Deep Science Publishing
https://doi.org/10.70593/978-93-49910-41-6_10

Chapter 10: The human-artificial intelligence collaboration: Augmenting analysts with automation

10.1. Introduction

Over the past several years, there has been a great deal of attention paid to artificial intelligence (AI), specifically in the field of machine learning and its application to a variety of domains. The best-known applications of AI, from autonomous vehicles to facial recognition to natural language interfaces, are often built to replace human effort entirely — the fate of many innovations throughout human history. Increasingly, however, a great deal of AI development is focused on creating tools that are designed to aid and augment human decision-making. This intelligent augmentation represents a significant evolutionary advance in the way we think about the role of technology in our world, and we believe that professionals will benefit most from AI tools that seek to enable people to do their work better, not replace them. Intelligence analysts, in particular, are poised to benefit significantly from advances in AI. The combination of highly skilled humans and highly skilled machines has the potential to leverage the strengths of each while mitigating the weaknesses. Intelligence analysis requires combining knowledge from sources, evaluating their credibility and value, and integrating the results to make recommendations on a complex, adaptive problem in a dynamic environment (Deng & Liu, 2018; Zhang & Zeng, 2019).

Human intelligence analysts are experts in many interrelated skills that go beyond rote information processing, including critically interpreting possibly incomplete data that is often noisy due to exaggeration or deception, predicting the future, and being situationally aware of local and global conditions. Despite the variety of sources of information available, it is increasingly difficult for these analysts to keep up with the deluge of data available at scale. They can be overtasked and work to exhaustion or disengagement, handicapping their efforts to provide proper analysis. Lastly, advancing their existing workflows can be challenging due to crippling bureaucratic and cultural issues. With the advent of modern AI/ML technology, we can offer partial automation of some tasks that these analysts currently attempt, as well as augmenting their existing capabilities. There are also increasing opportunities to present data results in natural, concise, and intuitive ways that facilitate faster and more accurate decision-making, improved staffing, and greater collaboration between analysts. Additionally, new data sources and task flavors are receiving attention or are arriving from new technological advances, opening new opportunities for impact. With proper steps taken, these benefits can be achieved across the range of intelligence tasks, domains, sources, and conditions (Amershi et al., 2014; Brynjolfsson & McAfee, 2017; Chui & Manyika, 2018).

10.1.1. Overview of Human-AI Collaborations

This report reviews different ways in which human intelligence and artificial intelligence (AI) systems have been integrated to perform analytical tasks, with a focus on using AI systems to augment human intelligence in ways that take advantage of the naturally complementary human and machine capabilities. The review covers diverse collaborations used to address many different types of analytical tasks, but with special attention to analyzing textual materials since that domain has been a focus of much recent work. In addition to collecting and providing examples of numerous human-AI collaborations reported in a wide variety of applied research literature, we develop a task typology that distinguishes different ways in which AI and human capabilities ought to be usefully combined to address analytical problems that require different types of data-intensive reasoning. Defining and systematically investigating different types of human-AI collaborations can provide a foundation for more general insights about how to match mixed-initiative and autonomous AI designs with different types of tasks and problem contexts. The next section elaborates on some distinctive features of each component of the unique human+AI combination. The following section discusses some higher-level issues related to conducting practical work that seriously attends to the collaboration rather than focusing on isolating one performer. The subsequent section introduces our typology that distinguishes different ways human-AI collaboration might be used for different types of data-intensive reasoning.

10.2. The Evolution of Human-AI Collaboration

While much of today’s excitement about AI brings with it rhetoric focused on automation and job replacement, we want to close out this section with some reflections on a larger frame of human-AI collaboration. In many ways, the history of IoT and the history of AI have been different segments of the same larger story, which begins with the development of binary electronic computing in the 1940s. For the first thirty years of this story, what we now call computers were used almost exclusively to serve as faster and more reliable replacements for older analog and batch digital operations, primarily inventory management, accounting, and scientific administration. During this period, AI was seen as a subset of computing more broadly, where the goal was to bring unabridged intelligent reasoning, independent of human input and control, to computers. Today, many of these same AI and computing tools commonly replace huge swaths of human jobs, from financial analysts to tax advisors to insurance underwriters. But at the same time, we're beginning to see a much larger dialogue around collaboration: how AI can help humans, paired at different functional and cognitive scales, to help process data and generate knowledge.

