Extending a Human-AI Collaboration Framework with Dynamism and Sociality

Michael Muller
michael_muller@us.ibm.com
IBM Research
Cambridge, MA, USA

Justin D. Weisz
jweisz@us.ibm.com
IBM Research
Yorktown Heights, NY, USA

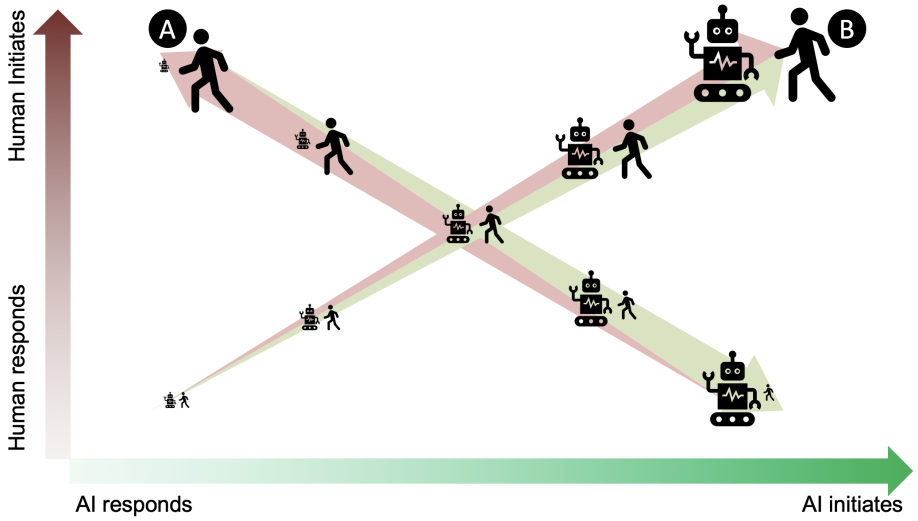


Figure 1: The Collaborating Humans and AIs (CHA) analytic space. Building on Shneiderman's Reliable, Safe, and Trustworthy model [83, 84], our space demonstrates how initiative may dynamically shift across different tasks in complex and organizational applications. In this figure, we identify two emergent axes. (A) The first emergent axis runs diagonally from upper left to lower right, and represents the Parasuraman et al. [73] zero-sum trade-off dimension of human-*versus*-AI initiative. (B) The second emergent axis runs from lower left to upper right, and represents human-*plus*-AI collaboration.

ABSTRACT

Abstract. We propose an integrated framework of Collaborating Humans and AIs (CHA) for the design, analysis, and critique of human-AI applications in current and future workplaces. Through four use cases, we explore different configurations of humans working with intelligent systems, informed by a 70-year history of frameworks of human-machine interaction. We make two primary contributions: (1) In contrast to prior analyses of mixed-initiative human-AI systems, we demonstrate that complex organizational applications involve dynamic shifts and exchanges of human and machine initiative. (2) We extend the analyses to include multi-user configurations, in which we can examine the sometimes-invisible human roles behind those structures, interrogate power relationships, and uncover mismatches between the desired and available levels of human and AI initiative.

CCS CONCEPTS

• **Human-centered computing** → **Computer supported cooperative work**; • **Computing methodologies** → *Cooperation and coordination.*

KEYWORDS

AI, frameworks, human-AI collaboration, autonomy, initiative, control

ACM Reference Format:
Michael Muller and Justin D. Weisz. 2022. Extending a Human-AI Collaboration Framework with Dynamism and Sociality. In *2022 Symposium on Human-Computer Interaction for Work (CHIWORK '22), June 8–9, 2022, Durham, NH, USA.* ACM, New York, NY, USA, Article 111, 12 pages. https://doi.org/10.1145/3533406.3533407

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHIWORK '22, June 8–9, 2022, Durham, NH, USA
© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9655-4/22/06…$15.00
https://doi.org/10.1145/3533406.3533407

1 INTRODUCTION

Relations between humans and their technological creations have been discussed in the Eurowestern traditions for centuries. Two hundred years ago, Wollstonecraft Shelley explored the responsibilities of Victor Frankenstein, the creator of an artificial sentience in *Frankenstein; or, the Modern Prometheus* [81]. One hundred years later, Čapek also questioned the relationship of created beings to their creators in *Rossum's Universal Robots* [11] - an early description of future workplace with an automated workforce.

The concept of the golem preceded either of these accounts by more than 1,000 years [56]. A golem is an autonomous laborer [55] that is formed of clay, and which must be activated through a code. In some accounts, the code is the inscription of the word *truth* (in Hebrew: \<'EmEt>) on the golem's forehead. Deletion of the *alef* (\<'>) character changes the word *truth* to the word *dead*, thereby de-activating the golem. Created for work as well as defense [55], the golem may be the earliest reference to a programmable artificial worker.

Contemporary research communities in AI, machine learning, HCI, and CSCW have generally aspired toward designing *human-centered* AI systems that are explainable [3, 61, 63, 67], equitable [22, 26, 26], and fair [7, 10], in order to preserve human control and minimize user harms [83]. However, not all automated systems ought to be designed to be overridden by human agency. In some cases, safeguards exist within automated systems to prevent their (mis)use by their human users. For example, anti-lock brakes and automated morphine pumps both effectively "lock out" their human users from overriding their behaviors: anti-lock brakes, when in action, respond to the driver's *intent* (stop the car) rather than their *action* (applying brakes continuously); automated morphine pumps dispense medication as requested, but do not allow a patient to over-medicate themselves, no matter how much they push the button [84].

All of these examples highlight a complex and nuanced space of human-machine interaction: to what extent is each party responsible for the other, and how much authority ought each party have over each other? The created being in *Frankenstein* bitterly complained about his creator's failure of care [82]. Clearly, there is no one "right" design for how much initiative or agency each party requires, as different applications have different requirements.

Our aim in this paper is first to summarize briefly one line of research that has examined different analytic spaces of human-initiative vs. machine-initiative. We will show that these simple trade-off analyses are not sufficient to describe contemporary organizational applications of AI technologies. We build on this earlier work by demonstrating how human-AI applications can be decomposed into their constituent steps - identifying tasks performed by human, AI, or both. This new analysis repairs weaknesses of the preceding work, particularly in representations of how much initiative, agency, or authority each party has at each step. This decomposition can lead to new insights about an AI-based application in workplace contexts, such as identifying steps in which either party has too little (or too much) initiative, agency, or authority. Previous work relied on simplifying assumptions that one human interacts with an AI application. However, this is an unrealistic assumption for organizational applications. Our analysis can help to understand AI applications that involve more than one human user in distinct roles. Using a stakeholder lens from Value Sensitive Design [30], we can also ask value-centric questions about the power of each human or AI entity in these complex configurations.

Our work makes the following contributions:

- We extend previous *static* models of Human Centered AI (HCAI) into the Collaborating Humans and AIs framework (CHA) to include *dynamically-shifting* initiatives and *multi-stakeholder* collaborations.
- We examine four complex and organizational human-AI applications in the CHA framework and show how the resulting space has additional descriptive and explanatory power.

2 BACKGROUND: THE EVOLUTION OF HUMAN-AI FRAMEWORKS

We briefly trace one line of thinking about configurations of humans and computer technologies in collaboration. In a paper of this scope, we do not compare models from the rich history of thinking about human-artifact and human-AI relations. Detailed reviews were recently provided by O'Neill et al. [72] and by Shneiderman [85]. Therefore, we refer interested readers to those works for a more extended version of the history.

2.1 Allocations

An early technical account of the differential assignment of joint work between human and computer was what has become known as the Fitts Allocation, published in 1951 [29]. The concept of "allocation" was as follows: Certain types of work should be uniquely assigned to *either* human *or* machine, respectively, according to a set of rules. In the sexist language of the times, Fitts and colleagues noted that "men [sic] are better at" certain types of tasks, and "machines are better at" other types of tasks.

Of course, no one thought that the allocation would remain fixed forever, because of changing computing capabilities and the creation of new human work-practices [8, 72]. About a quarter century after Fitts et al.'s foundational paper, Sheridan and Verplank [82] proposed a more flexible allocation of how humans and computers might jointly assign work tasks. Their innovation was that the degree of human initiative and computer initiative was not a matter of *fixed* allocations of tasks or functionality, but rather a *design decision* that could be selected on a per-application basis. While this new conceptualization assumed greater design flexibility (see Fiture 2), it retained the zero-sum nature of the Fitts Allocation: increases in human autonomy necessarily implied decreases in computer autonomy, and vice versa.

