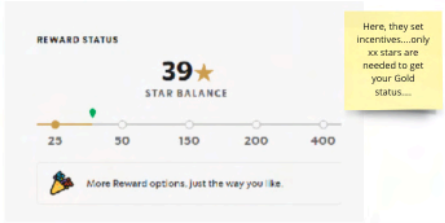
I share my card with my partner. So, I am not sure what they will get out of my data. I will give this another thought.

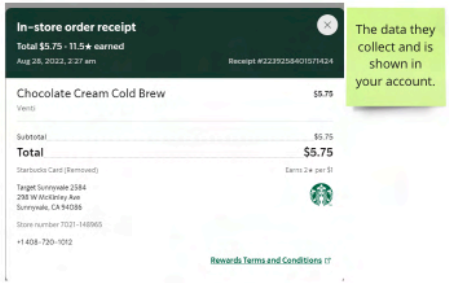
At the end - it's a money-making company. Think about how much money you pay for coffee beans, how many cups of coffee you can brew, and how much you pay for it.

A few years ago I did a little competition in the classroom and the main prizes were two Starbucks Cards with a little money on it.

Starbucks collect data from their shops (purchases), buy data from business partners, and collect data which







An example of recording thoughts from task three on a Miro board

What do we mean by Trust and Confidence

Definitions

We found that up to this point we were using the terms trust and confidence interchangeably. To help us focus our thinking we wanted to define these two terms to clear up any mis-communication of our thoughts and ideas. These two definitions are framed through the lens of Voice Assistants.

trust |truhst|
- noun

1. In the context of Voice Assistants, we describe trust as "a feeling of assurance or security based on an expectation of good intentions" which involves understanding the values and principles of the VA company, as well as its commitment to transparency and ethical practices. Of course, this requires that the VA companies should demonstrate their commitment to protecting user data and respecting user privacy, provide privacy policies to ensure users understand how their data is being used and what rights they have to protect it.

So, trusting a VA, hence in a VA company, requires an understanding of how the device and the system work and the potential risks associated with its use.

confidence |kon-fi-duhns|
- noun

1. In the context of Voice Assistants, having confidence implies “a sense of assurance that the device will complete tasks successfully; fulfil its function correctly without any errors” especially when giving complex commands or important tasks.

So, it means believing in the reliability and accuracy of the VA and trusting that it will work as expected. It is about the skills and competences of the VA device.

In conclusion, both trust and confidence are essential components of successful implementation of Voice Assistant technology into everyday life applications; they provide reassurance to users regarding data security as well as accuracy with regards to interpreting user commands accurately so desired results can be achieved quickly and efficiently every time they interact with a Voice Assistant. These two definitions are framed through the lens of Voice Assistants.

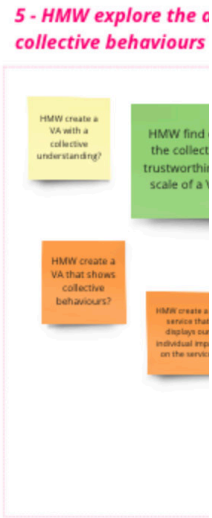
Provotyping to Explore Confidence

Task Four - Design a series of provotypes that explore notions of confidence when interacting with voice assistant.

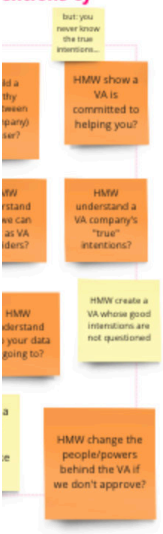
To help us explore notions of confidence we built a series of provocative prototypes. The purpose of this exercise was to open up engaging dialogues with the VA-PEPR team, experts and the general public.

The process of developing these provotypes followed a very similar process to the previous round of provotype building. We began with a series of short presentations from the team highlighting insights from the previous tasks, whilst these presentations happened the rest of the team wrote 'How Might We' statements to capture our notes as potential opportunities.

These 'how might we' statements were then clustered into seven themes for us to generate ideas around. The them 'Affordances' ended up folding into the other themes as it seemed to be an over arching principle that we wanted to explore.



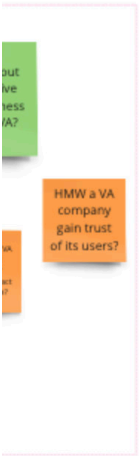
A screenshot of the Miro board where the team generated and organised their 'How Might We' statements



2 - HMW create VA competences that foster trustworthiness?

3 - HMW create a VA that allows a two-way relationship?

ffect of on VA services?



6 - HMW explore our trust relationship with a VA over time?

AFFORDANCES

Ideas were rapidly generated for each of the previous themes. These ideas were filtered and developed by the team to produce the following provotypes. These provotypes were presented at a workshop at IxDA 2023, Zurich. The following pages describe the insights gained from these provotypes.

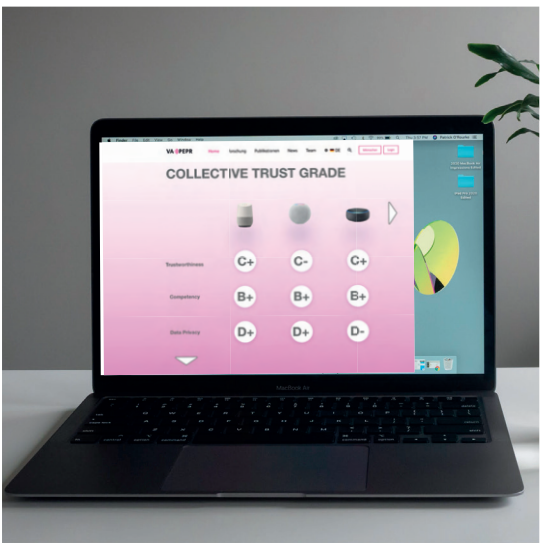
A selection of the rapid ideas generated around the decided themes

The final 3D printed provotypes. Here we used the same pink filament to keep a consistent visual language for this stage in our research

Independent Hardware

What if we could separate voice assistant hardware from voice assistant service providers?

What reasons would users change service providers for? Would users use different providers for different tasks or situations?



Collective Trust Metrics

What if we could create a trust/ confidence rating system for voice assistants?

What and how would they collect data to create collective trust metrics? What factors do users think are important for trust/ confidence? What affects the perceived trustworthiness of a voice assistant according to users?

A Machine's Confidence

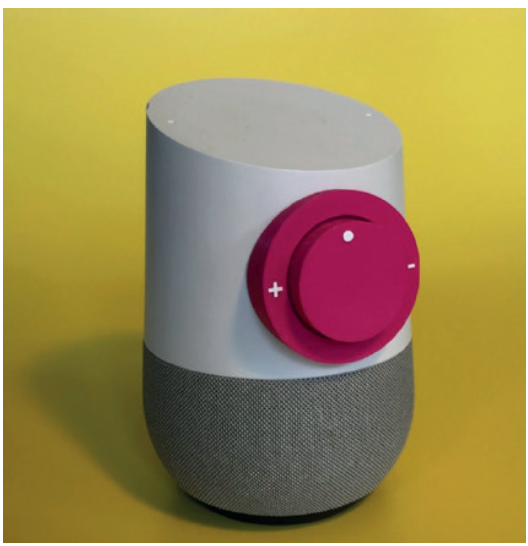
What if we could see how confident a voice assistant was in its own abilities?

What are users existing mental models of how voice recognition technology works? If we open up this black box technology and show more of it's process (including confidence) does this make people more or less confident in using the device?

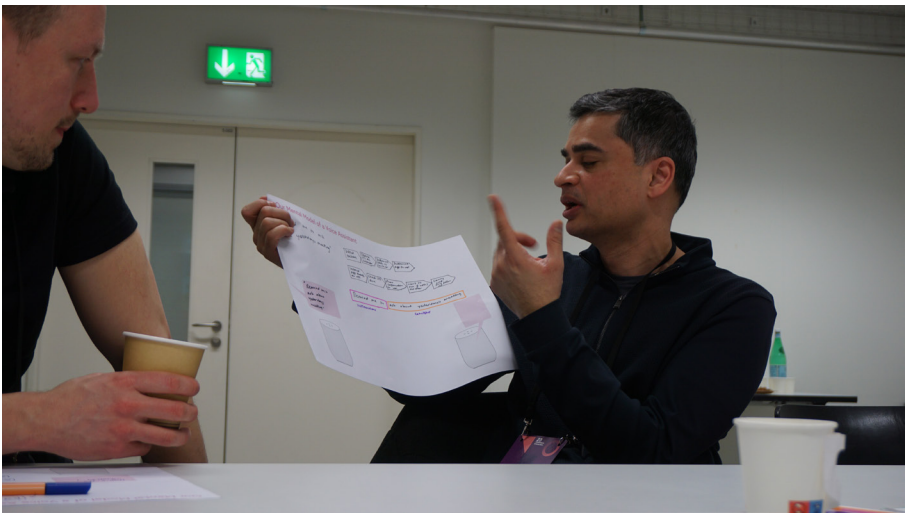
Community Range Filter

What if we could fine tune a set of filters to improve our confidence in our voice assistant?

What filters do people want to apply to their voice assistant? Would they change their filters depending on the request?



The Self Confidence Machine in action



Discussing the Self Confidence Machine at IxDA, Zurich

The Self Confidence Machine

Insights

- Everyone was very playful and fascinated by this device
- Mumbling into it to understand how they were heard and how confident the technology was in what it heard.
- Trying different languages/dialects - “bottle of water!”
- Using the provotype to prove articulation: My Chinese-American sometimes really struggle to understand me. “Look I have a very high understandably according to this device! It’s not me, it’s you!”
- Trying Swiss German: The device is 80% confident that ‘hoi’ is ‘hi’, ‘wetter’ is ‘weather’, shows what it thinks it is.
- Trusting the device if the estimation/confidence level is higher than XX%. Then do not show me anything, do not ask me back for confirmation! Hard to set a strict figure?
- You need to live with this for a while to set a figure.
- VAs are so confident even when they make mistakes, “here take this!” This is what you wanted!
- Would you rather want to know when it’s not confident? “Err, take this, if you like, is this what you wanted?”
- VA is so confident in playing the most popular cover version of the song I requested. That’s not I wanted! In this case I want confirmation!
- Similar confirmation conversation with a human would be very annoying after the 2nd one.
- Repeating/confirmation would be important in some certain scenarios, not always.

Feedback on the Provotype

- Re-think coloured interface for better readability/contrast.
- Add in scrolling feature
- Add in pause feature.
- Participants were hunching over to read it.
- Pink gives a playfulness to it.

The Community Range Filter on a Google Home

Discussing the Community Range Filter at IxDA, Zurich

Community Range Filter

Insights

- Apple does not collect everyone's data, putting in the cloud, analyse it all, use that to make a better interpretation of what you said. For privacy reasons, they keep things locally. But it might benefit from looking how everyone says something, which words do the system commonly get wrong? And learning from that... I'd be happy if it takes my data, collect it in a bigger data pool, to make the voice interaction more accurate, I am happy to contribute to that if in turn I am receiving a better voice service. Plus on the dial means everything that is available: I might want that. But it could be wrong information as well. If I want a specific thing about a specific town, I would expect that local people living there would know better, not someone on the other side of the world.
- Live results on a screen would give you feedback, "oh! This is getting wrong, I will turn it down a bit"
- Turn the dial as the input gets better or worse > fine-tuning - but you need to know the variables!
- You trust the locals for a local recipe, very specific recipe. If you go broader you get the varieties of that recipe / adapted versions. Not the one you are looking for.
- Local expressions, getting correct language use from locals
- Literal vs. Figurative dial
- Your community does not have to be geographically close to you, can be somewhere else and still be your community. How would a button know YOUR community? Assessment?
- English search for general things, German even Swiss German search for local things.
- Do I set it up in the beginning, provide all input all about me? If the privacy would not be an issue, it could learn it. Based on your search history, where your parents travel, etc.
- Search for News: Dial it to conservatives, dial it to liberals.
- Creating bubbles. Do you trust your bubbles?
- I do not trust the bubble assigned to me! >> YouTube suggesting [irrelevant] videos to me.
- Humans change! Your interests/hobbies change! These systems are not aware of it, and keep suggesting things based on your history.

- NEWS: I want to see news from not the usual web site I go to, news are too polarised, I start reading from the other website and I start to understand the other angle >> but always going back to the same website >> EDGE me out from my Comfort Zone! Getting a glance of the opposite opinion.
- Multi-language households > Dial set to different languages. Or, e.g., %15 Urdu - %75 English for people who mix languages in the everyday speaking.
- The dial needs your input. It cannot work it out itself. It cannot know how local you want to be.
- Where does my data go? Can I have control over my data with the dial? Trade off that for more accurate results?
- Dial: (-) Private search (sensitive, medical) / (+) general search (news)
- What do I want my neighbour to know vs. what do I not mind a person in China to know? Association. How likely that someone who knows me will see that.

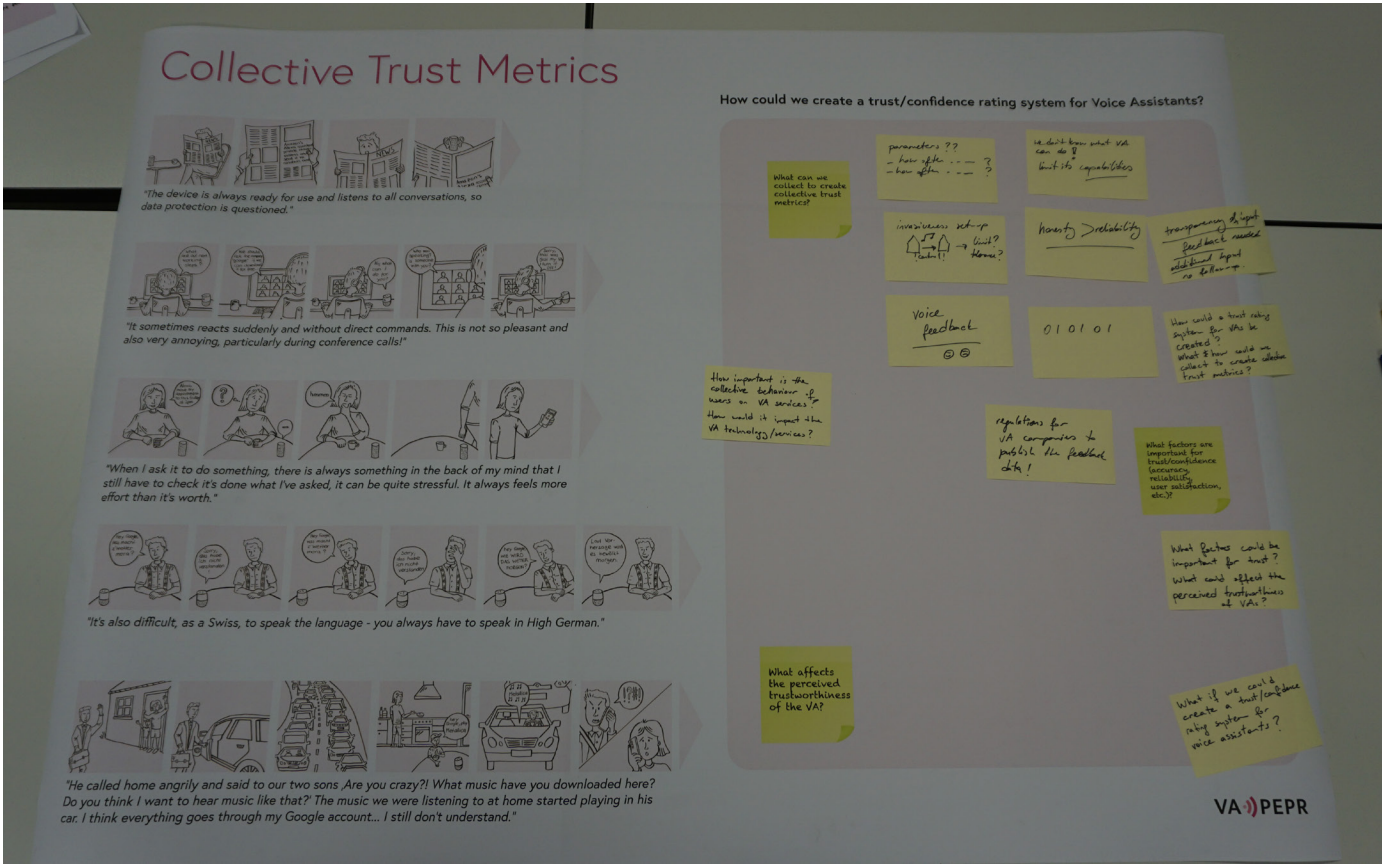
The Independent
Hardware Provotype

Independent Hardware

Insights

- I like the idea that you can take it out completely!
- No card then maybe it's just a speaker.
- What if the SimCards were tasks?
- VA is like a Swiss Army Knife. Here is my digital banking, here is my bla bla... Like an app, the card becomes your app.
- I deliberately bought a Sonos speaker model which does not have the voice input. You can switch it off but I did not trust that button.