This is a significant change in how AI is perceived and used. And it parallels a similarly significant change in how computing broadly is framed, away from a focus on automation and replacement and towards the idea of an open-source technology, focused on increasing the native capacity of humans to do ever more complex tasks. Ultimately, we believe that Augmented AI, grounded in universal serviceability and focused on the design and production of always-growing degrees of human-AI collaborative competence, will yield far more positive outcomes than we can possibly imagine. And it rests on a foundation of three essential points: meaningful human–machine feedback, transparent and interpretable tasks, and holistic, universal time support.

Fig 10 . 1 : Human-AI Collaboration

10.2.1. Historical Context

Traditionally, the fields of artificial intelligence (AI) and human-centered computing have evolved separately. They developed primarily along parallel but non-intersecting historical trajectories, with AI seeking to automate human cognitive tasks, while human-centered computing focuses on the invisible and often disposable technologies that surround us. However, rapid advances in AI research and emerging platforms that make cutting-edge AI accessible have brought significant changes to this historical status quo. Today, AI and humans operate in largely separate domains, communicating mostly through fixed interfaces, such as queue-like request-response interactions. We are on the cusp of a new era in human-AI interactions, as AI technologies start to focus not on replacing human tasks, but on complementing humans, augmenting rather than automating, taking a more indirect role in interacting with humans, adopting simple forms of low-bandwidth communication. There are several reasons why we need to change the nature of human-AI interaction, transform task automation into task augmentation, and enable humans to work closely with AI systems, learning from AI, but also customizing, instructing, and gaining confidence in AI. First, the complexity of domains we are asking AI systems to tackle has come to exceed the capabilities of either humans or AI-based systems in isolation, but the strengths of each are complementary and potentially mutually reinforcing. Second, while the historical context has mainly involved using AI to support mundane tasks, such as email sorting, search, and routing, unauthorized and often uninformed applications of AI to major, high-risk domains are now underway. Third, there is a rapidly increasing desire for transparency in AI reasoning and thus for AI systems to defer to human judgment in critical cases. Collaborative human-AI systems can in this way make AI systems more transparent and accountable, thereby enhancing confidence in AI and fostering broader acceptance of AI-based applications.

10.2.2. Current Trends

In addition to realizing the inherent potential of machine assistance, recent advances in AI have begun to manifest in products and applications that suggest capabilities that extend beyond the mere augmentation of users. Personalized recommendation has expanded beyond content to include action and process recommendations. In areas like drafting and writing, AI has graduated from assistance to co-authorship. More recently, tools and techniques aimed at surfacing the rationale behind AI's black-box decision-making bring users and AI to a footing where AI operates less as a tool or assistant and more as an analyst of its own.

Building on the observations, even supposing our interest in these phenomena is academic, we believe that this is a trend worth calling out and exploring further for two reasons: First, direction and intent matching aside, augmentation is only powerful in areas in which humans and AI present different skill sets. As Amdahl's Law reminds us, speedup from improvement of individual components diminishes when the components are partially effective and performance is bottlenecked by the relatively poor-performing components. In the context of human-AI interaction, working with an AI powerhouse in tasks where the AI shines and the human is fluent is key. Two lessons of historical perspective are that the very notion of efficient execution is broad enough to accommodate different forms of problem decomposition and parallelization.