2.2 Dimensions

After another quarter of a century, Parasuraman, Sheridan, and Wickens [73] updated the Sheridan and Verplank trade-off function in their description of a *design dimension* that ran between complete human autonomy and complete computer autonomy (Figure 2). Their dimension provided an even more flexible and analytic way to think about human and computer task-assignment as a series of design choices. Similarly to Sheridan and Verplank's concept, the Parasuraman et al. dimension retained the zero-sum assumption

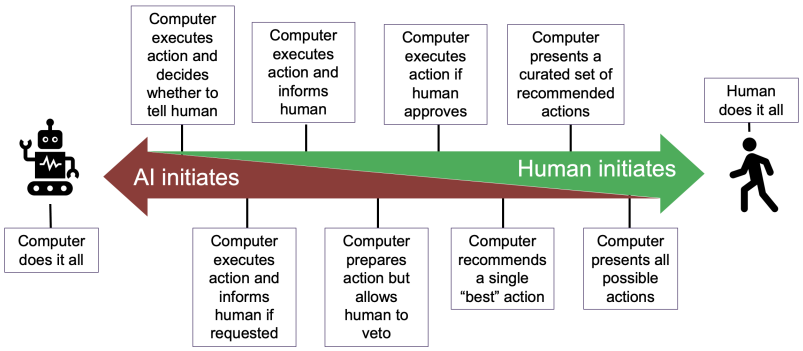


Figure 2: The human-computer trade-off allocation of Parasuraman et al. [73]. The dimension runs from complete computer autonomy (on the left) to complete human autonomy (on the right). The ten relative assignments of human-computer autonomy from Table ?? can be understood as points along this dimension.

that humans and computers were in a kind of competition for initiative.

2.3 Spaces

Twenty years after Parasuraman et al.'s uni-dimensional dimension of human-machine automation [73], Shneiderman [83] proposed a major expansion of the design space by decoupling human autonomy from computer autonomy. Shneiderman defined a two-dimensional space in which some applications might involve high human initiative *and* high AI initiative, while other applications might involve high values on *only one* of those dimensions, but not the other. The horizontal and vertical axes of Figure 1 correspond to Shneiderman's two dimensions, with human initiative along the vertical axis, and AI initiative along the horizontal axis (see also [65]).

3 A DYNAMIC, MULTI-STAKEHOLDER VIEW OF COLLABORATING HUMANS AND AIS (CHA)

In this paper, we build on Shneiderman's space of Human-Centered AI (HCAI) [83, 85]. We extend that framework from static descriptions of human-AI configurations, toward more dynamic representations of human and AI actions that take place responsively over time. We also consider AI systems that implicate more than one human, and we show how the extended representation can help us to find hidden human actors in a values-informed analysis of stakeholders in design (e.g., [30]. To strengthen the clarity of our analysis, we apply the concepts of multiple and distinct stakeholders from Value Sensitive Design (VSD) [30].

To meet the length expectations of CHIWORK, we describe only four examples from a larger set of unpublished analyses listed in the Appendix.

3.1 Emergent Dimensions for Complex Work Systems

We begin by integrating much of the history from Section 2 into a more powerful, unified concept. We then apply that concept to understand, diagnose, and critique a diverse set of CHA applications.

First, we return briefly to Parasuraman et al.'s zero-sum trade-off model, in which increases in human initiative implied decreases in AI initiative, and vice versa (Section 2.2). This dimension fits into the Shneiderman space of Section 2.3 as a minor diagonal from the extreme of high-human-initiative/low-AI-initiation, to the opposite extreme of low-human-initiative/high-AI-initiative, as shown in Figure 1. In some senses, this can be considered a kind of "human-vs.-AI" dimension.

The integration of the Parasuraman dimension into the Shneiderman space, suggested that there might be a second, *major* diagonal. That diagonal describes various degrees of human-AI collaboration, from low-intensity collaboration (e.g., a musicbox [84]) to high-intensity collaboration (e.g., a digital camera [84]). We added this emergent dimension to the Shneiderman space in Figure 1. In some senses, this can be considered as a "human+AI" dimension. We will use both the original dimensions of this space (human initiative, AI initiative) and the two emergent dimensions to anchor the process descriptions and social descriptions that follow.

3.2 Analyzing Influence, Interests, and Power through Value Sensitive Design

The preceding analyses have been applied primarily to simple configurations of one human and one piece of technology. When we analyze complex AI applications in organizations, the configurations often involve more than one human in different roles. Some of those roles look like what we would conventionally call "end-users," while other roles seem more remote from the situation of usage. In many organizational settings, the more remote roles have more organizational power: They structure and determine the conditions under which the end-users work. We need a language and analytic lens to understand how these different roles interact (differently) with a technology that is in some ways "common" to multiple roles.

Value Sensitive Design (VSD) provides a useful theory for considering these multiple-human roles and positions. Friedman and Hendry [30] proposed two categories of humans who were engaged with a design, technology, or policy: *direct* stakeholders who directly interact with the artifact (e.g., end-users), and *indirect* stakeholders who are involved in the technology in different ways. One

example occurs when the end-user mediates between technology and customer. Two cases below involve an indirect-stakeholder relationship of a customer to the direct-stakeholder end-user: the telephone operators in Section 3.5.2 and the bank loan officers in Section 3.5.3.

However, as we will show, there are often multiple stakeholders who are involved in applications that seem to interact with only one user at-a-time. For example, the Wikipedia editor in Section 3.4.1 who looks for instances of article-vandalism using the Huggle software bot, has impact on vandals, and works within a community of other editors. Similarly, the data-labelers in the active learning example of Section 3.4.2, are using a workflow-like application through which a data science worker guides their labeling activities.

In summary, many AI-based applications involve multiple human roles. Some humans are direct stakeholders who interact with the application (end-users) , and other humans are indirect stakeholders who are affected by the actions of the direct stakeholder, or who structure and control the working conditions of the direct stakeholder. These complex roles and relationships may be less apparent in simple AI-based consumer products. In this paper, we explore these complications when we move from consumer products to organizational systems.

3.3 Dynamic Shifts of Initiative and Autonomy

Shneiderman applied his conceptual space primarily to analyze familiar consumer or medical products, such as cars, elevators, and pacemakers [84]. In this paper, we apply the revised CHA framework of Figure 1 in two different directions: complex applications with dynamically-changing initiative between human and AI; and multi-stakeholder applications that involve two or more humans along with at least one AI. We recognize that these directions represent different choices from those that Shneiderman made, and we hope that other researchers will add their own directions and analytic representations over time.

In the first examples, we show how initiative changes dynamically within the steps of each of a single human-AI application. The changes in initiative are indicated by the position of each step in the analytic space. For readers with conventional vision, we encourage reference to the figures. For readers with reduced vision, we have written detailed descriptions of each figure in its caption, and we also add the following summary indicators of approximate semantic locations in the two-dimensional space of the figures:

- **H+** (high human initiative), **H0** (intermediate human initiative), and **H-** (low human initiative)
- **A+** (high AI initiative), **A0** (intermediate AI initiative), and **A-** (low AI initiative)

Our purpose is not to provide authoritative examinations of each use case, but rather to provide a common frame-of-reference from which interesting and productive discussions (and disagreements) can be had around existing and future human-AI systems. We hope to examine not only dynamics of initiative, but also to use feminist technoscience [40] to question common assumptions about how initiative, and hence aspects of power, may be configured.

3.4 Use Cases

3.4.1 Software Bots in Wikipedia. Historically, Wikipedia required a large amount of human labor to create and improve articles [39]. However, Wikipedia volunteers (often called "editors") became fatigued and sometimes withdrew from editing [28].

Because of the public access to editing, many editor actions have been focused on reducing acts of vandalism (i.e., damage to the text of Wikipedia articles). Editors have developed multiple software robots ("bots") that are designed to search for evidence of vandalism, to restore damaged text, and to keep records related to these events.

In this example, we describe the Huggle bot, which has been studied extensively in HCI and CSCW [31, 38, 39]. As described in Figure 3, Huggle operates under nominal human control (**H+A-**), but takes many actions autonomously (**H-A+**). The Huggle scenario also includes significant collaborative work by the human editor and the Huggle bot together (**H+A+**), to determine whether vandalism has actually taken place, and to decide on courses of action. Humans exercise considerable discernment and judgment to reduce the risk of overuse of bots *as decision makers* (e.g., [50, 51]), which might otherwise harm the moral order of Wikipedia [17, 28], thus worsening problems of bias and burnout.