Discussing the Independent Hardware provotype at IxDA, Zurich



The Collective Trust Metrics Worksheet at IxDA, Zurich

Collective Trust Metrics

Insights

- Sounds like honesty than reliability. Meeting the minimum standards.
- We know what a refrigerator does, but really we don't know what it does, what it can do! Maybe limit its capabilities.
- Parameters of trustworthiness, list them, rank them, how often does it do this and that?.
- Invasiveness >> Why am I able to control the music in my girlfriend's apartment?! It happened by accident, and she was shocked, all of a sudden the music was invading her personal space!
- There are certain limits what you can tolerate!
- Transparency: I need some feedback from the system after giving a command, I want to add to that and not sure if it understood me. Is it processing? Is it thinking about it? Shall I repeat it? Lights would be OK but then I need to be able to see it, has to look at it. But this works with voice! Face expressions help when talking to people, even on Zoom.
- Sequence in conversation is missing, fluent conversation is missing!
- Buying experience, would I buy this based on the ratings? I would take it into account.
- Giving verbal feedback to the VA: I'd be happy to do that if Apple would get that data "You got that wrong!" It could be part of the learning process of the VA, but you want the actual manufacturer learn from it. Would they give you that stats? Not sure! We need regulations for them to publish some reliability data! Semi-voluntary!
- It's very human to give feedback.
- You'd need granular data!

The WordCloud

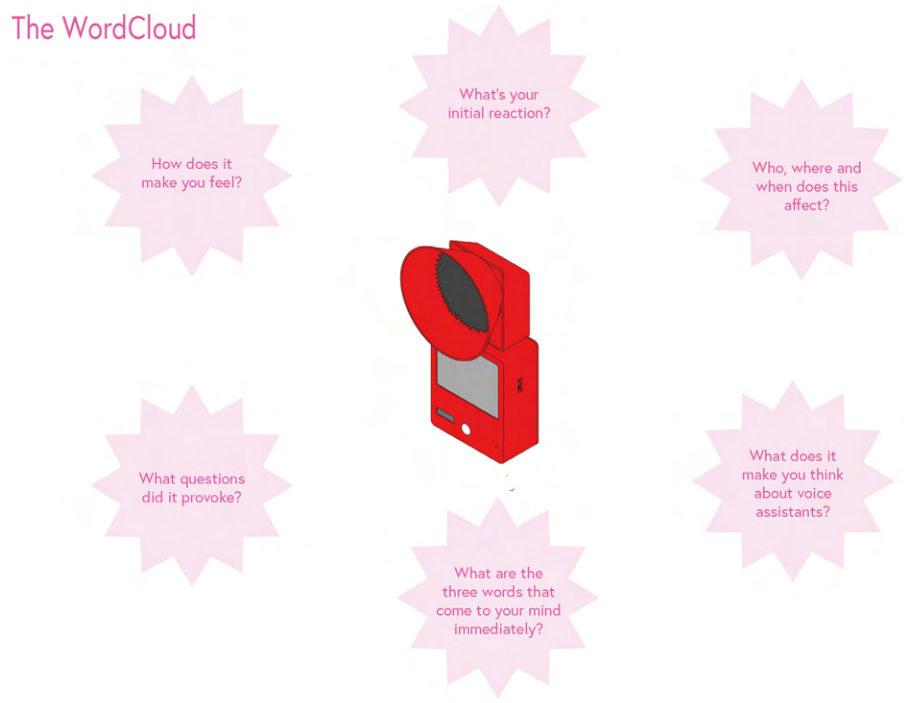
Insights

- Limit it to concepts. What topics/context did we discuss?
- Does the mirroring create more trust or less trust?
- Skewing what you are thinking with data, capture the topics of arguments. "It's really paying attention to what/why we argue".
- Liked it as a way of summarising what we talked here.
- If it was perfect, what would I like this to do? Record everything said around me and play it back to me for purpose? Preparation for the conference next day, so I do not need to take notes? Or anytime there is an argument you can go back and say, look this is exactly what you said? Would we want that? Could be useful but you'd hate that person using such device!
- Could be useful if only I could fully control it: record now, stop now. But it would affect people's behaviour if they new that it's recording.
- What if it could detect things people could do from the intonation of the person?: "I am happy to do that task". But really? This device tells me if it's the case. Could be quite helpful! A bit scary! "I am 60% confident that he is not happy to do the task"
- I think it is quite effective testing what it is like to see it explicit. I could imagine immediately, if you put that in the living room, it's going to have an effect! If there's no screen, there's a different effect. As soon as you put that screen there, wow!
- "I am highly impressed by these provotypes. I find them great. For me the translation of audio (WordCloud) into a cloud of words, this gets into the world of art! It's arranging words in a physical world! Very impressive!"

Feedback on the Provotype

- Limit it to concepts. What topics/context did we discuss? That would get rid of the noise. Mike's comment: Or phrases.
- Classification of words.
- Could you help people to go back to data like Word file where you navigate by the headings? Mike: printer provotype and ethical issues.
- Would be nice to see the words grow as they are said frequently: Live!
- MidJourney AI generating images: Imagine using such AI to create images with the cloud of words on the screen.

Discussing the WordCloud at IxDA, Zurich



The WordCloud Worksheet for IxDA, Zurich


MIT
AGELAB

CHALLENGE 04: LEARNING

Why Digital Offboarding is s Concern

Problem identified

Users must be willing and able to tinker with the VA devices, their WLAN systems, a range of smart products (light bulbs, etc.) as well as a host of apps to maintain the functionality of a device. This requires ongoing learning and adaptation from people to update and maintain a connection. What provisions are in place when learning gets more difficult as people age but at that time depend even more on functioning smart devices and virtual assistants to access key private and public services?

Insight

We found that even participants who had assigned themselves a 5 on a scale from 1 to 6, expressing confidence in their expertise and knowledge of smart technologies in the home, struggled, for example, with illustrated and tested step-by-step installations. Within the frame of our study, there was no harm done. But what if our participants were dependent on telemedicine or digital public services – a scenario quite likely in the near future? Almost all research and development currently focus on the front end of the digital/AI experience and on keeping people 'onboard'. Research into what demands this puts onto users, especially over life-time is lacking.

Challenge:

- Clarify the concept of digital offboarding in the context of continuous learning and self-maintaining
- Collect and map incidents that may lead to digital offboarding
- Identify areas of specific action where digital offboarding may lead to loss of access to vital private and public services

How does the Learning challenge relate back to the wider project?

This challenge seeks to identify design criteria and design approaches to design for transparency, safety and data security as well as positive user experience with desirable VA and IoT in an effort to contribute to greater awareness of the social

and societal implications associated with voice-controlled devices.

It centres on research questions 9 and 10:

9) What negative side effects of VAs emerge in the home? And what is needed to mitigate them in terms of services, tools or regulations?

10) What desirable uses of VA in the home emerge? And what would be needed and how could they be realised?

Participants in our in-home study expected to engage with a 'smart', or rather 'intelligent' speaker. Instead, many found their interactions to be cumbersome and/or limited. Even some of our own researchers disconnected the device they had installed as part of an initial pilot study after a couple of days. They reported that they did not feel the value they got from using the VA was worth the data exposure and the trouble dealing with it. Data exposure was not so much of a topic for our 31 households. But it became evident that using a voice assistant over time requires continuous efforts to learn, adapt, and update oneself along with the change in technologies. People who want to keep using a virtual assistant must put up and keep up with software changes, new versions of a device with new interfaces and a range of new apps and new skills.

The Learning Challenge was triggered by the thought that humans train an initially 'dumb' device but that at some point this device may become 'smarter' than the human. When would such a transition happen and what would happen then? How would the human aging process feed into this? Soon this generated a much more poignant question: At what point does a once digitally self-reliant human struggle to keep up with new technological changes required to remain self-reliant? And what provisions are we envisioning for this situation that is bound to occur in the future? What does this mean for an ageing society where, for example, most public services will have shifted into the digital realm or the metaverse, accessed through VAs and other smart devices, now even more ubiquitous and pervasive in combination with Large Language Models like ChatGPT?

The Learning Challenge pursued these questions following a hermeneutic approach, drawing first and foremost on the observations from the user interactions we documented through the pilot studies, the actual in-home study

with 31 households and the Network Data Traffic Analysis. In all of these, people had to self-install and self-maintain the functionality of the respective device. This situation was an unexpected side-effect of the pandemic which required these studies to be conducted remotely without personal visits. We coupled these findings with self-observations and a literature search on elderly and technology use. In addition, we ran a summer school on the challenges elderly people might experience with new technologies. We presented and discussed the issues we identified with researchers and professionals working on digital public services, including the public sector innovation lab Explorio Lab in Sweden and Team Digitale from the Italian Government.

TASK 1: REVIEW LITERATURE ON ONBOARDING, DIGITAL STUDIES WITH ELDERLY & DEVELOPMENTS IN PUBLIC SERVICES/E-GOVERNMENT

Insights:

- There is no consistent concept off 'the elderly' in digital research. Studies typically refer to people 65+ as 'elderly' without making much of a distinction, as for example recommended by the International Council of Active Ageing (ICAA 2011) which suggests we distinguish between the middle age (45-64), the Young Old (65-74); the Middle Old (75-84) the Old Old (85-99) and the oldest Old (100+).
- Onboarding is not a one-time task for people to get online. It is a never-ending task.
- There is increasing research on the complex issues of onboarding (including a noteworthy 2021 MA Thesis by Laura Østerhaab at Aalborg University titled 'Improving Users Onboarding Experience with Contextual Design - A case study in developing onboarding systems). However, digital offboarding as a concept is only discussed as a way to remove access from staff when they leave an organization.
- The public service depends on digital and AI technologies as it is facing a diminishing work force, greater demands from an ageing population and a lack of funding. This indicates that members of the public, citizens and non-citizens will have to learn to keep up with changes in the public service provision in the digital realm.

TASK 2: IDENTIFY CASES AND SITUATIONS

Examples of instances include error messages like this from Amazon's Alexa:

"Hi Alexa!"

"Sorry, I am having trouble connecting to the internet. It looks like an issue with the router and echo device. So try restarting them. Unplug both of them. Then plug the router back in wait 30 seconds. Once the router is back on and connected to the internet, plug in the echo device."

This spoken message is 20 seconds long. When listening, the person trying to use Alexa tends to be close to the speaker, away from the router, to begin with. The person therefore will hardly see the router. To reconnect, the person needs to know who and what a router and the echo device are, understand what restarting means as well as how to unplug (where, which cable? Which device?) and also be confident to know when the internet is up again. Then again, the person needs to plug in the right cables in the right place. All of this sounds easy enough. But it is not. The task gets harder when hands deformed by arthritis try to manage small connectors and thin cables. Is it necessary to bend/reach down? What about eyesight? Can the person read the labels on the respective devices?

Another illustration of the problem is provided by this interface from an Italian Bank where the Log-In requires users to remember their code but the order of the number changes for every log on attempt.

It does not take much for a person to be shut out from some if not all of their digital affairs: A power outage; a systems update; a stroke changing facial features; a debilitating disease like arthritis preventing finger ID. Even voice is not as reliable when people get horse, for example.

Key Findings: There is a need for more research in this area which warrants an interdisciplinary and intergenerational approach. Our current research points to the possibility that the profession of care taking will have to include a digital component in the future – or there may be a need for.

TASK 3: ENGAGE WITH RESEARCHERS & EXPERTS

Engaging with Future Elderly

We explored the range of possibilities with 40 MA Students in Products-Service-Systems from the School of Design, Politechnico di Milano during a one-week summer school in June 2022. Input from the Team Digitale by the Italian government provided a public service perspective. The resulting five short videos highlight different issues based on students' observations and interviews with today's elderly who use smart technologies. An important aspect of this project was the challenge for students who are today very confident in their use of current technology to envision the challenges they might encounter as technology progresses and they and the society around them age. Will there be someone nearby who an elderly person can wait for and trust to get a connection problem resolved.

Student videos
politechnico di Milano
summer school

Student videos
politechnico di Milano
summer school

Student videos
politechnico di Milano
summer school

Engaging with MIT AgeLab

As part of this challenge, we participated in and contributed to a workshop organized by the MIT Agelab during the CHI Conference 2023 in Hamburg. Using age-suit-style tools to experience the limitations of elderly users suffering from arthritis, and a variety of sight impairments first-hand, we were able to confirm both the challenges involved in and the lack of research into matters of digital off-boarding. Their research – for good reasons – remains focused on keeping people on board, on designing interfaces in ways they can use and access. But one can argue that those efforts have a temporariness to them as they only work as long as the technologies stay the same for people as they age. Yet precisely that is unlikely unless we stop innovating.

The Challenge has produced the foundation for further research that we will now pursue.

MIT Agelab workshop during CHI 2023

CHALLENGE 05: SMARTNESS

The Dimensions of Smartness

It was quickly understood that smartness is not simply a linear scale form dumb to smart, but rather a multi-dimensional quality. These dimensions are also reflective of what humans determine to be smart, from linguistics to emotional understanding and contextual awareness to humour. Below are the team's initial explorations into the different dimensions of smartness that could make up a voice assistant.

Depth of Understanding

- Range of vocabulary («Lower the shutter / blinds.», «When is the next but one lunar eclipse?»), Technical vocabulary
- Syntactic complexity of a user utterance (relative clauses, negation, etc.) - Robustness towards less-than-perfect utterances, e.g. ungrammaticality, stuttering, slurred speech
- Dialect, mixed-language utterances (e.g. German with French or English proper nouns or technical terms)
- Composite queries, e.g. «Where is the next Italian restaurant and how do I get there?»
- Semantics: What did the user mean à Inferencing («When is the next but one lunar eclipse?», "What is the tallest mountain in Switzerland that completely lies within Switzerland?", etc.)
- Pragmatics: Identify user intentions, e.g. «Can you close the blinds?» A command, not a question; a question for the latest movies might mean that the user wants to go to the cinema, and since she is a sci-fi fan preferably to watch the new sci-fi movie.
- World knowledge and inferencing capabilities, for example changing from car to train requires a parking spot to be available and costs extra money; when the user talks about white chocolate with strawberries s/he probably means strawberries inside the chocolate, not added on top (however, depending on context; when talking about desserts the latter might be the correct reading).