10.3. Understanding Automation in Analysis

As we move from simple task automation to more sophisticated task assistance and finally to humans augmenting machines, the specific shape of the collaboration and the stages toward achieving it also change. At the simplest level are tasks that have relatively simple inputs, outputs, and rules of inference. Automation algorithms are often sufficient to achieve valuable results in some paltry, repetitive, and tedious tasks. However, given that human analysis often involves complex and undetermined reasoning under conditions of uncertainty, many analysis tasks are poorly suited for algorithmic automation. Errors at this stage usually require human intervention to fix, and errors can cascade if humans are not closely involved. The tools available at this stage usually require analysts to be adept at understanding exactly the capabilities and limitations of the automation algorithms they use.

Granularity of conceptualization is the second concept that informs the potential complementarity of human and automated analysis. Granularity refers to the level of abstraction or complexity of a particular dimension of the analysis. For example, asking a team of analysts to create a map of a particular piece of software might involve a plan to create a simple flowchart at the simplest level, but the analysts might need to delve into more specific issues involving individual functions or modules in the execution phase in order to get a feel for how the software actually behaves.

10.3.1. Types of Automation

Broadly speaking, cognitive automation can be divided into three types, each with a different purpose and focus. Process automation is designed to automate routine, predefined, and structured cognitive tasks of business processes across an organization. It enables businesses to optimize process efficiencies and implement best practices by freeing up workers from repetitive and administrative tasks to focus on higher-priority work. Program automation provides software developers with intelligent programming assistants that help them create entire new lines of knowledge work, implement complex knowledge or business rules, and integrate advanced technologies such as machine learning, natural language processing, and computer vision. Analyst automation is designed to enable and augment analysts with a new class of intelligent tools that help them rapidly make sense of large data and complex information, identify and solve complex business problems, and accelerate decision-making.

With the continuous advancements in AI technologies and our increasing knowledge of how to create AI systems, analysts of various disciplines from business intelligence and social media analytics to risk management and fraud detection, knowledge management, and cyber and intelligence analysis can be increasingly augmented with AI systems to tackle the complexity and dynamism in their knowledge work. The result is the creation of a class of intelligent analysts who are augmented by AI, enabling them to make sense of the diverse data available, accurately model complex issues, and rapidly solve complex problems in a way that simply was not possible before. For example, intelligent analysts can now answer 'why' and 'what-if' type of reasoning questions to derive deeper usable insights from the data, support more advanced hypothesis testing and experimentation, as well as make more powerful and predictive decision models, provide more accurate risk assessments and high-value investigations, and become significantly more effective for organizational decisions.

10.3.2. Benefits of Automation

An obvious benefit of automation is to enable analysts to handle more data, process a pipeline of images or signals in less time, create and explore more complex hypotheses, or handle more functionally complete tasks in the face of a clearly increased cognitive load. This was illustrated in an initial study, in which the automated approach discovered 30% more localization events with much less demand on the analyst's time. For security operations regarding long durations or extensive networks of surveillance video, the labor savings could be far larger. However, ten long-term pilots of automated systems for imagery analysis have recognized a far larger range of benefits that could be achieved in security operations by enhancing teaming, coverage and sensitivity, contextual understanding and self-modeling, anomaly detection and hypothesis discovery, assessment, response recommendation, and export order, solicitation and targeting.

By providing experts the opportunity to focus on tasks they are uniquely equipped to do amid observations especially suited to inputs received from sensors at the edges, the automated support increases the analyst's effectiveness and competence. These effects achieved by increasing the analysts' opportunities "to perceive, understand, and make informed evaluations and decisions" are categorically challenged by proposals to re-task their brains completely based shoulder to shoulder as if the prior process of active analysis could be wholesomely replaced by symbolic inferences that interpret actions into human language. The value of a deep understanding of a situation is evident as systems that have a darkly opaque and shallow presence in crowds whose activities and culture they are tasked to "disinfect" are often disproportionately forceful and destructive, particularly upon soluble linguistic interfaces extant in the professional conduct of their mission.

