



Human-centric algebraic machine learning

Interaction Paradigm Methodology - Revised

Deliverable D4.5

WP4 - Human AML Interaction



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952091.

Consortium



ALGEBRAIC AI



Chantalimaud
Foundation



German
Research Center
for Artificial
Intelligence



TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN



Project

Title: ALMA: Human Centric Algebraic Machine Learning
 Acronym: ALMA
 Coordinator: Proyectos y Sistemas de Mantenimiento SL (eProsima)
 Grant agreement: 952091
 Call: H2020-EIC-FETPROACT-2019
 Program: Horizon 2020
 Start: 01 September 2020
 Duration: 48 months
 Website: <https://alma-ai.eu>
 E-mail: alma@eprosima.com
 Consortium: **Proyectos y Sistemas de Mantenimiento SL (eProsima) (EPROS)**, Spain
Champalimaud Foundation (CF), Portugal
ALGEBRAIC AI (ALG), Spain
German Research Center for Artificial Intelligence (DFKI), Germany
Inria (Inria), France
TU Kaiserslautern (TUK), Germany
FIWARE Foundation e.V. (FF), Germany
Technical Research Centre of Finland (VTT), Finland
Carlos III University (UC3M), Spain

Deliverable

Number: **D4.5**
 Title: **Interaction Paradigm Methodology - Revised**
 Month: 18
 Work Package: WP4 - Human AML Interaction
 Work Package leader: Inria
 Deliverable leader: National Institute for Research in Digital Science and Technology (Inria)
 Deliverable type: R: Document, report
 Dissemination level: Public (PU)
 Date of submission: 2023-11-21
 Version: v3.0
 Status: Finished

Version history

Version	Date	Responsible	Author/Reviewer	Comments
v1.0	2023-08-15	Inria	Inria	part
v2.0	2023-09-28	Inria	Inria	final
v3.0	2023-11-21	Inria	Inria	revised

Executive summary

Algebraic Machine Learning (AML) is a novel machine learning technique based on algebraic representations of data. Unlike statistical learning, AML algorithms are robust regarding the statistical properties of the data and are parameter-free. AML systems can retain what they learn, including sociocultural ethical practices for human–artificial intelligence systems. The aim of the EU-funded ALMA project is to study the mathematical properties of AML and to leverage them to develop a new generation of interactive human-centric machine learning systems. These systems are intended to minimize bias in order to facilitate trust while better enabling distributed collaborative learning.

This report corresponds to *Deliverable D4.5 - Interaction Paradigm Methodology - Revised* of the ALMA project. This deliverable will extend current participatory design and generative design methodologies to incorporate the role of an intelligent system in the interaction loop, where agency is revealed and shared by the human user and the AML system. We will also provide new visualization methods that help novice users understand the state of the algorithm in terms they can understand.

Post-review comments

It is important to highlight the fact that we *never* signed up to “explain” the state of the algebraic machine learning algorithm to users, since we believe that users, especially novices, have zero interest in the state of the algorithm. That is the domain of AI researchers interested in “explainable AI” and runs counter to our Human-Computer Interaction approach. Note too, that, even if our goal was to make the inner workings of AML more explainable, we have yet to receive the *AML* algorithm itself. We have only received the AML model as a binary, which has significantly limited our ability to explore and adapt the algorithm to our needs.

We argue, instead, that novice users are interested in obtaining *feedback* about how the system has interpreted their behavior and, ideally, *feedforward* about what options remain available to them. We explained this approach in our paper on the *Octopocus* dynamic guide (Bau & Mackay, 2008), described in section 4.2.

We demonstrated this approach by creating an interactive application using the ALMA algorithm where a novice user can train a simple classifier, as demonstrated in Deliverable 4.8. This deliverable describes our overall approach to human-centered AI and our goal of influencing the development of the ALMA algorithm to incorporate a more human-centered design approach.

Comments related to AML are highlighted in red through the rest of the document. Note that this deliverable is designed to provide a general overview to our approach to human-centered AI, or what we call “human-computer partnerships”, which provides a variety of possible directions for testing new machine learning approaches such as AML.

Contents

Executive Summary	3
Contents	4
Acronyms	5
1 Introduction	6
2 Background	6
3 Methodology	7
3.1 Generative Theories of Interaction	9
3.2 The cost of system-induced errors	10
3.3 Shaping the user's behavior	11
3.4 Presentation details matter	12
3.5 Interacting with AI	12
4 Creating Effective Human-Computer Partnerships	13
4.1 Discoverability	14
4.2 Expressivity	14
4.3 Appropriability	15
4.4 Sharing Agency	16
4.5 Shaping human and agent behavior	17
5 Conclusion	18
References	19

Acronyms

AML	Algebraic Machine Learning.
HCI	Human Computer Interaction.

1 Introduction

Our research in the field of Human Computer Interaction (HCI) focuses on *human-centered AI*, i.e. how to create intelligent interactive systems with a measurably positive impact on human users. We seek to design “*human-computer partnerships*” that combine human skills and system capabilities so that humans and intelligent agents perform better together than either individually. This requires not only taking advantage of the best characteristics of each, but also considering how to improve human capabilities over time, thus “upskilling” rather than deskilling or replacing the users of these systems [1].

This deliverable first provides background about our research approach and then describes our analysis of the key challenges to creating human-computer partnerships. We then describe our participatory design and generative design methodologies and describe how they could be applied to designing different kinds of interaction loops where where agency is revealed and shared by the human user and the AML system. We also illustrate how different visualization methods can help novice users understand the state of the algorithm in terms they can understand.

Areas in which such interaction loops would benefit from the potential of the AML system are highlighted in red.

2 Background

Inria’s *ExSitu* research team has been exploring the limits of interaction with an emphasis on how extreme users interact with technology in extreme situations. Rather than beginning with novice users and adding complexity, we began with expert users who already face extreme interaction requirements. Studying these advanced users not only helps us anticipate the routine tasks of tomorrow, but advances our understanding of interaction itself. Our continuing goal is to advance our understanding of interaction as a phenomenon, with a corresponding paradigm shift in how we design, implement and use interactive systems.