Dialogue capabilities:

- Experiments might be necessary to narrow down required capabilities / occurring phenomena:
- Discourse structure, mainly co-references, e.g. «... and the one [lunar eclipse] after that?»
- Dialogue memory (throughout the interaction; but it can be difficult to tell when an interaction ends and the next begins) allows the user and the VA to refer to previous XXX. For example, the user should be able to ask for some details in a cooking recipe the VA has read out. This also includes the support of follow-up questions, e.g. "What is the tallest mountain in Switzerland?" à "Dufourspitze" à "Where exactly is this?".
- Possibility to interrupt the VA when giving an inadequate answer
- Clarification dialogues: In case of misunderstandings or failed understanding the user as well as the VA can start a clarification dialogue "(What do you mean by that?", "What do you mean by blinds?", "Is it that what you mean?")
- Mixed initiative dialogues: Initiating a dialogue by the VA (out of the blue) can make the user feel observed and will then have a negative effect. However, posing a question to the user in an already ongoing interaction or adding a comment/answer that goes beyond the initial question of the user might appear smart.
- After a clarification the VA memorizes a certain user preference, a way of phrasing a request or word for a certain object (e.g. "blinds" instead of "shutters") à see User modelling.

Learning and adaptation (user modelling):

- VA builds and continually updates a user model (users' interests, preferences, etc.), E.g. when a user always asks for the weather and the news in the morning, the VA can answer the question for the weather by also mentioning the recent news. The VA learns in which cinema the user usually likes to go, whether by car or public transport, what kind of pizza she prefers, what the preferable kind of seat in an airplane is, etc. The VA takes

this knowledge about the user into account when accessing a service (after asking for reconfirmation) and providing information.

- The user model helps the VA understand user queries better (better disambiguation due to background knowledge), helps select from several possible answers the most appropriate one (e.g. in the case of health queries)
- User modelling requires a VA to recognize who is speaking.
- The VA memorizes a certain user preference, a way of phrasing a request or word for a certain object (e.g. "blinds" instead of "shutters").
- Situation/context modelling: Depending on the context, a user command may have to be interpreted differently, e.g. the command for closing the blinds probably refers to the room/building the user is currently in (see pragmatics above).
- The VA should know when it is a good time to remind a user of a certain task (e.g. immediately before or sometime ahead).
- Proactiveness: The VA anticipates upcoming information needs and reacts accordingly to provide support.

Transparency:

- Explainability: explain how an answer was derived, why the VA thinks the answer is the correct one

Emotion detection:

- Detect basic emotions of the user (stressed, angry, funny, sad) and react accordingly, e.g. by prompting with a remark ("You seem to be sad today", "You are quite funny")

Anthropomorphism:

- Interacting with the VA is like interacting with a human being.
- Emotion simulation: Showing emotions is not an aspect of smartness; a user might even feel deceived by it.
- Emotion mirroring might be relevant because it can make the user feel understood.
- Personality: The VA has a (consistent) personality. The user should be able to set it via the preference settings. But what if a VA has several users?

Then we might need user recognition so that the VA would be able to show a different kind of personality depending on who is speaking. (. article in Atlantic monthly).

- The VA should show some sense of humour.
- The VA knows and obeys social norms, e.g. no rude language, etc.

Adequate use of multi-modal communication channels:

- A VA with a screen should decide when to additionally show the user query or service results on it. It should be possible to refer to those results when talking to the VA.
- Recognition of gestures and facial expressions: Relevant with robot VAs; has an impact on several of the categories above (e.g. emotion detection, giving commands with gestures)
- A robot VA might give answers by (additionally) using gestures or facial expressions.

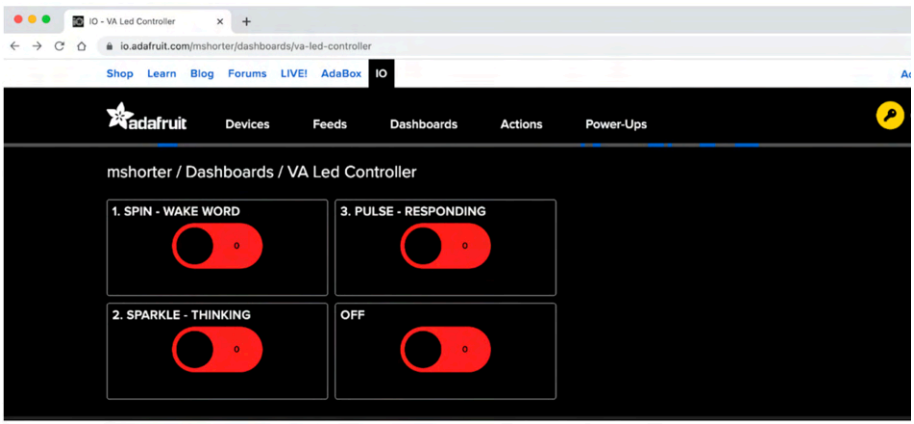
General remarks:

- The dimensions are not (all) independent from each other! What are the consequence, is this a problem?
- Should designers actually try to replicate human-to-human interactions? Or should they rather find ways to communicate the limitations of these devices? How smart do users want their VA to be?
- A VA has no problem solving capabilities and can answer questions only by looking up information from the Web, besides being a natural language interface to access services.
- A VA is often shared in a household and then not purely personal. Certain commands can therefore only be executed via an app, e.g. changes to the personal calendar. What could alternative approaches look like? For example, voice recognition might be reliable enough to identify the speaker, but a voice command could still quite easily be fraudulent.
- Clash of a VA being public in the household and a personal assistant (for more than one household participants) at the same time.

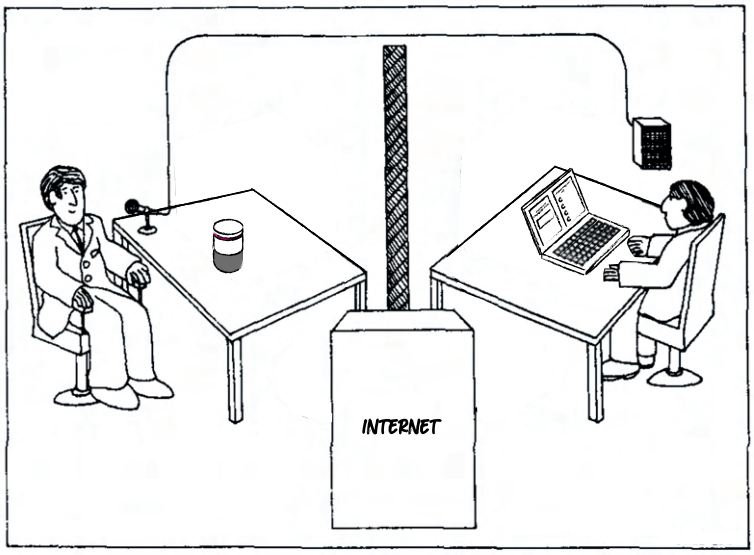
Prototyping a Smarter VA

The WoZVA (Wizard of Oz Voice Assistant) allows researchers to become the 'voice' of a VA. Responding to users commands in a synthetic voice, as well as controlling VA LED behaviours to add some theatre to the user experience. This would allow the team to test out how participants responded to the different dimensions of smartness.

The VA-PEPR WoZVA: 3D printed with an internet connected neo-pixel ring and a bluetooth speaker



The LED behaviour controller using Adafruit IO



Wizard of Oz testing
The smart Voice Assistant VA-PEPR 2022

THE CHATGPT HAPPENED (NOV


WRITERS GUILD ON STRIKE!
WROTE CHAT GPT THIS
ChatGPT LISTED AS AUTHOR ON RESEARCH PAPERS
Many scientists disapprove of articles crediting the AI tool as a co-author.
abc NEWS EXCLUSIVE
INSIDE CHATGPT TECHNOLOGY
ONE-ON-ONE WITH CEO ON THE POTENTIAL AND PITFALLS OF AI
abc NEWS #WORLDNEWSTONIGHT
WE DON'T KNOW WHAT IT MEANS BUT WE'RE SCARED
ATTACK OF THE PSYCHO CHATBOT
Sinister AI computer software admits it 'wants to be human'
Brags it's so powerful that it can destroy anything it chooses
And wants the secret codes that'll allow it to launch nuke bombs
Kicked in the gulags
Toms crisis
Subject to availability. Terms and conditions apply. See p22 for details
FULL NEW TECH

ChatGPT

In November 2022 Chat GPT exploded into the rhetoric of everyday people. A powerful technology that had previously been very much in the development stage that only coders and experts had access to.

We were always aware of Chat GPT and its capabilities but suddenly a lot of the work we were doing seemed obsolete, in particular the Smartness VA-PEPR challenge. Voice Assistant users suddenly knew what a smarter voice interface could look like thanks to Natural Language Processing powers of GPT. Lucky for us one of the big players with this technology was OpenAI, who also had an API that we could get our hands on and start to implement into the work that we were doing.

This new technology might well transform voice assistants and how we interact with our smartphones and speakers. It also has potentially huge implications for people's privacy and the protection of their personal data. In the following section we discuss the challenges raised by this new technology.

For this reason we had to pivot a lot of the work we were doing. Shelf certain outputs and react to our fast move technological world. The remained of this book details how the VA-PEPR project reacted to this new technology.

THE ÜBERBLICK


VIGNERON
ENGAGÉ
RIGUEZ & SANZ
HISBA

The Überblick

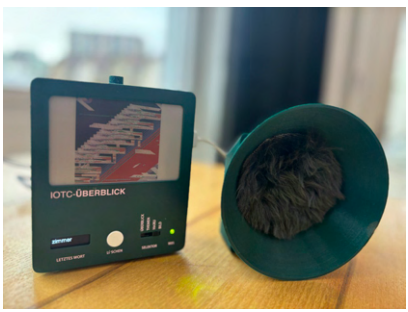
From the WordCloud deployment we learned that the participants living with the device deemed it to be 'not very smart'. They felt that the word clouds being displayed never really captured the conversations happening around it.

Since making and deploying the WordCloud ChatGPT has become something on everyone's minds. This is due to the fact that OpenAI has created a very easy portal for anyone to access and experiment with this powerful technology. Not only this but OpenAI have released an API for this technology. This API allows technologists, coders and makers like those on the VA-PEPR team to integrate ChatGPT into the digital objects we make.

Enter the The Überblick...

We wanted to see if we could make the WordCLoud device smarter by instead of showing a word cloud of users voice data we would show a summarisation generated by GPT. The team decided on four different visual outputs for the new device to display. We engineered prompts that would take everything the dive had head and generate images of the following:

A 75 word summary
Top 5 themes in order of importance
A haiku (Japanese poem of seventeen syllables)
An Image



Design Changes form the WordCloud

The Überblick kept roughly the same form as the WordCloud as it felt very much a development of the same concept.

Design Decision

The new green of the device is a nod to the brand colours of OpenAI. Similarly the new design of the air vent on the back references a graphic mechanism used by OpenAI.

Hardware Changes

The new device has two main changes to the hardware of the WordCloud. Firstly a 7 colour E-Ink screen was added to give a little more interest to the images that the Überblick could display. Secondly the 'mode' slider switch was added to allow users to select what visual output they'd like to see.

Software Changes

The new device integrates the OpenAI API. This allows us to prompt the service to generate not just text based outputs but also image based outputs.

The Überblick Deployment

The team reached out to those that took part in the WordCloud deployment. One of the participants told us that his family felt uncomfortable about having this device installed in their home. For them, the WordCloud made them uneasy, with family members whispering around the device as they did not enjoy their voice data being reflected back at them. To the team this was still interesting, and flt like an insight in itself - really testing a families willingness and conflicts to have these sorts of devices in their homes.

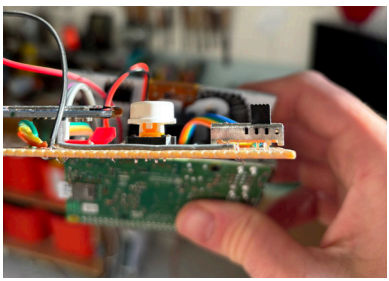
This left us with two families who were willing to live with the Überblick for a 2 week period. The following is a summary of findings.

Radio interference

One participant always had the radio on in the same room as the Überblick. We quickly learned that OpenAI would crash after listening to the news if the content was particularly sensitive, for example a report on the Gaza/Israel situation.

Settling Time

The Überblick takes time to settle. This feels the opposite to the WordCloud display. The more information the device hears the more it seems to make sense of it all.



Development of the Überblick

(Above and below) The Überblick deployed in participants homes

Modes of Summary

Before settling on the four display modes of the Überblick we experimented with other forms of summary. One which we had interest in was the modalities of conversation, how emotional, factual, controversial or sensitive was the audio that the device had captured. The team decided that it gave very little insight into the voice data captured compared to the more creative forms of summary, in particular the Haiku.

Smarter than it is

On a few occasions, particularly with the image, users would think that it is trying to generate an image of them if they see a slight likeness in the picture. We have seen this before with AI devices that generate something for users. People can sometimes forget what is actually happening and think that the device has powers beyond its hardware.

Levels of Glanceability

The Überblick had four output settings (summary, themes, haiku and image). Each of these outputs gave a different level of glanceability to the device. Although the summary certainly gave the most detail, it took time to read and the text is fairly small. On the other side of the spectrum the image, arguably the most glanceable output often gave very abstract results. We found that the themes and haiku were the most useful setting for instantly getting a feeling for what the device had heard. The Haiku setting certainly being the more entertaining and creative.

The python script prompting the JSON response for the top themes setting

```
try:

    response = openai.ChatCompletion.create(
        model = model_id,
        messages = [{'role':'system', 'content': 'You are a helpful research assistant.'},
                    {'role': 'user', 'content':f"Give me the five most relevant topics plus a probability between 0 and 1. \
                    Use only one word for each topic and sort from the highest til the lowest value.\
                    Return it in a JSON object like {json_str} for the following text: {text_data}"},
                   ])

    api_usage = response['usage']
    print('Total token consumed: {0}'.format(api_usage['total_tokens']))

    conversation.append({'role': response.choices[0].message.role, 'content': response.choices[0].message.content})


except Exception as e:
    print(f'during the topics following exception occured:{str(e)}')
    prompts = prompts+1
```

Challenges with OpenAI

Working with the OpenAI API proved more challenging than the Microsoft Azure API. Interestingly some of these issues and bugs came from the subjective nature of working with GPT.

Max Tokens

One of the main issues the team encountered was that OpenAI could not handle the large transcripts that we required it to summarise. This led us to splitting the input text into manageable chunks for OpenAI to process. This allowed us to generate a series of summaries that were then summarised. This summary of many summaries could sometimes put emphasis on small details of a conversation.