With respect to teams' core systems which are deliberately designed to be "team players who expect to share control with human teammates," they derive from investment in aspects of situational understanding, explanation, and adaptive reasoning. This is not a product of chance, as automation that is not sufficiently transparent and human-understandable will lead to coding problems it may experience and, more generally, may derail attempts to control its decision agenda, and force explanation or expect collaborative intelligence to little avail. These effects derive partly from the inherently ambiguous, constraining, variable, uncertain, and unexpected nature of the underlying situation.

10.4. The Role of Human Analysts

Our models have been developed with human analysts in mind. The aim of our work is to create a refinement cycle: model predictions are used to assist the analyst in making sense of large and complex data. The analyst provides feedback to better understand their changing needs, constraints, and abilities, and models are tailored and improved to support the next analysis. The human analyst drives interpretation and decision-making and is responsible for critical insight, creativity, and ethical considerations.

A complementarity of human and machine abilities also underlies our focus on high-level, rather than low-level cognitive functions. In early work, AI research focused on issues such as perception, planning, and natural language processing. These are important aspects of cognition, but they are defined by their biological instantiations rather than their functions. In recent years, a growing consensus has emerged from these specialist areas that a more general characterization of high-level cognition is that it is iterative, non-discriminative, bounded rationality, and optimize-and-learn. These are also characteristics of intelligence analysis, and it is these functions that our AI research seeks to support. At the high level, the issue is not whether AI can outperform human analysts, but rather whether it is possible to design systems with useful high-level capabilities that complement and augment the analyst.

10.4.1. Skills Required

During our interviews and over the course of our many years of collective experience creating and deploying analytical tools for the intelligence community, we have developed a sense of three distinct but interdependent strategic skills every analyst should possess:

—Analytical critical thinking: Piecing together evidence, determining its relevance, recognizing gaps, and jumping to logically sound conclusions in time-sensitive environments. Also: lateral thinking and creative thinking, which can aid in coming up with sound alternative viewpoints not yet represented by available evidence. —Research and analysis: Command of data with rigor gained through education and possibly also experience in data science. Data can take many forms, from numeric representations and unstructured text to geospatial or network graph data, to name just a few. —Communication: Contextualizing and succinctly describing findings in a way that decision-makers can easily understand. Also occasionally involved—storytelling to non-technical audiences to simplify complex problems and concepts that are not easily summarized mathematically or visually.

This presents a challenge. While it is one thing to show in a laboratory experiment that a model can perform certain quick-thinking critical thinking tasks accurately—such as image, text, or audio classification—real-world analytic work is difficult to validate in the short term. An analyst may not see products for years. To be of highest utility to the intelligence community, AI methods should support analysts in a job space beyond what nearly any individual is likely to encounter, deepening their training and experience. Start-up AI firms and other research-intensive organizations that can afford to utilize and manage such specialists are few, leaving the U.S. government in need of ready access to capabilities that it does not have the resources in-house to acquire.

10.4.2. Human Judgment in Analysis

For centuries, philosophers have long debated the power of human judgment as a unique feature that has eluded computational models. Computers can carry out complex and tedious calculations at incredible speed and can store vast amounts of knowledge. Yet human judgment, conceptualized as rational, while often imperfect, is what supposedly sets humans apart from machines. Human judgment is "that supreme effort of reason" that cannot be mechanized because it "involves a choice that is not the result of a computation or deduction made while seated at a desk." In the context of national security, it has been stated that "the real key to doing national intelligence isn't the information. It isn't the secrets. And it's not the camera clicks, either. Rather, it is the activity that comes between information and secrets, between secrets and wise policy - the act of decision making. That key is judgment."

It has been stated that the way to build AI is to devise means that would allow a machine to copy (or mimic) human judgment. It was discovered that judgment is highly susceptible to predictable biases and distortions, simply because people are forced to rely on heuristics, which are prone to error in complex environments. As a result, it is possible for computational models such as AI to augment or even surpass human judgment on a number of well-bounded tasks. Rapid advances in machine learning, the subfield of AI that involves the construction of general-purpose algorithms that can be trained on large datasets to carry out a variety of tasks, are now blurring the distinction between judgment and computation to the point where the debate is shifting from a philosophical to an empirical one. It is no longer a question of what judgment is, but how to reliably model judgment in a computational system. The issue is less about whether machines would exhibit human judgment and more about what roles judgment would play in the resulting human-machine system. This shift reflects a broad evolution in the approach to AI, away from questions of the ontology of thought and toward questions of the epistemology of computation.