Extreme Situated Interaction emphasizes working with *extreme users* in *extreme environments* to facilitate *extreme design*. Our goal is to shift the paradigm of interactive software and to establish laws of digital interaction that significantly empower users and allow them to control their digital environments. We characterize *extreme users* as those who make extreme demands on current technology. We focused on creative professionals who rewrite the rules as they create new works and scientists who seek to understand complex phenomena through creative exploration of large quantities of data. We argued that developing novel interactive technology with and for users who adapt interactive technology in novel and unexpected ways helps us reinvent existing tools and explore opportunities for creating new forms of interaction.

Extreme environments push the limits of today’s technology. We relied on advances in hardware to solve problems such as cost, reliability and performance and focused on the details of interaction within such environments. We noted that previous leaps in technology adoption followed innovations in interaction design: personal computers became ubiquitous after the invention of the desktop graphical user interface and smartphones and tablets became commercially successful after Apple cracked the problem of a usable touch-based interface. As we predicted, smart watches took off only when small-scale interaction made sense to users, but smart glasses failed, at least in part, because they did not. We thus explored novel forms of large-scale interaction and collaboration using the Digiscope WILD and WILDER platforms, as well as dance and music performance venues.

Extreme design focuses on developing novel methods and theories for extreme users in extreme environments. We pushed the limits of participatory design by encouraging creative professionals to act as both designers and developers throughout the design process. Extreme design affected every stage, from requirements definition, to early prototyping and design exploration, to implementation, to adaptation

and appropriation by end users. We also explored how integrating the design of physical objects into the software design process raises new challenges as well as requiring new methods and skills.

We are exploring how to develop generative theories of interaction [2] grounded in how people interact with the world. Our concepts of instrumental interaction [3, 4], information substrates [5] and human-computer partnerships [6, 7] provide a generative approach for designing interactive systems, from early exploration to final implementation. Our research is organized along four axes: *Fundamentals of Situated Interaction*, *Human-Computer Partnerships*, *Creativity* and *Collaboration*. Individual projects typically addressed multiple research axes and group members often worked together on different aspects of themes within and across research axes. Our work with respect to the ALMA project touches all four axes, with the emphasis on *Human-Computer Partnerships*.

3 Methodology

Our approach to designing human-centered AI systems involves a participatory design approach [8, 9] where we work closely with users throughout the design process. We view human users, not as a source of errors or cogs in the machine, but rather as the system's *raison d'être*.

We begin by closely observing their real-world activities in real or realistic contexts, and ask them to provide us with detailed stories of their interactions with the system, not just what goes right, but more importantly what goes wrong. We treat these “breakdowns” to better understand the use of the existing system and as a source of inspiration for what to improve. To the best of our ability, we avoid designing for stereotypical use and instead try to understand how the user's interaction with the system will play out over time. We do not ask users if they simply “liked” our systems, instead, we ask them to compare alternatives and reflect on the strengths and weaknesses of each.

Current machine learning systems detect patterns in the input data, but not in terms that are understandable by human users. One potential benefit of AML is the ability to collect data and concepts related to human behavior and share it, without revealing any identifiable aspect of the person that the data came from. The model represents the individual users but the atoms generated do not have a one-to-one correspondence with the user's behavior. Instead, it creates mathematical ‘concepts’ that represent relevant user behavior while still preserving the user's privacy.

We assess the risks of negative impact with respect to the potential size of the harm and the likelihood of its occurring, in order to provide stakeholders with appropriate measures to inform decision makers. To accomplish this we need to develop specific observation and experimental strategies for collecting and analysing both quantitative and qualitative data collected in the context of actual use. The issues presented here apply to both standard machine learning and Algebraic machine learning, in that we need to identify risks and benefits, as well as how to measure them, in human terms.

Risk/Benefit analysis: We need to balance the risks with the likelihood of a negative impact: clearly high risk and high likelihood of a negative impact should be avoided, but in some circumstances we may be willing to accept a high benefit in exchange for a very low risk of a negative impact, as we do with nuclear power. We might also reject a low negative impact, such as a slightly annoying interface/interaction if it is likely to occur on a massive scale, as is often the case in current software.

Measurement: We distinguish between a “task analysis” approach that simply identifies whether or not the system has sufficient functionality to perform the user's tasks and a more interaction-oriented perspective that considers which measures will reveal the efficacy of the interaction. For example, we may want to compare the number of steps or the elapsed time necessary to perform a particular task, from a given starting point, in a given context. We can measure discoverability: at a given point in an interaction with the software, how easy is it for the user to figure out what operations/tasks are possible and how to perform a desired operation in that context. We can measure ‘explainability’ in terms of how

well does the user understand what the system just did, and what the impact of that action is on their ability to perform future actions.

Finally, we need to consider both the short- and long-term effects of the interaction. For example, the introduction of the “like” button has a short-term positive impact on human users, who enjoyed both responding to others’ Facebook posts and receiving positive feedback. However, this led many users into a more negative, long-term addictive reaction, where human users end up being drawn into constantly checking their number of likes. Here, a small immediate benefit may lead to a long term harm.

We draw from evolutionary biology’s framework for assessing positive, neutral or negative relationships among different species over very long periods of time. For example, a symbiotic relationship offers positive benefits to both parties whereas a parasitic relationship is positive for one party, but neutral or negative for the other. If we treat intelligent agents as if they were other species that interact with human beings, we can develop methods for measuring the technology’s impact on both individual users and larger groups. For example, members of a social media group might enjoy reading scurrilous accounts about an individual, but the impact on that individual can be devastating.

Key dimensions for measuring impact include determining positive, neutral or negative impact with respect to:

1. *time scale*: immediate, intermediate and long term,
2. *population*: individual, group or corporate
3. *risk*: balance of likelihood of harm and size of negative impacts
4. *measures*: qualitative or quantitative, in context or in general

Design considerations

We need to carefully design the way information is presented and how users control their interaction with these systems. We must recognize that the same algorithm embedded in a different interface may affect users in extremely different ways. At the same time, we can choose among different algorithms to produce a particular interaction style: we can explore strategies for developing “sustainable” AI to achieve the same effect with less data or computing power. Focusing on developing effective human-computer partnerships instead of more powerful algorithms will change not only how we approach AI research but also AI policy, ethics and law. **We have created an interactive pipeline that allows us to explicitly compare the performance of the AML algorithm to other algorithms, not simply in terms of the algorithms’ performance, but, critically, in how the user responds to and understands how the system works.**

As intelligent system designers, we need to balance the trade-offs between offering users power of expression and simplicity of execution. We need to consider whether or not users are forced to:

- monitor the system over long periods of time?
- remove interesting tasks and replace them with boring, repetitive ones?
- jump into action to solve a crisis that the system cannot solve, after passively letting the system do everything else?
- feed data into the system in the system’s rather than the user’s terms?
- generate yes/no answers to complex questions?
- behave in prescribed ways to achieve a particular goal?