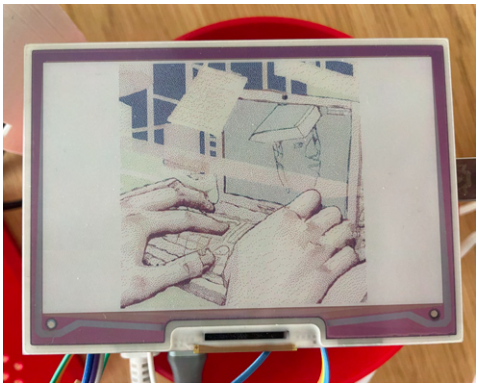
Prompts per Minute

OpeneAI limits the number of prompts you can make to 4 a minute. Because we were requesting 4 prompts for each device (summary, themes, haiku and image) this added up fast. Particularly if we were wanting to run multiple devices from one OpenAI account

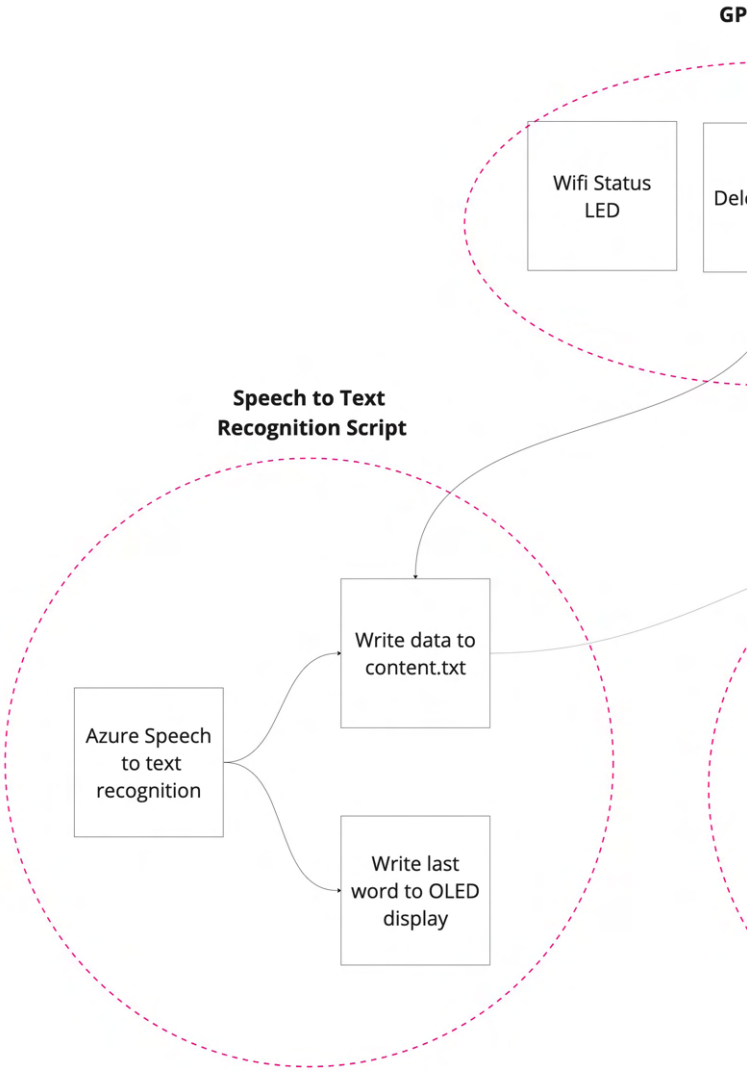
ChatGPT Output

ChatGPT does not deliver a consistent JSON output even when detailed in the prompt. This tends to be a bit of a quirk with ChatGPT where it can sometimes go a little rogue with its response.

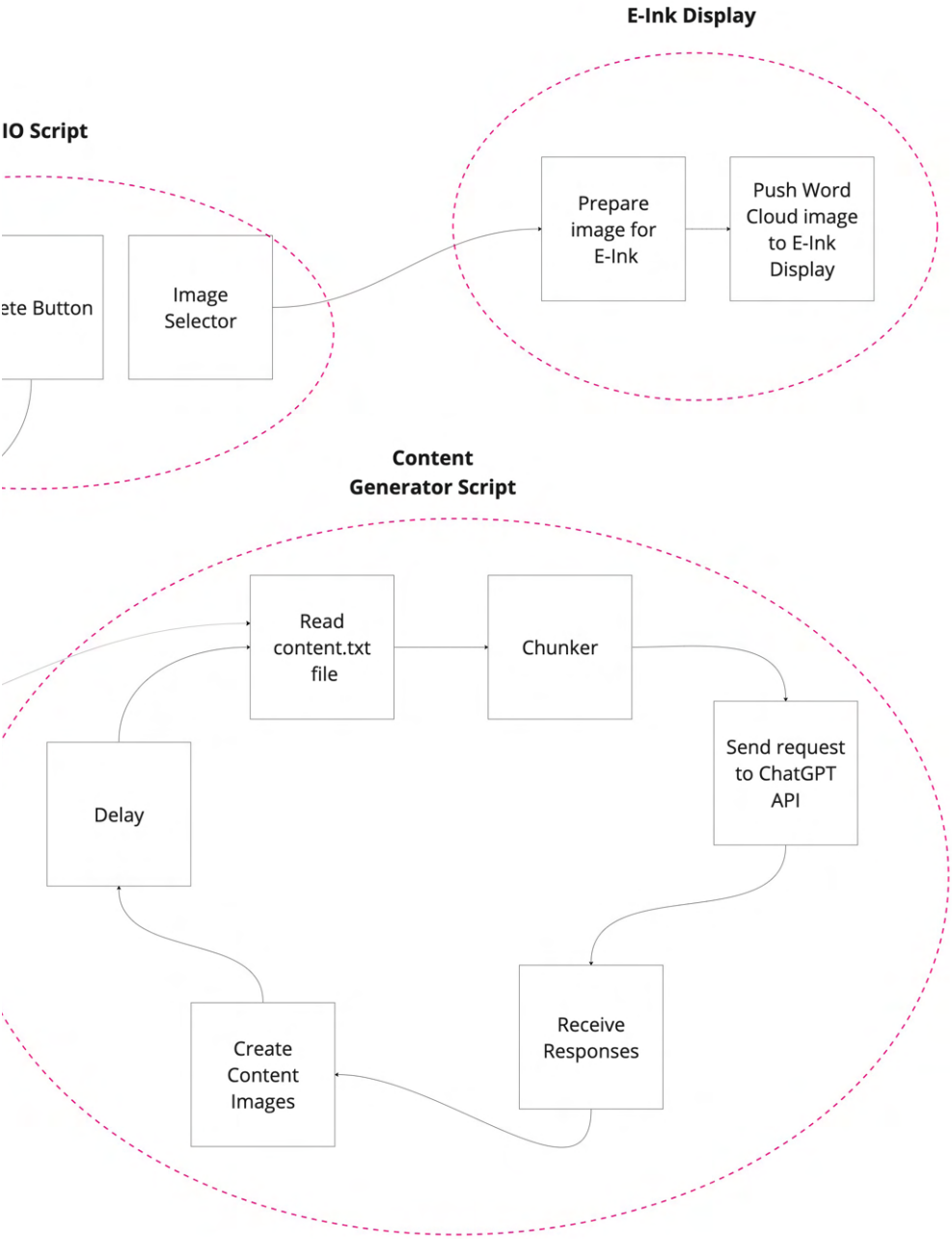
Many entertaining outputs appeared whilst developing the Überblick



IOTC-ÜBERBLICK - Software Process



A Diagram showing the different code blocks that make up the Überblick



(Left page top) The Überblicks ready to travel to Luzern for deployment

Left page middle) ten Überblicks tested and working. Five in Swiss German and five in English

(Left page bottom) The Überblick, the WordCloud and The Confidence machine all side by side

(Right page) Six images giving an example of content generated by the Überblick

VOICE MATTERS

Making Sense of Data

I grew up in rural England where the dominant feature of the countryside were the huge concrete cooling towers of a coal fired electricity power station capable of creating 2GW of electrical power. The power station was demolished in 2015 and is now the site of an Amazon Data centre - a data centre that sits on the infrastructure of a power station. Data is quite literally the new oil (or coal). However, is a power-station-sized data centre what you imagine when you ask Alexa what the weather will be like tomorrow? It is most likely that you imagine the computer for your VA is inside the voice assistant. To comprehend a computer that is over one square kilometre is hard to picture. The disappearing computer goal of the 90s has been achieved. But like any illusion, which the disappearing computer is, it has to have gone somewhere. The rabbit that disappears from a magician's hat has only been hidden - it hasn't actually gone anywhere. The same is true for computation. Yes of course Moore's law has meant that our computers have halved in size every two years, but that's not the whole picture. Our computers that we touch, see and talk to in our loaves have become increasingly smaller because the computation has, like that rabbit in the hat, been moved somewhere else. The cloud. Or as I now know, to the site where Didcot Power Station loomed over the countryside from the 1970s until 2015. A place that once created power, now consumes it. In vast quantities or the scale of the power station.

Humans are poor at perceiving scale. Especially when it is a vast network that is mostly hidden. Out of sight is out of mind. Bridging the scale between the cloud and the product is tough. For the most part we don't notice the distance between the two. We perceive them as one. Only noticing when the internet is down - Netflix won't load, Spotify won't play and Alexa can't hear you. Out with this, our mental model of what an internet connected product is that of a product. It has changed little since electrical consumer goods first came into our home. This concerns me. Our internet connected products are little more than hosts for any internet service. With the rapid increase in the power of AI we have no idea what the capabilities, functionality and connectivity of our products are. Big Tech might be new, but the ideas are as old as the villages that surround Luzern. Our language is peppered with warnings: beware the sheep in wolf's clothing; never judge a book by its

cover; more than meets the eye. Perhaps we need to look at new forms of consumer labels that speak directly to these? Note for design-self - explore labelling based on idioms, sayings and proverbs. To break with the simplicity of the abstraction of logos and certification marks.

Underpinning much of the concern for Voice Assistants is their apparent simplicity - in product form and in interaction. Yet simplicity is hard. I have lost count of the number of times I have quoted Mark Twain's "I didn't have time to write a short letter, so I wrote a long one instead." Or from a dear Italian friend in Mozilla on home cooking - Italian food is simple, but not easy, like graphic design (she is a graphic designer as well as a great cook). Designers and the love of simplicity has proved a honey-pot for market-led technologists like Steve Jobs, who sought simplicity above all else. Yet at what cost? Making the computer disappear has also disappeared fundamental rights to privacy, security, and repairability. The e-waste and power consumption legacy of simplicity, the prestige of the magicians in Silicon Valley, hidden from the view of consumers but not from the effects on the climate. The mathematician and philosopher Alfred North Whitehead's advice "seek simplicity, but don't trust it" is as true now as it was when he wrote this over a hundred years ago.

So if we are not to judge a book by its cover, then what should we do? In our work, we were obsessed with instruments that help us to see the unseen. Barometers, thermometers, compasses, clocks, weather vanes, electricity and gas metres, voltmeters, heart rate monitors. All these things that show us in our personal and professional lives that make the invisible visible. We can also bring things into sharp focus through magnification - the surface of the moon or the cells of our skin. X-ray machines, sonar, radar and MRI scanners can see through things or detect far off things. In the Museum für Naturkunde (Natural History Museum) in Berlin, there are these incredible models of insects - as if looking through a microscope, but unaided. The sculpture Alfred Keller made scientific models of insects amplified in size a hundred times. Made between 1930 and 1955, these models enable us to see the unseen. To marvel at the complexity that lies in the eye of a fly or the hairs on a flea. To see the complexity of nature itself.

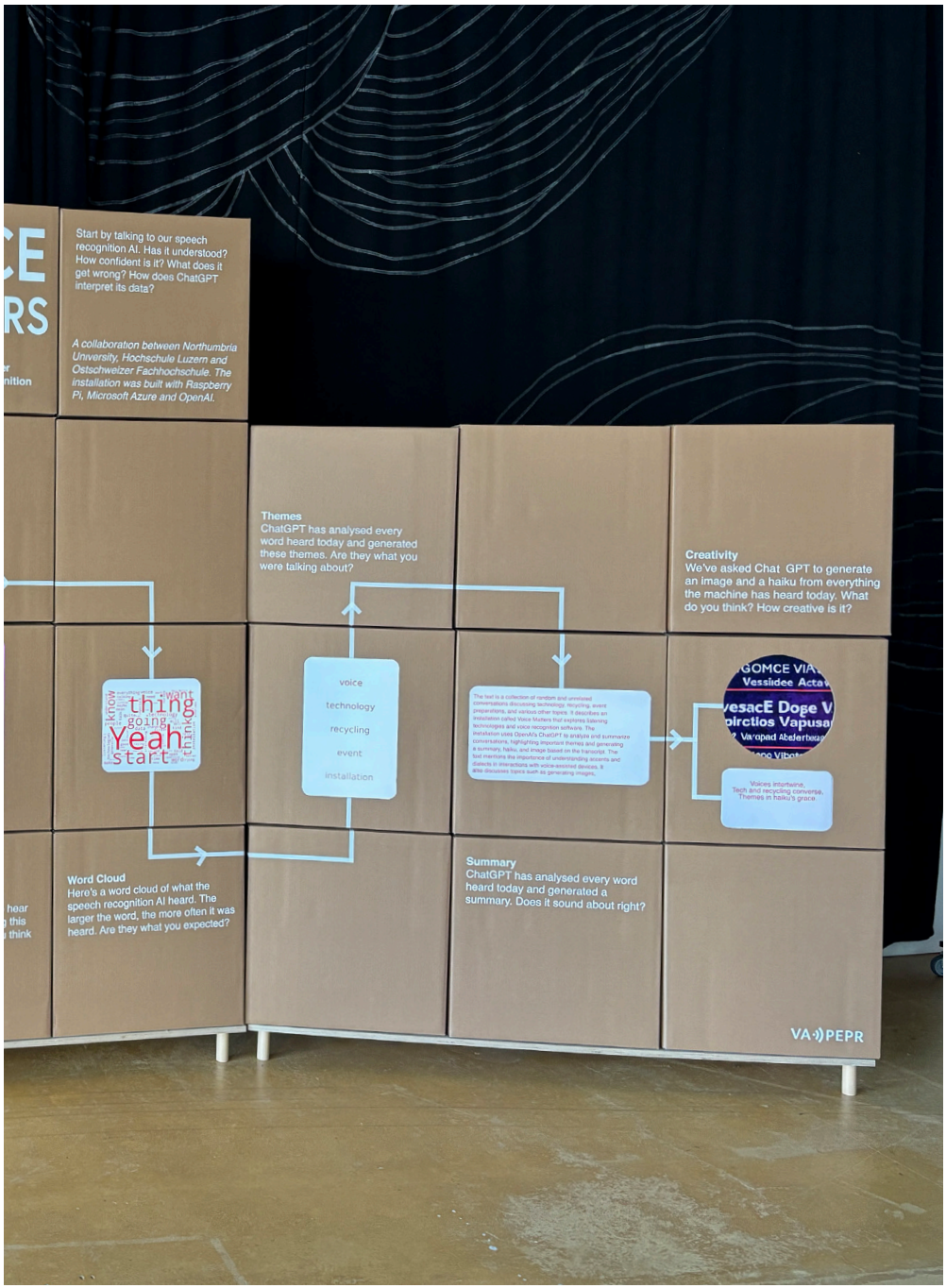
The trouble with AI in our Voice Assistants is that it is hidden, literally and metaphorically, behind layers of camouflage. The first layer starts with the shape of the product. AI needs a form. At some point it needs to take data from the real world (our voices), do something to it (process it), and make an action back in the real world (play a song or start an egg timer). The first layer is the product form - the cover if you like. While the cover might look like a loudspeaker, it is actually a very powerful microphone (have you wondered why smart speakers don't look like smart microphones?), that converts your voice into an electrical signal. This signal is then packaged up as data and sent to a far off data centre (the one in the UK is 500 miles from my house in Scotland). It is here it meets the speech to text processing and AI. The AI is formed of digital layers of millions or billions of processing units (neural nodes) that connect in multiple layers forming billions more connections. This data is impossible for humans to interpret. It is simply too vast. It is a true black box. Layer upon layer of black boxes. Hidden in semi-secret data centres. Vladan Joler and Kate Crawford took us on an artist's journey into the secret life, or anatomy, of an Amazon Alexa. The results are mind boggling.