In this example, the initiative shifted from (1) human activation of the bot, to (2) autonomous actions by the bot, and then to (3) shared actions by human and bot, before (4) escalation to even more intense shared action, and finally to (5) autonomous record-keeping by the bot. Clearly, the Huggle bot does not occupy a single location in the Shneiderman space of Figure 1, but in fact moves through three of the four quadrants of that analytic space.

The Huggle bot appears to be a single-user application. While that is nominally true, we can use the analytic lens of VSD to consider other stakeholders. The potential vandals are indirect stakeholders: They actions are analyzed by Huggle; their Wikipedia profiles may be modified by Huggle; and they may receive messages from Huggle. If they are acting as vandals, then their values and goals are oppose to those of the Wikipedia editors who use Huggle. Huggle thus intermediates communications and actions among humans with very different interests and motivations.

There is a second way in which Huggle has multiple stakeholders. There are many Wikipedia editors, and many of them use Huggle to find and repair acts of vandalism. Huggle is thus a shared tool among those editors, and they may need to agree on common usage protocols.

Finally, someone created Huggle, and someone maintains it. These roles do not appear to to interact or affect the editors' daily activities [31, 38, 39], so we did not include them in Figure 3. As we will see below, sometimes these roles do need to be considered in more complex applications.

3.4.2 Active Learning for Ground Truth Labels / Human-in-the-Loop. A second example from contemporary work practices is the work that humans do to write labels (annotations) onto machine-learning records. In machine learning, the assignment of ground truth labels (i.e., outcomes) to each data record can present challenges of knowledge, time, and expense "because it requires the participation of human experts" [45]. In some cases, crowdmarket workers have sufficient knowledge to assign labels [16]. In other cases, expert or domain knowledge is required to assign accurate labels [5, 20, 70]. Because the human is asked to provide, as it were, labeling services

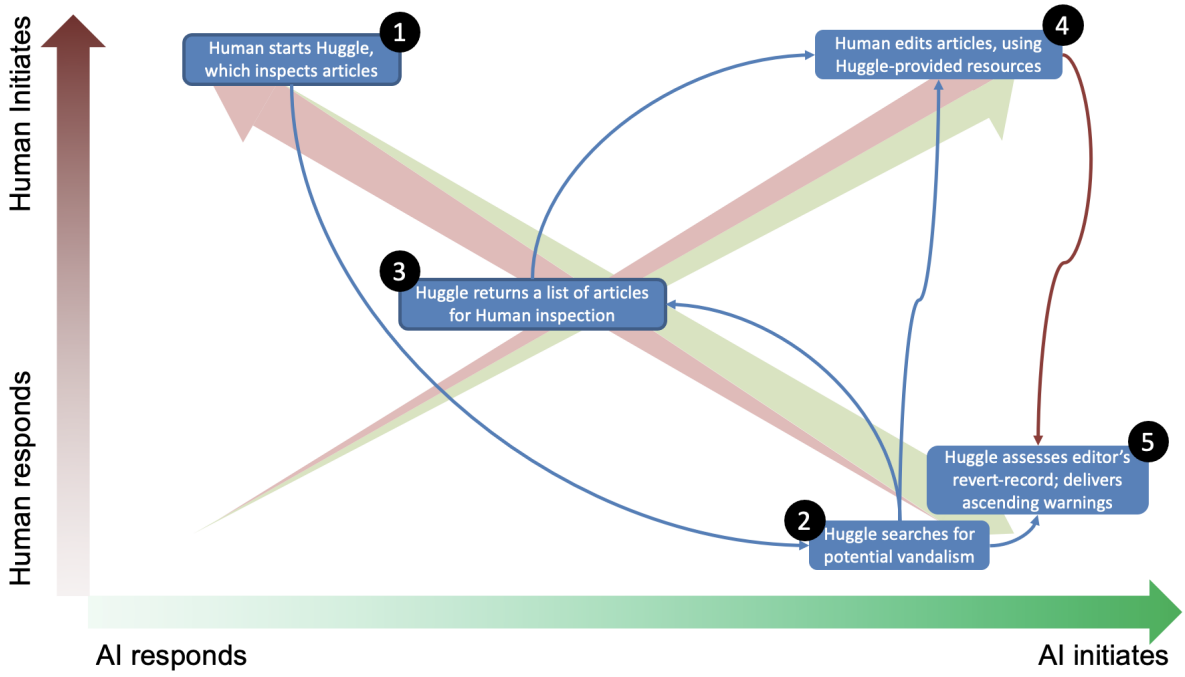


Figure 3: The Huggle software bot in Wikipedia. In Step 1, a human editor initiates Huggle (H+A-), which begins in Step 2 to search autonomously for potential instances of vandalism upon Wikipedia articles (H-A+). In Step 3, if Huggle finds potential vandalism, it consults with the human editor (H0A0). The human editor can then work collaborative with Huggle in Step 4 (H+A+) to understand what was done to the article's text, and optionally what other actions have been done by the editor who changed that text (i.e., in the language of the application, the potential vandal). Huggle may then record the actions taken by that potential vandal, and may also issue an automated warning to that person (H-A+).

to the AI, this configuration is sometimes called "human-in-the-loop" [94] or "human teaching" of the AI [86].

Researchers have turned to methods based on active learning to reduce these costs. Active learning involves a combination of human labeling and algorithmic labeling. At some points in this complex process, the human dominates the labeling process, while at other points the AI is in control, and in yet other cases, the labeling authority is shared between the human and algorithmic labelers.

In Figure 4, we illustrate one version of an active learning approach in the CHA space, based on the account provided by Hossain et al. [44]. We note that the actions begin with high human initiative and low AI initiative (**H+A-**), and proceed through human-AI collaboration (**H+A+**). During production mode, humans fade out of the action (**H-A+**), and return to active involvement (**H+A-**) only if a monitoring process detects problems (**H0A0**).

Most papers about active learning are written by data scientists (e.g., [5, 16, 20, 45] who create the algorithms, design the user interfaces, and recruit labelers to create the labels [70]. A separate and distinct literature has documented the experiences of the labelers [19, 46]. When more comparative studies are available, it will be useful to revise Figure 4 to include both the labelers (direct stakeholders) and the people who structure their work (indirect stakeholders.

3.4.3 Summary: Dynamic Initiative. Shneiderman's HCAI analysis [84] treated technologies and applications as "point locations" in a two-dimensional analytic space of autonomy. In CHA, we have added analytic power to highlight choices in complex and/or organizational AI applications, showing dynamic changes in initiative during ordinary usage *within a single application.*

3.5 Multi-User Applications

Earlier in this paper, we said that we would need to challenge certain aspects of Shneiderman's framework. In the Introduction to his already-influential book, *Human-Centered AI*, Shneiderman states that

> Researchers and developers for HCAI systems value meaningful human control.... HCAI systems... emphasize human control, while embedding high levels of automation by way of AI and machine learning. Examples include digital cameras and [automotive] navigation systems, which give humans control yet have many automated features. [85, p. 8]