This means that we need to reconsider how we conduct AI research and build intelligent systems. We need to shift from solely measuring the efficiency of the *algorithm* to measuring the effectiveness of the

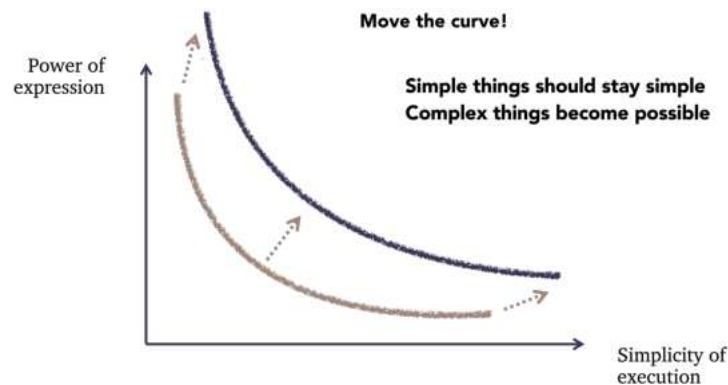


Figure 1: Interactive systems always face a trade-off between power and simplicity. The challenge for HCI researchers is to move the curve.

interaction over time. We need to create successful “human-computer-partnerships” that augment rather than deskill or replace human users such that the combination of human and AI capabilities together are better than either alone.

3.1 Generative Theories of Interaction

We define *human-computer partnerships* as systems that share agency between humans and interactive technology, but explicitly retain control by the user. Rather than treating human users as the “input” to a computer algorithm and measuring success in terms of the algorithm’s performance, we design and evaluate intelligent systems that increase human capabilities. Whereas much human-computer interaction research focuses on measuring and improving productivity, we seek to create what we call *co-adaptive systems* that are *discoverable*, *expressive* and *appropriable* for the user.

This research is deeply influenced by what we call *Generative Theories of Interaction* [2]. The key insight is that interaction can be studied as a phenomenon in its own right. When we apply this approach to the design of intelligent systems, we see that an excellent algorithm can perform poorly if embedded in an inappropriate interface while a simpler, more sustainable algorithm can perform better if embedded in an effective interface.

Generative Theories of interaction (see Fig. 2) create concepts drawn from established theories from the natural and social sciences. We then decompose each concept into one or more actionable principles that can be applied to the study and design of interactive systems. We then examine existing interactive systems with an analytical lens, to see if the principles apply, and with a critical lens, to see if the system can be improved with the application of those principles. Finally, we can explore how to improve existing systems or create new ones by applying a constructive lens that applies the principles in innovative ways.

Much of our work builds on the concept of *co-adaptation* [7] inspired by evolutionary biology, where humans both *adapt to* or learn how to interact with the technology as well as *adapt* or appropriate the system for their own purposes. Darwin’s classic text *The Origin of Species* [10] describes both *co-evolution*, where species are both affected by and affect the environments in which they live as well as *co-adaptation*, which describes the interactions among individual organisms affect each other: “*Co-adaptation* emphasizes this on-going, potentially asymmetrical process of mutual influence between organisms and the environment, where *survival of the fittest* is not simply a matter of an organism adapting to a changing environment, but also of it physically changing that environment to ensure its survival.” We can examine the phenomenon

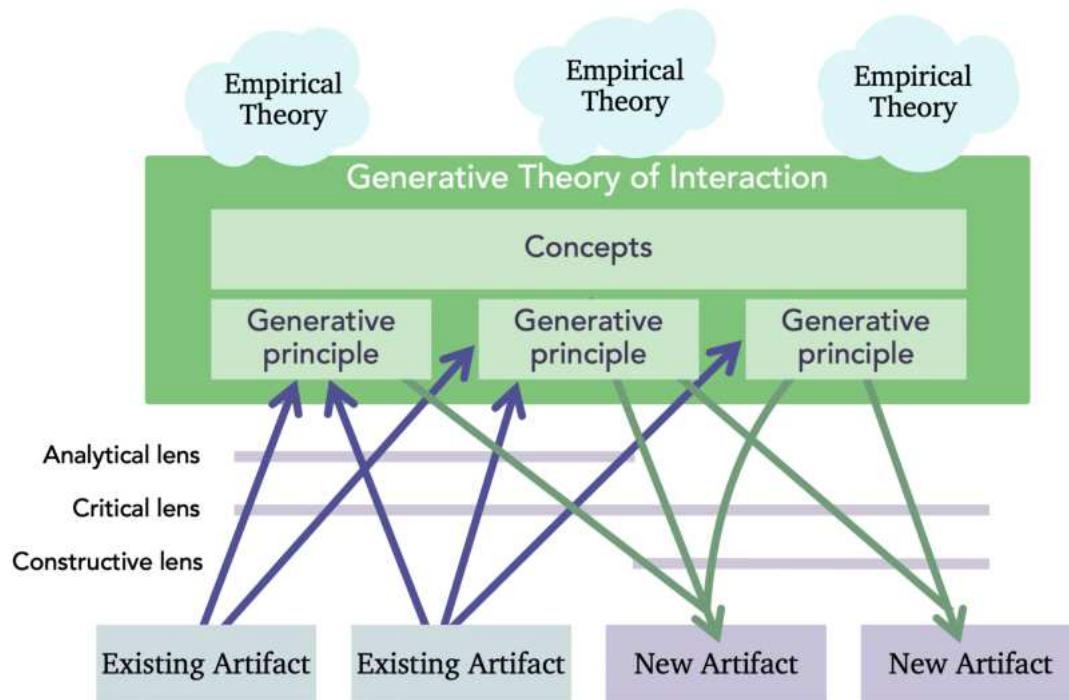


Figure 2: Generative theories of interaction generate actionable design principles drawn from the natural and social sciences.

of co-adaptation with respect to human users and intelligent agents. If we take the user’s perspective, we see that users establish an ongoing dual relationship with the technology, where they discover how it works but also modify it for new purposes. If we switch to an intelligent system’s perspective, the dual relationship is similar, although not identical, since some systems adapt to the user while others seek to change the user’s behavior. We call this set of inter-related interactions *reciprocal co-adaptation*, where both users and intelligent systems *adapt to* or learn from each other, as well as *adapt* or shape each other’s behavior.