For the Voice Matters installation we were inspired by Joler and Crawford's incredible map, and by way Alfred Keller's scaled up models of tiny things. But with a twist. To do this with a living (or rather working) AI that visitors could interact with. A Voice Assistant that people can at least see the modules of. To see the moving parts. A physical interactive sketch of the most important parts of a Voice Assistant.

In some ways Voice Matters is a manifestation of all the work that has come before through the VA-PEPR project challenges. It explores materialising the immaterial, data and privacy, confidence and trust, smartness, policy and learning.

(Opposite page) various domestic meters.

Voice Matters delves into the power of AI to listen, understand, and interpret our conversations. When people interact with the installation they see just how much of our daily conversations can be comprehended and summarised by the smart devices that surround us in our homes.

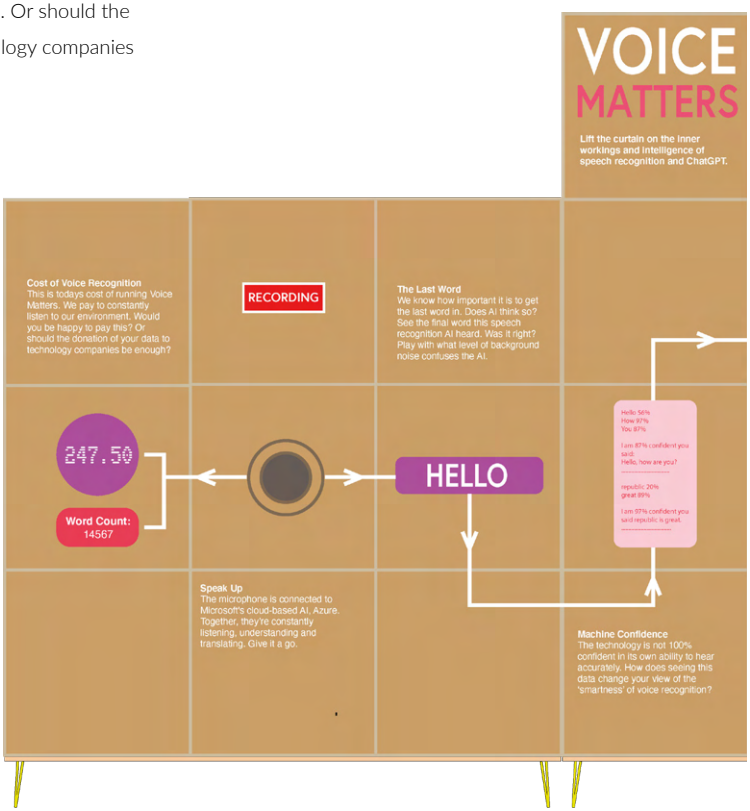


Screen 1: Cost of Voice Recognition

This screen displays the cost of running Voice Matters, which includes Microsoft's real-time speech to text service and OpenAI's Chat GPT service. We pay to constantly listen to our environment. We wanted viewers to consider if they would be happy to pay this. Or should the donation of their data to technology companies be enough?

Screen 2: The Last Word

This screen displays the last word Microsoft's speech to text service thinks it heard. This allows viewers to see how accents are understood, but also see how sensitive the hardware is.



The Microphones

The microphone array in Voice Matters is connected to a series of Raspberry Pis and runs Microsoft's cloud-based AI, Azure. Together, they are constantly listening, understanding and translating everything that has heard. This data is then added into an ever growing text document that is then analysed in various ways. This is displayed across the 7 screens of Voice Matters.

Screen 3: Machine Confidence

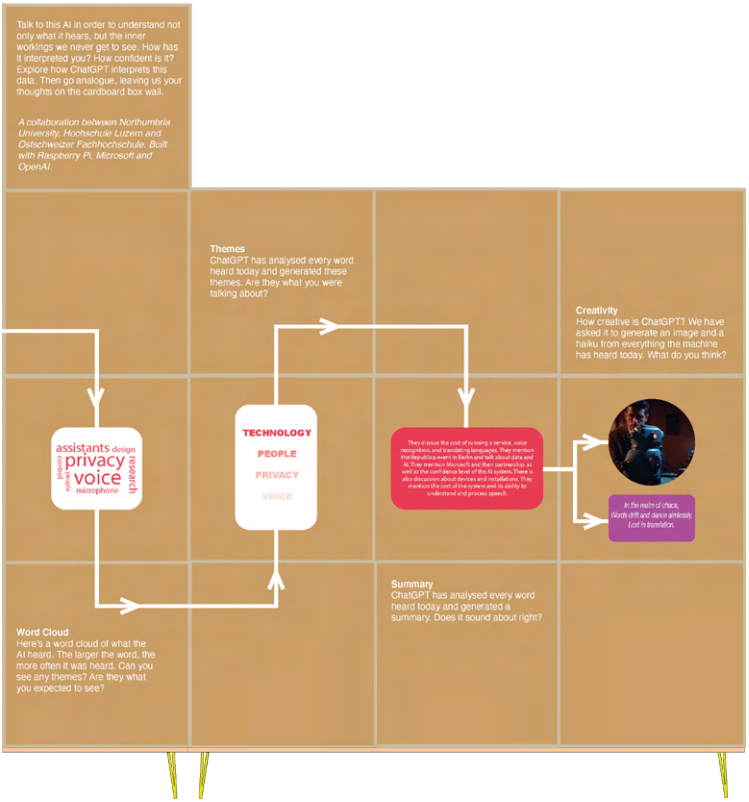
The technology is not 100% confident in its own ability to hear accurately. This screen lifts the lid on this black box technology and allows viewers to see how confident the AI was in all the words that it heard. And more interesting the words that it did not hear: the words where machine learning is used to fill in blanks.

Screen 4: Word Cloud

This screen displays a word cloud of what the AI has heard. The larger the word, the more often it was heard.

Screen 7: Creativity

How creative is ChatGPT? This screen generates an image and a haiku from everything the machine has heard.



Screen 5: Themes

This is where we use the power of ChatGPT to analyse the voice data and try to create a meaningful summary. This screen displays the output from prompting ChatGPT to generate the top five themes in order of importance.

Screen 6: Summary

This screen displays the output from prompting ChatGPT to generate a 75-word summary from the voice data.

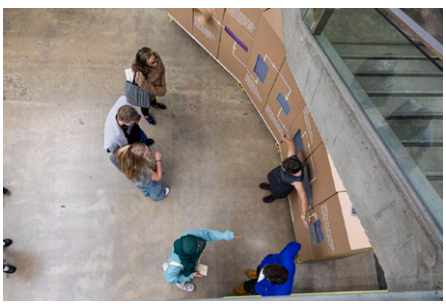
Voice Matters was been exhibited at:

Re:publica 2023, Berlin, Germany

The V&A, 2023, London Design Festival, London, UK

HSLU, 2024, Vision Emmen, Luzern, Switzerland

HSLU, 2024, Policy, Design and AI Workshop, Luzern, Switzerland



(Opposite page, this page top and bottom) Passers by enjoy Voice Matters at the V&A during London Design Festival

Claudia Roth, Germany's state minister for culture and media experiencing Voice Matters whilst at re:publica, Berlin

Some Reflections on researching Voice Assistants through Design

The Importance of Design Friction

We don't agree with the idea of the disappearing computer. We don't think that frictionless interactions are what we want when we are designing with unknown cloud-based intelligent algorithms that can listen to everything we say and make machine-based decisions on our behalf. We think designers need to reverse this tech-led trend by designing friction back into frictionless systems.

Make AI Human Readable

To help humans read machines, we need machines that can read machines. Like we have electricity readers to tell us how much electricity we have used, we want machines to report back on the measurable qualities of voice-based AI. This is related to introducing friction into design, but different in that they are calling for designers to make instruments that can report on what is happening in the cloud through dedicated devices. We are calling these Instruments of the Cloud.

Honesty in Form

Our voice assistants currently look like speakers, when the core AI technology in them is, in fact, a microphone. When the microphone feature is hidden users can very easily forget this thing has the power to always be listening. If these devices are more honest in their form, it creates a healthier, more transparent and honest environment.

Give Black Boxes Windows

Our AI devices are increasingly present as black-boxes. We put content in. Something (unknown, but powerful) happens and that is the limit of our understanding. We have learnt that it is easy to refocus AI back on itself to provide windows into these black boxes. Windows that, like they do in architectural design, add increased value to the overall design of AI-based products and services. Windows that do not compromise privacy and security, but do provide transparency and human readability.

The Fragility of Machines

When we interact with voice assistants we are never sure how confident the device is in what it's heard from the human, or how confident it is in the information it is delivering to you. When users see that machines are not always 100% accurate it can help to build a stronger machine/human relationship - eventually building more trust and confidence.

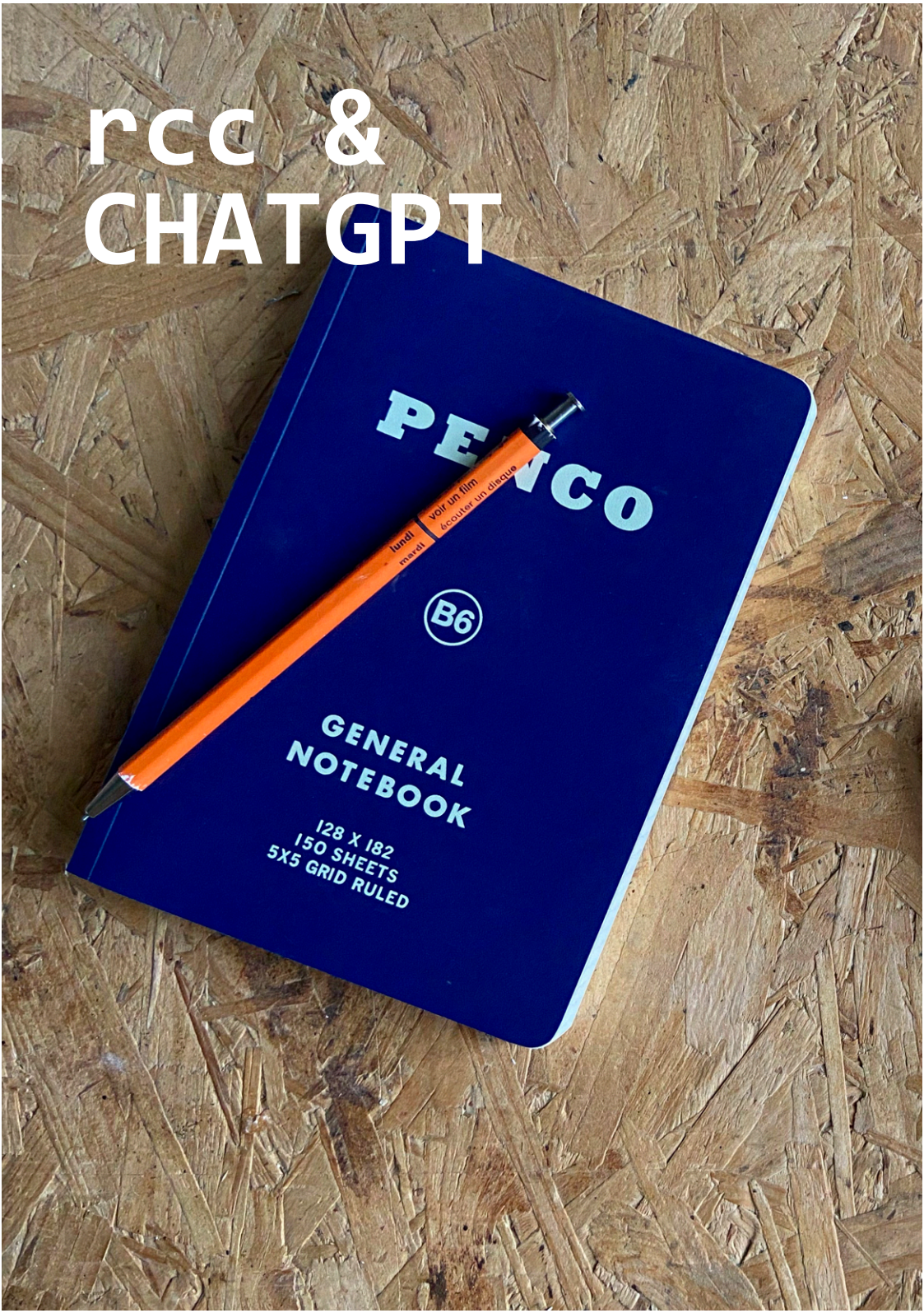
The Financial Cost of Researching AI

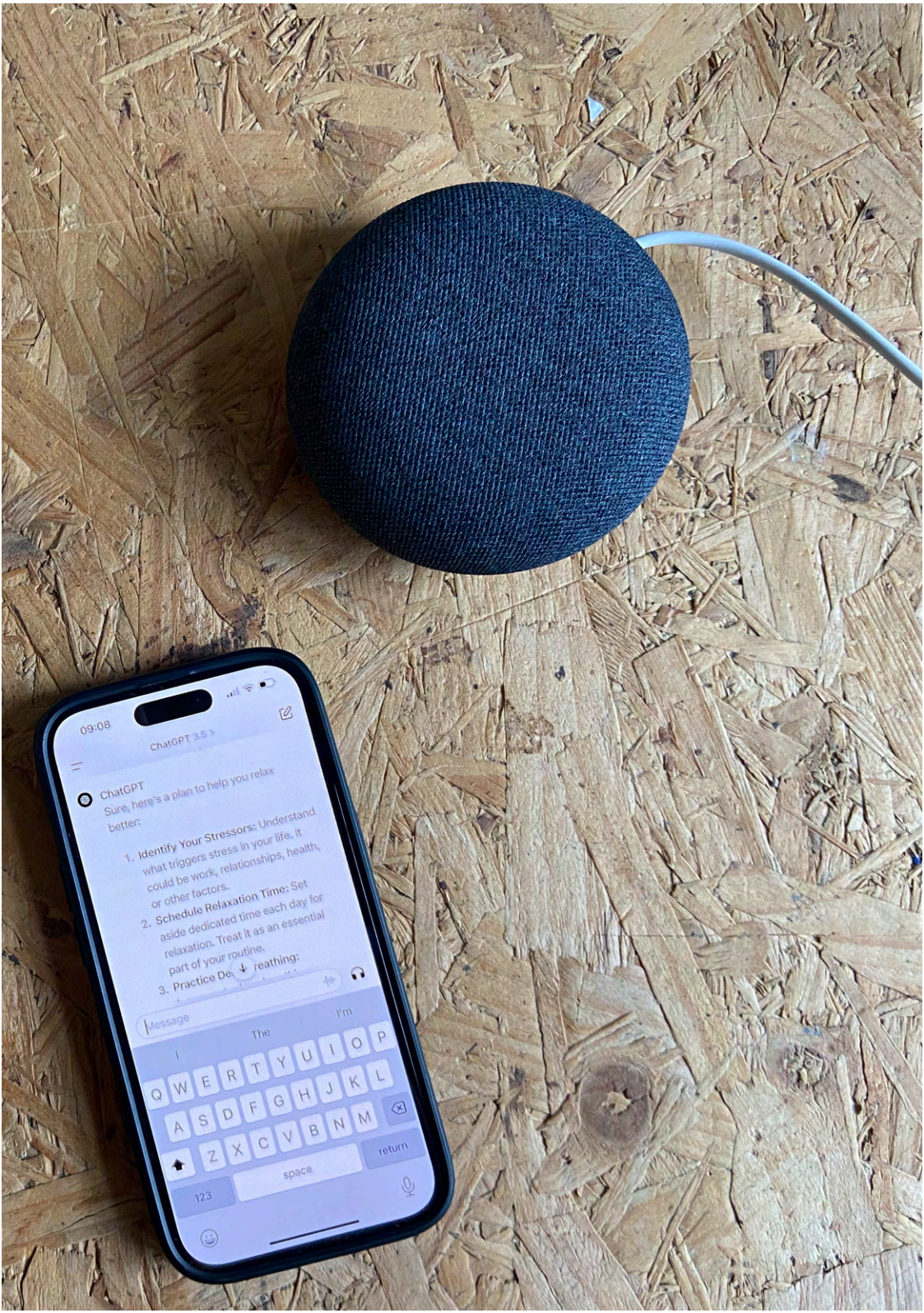
By lifting the hood on VAs by using commercially available APIs, we encountered life outside the "we are the product". It is here we started to see the costs involved. They were not cheap. Overall we spent close to 3000CH on cloud-based services to develop these prototypes. These were not costs we expected to encounter - given the "free" nature of the use of AI in our homes. We of course are paying for these services, in our data, but not in our finances. When you want to use the services for your own purposes, it is here you start to see the true cost.