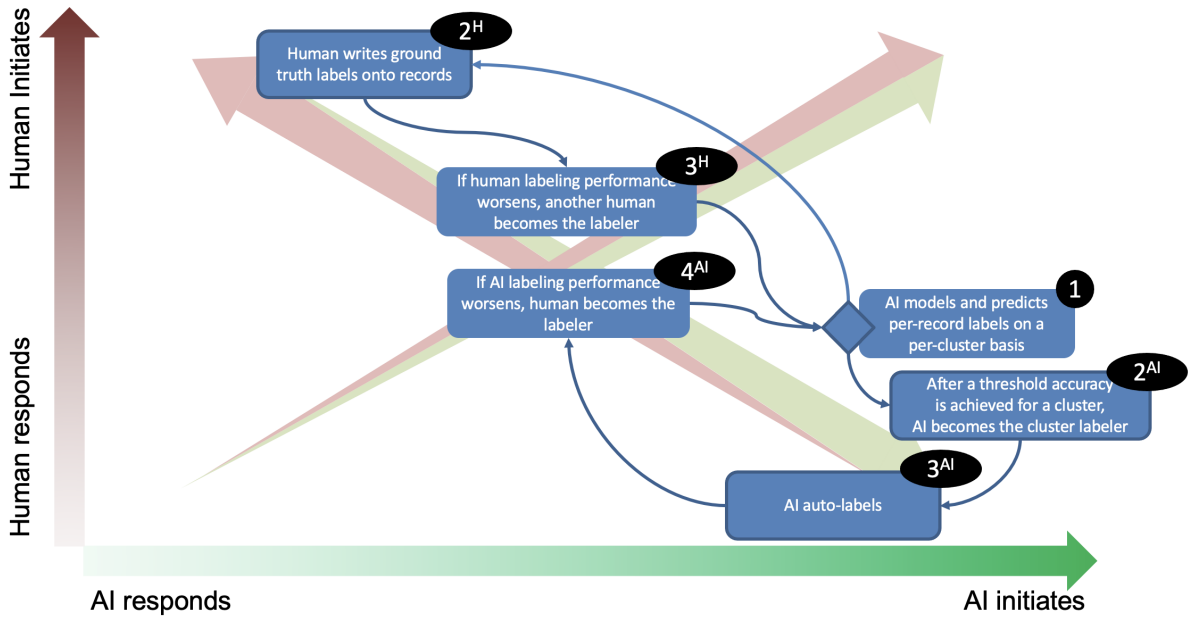


Figure 4: Active Learning process. In Step 1, the work begins as the human provides labels for an initial set of records (H+A-). In Step 2, the AI begins to create a model of the human's process, and begins to predict labels (H+A0). In Step 3, the human monitors the performance of the AI labeling model (H+A+). In Step 4, the AI model becomes accurate enough to replace the human labeler for many records (H-A+), and in Step 5, commits labels automatically (H-A+). However, it is possible that the stream of data may change, or that there may be other problems with the predictive labeling model. Therefore, in Step 6, if the automatic labeling performance worsens, the human expert may again need to take over the task of labeling (H0A0).

However, many workplace AI applications are designed to constrain human actions into approved patterns, rather than to empower those humans. Perhaps we should not be surprised, because workflow systems have channeled human work for decades [6, 9, 24], including workflows configured as crowdwork [15, 34] and as citizen science [88, 93].

Russell et al. note that "In the main, the attention of workflow researchers and workflow developers has focussed on the process perspective, i.e., control-flow" [79, p. 216]. Their contribution reconsiders the process-control assumptions that drive workflows, by proposing "push patterns ('the system pushes work to a worker') and pull patterns ('the worker pulls work from the system') to describe the many ways in which work can be distributed." We add that even the "pull patterns" receive system-structured workflows and system-designated assignments for generic or particular humans.

Conventional AI-based systems for planning, optimizing, and conducting workflows typically treat humans as resources to be allocated for efficient workflows [15, 62, 75, 78, 95]. This may not be the world that we would prefer. It is, however, a powerful attribute of the world in which we live, where humans' work is channeled by systems that have been designed by other humans. In the next examples, we consider how algorithms structure work that is explicitly done by groups of interdependent humans.

3.5.1 Bots and Labels, Revisited. As discussed in Section 3.4.1, the use of Wikipedia bots for vandalism-detection, actually involves at least two humans: the Wikipedia editor and the (possible) vandal. In a more subtle way, the Active Learning example in Section 3.4.2 also involves at least two humans in different roles: the human labeler and the humans who set up the labeling task. In both cases, the AI performs an explicit or implicit mediating function, structuring the work of one human in relation to one or more other humans. We now take up two cases in which the roles of *multiple* humans are in the foreground.

3.5.2 Telephone Operators. The field of HCI has been interested in the work of telephone operators for several decades [35, 36, 48, 60, 71]. The analyses in this paper allow us to take another look at this well-studied domain. Before easy online searches were available, customers of telephone companies often contacted telephone operators (employees of those companies) for assistance in finding a telephone number [35]. Sometimes the telephone number was obvious, such as a specific person at a known address. But, sometimes the customer brought a problem rather than a name to look up, such as "my dog ran away" [71]. In this case, the operator's job included refinement and clarification of the customer's request, until the operator was able to find a telephone number to meet the customer's need.

Because of the volume of calls and the staffing requirements, telecommunications companies attempted to reduce the human

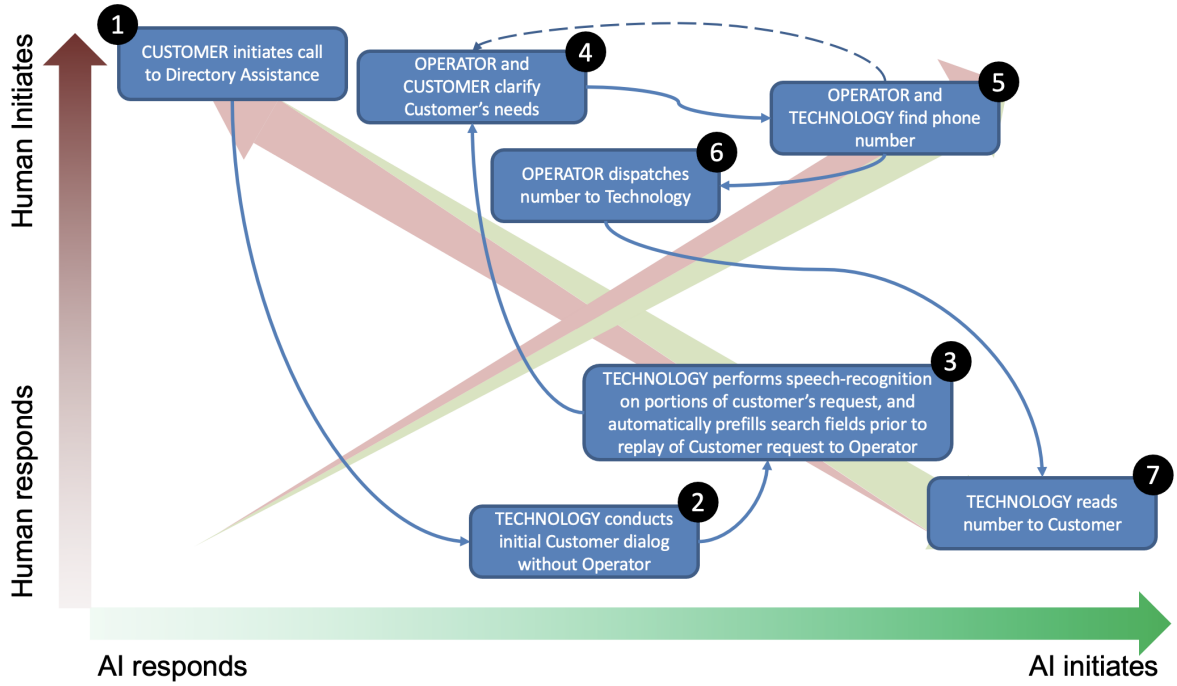


Figure 5: Telephone operator and customer in collaboration with technology, told primarily from the operator's perspective. In Step 1, a customer places a call to the telephone operator (H+A-). The initial dialog with the customer takes place automatically in Step 2, using speech-recognition and audio-storage technologies (H-A+). In Step 3, the technology pre-populates the operator's workstation with speech-recognized data, and plays back any audio to the operator (H-A+). In Step 4, the operator and the customer collaboratively refine the customer's request (H+A-), until in Step 5 there is sufficient information for the operator to collaborate with technology (H+A+) to find the number that the customer needs. In Step 6, the operator can then dispatch the number (H+A0) for replay to the customer in Step 7 (H-A+).

work-time for each call placed to a customer-assistance bureau that provided telephone numbers on-demand [14, 35, 36, 48, 60], sometimes applying early forms of modeling algorithms to compute efficiencies [35], or early forms of speech recognition to reduce the number of conversational interchanges between telephone operator and customer [76]. Despite the fact that this research took place more than two decades ago, the work involved two humans, one of whom was collaborating with computerized technology, including primitive forms of AI (i.e., speech recognition). Here, we apply the CHA framework to bring new insights regarding the locus of control in customer-operator interactions, and the potential to "tune" the application of AI technologies along a continuum of service quality and labor savings.

Figure 5 describes a typical workflow for a telephone operator interacting with a customer. It is notable how many times the initiative passes from human to computer and back again. Significant work is done in three of the four quadrants of the analytical space (**H+A-**, **H+A+**, and **H-A+**). At first glance Figure 5 suggests a graceful intermingling of human work and algorithmic work. However, is also notable that, in Step 3, portions of the customer's request are analyzed via speech recognition, and the results are used to pre-populate data fields for the operator. Usually, a dictionary of city names was sufficiently small to support high-accuracy speech recognition. However, a dictionary of human names may be extraordinarily large. Further, if a customer has a need that cannot be reduced to "first-city-then-listing," then the speech-recognition will almost inevitably fail (e.g., when the customer says "my dog ran away"). The analysis in Figure 5 could be used to evaluate trade-offs between service quality and labor savings by tweaking operators' reliance on speech recognition, depending on corporate goals and regulatory requirements.