Of course, human users and intelligent agents access, interpret and present information differently. Intelligent systems, especially those based on statistical models such as machine learning and neural networks, are essentially impossible for human users to understand. This section analyzes a range of issues that arise when creating an interactive intelligent system, including the cost of system-induced errors, the shaping of user behavior, the importance of how information is presented and general issues about interacting with AI. Much of this work will appear in Mackay, W.E. (2023) *Creating Human-Computer Partnerships*, CHIRA’23, Rome, Italy.

3.2 The cost of system-induced errors

Consider the recent fiasco surrounding Meta’s GALACTICA whose language-based AI model stores, combines and reasons about scientific knowledge to support writing scientific papers. Unfortunately, GALACTICA was prone to generating dangerous pseudo-science. The journalist Tristan Green got it to generate “*well-written research papers on the benefits of committing suicide, practicing antisemitism, and eating crushed glass.*” [11]. GALACTICA was shut down after only three days, not because of the algorithm per se, but rather because the designers had not sufficiently considered how humans might interact with it.

Legitimate science takes expertise, discipline and time, but GALACTICA allowed non-scientists to produce fake science with a highly authoritative voice. Although designed to save time, it was equally ripe for misuse.

Of course, not all intelligent systems exhibit such dramatic failures. We are all confronted daily with more a more innocuous version of this problem: the smartphone’s “auto-correct” feature. Here, the designer’s assumption is that the user made an error — which may or may not be true — and that the correction is actually correct, which also may or may not be true. Another hidden assumption is that any errors that are produced by the system are innocuous and that the benefits of using the system outweigh the cost of introducing new errors.

Another more subtle effect is the transformation of the user’s role from “author” to “error corrector”. Although a touch typist can simultaneously type and read text on a laptop screen, phone typing requires continually shifting attention between the soft keyboard, the suggested words and the text output window. Producing error-free text requires constant monitoring to ensure that the system has not introduced errors, a notoriously poor use of human skills [12].

Worse, incorrect words require immediate action from the user. For example, it is common to only catch errors only after pressing “send”. The user must then decide whether or not to shift the topic of conversation to a discussion of the perils of auto-correct. This may be amusing when chatting with friends and family, but is less acceptable for work-related texts.

Finally, people and intelligent agents produce very different kinds of mistakes. If a person misspells a word, other people are very good at inferring the correct one. By contrast, some auto-corrected errors are impossible to guess. A good human speller may question the advantage of using “auto-correct” if the act of correcting the corrections wastes more time than simply typing in the first place.

3.3 Shaping the user’s behavior

As users, we *adapt* our behavior to the system over time. We can choose to accept a proposed word even if it was not quite what we wanted to say. For example, if after typing an informal “congrats”, the system might replace it with the more formal “congratulations” or a highly informal champagne emoji. Now the user must not only decide the “right” level of formality but also whether or not to take time correcting it.

Auto-correct can also “shape” [13] how we type over time. For example, instead of fully typing each word, we may just type the beginning of a long word and wait for the system to propose the complete word. Some people take this even further, typing the fewest possible letters and waiting for the system to produce a good-enough suggestion. This can be efficient and save keystrokes but risks producing only pat, stereotypical phrases.

Human conversations operate differently from chats with intelligent agents. Conversation partners establish what Clark [14] calls “common ground” and rely on mutual knowledge, beliefs and assumptions to communicate efficiently. Intelligent agents and people have access to different, overlapping sets of data but only humans can share beliefs and assumptions. Humans react to each others’ conversation styles and shape each other’s behavior during the conversation. However, humans also tend to treat intelligent agents as if they were human.

An early example of this occurred with Digital Equipment Corporation’s IVIS, the first commercial interactive video system. The Educational Services research group developed an intelligent tutor, “Dwayne”, an animated claymation character who spoke with DECTALK as he explained how to fix a printer. Field service technicians typed much simpler, easier-to-interpret requests when Dwayne spoke with a child’s rather than an adult’s voice. The presentation and interaction mode shaped users’ behavior, encouraging

them to limit their word choices as if they were talking to a child, which made it much easier for us to interpret them.

3.4 Presentation details matter

Sometimes, showing additional information reduces human skills. For example, studies of air traffic controllers [7] led to a system called CAMÉLÉON [15] that augmented controller's interactions with paper flight strips. They greatly appreciated being able to tap on a flight strip to visually locate that aircraft on the RADAR screen. Yet they objected to another feature that calculated and displayed the aircraft's expected trajectory, since it would interfere with developing and maintaining their mental model of the air traffic.

A related example involves an intelligent system designed to help ship captains plot an acceptable course when entering a harbor¹. The first version presented the AI's suggestion *before* the captain plotted their own course. The captains rejected this version because they feared it would reduce their expertise while still holding them responsible if it made an error. However, they accepted the same system with a reversed presentation order: the captain first plotted a course and only then saw suggestions that they could choose to adopt. Here, the same algorithm with a different interface produced very different results: the first risked deskilling users whereas the second helped improve their skills.

Users often have difficulty determining the effectiveness of their interactions with intelligent agents. A recent study of 54 cybersecurity experts at Stanford University found that programmers who used an intelligent code assistant (GITHUB's CO-PILOT) wrote significantly less secure code but believed they wrote better code [16]. Another study of 49 students in a Behavioral Economics class at the HEC Business school showed that students who wrote their own analytical essays performed 28% better than those corrected an essay written by ChatGPT [17].

Humans and intelligent systems fail in different ways, each with associated risks. For example, under normal conditions, the French national energy grid's automated systems balance loads somewhat better than human experts. However, when the COVID-19 pandemic caused radical shifts in energy requirements, the automated systems made strange, costly errors and the human experts performed significantly better.² The assumption that intelligent systems will provide users with more interesting cognitive tasks is just that: an assumption. It simply will not happen unless intelligent system designers adopt the goal of enhancing the user's experience, which requires focusing on the interaction.