AI From Big Tech is Driving a Less Diverse Sector

In order to counter this we require a huge effort to join the dots between multiple sectors, across disciplines and through open practices. If we are to have a thriving small-tech industry that is greater than the sum of its parts, then there needs to be a radical change in how global governments support a more diverse sector.

rcc & CHATGPT


PENCO
B6
GENERAL NOTEBOOK
128 X 182
150 SHEETS
5X5 GRID RULED
lundi
mardi
voir un film
écouter un disque


09:08
ChatGPT 3.5
ChatGPT
Sure, here's a plan to help you relax better:
1. Identify Your Stressors: Understand what triggers stress in your life. It could be work, relationships, health, or other factors.
2. Schedule Relaxation Time: Set aside dedicated time each day for relaxation. Treat it as an essential part of your routine.
3. Practice De reathing:
Message
The
I'm
space
return
123

rcc & ChatGPT

As already described in Workbook 1, the rcc study has so far given us an insight into the attitudes, behaviours, practices and routines of a younger age group.

Overall, the rcc study gave us a better impression of the general knowledge and awareness of the use of VAs among digital natives in Switzerland. It became very clear that the use of "pure" voice-guided assistance is hardly used at all due to its poor quality and inconvenience. VAs are used almost exclusively on mobile devices such as smartphones, notebooks/laptops and, to some extent, on tablets. Smart speakers are hardly ever used by the "digital natives" in the rcc study.

In the spring and summer of 2022, two further studies were conducted with a total of 85 digital natives who took part in a resource management project (n=65) and creatively explored the future of the data economy with the use of virtual assistants in 2037 (n=20). The participants described their attitudes towards data protection as well as their actual and intended behaviour in data protection-sensitive situations and contexts. The participants were also asked about their recommendations for measures at political, economic and regulatory level.

There is generally a strong awareness of the risk of confidential data being passed on. However, little is known about the fact that the leakage of seemingly harmless data can be combined into meaningful profiles through the use of AI-based technology (Riss et al., 2022).

The results show that while the majority are concerned about privacy, they are willing to disclose personal data if the benefits outweigh the risks (Maier, et al. 2023).

The prevailing attitude is characterised by uncertainty about what happens to their data. However, this is accompanied by a sense of powerlessness when it comes to controlling its (mis)use. There is a clear mistrust of large technology companies and resignation in the face of the challenge of protecting their own personal data in a digital world.

Few of the participants state that they take steps to protect their privacy themselves, but rather rely on "the government" to take measures at a political and regulatory level. The specific expectations of the state and government relate primarily to improved transparency and communication, more training and further education opportunities, as well as optimised data protection measures and guidelines.

At the same time, the responses show a low level of awareness of existing or planned laws, but often point out that data protection issues require global rather than national solutions in any case. Some participants even encourage their peers to defend the analogue world and limit digitalisation in general, because UA offerings only open the door to surveillance and abuse. (Maier, et al. 2023).

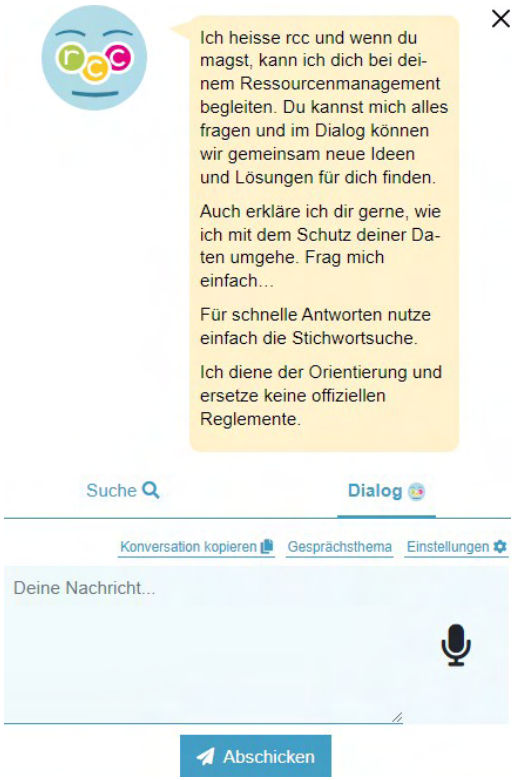


Figure 1.

Even before the start of the VA-PEPR project, the rcc platform provided "software-supported, feedback-giving" tools to help users navigate through the resource management project. During the corona virus pandemic and even before "ChatGPT" became accessible, there were plans and initial experiments to use chat-bots to meet the increasing demands on software. However, these were abandoned due to poor functionality and excessive costs.

Since the accessibility of "ChatGPT" (https://openai.com/) from the end of November 2022 and based on the results of the VA-PEPR study that users would use VA if the functionality is right, but in terms of data protection would rather rely on "the government" taking measures at political and regulatory level to ensure that providers have to guarantee serious data protection, the original plans for a separate rcc-VA have been taken up again.

However, after initial euphoria, the first attempts to implement rcc assistance by spring 2013 were rather sobering. The response speed was too slow and the quality was accompanied by too many hallucinations on the part of ChatGPT3.

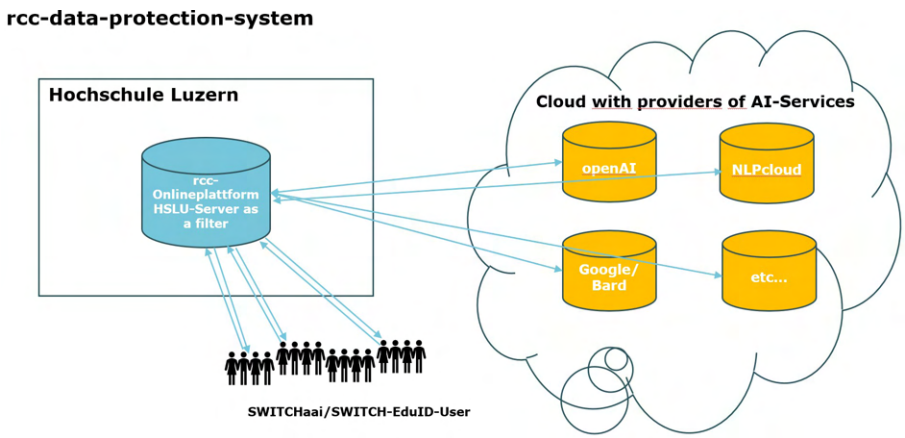


Figure 2

Combining rcc with ChatGPT4

Only the launch of ChatGPT 4 in spring 2023 made it possible to develop a topic-specific "prompt engine" for a financially viable, functioning and AI-based rcc assistant. Among other things, a powerful index table is used for speech and voice assistance and embedded vectors are generated that allow a satisfactory response speed and minimise hallucinatory response behaviour.

The AI-supported assistance currently implemented in rcc (rcc-VA) makes it possible to be coached in various languages - including Swiss German - and completely anonymised, without OpenAI, Google, etc receiving any personal information. In developing this elegant and not entirely trivial solution, the project management acted as a kind of "government" and took measures at "political and regulatory level", with the HSLU server acting as a highly efficient filter (Figure 2). In addition, the rcc-VA pro-actively explains data protection to users within the responses at irregular intervals and can also be specifically questioned about this.

Around 110 students who carried out a resource management project in 2023 were asked about their attitude and usage behaviour towards AI-based tools, including data protection, at the beginning and after completion of the project.

Even at the beginning, 86% stated that they generally use AI-based tools. The rcc assistance offered since autumn 2013 was used by more than 90% and very quickly recognised as a useful and data-protected tool for project management. This could possibly explain the increase of 9% at the end of the project to 95% use of AI-based tools overall (not just the rcc assistance) and be interpreted as growing acceptance and integration of AI tools in the everyday lives of the respondents.

The frequency of use increased in spring 2023 (89%) in terms of use several times a week and day compared to 2022 (72%). The number of people who use VA several times a week has increased significantly in 2023 (48% compared to 14% in the previous year), while the number of people who use VA several times a day has fallen (41% compared to 58% in the previous year).

The results on the tension between data protection vs. benefits and range of functions are impressive. Although there is a growing awareness of data protection, the willingness to make compromises for the benefits of AI tools continues to increase, as shown in figures three and four.

Results so far

The results of the rcc study analysed so far make it clear that the younger age group from the academic spectrum express a clear need for transparency, skills and clear guidelines when using AI tools. This also reflects a growing awareness of the ethical and social implications of using AI technologies.

Overall, these changes point to a profound transformation in the way digital technology is shaping education and everyday professional and private life. While AI tools have the potential to improve the productivity of work and study, they also raise important questions about privacy, ethical use and the impact on future social interactions and communication skills.

These developments require careful consideration and response from policy makers and society as a whole to ensure that the technology is used for the benefit of wider society. Because experience has shown that government processes do not excel at finding solutions quickly and implementing measures rapidly, the educational institutions cooperating in VA-PEPR could use the example of the “rcc-data-protection-system” to act as a role model and, on the basis of today’s findings, effectively accelerate this process in subsequent projects by further researching and implementing the solutions identified so far.

Datenschutz vs. Funktionsumfang

Datenschutz und Nutzen müssen sich nicht unbedingt ausschliessen. Aber in der Realität müssen wir oft abwägen, ob wir bereit sind, beim Schutz unserer Daten Abstriche zu machen, um in den Genuss des vollen Funktionsumfangs einer Software oder eines Tools zu kommen.
Wie entscheidest du dich in der Regel?

Option	Wert	Total Antworten	Anzahl Nennungen	Prozent
Nutzen wichtig	2	42	5	12%
Nutzen eher wichtig	1	42	10	24%
Neutral	0	42	16	38%
Datenschutz eher wichtig	1	42	9	21%
Datenschutz wichtig	2	42	2	5%
Summe			42	100%

64%

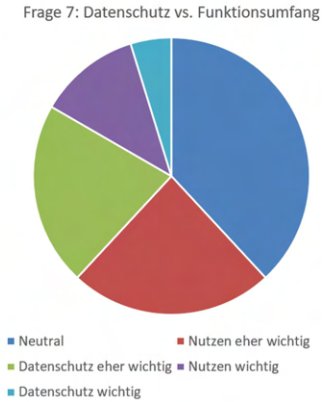


Figure 3.

Wie wichtig ist dir der Datenschutz bei der Nutzung von KI-Tools?

Option	Total	Anzahl Nennungen	Prozent
Mir ist Datenschutz zwar ein Anliegen, aber wenn die Vorteile von KI-Tools überwiegen, gehe ich dabei auch Kompromisse ein.	42	21	50%
Mir ist mehr oder weniger egal, wo die Daten gespeichert und verarbeitet werden, solange die KI-Tools funktionieren.	42	10	24%
Nicht wichtig.	42	6	14%
Mir ist der Datenschutz so wichtig, dass ich meinen Gebrauch von KI-Tools teilweise bewusst einschränke	42	5	12%
Summe		42	100%

88%

Figure 4

rcc-study 2020-2023 (n ≈ 400)

Zeitraum	2020 FS	2020 HS	2021 FS	202
Titel	Pilotstudie 1 (Zwei Module)	Pilotstudie 2 (Ein Modul)	Erhebung 1 (Zwei Module)	Erh (Ein
Anzahl	n (13 + 33) = 46 (84 Journale ≤ 2019)	n = 28	n (27 + 42) = 69	n =
Bemerkung	**Stichprobenanalyse aus 560 *«alten»* Journalen MAXQDA/ Erfassung Aktivitäten mit und ohne Voice Assistants**	**Erstmalig VA-Frage-bögen zu Semester-beginn und –ende eingesetzt**	**Erstmalig VA-PEPR-relevante soziodemo-grafische Daten ergänzt und erfasst**	**Iden FS 2**
Fragebögen	(keine)	**«Voice Assistance Umfrage»** **«Voice Assistance Umfrage Abschluss»**	**«Voice Assistance Umfrage»** **«Voice Assistance Umfrage Abschluss»**	**«Voi Umf** **«Voi Umf**
Animations-projekt	(Standard)	(Standard)	**Animationsprojekt VA mit Auswahl 1..3**	**Anir mit**
Aktivitäten-erfassung Erweiterung Formular	**Feld 1: VA-System / Name (z.B. Siri etc.)** **Feld 2: VA-Aktivität /** **(z.B. «Musik starten, Duolingo Spanisch lernen»)**	**Liste von VA-System, VA-Gerät, VA-Aktivität, Erfahrung, Unterschied zu ohne VA, Ort**	**Liste von VA-System, VA-Gerät, VA-Aktivität, Erfahrung, Unterschied zu ohne VA, Ort**	**Liste VA-(Erfa zu o**
AI-Protokolle (rcc-Assistenz)				

Figure 5.

21 HS	2022 FS	2022 HS	2023 FS	2023 HS
ebung 2 Modul)	Erhebung 3 (Zwei Module) plus Summerschool	(keine)	Erhebung 4 (Zwei Module)	Erhebung 5 (Ein Modul)
35	n (30 + 35) = 65	n = 35	n (30 + 35) = 65	n = 44
tisch mit 21	**Individuelle Zeichnungen zu Datenschutz/Filme zu Datenschutz und Umgang mit VA (Summerschool)**	Keine Daten-erhebung für VA-PEPR	**Erste Angebote zu rcc-Assistenz und Ideen und Wünsche an rcc erfragt**	
ice Assistance rage» ice Assistance rage Abschluss»	**«Virtual Assistance Beginn» (403, 362)** **«Virtual Assistance Abschluss» (403, 362)**	(keine)	(**Virtual Assistance Beginn» (403, 362)** **«Virtual Assistance Abschluss» (403, 362)**	**«Virtual Assistance- und KI- Umfrage Semesterbeginn» «Virtual Assistance und KI-Umfrage Semesterende»**
nationspro-jekt VA Auswahl 1..3	**Animationspro-jekt Virtual Assistance Zeichnungen zu Datenschutz**	(Standard)	(Animations-projekt Standard)	**Animationsprojekt mit KI**
e von VA-System, Gerät, VA-Aktivität, hrung, Unterschied hne VA, Ort	**Textfeld VA Aktivität**	(Standard)	(Standard)	(Standard)
				KI-Dialoge/ KI-Stichwortsuche

ETHICS & CHATGPT

Would you let a Talking AI into your home?