This kind of analysis could be used diagnostically to find points of breakdown; for example, a strict use of speech recognition for a *location* could interfere with a customer's ability to state a less well-defined need such as a lost dog. Thus, the CHA approach can be used for critical reflection on the technologies and practices that shape jobs. It is particularly significant that so much of the work is done according to computer dominance **(H-A+)**. The work of the operator, including their work with the customer (**H+A-**), is strongly shaped, and constrained, by the technology [21]. A multi-party VSD analysis of the needs of both operators (direct stakeholders) and customers (indirect stakeholders) might reveal opportunities to increase collaboration possibilities between them). Alternative designs might have foregrounded the skills [59] and knowledge [71] of the operator workforce, to promote higher service quality to customers who were looking for help with life challenges and

urgent needs. Such a design might have put powerful search and disambiguation tools at the disposal of operators, shifting more of their work into the human-AI collaboration quadrant (**H+A+**).

3.5.3 Bank Loan Applications. AI support for bank loan approvals provides an even stronger case of multi-stakeholder AI systems. The process of applying for a bank loan initially appears to be a two-human use-case; a simple version of this process is sketched in Figure 6. After a customer applies for a loan, "almost all the applications are manipulated and controlled by machine learning algorithms" [52]. The bank loan officer makes relatively few decisions, and the customer makes even fewer decisions [92]. Human action and agency have been replaced by algorithmic action and agency.

In the language of VSD, the customer may be a direct stakeholder if they enter their own data into the bank loan system. More importantly for the use of the algorithm, the loan officer is clearly a direct stakeholder, because they interact with the loan-rating system, and they deliver news of the outcome to the customer. If this technology involves "human control" [85], neither the customer nor the bank officer is in a position to control the outcome. We must look further to find which human is in control, and how they exercise that control.

In the CHA context, feminist technoscience [40] informed by VSD allows us to consider two additional actors. While each of these actors is invisible in the conventional analysis, they are both *direct* stakeholders. One actor (or more likely, a group of actors) is the developer of the loan-rating algorithm, who choose the factors that predict loan-repayment, and who choose how those factors are combined into a predictive model and pipeline [68]. A second actor may be a bank software engineer who parameterizes the algorithm for the specific needs of a particular bank. Thees two hidden actors exercise more power over the outcome than either of the two visible actors.

The algorithm functions as an operationalization of the bank's policies. The developer(s) and the engineer(s) have greater ability to determine how those policies are enacted, in comparison to the loan officer or the customer. A deeper, VSD-informed examination of the human *policy-making* roles among developers and bank engineers, may show that their decisions have been constrained by other, more powerful *indirect* stakeholders, such as executives who choose among predictive models and other executives who specify the parameters for a particular bank. Further indirect-stakeholder influence may be exerted by legislators and regulators.

Researchers have explored the use of AI-based explanations (XAI) to help customers to understand certain loan outcomes (e.g., [23]). Even if there were solutions to the current problems of intelligibility and actionability in relation to bank loans (e.g., [18, 67]), a technical solution does not markedly increase the action or agency of the customer or the bank officer. The algorithm would remain dominant, and an explanation would occur in the AI-initiative quadrant (**H-A+**).

This brief CHA analysis of an organizational commercial application can be used to promote critical reflection on how policies become designs, and how design decisions can have profound impacts on stakeholders. As with the telephone operators example in Section 3.5.2, we can use this analysis to inquire why all of the human experiences with the loan process take place in the quadrant of AI domination (**H-A+**). There may be good reasons to *attempt* to govern portions of complex decisions via algorithms, such as an aspiration to reduce bias [32] - although there is strong evidence that such algorithms tend to embody and transmit bias, rather than reducing bias [37, 41]. Nonetheless, in the bank loan example, it was a series of policy and design decisions to implement the algorithm as an inflexible authority that drove all important decisions into algorithmic dominance (**H-A+**), rather than as a source of advice that could contribute to human-human negotiations (**H+A+**). It is a human and a society decision when we choose a position along a trade-off between human initiative and algorithmic initiative. CHA analysis, supported by VSD, can help us to clarify those decisions, and to interrogate the basis for those decisions.

3.5.4 Summary: Multi-Stakeholder Applications. In Section 3.3, we analyzed dynamic changes in initiative between humans and AIs as collaboration partners. The current section (Section 3.5) used CHA to consider AIs as an additional collaborator (or constraint-imposer) in human-human collaboration. By applying VSD stakeholder analysis to the expanded framework in this paper, we could focus on critical design decisions that narrow the range of human collaborative decision-making.

4 DISCUSSION

We described an expanded analytic space for frameworks of Collaborating Humans and AIs (CHA). We used this space to go beyond the prior "static" analyses of mixed human-AI initiatives [29, 43, 72, 73, 82, 83, 85], demonstrating the dynamically-changing nature of human-AI relationships in complex and organizational applications. We also used the CHA space to analyze how multiple stakeholders may interact with an AI system, and indirectly with one another as mediated by that AI system. We informed our analysis with principles from Value Sensitive Design [30].

4.1 Commonalities and Contrasts

We showed that many different kinds of AI applications involve dynamically-changing configurations of human and AI initiative. The operational steps of all of our use cases involved at least two regions of the CHA framework, and all but two of them involved three regions. The analysis, diagnosis, and critique of contemporary AI applications require the ability to track persons, agents, actions, and outcomes across fluid and situational processes [33, 43, 50, 51, 65, 77]. We hope that the CHA representation can support and extend those analyses.

One possibility for future work is to refine the steps in the CHA analyses as potential generic actions. These actions might then be used to analyze across task- and application- domains, seeking commonality of component actions. Spoto and Oleynik [87] and Muller and Weisz [69] have proposed basic vocabularies of actions that can, in principle, be performed by either human or AI. These actions were analyzed primarily into *sequences* of elemental operations. A more complete collection of CHA analyses could extend those analyses, but with a greater emphasis on how *collaborations* occur at each stage of the action. The appendix provides a possible starting set of such a collection of cases.

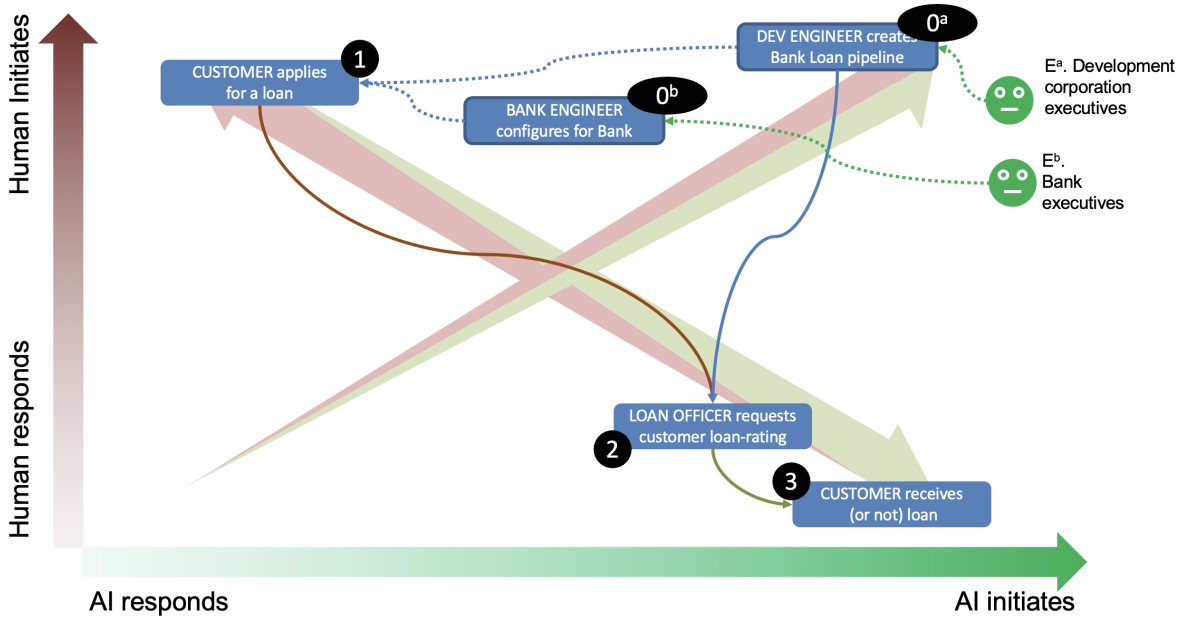


Figure 6: Bank loan process. In Step 1, a customer applies for a bank loan (H+A-). In Step 2, the loan officer requests a loan rating or approval from the AI (H-A+). In Step 3, the customer either does or does not receive the loan approval. We note that the policy that controls the loan approval is determined in part in *prior* Step 0^a by the developers of the AI (H+A+, where H is the developer). The policy may be further configured, within the parameters allowed in a *prior* step 0^a by the AI systems developer, which are specified in a *prior* step 0^b by a bank engineer (H0A+). These two roles (direct stakeholders) receive their instructions from development corporation executives E^a and bank executives E^b, who are *indirect* stakeholders (i.e., they do not interact with the loan-approval system). Neither the customer nor the loan officer can change the AI implementation of those policies, because the implementation and policies are controlled by parties who are absent from the negotiation (E^a, 0^a, E^b, 0^b) between the customer and loan officer.