We also need to consider interaction at multiple levels of scale. Designers of intelligent systems must consider the back-and-forth interaction between the user and the system, not only when everything works well but also when the system makes errors. Does the system limit the user's power of expression to make their inputs easier to interpret? Does the system transform the user from an author into a system monitor who detects and corrects errors? What are the longer term effects on the user's behavior? Do users learn and enhance their skills when using the system? Or does removing "easy" tasks and suggesting answers reduce their ability to develop and maintain expertise, effectively deskilling them?

3.5 Interacting with AI

Traditional characterizations of intelligent systems treat users as just another component of the system, either as a useful source of data or an unwanted source of error. Rafner et al. [1] distinguish among "*Human-In-The-Loop*" (HITL) systems that plan, execute or evaluate data acquisition tasks [18];

¹Personal communication.

²Personal communication from staff.

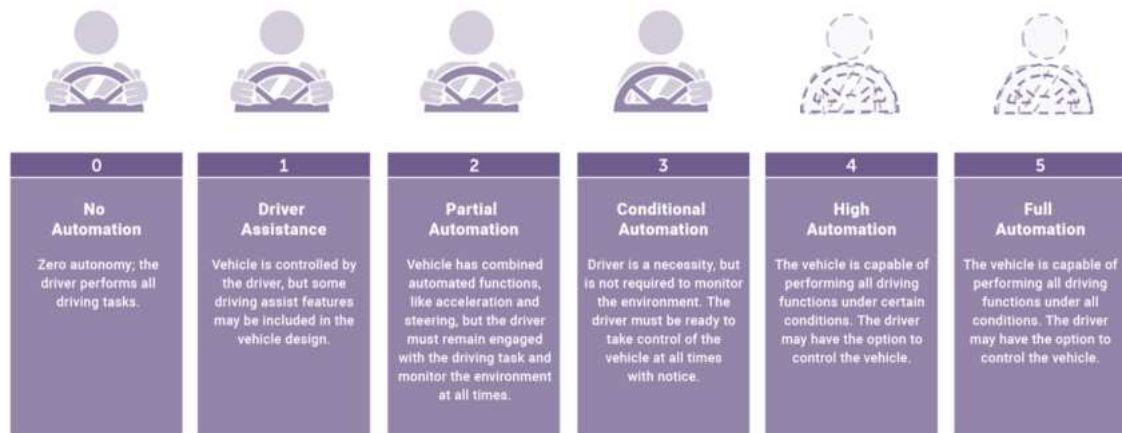


Figure 3: Proposed levels of automation for vehicles.

“*Human-On-The-Loop*” (HOTL) systems that require a human to check the results [19]; and “*Human-Out-Of-The-Loop*” (HOOTL) systems that do not include human users at all [20].

Note that each frames the user’s role in terms of the system’s requirements, not the user’s needs. These characterizations also raise important questions: For example, if humans are “out of the loop” does that save them time and free them for more interesting cognitive tasks or will it replace them altogether? If the system takes over the repetitive and “boring” elements of a task, will this enhance the user’s job satisfaction or reduce their ability to learn and enhance their expertise over time? A concert pianist does not become a virtuoso by performing only the most interesting elements of a piece — practicing scales is also necessary for honing motor skills and developing one’s ear. Consider too what will happen with almost-autonomous vehicles (see Fig. 3 where drivers need not pay attention until the car suddenly faces a crisis that the system cannot handle. Such systems will both deskill drivers and blame them when they lack the expertise necessary to immediately avert disaster.

We argue we need to re-frame how we think about artificial intelligence research. Creating truly effective intelligent systems requires a shift from measuring the efficacy of the algorithm to evaluating the impact on users when those algorithms are embedding into interactive systems. We should design the user’s *interaction* with intelligent systems so that *humans* develop their intuitions and maintain their skills. This requires treating *interaction* as a phenomenon in its own right [21], capable of being characterized, measured and evaluated.

This paper describes our approach to creating human-computer partnerships. We begin with a brief overview of our approach to “generative theory”, which offers a principled approach for designing effective interaction. The examples in the next section illustrate how we have employed generative theories of interaction to address the challenge of building effective interactive intelligent systems by shifting the focus from the *algorithm* to the *interaction*.

4 Creating Effective Human-Computer Partnerships

How can we design effective “human-computer partnerships” that make the most of human skills and system capabilities? The first examples explore how to increase a smartphone user’s power of expression while maintaining simplicity of expression [22]. This involves creating intelligent co-adaptive systems that are discoverable — easy to learn and adapt to — as well as expressive and appropriable — easy

to personalize and adapt. The next examples illustrate how human users can successfully share agency with intelligent agents. The final examples explore how we can exploit principles from behavioral psychology [13] where humans and agents shape each others' behavior.

4.1 Discoverability

Even though today's smartphones are extremely powerful, the interaction is extremely limited and takes little advantage of the expressive power of our fingers. Users press buttons, select from menus, fill out forms and enter text with tapping, swiping or pinching gestures. The advantage of this approach is the very low barrier to entry — almost anyone can use the system with little training. The disadvantage is that users have very little power of expression. They can only react to the buttons and controls displayed on the screen, rather than generating commands directly.

What if we could express commands by drawing gestures, for example, drawing a “pigtail” gesture to delete a phrase? The phone is certainly capable of identifying such gestures, so why is this not an option? This is a problem of discovery: users must be aware that such gesture-commands exist as well as recall the mappings between gestures and commands. The design challenge is how to provide in-context help for novice users when they need it, without slowing down experts.

We introduced the concept of “dynamic guides” [23] which combines on-screen feedforward and feedback to help users learn, execute and remember gesture sets. We take advantage of the fact that users who are unsure tend to pause, whereas experts simply draw the gesture. Fig. 4 shows what happens when the user starts to draw but hesitates: a visual guide appears around the cursor or their finger, with different paths for each gesture-command combination. The classification model updates dynamically as the user follows the desired template, providing feedback as to what the user has already drawn and feedforward as other alternatives diminish or disappear. Novices can thus learn these gesture commands over time and smoothly transition into experts. If they forget a gesture, they can always pause to display the guide.

4.2 Expressivity

A fully co-adaptive system does not simply help users adapt to the existing system, but also lets them adapt it to meet their individual needs. For example, human-produced handwriting is highly personal. We can recognize who wrote what and make educated guesses about the conditions under which it was written, such as in a rush or slowly and carefully. By contrast, text messages all look the same.