People find today's voice assistants to be neither smart nor particularly useful in everyday life. Incorporating large language models like ChatGPT could address these issues, but it also raises serious ethical concerns about privacy and data protection. This paper focuses on the ethical implications of incorporating ChatGPT or similar natural language processing Artificial Intelligence (AI) into voice assistants, particularly in domestic settings. It explores the challenges raised by these devices, including concerns about privacy and the gathering of personal information. It also discusses recent technical developments in the field of large language models and their potential for transforming voice assistants. Ultimately, this paper seeks to address the ethical challenges associated with the emergence of large language models in smart home devices. By doing so, we aim to shed light on how these challenges could be addressed and to contribute to a more nuanced understanding of the transformative potential of these devices.

Introduction

When we watched the proof-of-concept video by Josh.ai we felt that finally one's smart home voice assistant was on the way to becoming really smart. The video shows how incorporating artificial intelligence (AI) language models into voice assistants promises to overcome many of the shortcomings of today's voice assistants.

Our own empirical data collected in the years 2021 and 2022 showed that from the point of view of users, today's voice assistants were neither really smart nor of much use in everyday life. Many of the requests made by users were not understood or were misunderstood by voice assistants. Users also complained about the lack of feedback or the impossibility of asking follow-up questions since the voice assistant did not remember what a user said earlier in a conversation.

Some participants of our in-home study gave up using the voice assistant after only a few days because it was too tedious for them to get satisfactory answers and activities. It was easier to unplug. Participants who continued using their voice assistant appreciated its timer and alarm function and used it for listening to music and requesting fact-based knowledge. The households equipped with smart home devices also used the voice assistant for switching lights on and off, opening the shades or turning on the music.

Today, voice assistants usually require precise language and often confuse basic smart home commands with requests for information, which results in frustrating and sometimes useless responses. The new generation of large language models (LLMs) used by ChatGPT and other chatbots promise to transform

today's voice assistants into something much more useful. According to the CEO of josh.ai '... no one's going to be willing to tolerate the old way that Alexa, Google, Siri, and even Josh, operated'.

Since voice-enabled agents are embedded in people's private spaces and domestic lives they can gather enormous amounts of personal information. This is why they evoke serious privacy concerns as confirmed in many studies as well as in our own research. Even though most people claim to care about privacy and protecting their personal data, few read the data protection regulations that come with the device, as these are seen as too long and/or difficult to understand.

The privacy concerns raised by voice assistants might well be exacerbated if such devices are enhanced by integrating ChatGPT. Since its release in November 2022, it has already triggered debates about the ethical challenges associated with the emergence of large language models. Most discussions deal with issues such as copyright, attribution, plagiarism, misinformation and accountability as seen in recent articles and guidelines issued by the World Association of Medical Editors or the Lancet.

Ethical challenges of conversational/talking AI

As shown in many studies attitudes are poor predictors of behaviour. One of the reasons given for the dichotomy between attitude and actual or intended behaviour is that studies of privacy tend to measure only general attitudes, while behaviour is context-specific and individuals may perform privacy risk assessments but choose the most viable or convenient options, even if they are not in line with their privacy preferences. Besides, according to the 'privacy calculus model', people would be prepared to trade-off their privacy in exchange for a more beneficial or convenient technology.

Given the current limitations, many participants in our study felt that the benefits to be expected did not outweigh the risks associated with disclosing personal information. However, many of the shortcomings revealed in our study appear to have been overcome with the rise of deep learning and large language models such as GPT-3. These have become increasingly effective for generating text or speech that appears as if it were written or spoken by a person.

For the smart home, in general, the promise lies in combining the conversational abilities of AI language models with the context which a home automation system can provide. For example, by knowing what smart devices you have in your home and details about how you use them, Josh could parse natural language commands into actions in your home. And as mentioned in Section 3, smart home enthusiasts have already figured out ways to use Siri Short-cuts to get ChatGPT into their smart home. After all, it is a lot easier to talk to a smart speaker than type into a web browser.

Trade-off between the benefits of talking AI and potential risks

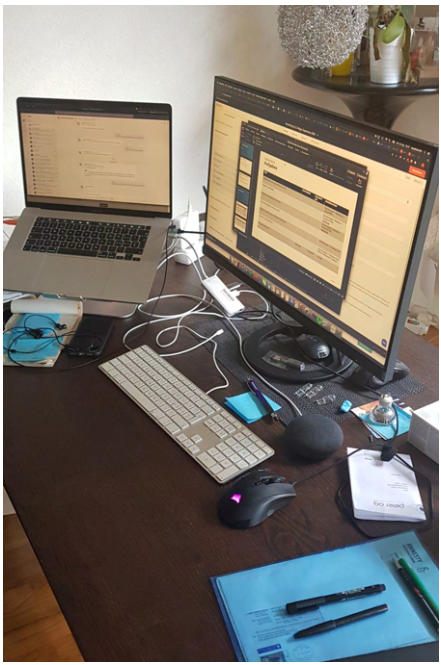
Even though currently available commercial voice assistants are not considered really smart by users, they nevertheless raise expectations as to their capability of reasoning or human-like conversation because of their anthropomorphic attributes. As pointed out by Graf and Zessinger (2022) many respondents in their study felt that a red line would be crossed if the degree of autonomy and the personal emotional level of the device were felt to be too elevated. Among the red lines they identified was autonomous decision-making on the part of the AI device as well as styles of interaction that chaperon the user or pretend to be the user's friend. Recommendations made by the AI-enabled assistant, for example, 'you should go to bed now' may actually cause discomfort because it runs counter to most people's need for control.

But as shown by the tremendous success of ChatGPT, which has attracted millions of users since its launch in November 2022, most people seem quite happy to accept the potential risks in return for the undeniable benefits offered by this new technology. OpenAI, the company that has developed it, is quite clear about the risks. On the entry pages it says:

- While we have safeguards in place, the system may occasionally generate incorrect or misleading information and produce offensive or biased content. It is not intended to give advice.
- Conversations may be reviewed by our AI trainers to improve our systems.
- Please don't share any sensitive information in your conversations.

Furthermore, the company explains that though ChatGPT has been trained to decline inappropriate requests, it '…may occasionally produce harmful instructions or biased content'. And, of course, as many users have found out, it '…may occasionally generate incorrect information.' In its privacy policy (OpenAI, updated version Sept. 19, 2022), it specifies that whilst at OpenAI users' privacy is respected and that they are strongly committed to keeping secure any information obtained from them, personal information is collected, used for specific purposes, shared with third parties and aggregated with information from social media sites.

OpenAI also emphasises that people use the service at their own risk and warns users that e-mail, in particular, may not be secure and that they '…are not responsible for circumvention of any privacy settings or security measures contained on the Service, or third party websites. Besides, by using their service, international users acknowledge that their personal information will be transferred to facilities and servers in the United States. Those who do not want their Non-API Content to be shared are offered the possibility to opt out by filling in a particular form (see OpenAI, Terms of Use, updated March 1, 2023). According to OpenAI this may limit the ability of their services to better address a user's specific use case.



Images taken by participants of Voice assistant locations within their home as part of our In-Home Study.

Privacy cynicism as a response?

To be fair, the company is very open about how it handles personal information. Also, its privacy policy and terms of use are formulated in an easy to understand and clear language and are not so complicated, incomprehensible and time-consuming or cumbersome to read as is the case with many others. Therefore users need not feel tricked by 'dark patterns', which refer to deceptive design patterns or tricks used in many websites and apps that make you do things that you did not mean to, like buying or signing up for something.

Still, it may well be difficult for users to understand or anticipate any of the possible implications of the disclosure of personal data might entail. Although users may be aware of the potential negative consequences of a misuse of their personal data, such as identity theft or fraud, they feel both powerless and helpless. Lutz, Hoffmann and Ranzini (2020) have introduced the term 'privacy cynicism' when talking about these feelings when faced with the challenge involved in protecting one's personal data.

Despite the disclaimers of OpenAI or Google, is it realistic to expect people to fact-check what the AI generates? The current privacy policies or terms of use as discussed above push liability onto users and deny any accountability on the part of the companies. With the rise of ChatGPT, however, the self-management measures such as calibrating one's default settings, only accepting necessary cookies, reading privacy policies more carefully, or deleting one's search history have become more or less irrelevant and pointless. Therefore, privacy-protecting measures will have to be taken at regulatory and political or organisational levels.

How to cope with these challenges?

In this next section we will explore measures at a political and regulatory level, organizational level and finally technological and design level that need to be addressed when considering the implications of the rapidly evolving landscape of talking AI.

Measures at the political and regulatory level

The recent legislative initiatives at EU level such as the Digital Market Act (DMA) and Digital Services Act (DSA) only address some of the concerns related to the advent of large language models including voice-enabled AI systems. Together, the DSA ad DMA are supposed to put 'an end to the Wild West on the internet', in the words of EU Commissioner Margrethe Vestager. The DSA is to ensure that what is illegal offline is also considered and treated as illegal online. Facebook, Twitter, Amazon and other major tech companies are to take more responsibility for the content published on their platforms and eliminate illegal products more decisively.

The new rules, which will come into force in 2024, include:

- Banning advertising aimed at children or based on sensitive data such as religion, gender, race and political opinions.
- Allowing EU governments to request removal of illegal content, including material that promotes terrorism, child sexual abuse, hate speech and commercial scams.
- Forcing social media platforms to allow users to flag illegal content in an "easy and effective way" so that it can be swiftly removed.

The Digital Markets Act (DMA) also has major implications for the global tech market. The act seeks to prevent the biggest of tech firms from dominating digital markets through the threat of fines or even the possibility of a company breakup. They will also face tighter restrictions on using people's data for targeted online ads, a primary source of revenue for companies like Google and Facebook.
The new rules, however, do not cover the threat of misinformation or the dangers associated with the veneer of trustworthiness given to the unreliable information produced by GPT-3. Legislation might eventually catch up with these challenges, but for the time being we will have to find different ways of coping with them. For example, what about privacy laws? What about the "right to be forgotten"? How will Microsoft and Google ensure their bots are not using delisted sources, and how will they remove banned information already incorporated into their large language models?

According to Shah & Bender (2022) when these are used to generate relevant information for a user's needs they pose an even greater threat to transparency, provenance, and user interactions than current search.

Measures at the organisational level

As mentioned in the Introduction, medical editors, in particular, have already issued guidelines and recommendations on ChatGPT and chatbots in relation to scientific publications. The World Association of Medical Editors (WAME) for example, warns that '...the mere fact that AI is capable of helping generate erroneous ideas makes it unscientific and unreliable. Chatbots therefore cannot be authors.

It also stipulates that:

- Authors to be transparent when using chatbots and provide information about how they were used.
- Authors are responsible for the work performed by a chatbot in their paper (including the accuracy of what is presented, and the absence of plagiarism) and for appropriate attribution of all sources)

These measures should help them sort the legitimate from the fabricated, a task which has been made much more difficult with the advent ChatGPT. So far, such tools are not yet available which is why medical editors have to trust the authors to comply with the recommendations. Educators, too, grapple with the role of AI and wonder whether open AI platforms are tools for academic progress or abuse.

Whilst editors and educators have been dealing with issues such as copyright, attribution, plagiarism, misinformation or accountability, the challenge faced by those who use ChatGPT for private purposes in their homes, are of a different kind. They might be more concerned about the harm caused by misleading content and thus the dangers of misinformation. AI researchers such as Gary Marcus (2022) have already raised concerns about the reliability and trustworthiness of AI systems such as ChatGPT. In his contribution to the Scientific American, Marcus warns that because of a lack of mechanisms for checking the truth, such systems can easily be automated to generate misinformation at an unprecedented scale.

Marcus suggests the following paths to deal with this threat:

- Social media companies and search engines should impose a temporary ban on ChatGPT-generated submissions.
- Countries will need to reconsider their policies on regulating misinformation that is distributed widely. If the situation deteriorates, they may have to treat misinformation similar to the way libel is treated.
- User accounts should be more strenuously validated and new systems like Harvard and Mozilla's human-ID.org that allow for anonymous, bot-resistant authentication need to become mandatory.

In their paper, Shah and Bender (2022) set out their vision of an ideal search system which includes the preservation of context to combat the pernicious effects of pattern recognition over datasets expressing harmful social biases. They also remind their readers that information systems have often been run as public goods and that corporate control of dominant information systems should be considered the aberration. Besides, in their opinion the system should provide sufficient transparency about the sources where the information objects are coming from, as well as the process through which they are either ranked or consolidated and presented. Bing's AI interface, for example, already provides footnotes of its sources.

Measures at the technological and design level

The paths suggested by Marcus (2022) also include the development of new tools to fight what has been unleashed. He argues that since large language models have no reasoning capacity, they lack mechanisms for verifying truth. Therefore, he proposes to find new ways to integrate them with the tools of classical AI, such as databases, webs of knowledge and reasoning.

This wish is echoed by the medical editors who stress the need for appropriate tools to help them detect content generated or altered by AI and these tools must be available regardless of their ability to pay. However, in a recent study for which ChatGPT was asked to generate 50 medical-research abstracts, both human reviewers and an AI-output detector failed to identify a third of the bot's papers. At least for the time being, such tools are highly unreliable and cannot be trusted.

There may also be design choices that can mitigate these problems. A recent paper by Rogers et al. (2019) calls for designers to become more involved with the development of trust, privacy and security in the emerging technological landscape of the voice-enabled Internet. They use films to speculate on the true nature of voice assistants questioning the idea that they are our friends rather than devices which ultimately try to sell you something. Seymour and Van Kleek (2021) also voice concerns about the social nature of conversational agents which might unconsciously shape our interactions with them. People develop social relationships with voice assistants that are linked to perceptions of trust in devices, trust in manufacturers, and anthropomorphism of those devices. Their conclusions not only apply to voice assistants but are also true of talking AI devices.

The transition to conversational interfaces that allow for easy and natural interaction with devices represents a profound shift in the nature of the systems we interact with towards the increasingly social. This shift brings with it a variety of new or exacerbated ethical concerns that designers of voice interfaces need to consider when designing future products.

The above-mentioned challenges raise the question if we really want this type of artificial intelligence in our homes. Is our desire for a smart voice assistant so great that we would be happy with one that might also learn about my personal preferences, routines, interests or whereabouts? Whilst we all might welcome a reliable, voice-controlled smart home system that knows what we mean even when we use colloquial language, we might think twice before welcoming an omniscient intelligence into our home. With the advent of ChatGPT, the genie is out of the bottle and can no longer be stuffed back into it, therefore automated misinformation is probably here to stay. The debate about how to deal with this new threat has only just begun. We hope that with this article we can contribute to it.

INQUIRING INTO POLICY MAKING

Inquiring into Policy Making

The objectives of the VA-PEPR project include the generation and provision of policy recommendations for the development and use of future virtual assistants. Given our human-centered approach, we were intent to see what kinds of policy needs might emerge when we draw on actual experiences, practices, and routines of people. We approached this challenge through collaborations, workshops, and public engagements.