The CHA analyses suggest another method for comparing and contrasting AI applications and domains along the emergent dimensions discussed in Section 3.1. Some applications, like active learning, tend to cluster along the huamn-vs.-AI trade-off dimension (Figure 4). Other applications, like the telephone operator and the bank loan process, are less tied to a single axis (Figures 5 and 6). With analysis of more cases (suggested in the appendix), we may be able to analyze in terms of more general principles, such as human-vs.-AI and human-plus-AI dynamics.

We also note that the active learning example in Figure 4 has relatively little activity in the high-collaboration region of the upper right of the figure (**H+A+**). We speculate that other human-in-the-loop applications may also lack high degrees of collaboration. CHA-based analyses can help to test that conjecture.

The very different analyses of Muller and Weisz [69], Parasuraman [73], and Spoto and Oleynik [87] agree in providing a trade-off analysis of which entity (human or AI) performs which action in a complex sequence of human and AI activities. Organizations are likely to re-evaluate their allocation of future work in terms of capabilities (e.g., [29] but also economics (e.g., [35, 36]). The "best" configuration of human and AI in complex tasks is likely to become a very temporary conclusion as technologies, work practices, and workforces change over time. If the configurations of humans and AIs are now considered as a socio-technical problem [4, 25, 66, 74], then organizations may also need to concern themselves with rapidly changing socio-technical-economic opportunities and choices. We hope that a CHA-derived representation - possibly combined with a per-activity analysis based on [69, 87], may help organizations to analyze their options.

4.2 Examining AI Applications as Instantiations of Policies

The preceding sub-section took a relatively organizational view of future of work. In this sub-section, we want to look instead from the worker's perspective.

Using VSD as an analytic lens, we showed that AI applications may intermediate between or among organizational stakeholders. A crucial question for these intelligent workflows is the assumed or enacted *location* of the intelligence. Historically, we have tended to assume that a technological process in the workplace should contain the knowledge of an organizational task, where that knowledge has been reduced to a sequence of actions, usually including a set of evaluative metrics to be applied to the performers of those actions

[15, 34, 88, 93]. Technology design has tended to assume that low-status workers need the guidance of a well-designed or intelligent workflow [35, 36], while high-status workers or consumers need to be able to be able to guide the intelligent technology [85]. We used VSD to focus on some of these design choices.

However, intelligent workflows are of course designed by humans. We also used the stakeholder concepts of VSD [30] to help to question how a complex system might be designed, and we began to ask questions inspired by feminist technoscience [40] about *who* the hidden designers of workplace AI applications might be. In this paper, we just began that analysis. We hope to continue this line of inquiry into other complex, social and societal applications.

4.3 Limitations

The text of this paper analyzed four use cases of human-AI configurations. We note in the Appendix that we have applied these analytic approaches to a larger and more diverse set of AI applications, with convergent results.

This work was presented in an analytic style, for a primary readership of researchers. We hope to discuss the themes and topics of these analyses with data science workers, to convert the academic language into checklists and other job-aids that can be applied in the fast-paced world of applied indusry applications.

5 CONCLUSION

In the Introduction, we considered manufactured beings and programmable workers. We considered four examples of complex created systems and their implications for workers in workplaces. Sometimes we program the artificial entities, but sometimes, it appears, they may program us.

We proposed the CHA framework to help us to think about diverse configurations of humans and AIs in current and future workplaces. The CHA framework integrates multiple prior frameworks in a unified analytic space, in which both human-*versus*-AI and human-*plus*-AI dimensions are explicitly represented. We used CHA to consider human-AI interactions as a sequence of shifting patterns of initiatives. Finally, we combined CHA, VSD, and feminist technoscience to represent visible and invisible stakeholders in human AI collaborations.

ACKNOWLEDGMENTS

We thank Michelle Brachman for helpful consultations to clarify domains for analysis.

REFERENCES

[1] Ankit Agarwal, Rachel A Freedman, Felicia Goicuria, Catherine Rhinehart, Kathleen Murphy, Eileen Kelly, Erin Mullaney, Myra St. Amand, Phuong Nguyen, and Nancy U Lin. 2017. Prior authorization for medications in a breast oncology practice: navigation of a complex process. *Journal of Oncology Practice* 13, 4 (2017), e273–e282.

[2] Melanie Arthur, Jerris R Hedges, Craig D Newgard, Brain S Diggs, and Richard J Mullins. 2008. Racial disparities in mortality among adults hospitalized after injury. *Medical care* (2008), 192–199.

[3] Vijay Arya, Rachel KE Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C Hoffman, Stephanie Houde, Q Vera Liao, Ronny Luss, Aleksandra Mojsilović, et al. 2019. One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques. *arXiv preprint arXiv:1909.03012* (2019).

[4] Aleksandre Asatiani, Pekka Malo, Per Rådberg Nagbøl, Esko Penttinen, Tapani Rinta-Kahila, and Antti Salovaara. 2021. Sociotechnical envelopment of artificial intelligence: An approach to organizational deployment of inscrutable artificial intelligence systems. *Journal of the Association for Information Systems* 22, 2 (2021), 8.

[5] Zahra Ashktorab, Michael Desmond, Josh Andres, Michael Muller, Narendra Nath Joshi, Michelle Brachman, Aabhas Sharma, Kristina Brimijoin, Qian Pan, Christine T Wolf, et al. 2021. AI-Assisted Human Labeling: Batching for Efficiency without Overreliance. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.

[6] LS Baumann and RD Coop. 1980. Automated workflow control: A key to office productivity. In *Proceedings of the May 19-22, 1980, national computer conference*. 549–554.

[7] Rachel KE Bellamy, Kuntal Dey, Michael Hind, Samuel C Hoffman, Stephanie Houde, Kalapriya Kannan, Pranay Lohia, Jacquelyn Martino, Sameep Mehta, Aleksandra Mojsilovic, et al. 2018. AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. *arXiv preprint arXiv:1810.01943* (2018).

[8] Jeffrey M Bradshaw, Paul J Feltovich, and Matthew Johnson. 2017. Human–agent interaction. In *The handbook of human-machine interaction*. CRC Press, 283–300.

[9] Steen Brahe and Kjeld Schmidt. 2007. The story of a working workflow management system. In *Proceedings of the 2007 international ACM conference on Supporting group work*. 249–258.

[10] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. PMLR, 77–91.

[11] Karel Čapek. 1920 and 2020. *RUR*. Standard Ebooks.

[12] Nancy J Cooke and William F Lawless. 2021. Effective Human–Artificial Intelligence Teaming. In *Systems Engineering and Artificial Intelligence*. Springer, 61–75.

[13] Sam Corbett-Davies, Sharad Goel, and Sandra González-Bailón. 2017. Even imperfect algorithms can improve the criminal justice system. *New York Times* (2017).

[14] Anna L Cox and David Peebles. 2008. Cognitive modelling in HCI research. Cambridge University Press.

[15] Peng Dai, Daniel Sabey Weld, et al. 2010. Decision-theoretic control of crowdsourced workflows. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*.

[16] Patrick M De Boer and Abraham Bernstein. 2017. Efficiently identifying a well-performing crowd process for a given problem. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1688–1699.