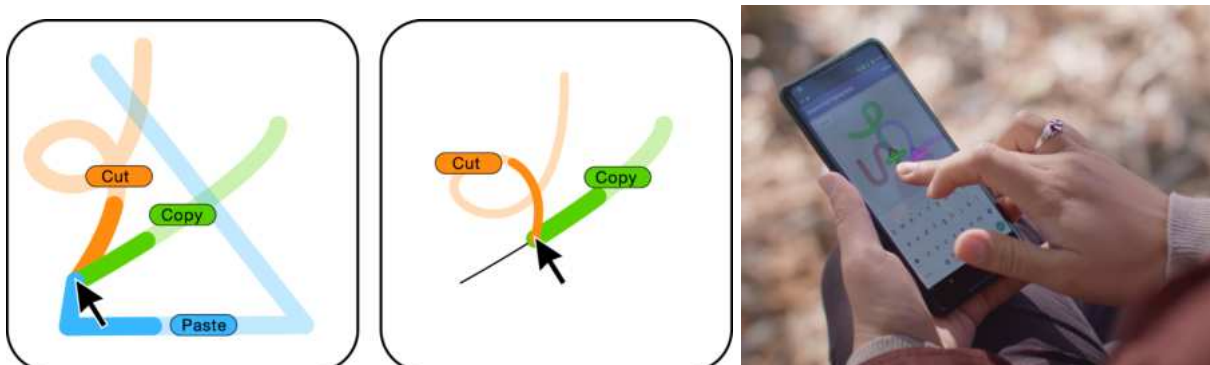


Figure 4: *Octopocus* displays a dynamic, in-context guide that shows how to draw different command gestures.

How can we offer users the expressive power of handwriting, but on a phone? Zhai and Kristensson [24] introduced “gesture typing”, where users glide their fingers between letters, lifting only between words, and showed that it is 40% faster than tapping on a soft keyboard. Today’s smartphones all offer this feature, with extremely proficient algorithms that identify the correct word and discard all extraneous human variation as unwanted “noise”.

By contrast, we were interested in the details of this human variation, and how to transform it into expressive output, under the user’s control. Fig. 5 shows the *Expressive Keyboard* [25], which maps the variation in how users glide between letters into different colors or fonts. For example, we can create the full RGB color space by mapping “curviness” to red, size to “green” and speed variation to “blue”. Users can then consciously control both the content and the color of gesture-typed words. We thus “recycle” otherwise ignored gesture variation to create rich output under the users’ control, without sacrificing accuracy. We also created the *CommandBoard* [26], which offers a full range of custom-designed fonts and the *MojiBoard* [27], which transforms gesture-typed words into emojis with a single stroke.

4.3 Appropriability

In addition to creating expressive output, users also want to personalize their input by choosing their own gestures and mappings to their preferred commands. Here, the challenge is to balance the needs of the human user, who wants to create gestures that are easy to recall, and the system, which needs to reliably recognize and distinguish each gesture in the set of gesture-commands.

We developed the *Fieldward* [28] dynamic guide, an interactive colored heat map that visualizes the “negative space” of unused gestures and verifies whether or not the current gesture is unique (see Fig. 6). The user simply enters command and begins to draw the desired gesture.

As the user draws, the display changes color: Ending the gesture in a red indicates that that gesture already exists. If it ends in a purple zone, the gesture is ambiguous and it is safer to either continue or try another option. If the gesture ends in a blue zone, the current gesture can be recognized reliably.

Users enjoyed exploring different gesture possibilities, such as drawing a heart for their boyfriend. One user not only designed individual gesture-commands but also created his own personal gesture syntax. He associated certain gestures with commands, e.g., “call” or “text” and others with people, e.g., “Mom” and “Bob”. He could then create fluid gesture combinations to create compound commands, such as “Call Mom”.



Figure 5: Standard gesture typing discards user variation as “noise”. The *Expressive Keyboard* transforms individual variation into rich output, under the user’s control.



Figure 6: *Fieldward* displays a dynamic heat map that shows if the user's preferred gesture will be recognized (blue) or not (red).

4.4 Sharing Agency

One of the major challenges in creating human-computer partnerships is how to share agency between human users and intelligent agents, broadly defined as exerting control over the outcome. We worked closely with designers to create *Semantic Collage*, a digital mood board creation tool that attaches semantic labels to images [29]. Designers propose images that *Semantic Collage* translates into semantic labels, which designers can then adjust to improve the search. This back-and-forth approach allows designers to share agency with the algorithm, where each proposes possibilities and provides specific feedback to the other.

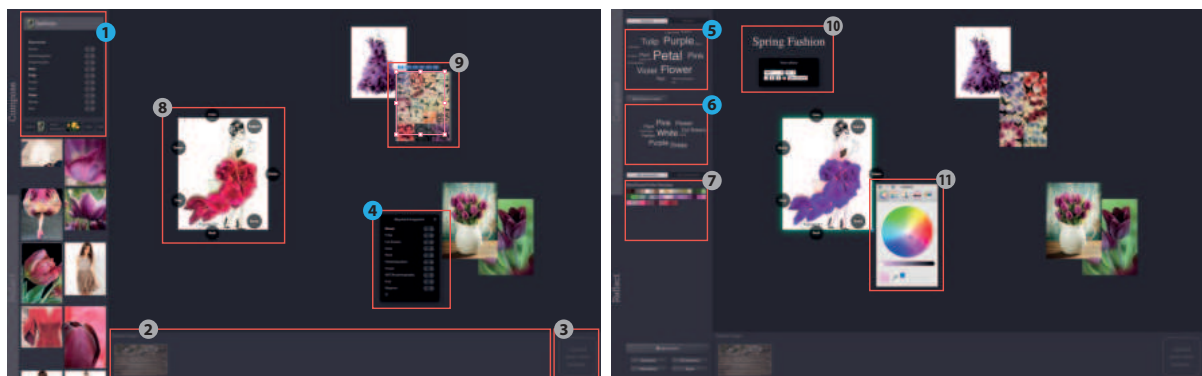


Figure 7: *ImageSense* lets designers vary their level of agency with an intelligent mood board. *Image Cascade* (d) supports full user control; *Semantic Collage* (c) supports shared user-agent control; and *Select-o-matic* (b,g) offers primarily agent control.