10.5. AI Technologies Enhancing Analysis

AI, both machine learning and rule-based algorithms, is now applied throughout the enterprise in many ways, particularly in the fields of automation, employee decision-making, human-to-machine communication, analytics and intelligence, employee assistance and advice, and cyber technologies. The specific AI and automation involved fall into three broad technical capabilities: advanced algorithms, natural language processing, and emerging technology. The specific types of AI and machine learning technologies used for analytics are many, and each has a specific purpose related to how it deals with and supports different tasks in the analytics lifecycle. Some of the key AI/machine learning technologies include robotic process automation, cognitive computing, machine learning, deep learning, text analytics, natural language processing, computer vision, natural language generation, and machine learning operations. Each of these has augmentative or replacement potential and can be combined into the complete AI pipeline. Natural language processing and machine learning are the most important AI capabilities used as an interface to both the machine and the machine learning/AI pipeline of setting up and acting on the source data, training AI models, model lifecycle management, and AI outputs back to operations.

To resolve the design challenge of AI platforms involving many elements from creating an automation-supportive, augmented analyst work environment, we work from the specific steps of the AI pipeline that frame the work on the analysts' tasks during the complete lifecycle of analytics: data, problem definition, model definition, model training, model deployment, and model operation. In each of these AI pipeline steps, tasks and steps interact with analysts who depend on the quality of information learned and how it is used by all participants. AI completes a task or solves a problem for stakeholders at different analysis task steps, and each involves different AI techniques. Quality information has to flow smoothly between analysts, AI systems, and daily stakeholder dialogues. Three areas can be targeted to help analyst-AI teamwork and streamline communication to increase the efficiency of learning and business operations.

Fig 10 . 2 : AI Technologies Enhancing Analysis

10.5.1. Machine Learning Applications

Machine learning, especially when the problem ceases to have unambiguous right answers, is fundamentally about making better predictions given data. This can be in the form of identifying previously unknown patterns in data, labeling data with categories in an intelligent way, or even creating queries or thresholds to separate data in a useful way. There are three important failure modes of machine learning in AI-augmented analysis, of which two are unique to this field. First, machine learning may make a wrong prediction. This can lead to missed opportunities or to wasted time. Not only must we strive for accurate predictions from machine learning, but we must also find ways to interact with machine learning results to increase the confidence that we have in their accuracy.

Second, while many categories used in supervised learning come from human-created taxonomies, sought-after categories may themselves have varied or axiomatic definitions. This makes it difficult to gather data and evaluate learning algorithms. Finally, for many organizations seeking to operationalize an AI-augmented analysis solution, infrequent wrong predictions can be significantly more damaging than a model that is somewhat useful but not amazing the rest of the time. We want a model that explains what it is basing the prediction or recommendation on so that, in hard cases or when significant risks are possible, we understand its provenance. And we need explicit confidence measures or multiple de-conflicting models that we can rely on to know when to disregard a prediction.

10.5.2. Natural Language Processing

On relatively short text inputs, unsupervised discovery of patterns in language, rather than the extraction of information from designated, pre-tagged parts of the text, may be more useful. Particularly in cases where some structure is available about the rough categories that might be useful to look into, some form of unsupervised clustering of words or phrases can be used as an initial step to identification. Furthermore, in text synthesis, non-traditional approaches may be useful. To ensure the generated text is coherent and makes sense, however, significant recombination and selection based on a final generative model is typically needed.