[17] Paul B De Laat. 2015. The use of software tools and autonomous bots against vandalism: eroding Wikipedia's moral order? *Ethics and Information Technology* 17, 3 (2015), 175–188.

[18] Martijn Demollin, Katarzyna Budzynska, Carles Sierra, et al. 2020. Argumentation theoretical frameworks for explainable artificial intelligence. In *2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence*. 44–49.

[19] Emily Denton, Mark Díaz, Ian Kivlichan, Vinodkumar Prabhakaran, and Rachel Rosen. 2021. Whose Ground Truth? Accounting for Individual and Collective Identities Underlying Dataset Annotation. *arXiv preprint arXiv:2112.04554* (2021).

[20] Michael Desmond, Michael Muller, Zahra Ashktorab, Casey Dugan, Evelyn Duesterwald, Kristina Brimijoin, Catherine Finegan-Dollak, Michelle Brachman, Aabhas Sharma, Narendra Nath Joshi, et al. 2021. Increasing the Speed and Accuracy of Data Labeling Through an AI Assisted Interface. In *26th International Conference on Intelligent User Interfaces*. 392–401.

[21] Barbara Brandenburg Diekmann. 1996. *The mutual shaping of task context and system interface in directory assistance conversations*. University of Colorado at Boulder.

[22] Catherine D'Ignazio and Lauren F. Klein. 2020. Seven intersectional feminist principles for equitable and actionable COVID-19 data. *Big data & society* 7, 2 (2020), 2053951720942544.

[23] Finale Doshi-Velez, Mason Kortz, Ryan Budish, Chris Bavitz, Sam Gershman, David O'Brien, Kate Scott, Stuart Schieber, James Waldo, David Weinberger, et al. 2017. Accountability of AI under the law: The role of explanation. *arXiv preprint arXiv:1711.01134* (2017).

[24] Paul Dourish. 2001. Process descriptions as organisational accounting devices: the dual use of workflow technologies. In *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*. 52–60.

[25] Upol Ehsan and Mark O Riedl. 2020. Human-centered explainable ai: towards a reflective sociotechnical approach. In *International Conference on Human-Computer Interaction*. Springer, 449–466.

[26] Karen Elliott, Rob Price, Patricia Shaw, Tasos Spiliotopoulos, Magdalene Ng, Kovila Coopamootoo, and Aad van Moorsel. 2021. Towards an Equitable Digital Society: Artificial Intelligence (AI) and Corporate Digital Responsibility (CDR). *Society* (2021), 1–10.

[27] CA Ellis. 2000. *An evaluation framework for collaborative systems*. Technical Report. Technical Report, CU-CS-901-00, University of Colorado, Department of ....

[28] Ryan Faulkner, Steven Walling, and Maryana Pinchuk. 2012. Etiquette in wikipedia: Weening new editors into productive ones. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–4.

[29] Paul M Fitts, MS Viteles, NL Barr, DR Brimhall, Glen Finch, Eric Gardner, WF Grether, WE Kellum, and SS Stevens. 1951. *Human engineering for an effective air-navigation and traffic-control system, and appendixes 1 thru 3*. Technical Report. Ohio State Univ Research Foundation Columbus.

[30] Batya Friedman and David G Hendry. 2019. *Value sensitive design: Shaping technology with moral imagination*. Mit Press.

[31] R Stuart Geiger and David Ribes. 2010. The work of sustaining order in Wikipedia: The banning of a vandal. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 117–126.

[32] Andra C Ghent, Rubén Hernández-Murillo, and Michael T Owyang. 2011. *Race and subprime loan pricing*. European Regional Science Association.

[33] Vlad Petre Glăveanu. 2013. Rewriting the language of creativity: The Five A's framework. *Review of General Psychology* 17, 1 (2013), 69–81.

[34] Shinsuke Goto, Toru Ishida, and Donghui Lin. 2016. Understanding crowdsourcing workflow: modeling and optimizing iterative and parallel processes. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 4.

[35] Wayne D Gray, Bonnie E John, and Michael E Atwood. 1992. The precis of Project Ernestine or an overview of a validation of GOMS. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 307–312.

[36] Wayne D Gray, Bonnie E John, Rory Stuart, Deborah Lawrence, and Michael E Atwood. 1995. GOMS meets the phone company: Analytic modeling applied to real-world problems. In *Readings in Human–Computer Interaction*. Elsevier, 634–639.

[37] Sumona Halder. 2020. *The Dangers of Algorithmic Autonomy: Efficient Machine Learning Models of Inductive Biases Combined With the Strengths of Program Synthesis (PBE) to Combat Implicit Biases of the Brain*. Ph. D. Dissertation.

[38] Aaron Halfaker. 2013. *Maintaining the efficiency of open production systems at scale: A case study of wikipedia*. Ph. D. Dissertation. University of Minnesota.

[39] Aaron Halfaker and John Riedl. 2012. Bots and cyborgs: Wikipedia's immune system. *Computer* 45, 03 (2012), 79–82.

[40] Donna J Haraway and Thyrza Nichols Goodeve. 2018. Femaleman© _Meets_Oncomouse™: Mice Into Wormholes: A Technoscience Fugue in Two Parts. In *Modest_Witness@ Second_Millennium. FemaleMan© _Meets_OncoMouseTM*. Routledge, 46–118.

[41] Thomas Hellström, Virginia Dignum, and Suna Bensch. 2020. Bias in Machine Learning–What is it Good for? *arXiv preprint arXiv:2004.00686* (2020).

[42] Andrew Hertzberg, Jose Maria Liberti, and Daniel Paravisini. 2010. Information and incentives inside the firm: Evidence from loan officer rotation. *The Journal of Finance* 65, 3 (2010), 795–828.

[43] Eric Horvitz. 1999. Principles of Mixed-Initiative User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 159–166. https://doi.org/10.1145/302979.303030

[44] HM Sajjad Hossain, Md Abdullah Al Hafiz Khan, and Nirmalya Roy. 2017. Active learning enabled activity recognition. *Pervasive and Mobile Computing* 38 (2017), 312–330.

[45] Sheng-Jun Huang, Jia-Lve Chen, Xin Mu, and Zhi-Hua Zhou. 2017. Cost-Effective Active Learning from Diverse Labelers.. In *IJCAI*. 1879–1885.

[46] Lilly Irani. 2015. The cultural work of microwork. *New media & society* 17, 5 (2015), 720–739.

[47] Liang Jie, Cao Jian, and Wang Lei. 2017. Design of multi-mode UAV human-computer interaction system. In *2017 IEEE international conference on unmanned systems (ICUS)*. IEEE, 353–357.

[48] Bonnie E John. 1994. *A database for analyzing sequential behavioral data and their associated cognitive models*. Technical Report. CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE.

[49] Bernadette Jones, Paula Toko King, Gabrielle Baker, and Tristram Ingham. 2020. COVID-19, intersectionality, and health equity for indigenous peoples with lived experience of disability. *American Indian Culture and Research Journal* 44, 2 (2020), 71–88.

[50] Anna Jordanous. 2016. Four PPPPerspectives on computational creativity in theory and in practice. *Connection Science* 28, 2 (2016), 194–216.

[51] Anna Kantosalo and Tapio Takala. 2020. Five C's for Human-Computer Co-Creativity-An Update on Classical Creativity Perspectives.. In *ICCC*. 17–24.

[52] R Karthiban, M Ambika, and KE Kannammal. 2019. A review on machine learning classification technique for bank loan approval. In *2019 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 1–6.

[53] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2020. BlindPilot: A robotic local navigation system that leads blind people to a landmark object. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.

[54] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2020. Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–22.

[55] Hillel J Kieval. 1997. Pursuing the Golem of Prague: Jewish culture and the invention of a tradition. *Modern Judaism* 17, 1 (1997), 1–23.

[56] Hillel J Kieval. 2000. 4. Pursuing the Golem of Prague: Jewish Culture and the Invention of a Tradition. In *Languages of Community*. University of California Press, 95–113.

[57] Atay Kizilaslan and Aziz A Lookman. 2017. Can Economically Intuitive Factors Improve Ability of Proprietary Algorithms to Predict Defaults of Peer-to-Peer Loans? *Available at SSRN 2987613* (2017).

[58] Masaki Kuribayashi, Seita Kayukawa, Hironobu Takagi, Chieko Asakawa, and Shigeo Morishima. 2021. LineChaser: A Smartphone-Based Navigation System for Blind People to Stand in Lines. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[59] Deborah Lawrence, Michael E Atwood, and Shelly Dews. 1994. Surrogate users: mediating between social and technical interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 399–404.