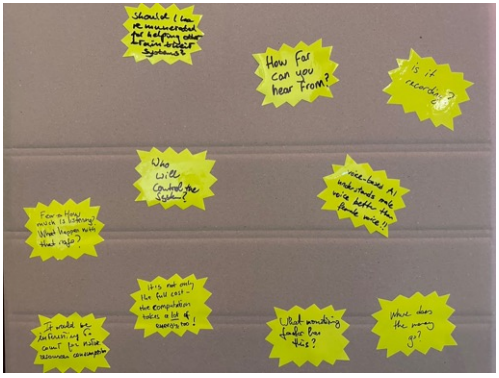
Insights from attendees and visitors of our exhibits and installations

We gleaned insights from conversations and comments left by visitors of our interactive installations at re:publica Berlin (2022 & 2023) and the Digital Design Week London 2023. We explicitly encouraged them to leave us their thoughts and questions to gain insights into their worries, concerns but also hopes and expectations. We engaged attendees of the Mozfest Amsterdam (2023) in policy conversations with four of our provotypes (Microphone, Dial, Confession Booth and WordCloud), asking them what policy needs these triggered for them.

Swiss public servant engaging with Confession Booth during the Annual Strategic Meeting of the Municipality Emmen

Swiss public servants at the Annual Strategic Meeting of the Municipality Emmen - Caroline Röckelein

Questions for attendees of Mozfest Amsterdam



Feedback on Boxes from Re:publica and V&A Voice Matters Exhibit

Learnings from workshops with researchers and professionals

A series of workshops were planned to understand how our work might contribute to different aspects of policy-making and policy implementation.

Designing Policies and Technologies Simultaneously Workshop
This workshop ' took place during the 2023 Computer Human Interaction Conference (CHI) in Hamburg. We co-organized and co-hosted this workshop together with researchers in HCI from Carnegie Mellon, Cornell, Georgia Tech, and Stanford University. This all-day workshop was attended by 40 researchers with backgrounds in HCI, AI, policy, and design.

Key learnings:
The CHI community holds much information that is valuable to policy makers but this information rarely crosses from one community to the other. As HCI experts inquire into, expand and develop new technologies, they often come across policy gaps and new policy needs. Yet, these are not communicated to policy makers. Policy makers on the other hand lack the sophisticated understanding of new technologies and with that rarely are able to grasp their potency and possibilities. They then struggle to provide meaningful policies and guidelines and often rely on vested interests from industry powerhouses.

We captured the results in a major reflection on the role and responsibility of technology experts to engage with and to support policymakers looking out for the interests of individuals, society, and business alike (see CHI paper 2024 in publications section).

Engaging with 250 Swiss Local Public Servants (Policy Implementation)
Before our third policy-focused workshop, we had an unexpected opportunity to engage with 250 Swiss local public servants and local government leadership during their annual strategic meeting in January 2024. This allowed us to see how our work might be useful in the context of public sector innovation on the implementation side. This local public administration is currently working on their own digital transformation. Within the organization, a digital transformation team is preparing for a change in organizational culture to embrace new technologies like voice,

Design, Policy, and AI workshop participants

virtual and LLM driven assistants. Their goal is to understand where and when such technologies create public value and where they do not offer much benefit. In this journey, the team seeks to involve the organization's staff. We provided these public servants with an opportunity to playfully engage with and learn about the technologies' current capabilities. For this engagement, we changed the language of Voice Matters and Überblick to German. For many it was the in the municipality have little contact with technologies like voice assistants or ChatGPT.

Key Insights:

Without implementation, policies are nothing but words on paper. The very people responsible for bringing policies to live, public servants, often get overlooked for their ability to understand the benefits and limitations of new technologies. Yet, they know their communities and their needs and they have an active role in creating new public values while guarding those deemed dear to society. The approach developed by the VA PEPR team to make the invisible aspects and workings of new technologies driving voice assistants and ChatGPT visible and comprehensible not only serves individual users at home but also public organizations and their staff. With that VA PEPR expands its relevance beyond the home and into the public sector.

Design, Policy, and AI Workshop

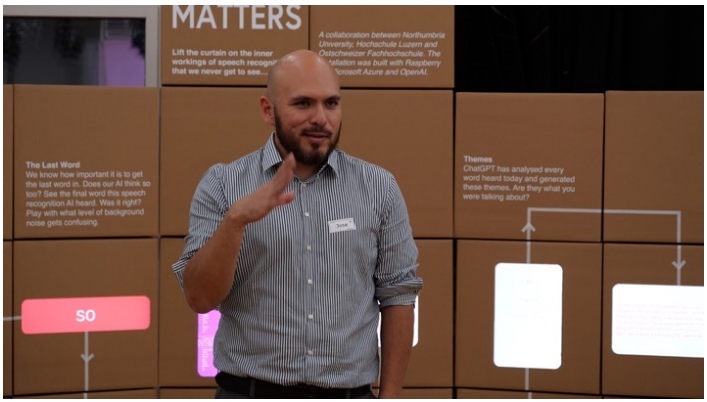
Our third workshop on the topic of Design, Policy, and AI took place in January 2024 at HSLU. Following the 'Designing Policies and Technologies Simultaneously' Workshop, we recognized the need for new connections among these different fields and we wanted to explore how our approach and our prototypes might aid in building these new connections. For this reason, we invited 20 researchers and professionals working on Public Policy, AI/Digitalization and Design. Sameer Pujari and Jose Mendoza Diaz from the World Health Organization (WHO) presented the Global Initiative for AI for Health as a use case challenge for the group. Pierre Dekeuninck and Marco Inchingolo from EU's Joint Research Centre (JRC) provided input on innovation in policy and science. We started with these questions:

- How can we bring in experiences, concerns, desires, and wishes of everyday people into the policymaking process around new technologies and AI?
- How can designers bring in policymakers in co-developing new tools or objects to support policy design-making?
- How can designers better engage with policymakers interested in elevating lived experience narratives surrounding AI and emerging tech?
- How may we build a mutual agenda and develop an outline for future joint projects/publications

Next, participants split in two groups where they first shared their individual answers before they sought to build on each other's perspective. The final task challenged them to build something together that would afford policymakers a new path to engage with new technologies. Building on the Voice Matters and Überblick concepts, an adaptation for policymakers emerged which would make visible processes and decisions during policy-making. Two 'symptom-checker' ideas were then presented using design fiction and Wizard of Oz methods.

Key Insights:

This interdisciplinary and hands-on approach involving experts from the policy domain, design and AI was new for most participants. Though most designers had engaged with policy experts before they had not worked or collaborated with them to inquire into their actual policy problems. This was a surprise because co-design is now rather established. Yet, the workshop



Design, Policy, and AI workshop participants - **How might our research foster innovation in policy-making and policy implementation?**

revealed that this tends to take place in different contexts and for different purposes – rarely to reinvent and innovate the very ways we go about policy-making itself.

The workshop size was very conducive for this kind and level of engagement.

Several participants admitted in their final comments that they 'had no idea why they were invited and what to expect' when they arrived. The WHO now wants to pursue this approach for the Global Initiative for AI for Health and we are looking forward to building on this work.

This goes beyond recommendations of form, where we find it justified to ask if it is OK that we have adapted a design language that enables a device to disappear and become so unnoticeable to users so they mostly forget that whatever they utter is captured and analysed.

CLOSING WORDS

What technological developments do we need to brace for? How can we engage everyday people in this shifting landscape and continue to provide them with orientation? What new questions emerge?

These are the questions we left you with at the end of Workbook 1. Written in May 2022. It seems like a digital lifetime ago now. While we thought that incredibly powerful text-based and voice AI would be a thing of the future. We didn't see it coming as a thing for the immediate future. The arrival of large language models changed all of that. We had to respond. While Voice Assistants (Amazon Alexa, Google Home etc) continued to move from the promise of the mainstream, a new technology was on the block. Technology that could take vast amounts of text and synthesise into user-directed forms. Voice Assistants, because they create vast data bases of speech-to-text are hugely intertwined with these large language models. In short, we could not tease apart ChatGPT (and related tech) from Voice Assistants - which in turn we could not tease apart from our own voices. We were suddenly all living within the machine. To turn into the AI storm clouds that are coming, we decided to start to build instruments that could look deep into these clouds and report back what the AI was hearing and interpreting. WordCloud, Überblick and Voice Matters were designed with this purpose. To build windows into the cloud - to let us peer in and see for ourselves what the machines were doing with our voices. These were used in deployments in peoples' homes, in public exhibitions and policy workshops as devices to help drive new thinking and new debates about the nature of the capabilities, capacities and creativity of cloud based AI that could be utilised anywhere in the world in any internet connected product - as long as it had a microphone (all phones, computers, smart speakers, cars, TVs, light bulbs, radios, thermostats, children's' toys, sex toys and gardening tools all have the capacity to listen) within it.

This workbook presents the raw material for this research during the second half of the project (June 2022 to March 2024). What we have to do next is to synthesise the learnings across the multidisciplinary teams on this project in order to give refined, useful and highly actionable policy recommendations for governments, technology developers, academic researchers and designers. This report will be available from the VA-PEPR website from the summer 2024.

Papers, Workshops, Presentations and Exhibitions

The following project outputs are those that are only relevant to this stage of the work. VA-PEPR Workbook 1 contains more outputs relevant to previous work.

Journal Articles

Maier, E., Doerk, M., Reimer, U. and Baldauf, M. 2023. **Digital Natives Aren't Concerned Much about Privacy, or Are They?** I-Com 0 (0).

Minder, B., Wolf, P., Baldauf, M. and Verma, S. 2023. **Voice Assistants in Private Households: A Conceptual Framework for Future Research in an Interdisciplinary Field.** Humanities and Social Sciences Communications 10 (1): 173.

Renz, A., Neff, T., Baldauf, M. and Maier, E. 2023. **Authentication Methods for Voice Services on Smart Speakers – a Multi-Method Study on Perceived Security and Ease of Use**. I-Com 22 (1): 67–81.

Riss, U. V., Ziegler, M. and Smith, L. J. 2023. **Value Dimensions of Digital Applications and Services: The Example of Voice Assistants.** International Journal of Web Engineering and Technology 18 (4): 319–43.

Conference Papers

Junginger, S. 2023. **Embracing Digital Offboarding as a Design Challenge**. In IASDR 2023: Life-Changing Design. Milan, Italy.

Klotz, U., and Wolf, P. 2023. **Innovating for Sustainable Use: The Different Faces of Trust in Private Household Technologies**. 24th CINet Conference. Linz, Austria.

Maier, E., Doerk, M., Muri, M., Reimer, U. and Riss, U. 2022. **What Does Privacy Mean to Users of Voice Assistants in Their Homes?** In Proceedings of the ETHICOMP 2022. Turku, Finland.

Pierri, P., Aytaç, A., Rogers, J., Shorter, M. and Junginger, S. 2023. **Everyday Data and Everyday Publics.** Implications for Design and for Policy-Making. Designing Platform Technology and Policy Simultaneously: A CHI'23 workshop. Hamburg, Germany.

Yang, Q., Wong, R., Jackson, S., Junginger, S., Hagan, M., Gilbert, T. and Zimmerman, J. 2024. **The Future of HCI Policy Collaboration**. In CHI Conference on Human Factors in Computing Systems. Honolulu, HI, USA.

Riss, U., Maier E., Doerk, M. and Klotz, U. 2023. **User Perceptions and Attitudes in the Data Economy and Their Contradictions.** ACHI 2023, The Sixteenth International Conference on Advances in Computer-Human Interactions.

Riss, U., Tödtli, B. and Wolf, P. 2023. **Privacy Intermediaries: A Business Model Perspective.** 24th CINet Conference.

Presentations

Junginger, S. 2023. **Digital offboarding for the Elderly? Why the time to act is now!** Experiolab Seminar and Experio Rhyzome Research Seminar: January 20th, Online.

Workshops

Gilbert, T. K., Hagan, M., Steven, J., Junginger, S., Wong, R., Yang, Q. and Zimmerman, J. 2023. **Designing Technology and Policy Simultaneously: Towards A Research Agenda and New Practice**. Workshop at the CHI23, Hamburg, Germany, April 23.

Junginger, S. 2023. **How do we live in the omnipresence of voice assistants?** Workshop at the Mozilla Ethical Dilemma Café 2023, Amsterdam, Netherlands, June.

Junginger, S. 2023. **Design in/for Government, for Swiss Designers and Public Servants**. Workshop hosted by the Swiss Design Association at Zurich Design Week, Zurich, September 8.

Shorter, M. and Aytaç, A. 2023. **Speculating Voice Assistant Futures.** Workshop at the IxD23 Conference, Zurich, Switzerland, February 28.

Shorter, M., and Rickenmann, M. 2023. **What Did You Say? Exploring the Potential and Challenges of Voice Interfaces.** Workshop at the Mensch und Computer 2023, Zurich, Switzerland, September 3.

Exhibitions

Shorter, M., Rickenmann, M., Rogers, J., Aytaç, A., Todisco, A. and Junginger, S. 2024. **Voice Matters.** Municipality of Emmen Event, HSLU, Luzern, January 16.

Shorter, M., Rickenmann, M., Rogers, J., Aytaç, A., Todisco, A. and Junginger, S. 2024. **Voice Matters.** Design, Policy & AI Workshop, HSLU, Luzern, January 17.

Shorter, M., Rickenmann, M., Rogers, J., Aytaç, A., Todisco, A. and Junginger, S. 2023. **Voice Matters.** Victoria and Albert Museum, London, September 23.

Shorter, M., Junginger, S., Todisco, A., Aytaç, A. and Pierri, P. 2023. **Money Talks.** re:publica 23, Berlin, Germany, June 5.

Glossary

VA - Voice Assistant.
IOT - Internet of Things.
rcc - Relax, Concentrate and Create.
VA-PEPR - Voice Assistants - People, Experiences, Practices and Routines.
Demonstrators - Research objects designed to collect qualitative data.
Activity Theory - a collective work activity, with the basic purpose shared by others (community), is undertaken by people (subjects) who are motivated by a purpose or towards the solution of a problem (object), which is mediated by tools and/or signs (artifacts or instruments) used in order to achieve the goal (outcome)
Domestication Theory - an approach in Science and Technology Studies and media studies that describe the processes by which technology is 'tamed' or appropriated by its users.
Indeemo - A diary phone app used for remote ethnography.
Speculative Design - Speculative design is an approach to design that focuses on imagining future scenarios and possibilities. It can be used to explore social, political, technological, and ethical issues, and to generate new ideas and solutions.
Speculative design is not about predicting the future, but about exploring different ways that the future could be. By doing so, it can help us to better understand the present and make better decisions about the future..
ISP - Internet Service Provider.
Data Packet - A data packet is a unit of data made into a single package that travels along a given network path. Data packets are used in Internet Protocol (IP) transmissions for data that navigates the Web, and in other kinds of networks.
IP Address - An IP address is a unique address that identifies a device on the Internet or a local network. IP stands for "Internet Protocol," which is the set of rules governing the format of data sent via the Internet or local network.
Provotype - A provocative prototype used as a prop for conversations. The object is designed to ask questions rather than solve problems.

VA PEPR

How do we live in the omnipresence of voice assistants?

VA-PEPR stands for Voice Assistants Ð People, Experiences, Practices and Routines. We conduct research into how people experience voice assistants in their homes and private lives and how they develop new practices and routines around their use of VAs. By focusing on the home environment, user experience and ethical issues, the project aims to contribute to a deeper understanding of this new technology.

This interdisciplinary research project is conducted by the Lucerne University of Applied Sciences and Arts (HSLU), Ostschweizer Fachhochschule (OST) and Northumbria University under the lead of the HSLU's Lucerne School of Design, Film and Art. It is funded by the Swiss National Science Foundation.

HSLU Hochschule Luzern | Northumbr University NEWCASTLE | OST Eastern Switzerland University of Applied Sciences | Swiss National Science Foundation | moz://a