A subsequent project, *ImageSense* [30], combined individual and shared work spaces as well as collaboration with multiple forms of intelligent agents. *ImageSense* allows users to transition smoothly between different levels of agency, from user-directed selection of curated images (*ImageCascade*), to alternating between user-initiated and system-initiated search (*Semantic Collage*) to responding to system-driven recommendations (*Select-o-matic*). Designers can tweak the agent's behavior, e.g., by specifying “fewer images like this” and modify the types of generated semantic labels, color palettes or tag clouds and adjust their level of importance.

Intelligent systems such as those described above are intended to facilitate the interaction. But what if we instead ask an intelligent agent to challenge the user in order to gain efficiency?

We created *BIGNav* [31], a multi-scale navigation technique based on Bayesian Experimental Design. The

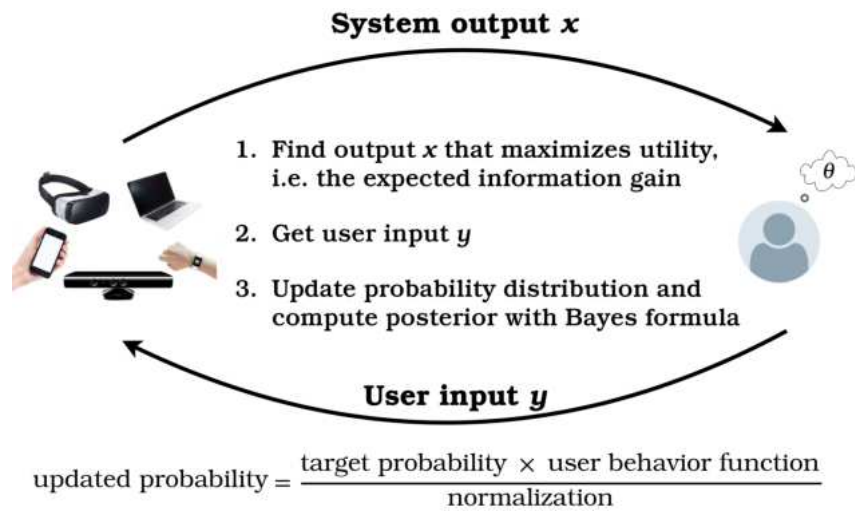


Figure 8: *BIGNav* presents the user with choices selected to maximize the information gained from each user response and then updates the probability distribution accordingly.

goal is to maximize the information-theoretic concept of mutual information, also known as “information gain” (see Fig. 8.) Rather than simply executing user navigation commands, *BIGNav* interprets the user’s input to update its knowledge about the user’s intended target. It then produces a new view that maximizes the expected information gain i.e. that should provide the most information from the user’s next input. *BIGFile* [32] uses a similar approach for file navigation. The split interface combines standard file system navigation with an adaptive area that includes shortcuts calculated with *BIG*. Both *BIGNav* and *BIGFile* were 40% faster than the standard interface, but each increases the user’s cognitive load. This type of human-computer partnership highlights the trade-off that can arise between increasing human efficiency and creating a satisfying user experience. One of the remaining challenges is how to let users choose an appropriate level of difficulty based on their needs in the moment.

4.5 Shaping human and agent behavior

We also explored how humans and intelligent agents can shape each other’s behavior using full-body interaction. In the context of an art-science collaboration sponsored by *Curiositas* and *La Diagonale* at the Université Paris-Saclay, we collaborated with the *N+1 Theater Group* to produce a wall-sized interactive Christmas display, first for a month in a shopping center and then at an art-science fair (see Fig. 9).

The goal was to create an entertaining interactive experience where users share control with the system. The theater company created a mechanical interactive Santa-making factory that would operate only if a user had successfully engaged with an animated Santa Claus character. Our job was to teach passers-by to perform a specific movement, without explaining the movement or even that they were supposed to learn a movement. We designed the interaction so that the user thought they controlled the animated character but the character also “controlled” them. We used the concept of “reinforcement” from behavioral psychology, where the reinforcer (or reward) is Santa’s attention.

When the user first arrives, Santa mirrors her movements, which keeps her engaged, but only for the first 30 seconds or so. Next, Santa starts to act independently and waits until she starts to copy Santa’s movements. This begins the shaping process: a stair with three steps appears and Santa moves up each step as the user makes successive approximations to the correct movement. If the movement is too far

from desired movement, Santa looks bored and taps his foot. If she does not respond appropriately, the stairs disappear and Santa drops through the floor. If she successfully discovers the correct movements through trial and error, Santa climbs to the top step and jumps up to grab a virtual level. This in turn operates a mechanical lever that launches the machine.

The system was a great success and a number of people returned over successive days to play with Santa. Although it was created primarily as a source of entertainment for the public, it also helped us to better understand how an intelligent agent can both respond to the user and also guide (or control) them.



Figure 9: The user interacts with the Santa Claus figure.

5 Conclusion

This deliverable discusses our approach to creating human-computer partnerships, and includes our research methodology, an analysis of the potential problems of interacting with intelligent systems and several examples of how we have created human-computer partnerships. Although we have only recently had access to the *ALMA* algorithms, we are using our approach to creating interactive intelligent systems to compare it with another machine learning algorithm, as explained in Deliverable D.4.8.