[60] Debbie Lawrence and Shelly Dews. 1992. Natural dialog in a time-sensitive setting: A study of telephone operators. In *Posters and Short Talks of the 1992 SIGCHI Conference on Human Factors in Computing Systems*. 103–104.

[61] Q Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. *arXiv preprint arXiv:2001.02478* (2020).

[62] Anthony Liu, Santiago Guerra, Isaac Fung, Gabriel Matute, Ece Kamar, and Walter Lasecki. 2020. Towards hybrid human-AI workflows for unknown unknown detection. In *Proceedings of The Web Conference 2020*. 2432–2442.

[63] Maria Teresa Llano, Mark d'Inverno, Matthew Yee-King, Jon McCormack, Alon Ilsar, Alison Pease, and Simon Colton. 2020. Explainable Computational Creativity.. In *ICCC*. 334–341.

[64] Alexandra Lyn. 2020. Risky Business: Artificial Intelligence and Risk Assessments in Sentencing and Bail Procedures in the United States. *Available at SSRN 3831441* (2020).

[65] Mary Lou Maher. 2012. Computational and collective creativity: Who's being creative?. In *ICCC*. Citeseer, 67–71.

[66] Erin E Makarius, Debmalya Mukherjee, Joseph D Fox, and Alexa K Fox. 2020. Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research* 120 (2020), 262–273.

[67] Avleen Malhi, Samanta Knapic, and Kary Främling. 2020. Explainable agents for less bias in human-agent decision making. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*. Springer, 129–146.

[68] Michael Muller and Angelika Stroymayer. 2022. Forgetting Practices in the Data Sciences. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. In press.

[69] Michael Muller, Justin D Weisz, and Werner Geyer. 2020. *Mixed Initiative Generative AI Interfaces: An Analytic Framework for Generative AI Applications*. Retrieved 13-Jul-2021 from https://computationalcreativity.net/workshops/cocreative-iccc20/papers/Future_of_co-creative_systems_185.pdf

[70] Michael Muller, Christine T Wolf, Josh Andres, Michael Desmond, Narendra Nath Joshi, Zahra Ashktorab, Aabhas Sharma, Kristina Brimijoin, Qian Pan, Evelyn Duesterwald, et al. 2021. Designing Ground Truth and the Social Life of Labels. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.

[71] Michael J Muller, Rebecca Carr, Catherine Ashworth, Barbara Diekmann, Cathleen Wharton, Cherie Eickstaedt, and Joan Clonts. 1995. Telephone operators as knowledge workers: Consultants who meet customer needs. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 130–137.

[72] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. 2020. Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors* (2020), 0018720820960865.

[73] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans* 30, 3 (2000), 286–297.

[74] Sun Young Park, Pei-Yi Kuo, Andrea Barbarin, Elizabeth Kaziunas, Astrid Chow, Karandeep Singh, Lauren Wilcox, and Walter S Lasecki. 2019. Identifying challenges and opportunities in human-AI collaboration in healthcare. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. 506–510.

[75] María Dolores R-moreno, Daniel Borrajo, Amedeo Cesta, and Angelo Oddi. 2007. Integrating planning and scheduling in workflow domains. *Expert Systems with Applications* 33, 2 (2007), 389–406.

[76] Lawrence R Rabiner. 1997. Applications of speech recognition in the area of telecommunications. In *1997 IEEE Workshop on Automatic Speech Recognition and Understanding Proceedings*. IEEE, 501–510.

[77] Mel Rhodes. 1961. An analysis of creativity. *The Phi delta kappan* 42, 7 (1961), 305–310.

[78] Mark O Riedl. 2019. Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies* 1, 1 (2019), 33–36.
[79] Nick Russell, Wil MP van der Aalst, Arthur HM Ter Hofstede, and David Edmond. 2005. Workflow resource patterns: Identification, representation and tool support. In *International conference on advanced information systems engineering*. Springer, 216–232.
[80] Vernon Schmitt, Gavin Jenney, and James Morris. 1998. *Fly-by-wire.* SAE.
[81] Mary Wollstonecraft Shelley. 1818 and 2018. *Frankenstein, or the modern Prometheus.* ARC, Amsterdam University Press.
[82] Thomas B Sheridan and William L Verplank. 1978. *Human and computer control of undersea teleoperators.* Technical Report. Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab.
[83] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction* 36, 6 (2020), 495–504.
[84] Ben Shneiderman. 2021. Tutorial: Human-Centered AI: Reliable, Safe and Trustworthy. In *26th International Conference on Intelligent User Interfaces*. 7–8.
[85] Ben Shneiderman. 2022. *Human-Centered AI.* Oxford University Press.
[86] Patrice Y Simard, Saleema Amershi, David M Chickering, Alicia Edelman Pelton, Soroush Ghorashi, Christopher Meek, Gonzalo Ramos, Jina Suh, Johan Verwey, Mo Wang, et al. 2017. Machine teaching: A new paradigm for building machine learning systems. *arXiv preprint arXiv:1707.06742* (2017).
[87] Angie Spoto and Natalia Oleynik. 2017. *Library of Mixed-Initiative Creative Interfaces.* Retrieved 19-Jun-2021 from http://mici.codingconduct.cc/
[88] James Sprinks, Jessica Wardlaw, Robert Houghton, Steven Bamford, and Jeremy Morley. 2017. Task Workflow Design and its impact on performance and volunteers' subjective preference in Virtual Citizen Science. *International Journal of Human-Computer Studies* 104 (2017), 50–63.
[89] Ufuk Topcu, Nadya Bliss, Nancy Cooke, Missy Cummings, Ashley Llorens, Howard Shrobe, and Lenore Zuck. 2020. Assured autonomy: Path toward living with autonomous systems we can trust. *arXiv preprint arXiv:2010.14443* (2020).
[90] Achala Vagal, Shaun Wahab, Shari Lecky, Erica Washburn, Robert Schwartz, Craig Vogel, and Mary Mahoney. 2020. Optimizing patient experience using human-centered design. *Journal of the American College of Radiology* 17, 5 (2020), 668–672.
[91] April Yi Wang, Dakuo Wang, Jaimie Drozdal, Michael Muller, Soya Park, Justin D Weisz, Xuye Liu, Lingfei Wu, and Casey Dugan. 2021. Themisto: Towards automated documentation generation in computational notebooks. *arXiv e-prints* (2021), arXiv–2102.
[92] Yuelin Wang, Yihan Zhang, Yan Lu, and Xinran Yu. 2020. A Comparative Assessment of Credit Risk Model Based on Machine Learning——a case study of bank loan data. *Procedia Computer Science* 174 (2020), 141–149.
[93] Marco Willi, Ross T Pitman, Anabelle W Cardoso, Christina Locke, Alexandra Swanson, Amy Boyer, Marten Veldthuis, and Lucy Fortson. 2019. Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution* 10, 1 (2019), 80–91.
[94] Xingjiao Wu, Luwei Xiao, Yixuan Sun, Junhang Zhang, Tianlong Ma, and Liang He. 2021. A Survey of Human-in-the-loop for Machine Learning. *arXiv preprint arXiv:2108.00941* (2021).
[95] Haoqi Zhang, Eric Horvitz, and David C Parkes. 2013. Automated workflow synthesis. In *Twenty-Seventh AAAI Conference on Artificial Intelligence*.

A APPENDIX: HUMAN-AI CONFIGURATIONS THAT CAN BE ANALYZED IN THE CHA FRAMEWORK

Configurations in **bold** were analyzed in this paper's text.

- Dynamic shifts in initiative
  - **Huggle (Wikipedia bot)** (e.g., [31, 38, 39])
  - **Active learning / Human-in-the-Loop** (e.g., [5, 20, 70, 86, 94])
  - Generative documentation for source code (e.g., [91])
  - Fly-by-wire aircraft (e.g., [80])
  - Drone operations by government agencies (e.g., [12, 47, 89])
  - Corporate travel approval (e.g., [27])
  - Intelligent assistive technology for low-vision or no-vision travelers (e.g., [53, 54, 58])
- Multi-human applications
  - **Telephone operator workflows** (e.g., [35, 36, 48, 60, 71])
  - **Bank loan approvals** (e.g., [42, 52, 57, 92])
  - Insurance approvals (e.g., [1, 90])
  - Carceral applications (bail, sentencing, early-release (e.g., [13, 64])
  - Medical procedure approvals (e.g., [2, 49])