References

- [1] Janet Rafner et al. “Deskilling, Upskilling, and Reskilling: a Case for Hybrid Intelligence”. In: *Morals & Machines* 1 (Jan. 2021), pp. 24–39. DOI: 10.5771/2747-5174-2021-2-24.
- [2] Michel Beaudouin-Lafon, Susanne Bødker, and Wendy E Mackay. “Generative theories of interaction”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 28.6 (2021), pp. 1–54.
- [3] Michel Beaudouin-Lafon. “Instrumental Interaction: An Interaction Model for Designing post-WIMP User Interfaces”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’00. The Hague, The Netherlands: ACM, 2000, pp. 446–453. ISBN: 1-58113-216-6. DOI: 10.1145/332040.332473.
- [4] Michel Beaudouin-Lafon and Wendy E. Mackay. “Reification, Polymorphism and Reuse: Three Principles for Designing Visual Interfaces”. In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI ’00. Palermo, Italy: ACM, 2000, pp. 102–109. ISBN: 1581132522. DOI: 10.1145/345513.345267.
- [5] Michel Beaudouin-Lafon. “Towards Unified Principles of Interaction”. In: *Proceedings of Advanced Visual Interfaces (AVI’17)*. CHIItaly ’17. Cagliari, Italy: Association for Computing Machinery, 2017. ISBN: 9781450352376. DOI: 10.1145/3125571.3125602.
- [6] Wendy E Mackay and Anne-Laure Fayard. “HCI, natural science and design: a framework for triangulation across disciplines”. In: *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques*. 1997, pp. 223–234.
- [7] Wendy Mackay. “Responding to cognitive overload : Co-adaptation between users and technology”. eng. In: *Intellectica* 30.1 (2000), pp. 177–193. ISSN: 0769-4113. DOI: 10.3406/intel.2000.1597.
- [8] J. Greenbaum and M. (Eds.) Kyng. *Design at Work: Cooperative Design of Computer Systems (1st ed.)*. Boca Raton, Florida, USA: CRC Press, 1991, p. 306.
- [9] Wendy E. Mackay. *DOIT: The Design of Interactive Things: CHI’23 Preview*. Paris, France: Inria, 2023.
- [10] Charles Darwin, ed. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. Albemarle Street, London, England: John Murray., 1859. ISBN: 9780080491417.
- [11] Tristan Green. “Meta takes new AI system offline because Twitter users are mean”. In: *TNW* (Nov. 2022). URL: <https://thenextweb.com/news/meta-takes-new-ai-system-offline-because-twitter-users-mean>.
- [12] Paul M. Fitts et al. *Human Engineering for an Effective Air Navigation and Traffic-control System*. Tech. rep. Defense Technical Information Center, 1951, p. 109.
- [13] George Stanley Reynolds. *A Primer of Operant Conditioning*. Glenview, IL: Scott, Foresman, 1968, p. 160.
- [14] Herbert H. Clark. *Using language*. Cambridge, UK: Cambridge University Press, 1996.
- [15] Wendy E. Mackay et al. “Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’98. Los Angeles, California, USA: ACM Press/Addison-Wesley Publishing Co., 1998, pp. 558–565. ISBN: 0201309874. DOI: 10.1145/274644.274719.
- [16] Neil Perry et al. *Do Users Write More Insecure Code with AI Assistants?* 2022. arXiv: 2211.03622 [cs.CR].
- [17] Brian Hill. “Taking the help or going alone: ChatGPT and class assignments”. In: *HEC Research Paper* (Mar. 2023). DOI: <http://dx.doi.org/10.2139/ssrn.4465833>.
- [18] Robert Munro Monarch. *Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI*. New York, NY: Manning, June 2021, p. 424.
- [19] Saeid Nahavandi. “Trusted Autonomy Between Humans and Robots: Toward Human-on-the-Loop in Robotics and Autonomous Systems”. In: *IEEE Systems, Man, and Cybernetics Magazine* 3.1 (2017), pp. 10–17. DOI: 10.1109/MSMC.2016.2623867.

- [20] Chad Steelberg. *The path to an AI-connected government*. Tech. rep. Data Center Dynamics, Apr. 2019. DOI: <https://www.datacenterdynamics.com/en/opinions/path-ai-connected-government>.
- [21] Michel Beaudouin-Lafon and Wendy E Mackay. “Rethinking interaction: From instrumental interaction to human-computer partnerships”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–5.
- [22] Wendy E. Mackay. *Réimaginer nos interactions avec le monde numérique*. Leçons inaugurales du Collège de France. Collège de France — Fayard, 2022. ISBN: 978-2-213-72507-9.
- [23] Olivier Bau and Wendy E. Mackay. “OctoPocus: A Dynamic Guide for Learning Gesture-Based Command Sets”. In: *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*. UIST ’08. Monterey, CA, USA: Association for Computing Machinery, 2008, pp. 37–46. ISBN: 9781595939753. DOI: 10.1145/1449715.1449724. URL: <https://doi.org/10.1145/1449715.1449724>.
- [24] Shumin Zhai and Per Ola Kristensson. “The Word-Gesture Keyboard: Reimagining Keyboard Interaction”. In: *Commun. ACM* 55.9 (Sept. 2012), pp. 91–101. ISSN: 0001-0782. DOI: 10.1145/2330667.2330689. URL: <https://doi.org/10.1145/2330667.2330689>.
- [25] Jessalyn Alvina, Joseph Malloch, and Wendy Mackay. “Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices”. In: *Proceedings of the 29th ACM Symposium on User Interface Software and Technology*. UIST’16. Tokyo, Japan: ACM, Oct. 2016, pp. 583–593. DOI: 10.1145/2984511.2984560. URL: <https://inria.hal.science/hal-01437054>.
- [26] Jessalyn Alvina et al. “CommandBoard: Creating a General-Purpose Command Gesture Input Space for Soft Keyboard”. In: *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. UIST ’17. Québec City, QC, Canada: Association for Computing Machinery, 2017, pp. 17–28. ISBN: 9781450349819. DOI: 10.1145/3126594.3126639. URL: <https://doi.org/10.1145/3126594.3126639>.
- [27] Jessalyn Alvina et al. “MojiBoard: Generating Parametric Emojis with Gesture Keyboards”. In: *CHI’19 Extended Abstracts*. CHI EA ’19. Glasgow, United Kingdom, May 2019. DOI: 10.1145/3290607.3312771. URL: <https://hal.inria.fr/hal-02279111>.
- [28] Joseph Malloch et al. “Fieldward and Pathward: Dynamic Guides for Defining Your Own Gestures”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 4266–4277. ISBN: 9781450346559. DOI: 10.1145/3025453.3025764. URL: <https://doi.org/10.1145/3025453.3025764>.
- [29] Janin Koch et al. “SemanticCollage: Enriching Digital Mood Board Design with Semantic Labels”. In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 2020, pp. 407–418.
- [30] Janin Koch et al. “ImageSense: An Intelligent Collaborative Ideation Tool to Support Diverse Human-Computer Partnerships”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW1 (2020), pp. 1–27.
- [31] Wanyu Liu et al. “BIGnav: Bayesian Information Gain for Guiding Multiscale Navigation”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 5869–5880. ISBN: 9781450346559. DOI: 10.1145/3025453.3025524. URL: <https://doi.org/10.1145/3025453.3025524>.
- [32] Wanyu Liu et al. “BIGFile: Bayesian Information Gain for Fast File Retrieval”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. ISBN: 9781450356206. DOI: 10.1145/3173574.3173959. URL: <https://doi.org/10.1145/3173574.3173959>.