This is a fundamental research issue, and there have been several interesting approaches to it. However, there are limited open tools that are widely applicable for non-researchers. If non-researchers want to work with NLP, they typically need to learn toolkits that expose linguistic primitives such as treebanks, parts-of-speech tagging, and named entity recognition. They also need to learn some intricacies about language itself. While it is useful that some researchers and practitioners are learning NLP methods, for wider demographic impact, it might be useful to spend more time making sure NLP tools are successfully transformed into sufficiently usable and customizable components. At a minimum, creating user-friendly and customizable wrappers around existing research implementations might make high-quality natural language interfaces more accessible.

10.5.3. Data Visualization Tools

Data visualization is one of the oldest and most powerful tools for human analysis. By joining the data to the perceptual and cognitive systems of the brain, visualizations allow us to detect patterns, spot anomalies, and identify trends in vast amounts of information. Thus, data visualization is a cornerstone of almost every form of data analysis. Modern data visualization systems support a wide variety of interaction techniques that allow analysts to adjust ranges, compare parts, and change perspectives, but several challenging problems remain. Here are some examples of problems that are the target of ongoing research. Today's exploration tools allow flexible tracking of many dimensions of information. They also allow the use of guidelines and other auxiliary displays that provide context for the data. However, there is little automatic assistance or automation in helping analysts make sense of the data and map it to the visual dimensions; nor are the datasets associated with the visualization endowed with any knowledge that might be used to provide richer, domain-specific visual presentations. The goal of this research is to explore the science of augmenting data visualizations with methods from statistical machine learning to facilitate effective data analysis.

We have examples of data exploration tools that automatically find and visualize interesting properties of the data. But these systems generally serve only one exploration context, are not flexible and general-purpose, and cannot use information that resides in external databases. Anomaly detection is really only a sub-area of an overall problem with the interpretability of model outputs, which is a challenge. It would be very useful to extend these ideas to facilitate exploration and sense-making of intermediate results during more complex data analysis tasks. In this research, we are exploring approaches in which we build and explore variants where statistical descendant modules can guide the model toward learning representations that have useful invariants. We wish to identify structural properties of the data when visualized that need to be enforced in the model learning process in order to make the learned features more aligned to human perception and understanding. The goal of this research is to allow us to use advanced data visualizations to learn representations of the data that can be easily understood and interpreted by analysts.

10.6. Challenges in Human-AI Collaboration

The design of a collaborative partnership between an analyst and an AI should simultaneously tackle several key challenges. First, the AI must be able to accurately capture the state of incomplete knowledge and uncertainty of an active analyst and be able to understand the feedback from the analyst that could change this state. The feedback may come in the form of a correction, clarification, or refutation of an AI-generated hypothesis or a question seeking more information on the AI's rationales. The AI should be able to accurately model the analyst's intent behind the feedback: is the analyst attempting to critically challenge the AI's finding, is the analyst refining model output as part of an ongoing exploration process, or does the analyst not have enough information about a particular topic and is asking the system to provide support? The feedback during model validation, embedded as it is in the feedback loops during reasoning and explanation generation, will help the AI understand the evolving understanding of the human analyst and the context in which feedback is intended.

Second, an operating AI should be able to assist the human in managing the collaboration. In addition to providing the requested support and explanation, the AI should monitor the processes that are enabled by its explanations and detect opportunities to intervene and improve the processes with efficient and effective automation. For such monitoring and intervention, the AI should understand the process steps that are being performed and be able to reason about failure modes and support the end-to-end process that includes human and autonomous contributions. Moreover, the

AI should recommend tasks that help refine its models and their relevance and suggest how to best use limited analyst resources to answer the more critical information demands.

10.6.1. Trust and Reliability

Generally, trust has been linked to an agent's capabilities, integrity, and consistency. Regarding the capabilities of a system, it might exhibit only a single competence that the observer trusts or a set of competencies that align with the observer's understanding of the trustworthiness of the system in relation to the competencies. Similarly, in human-human interaction, a helper is trusted more if they keep up-to-date records and if their opinions deviate less from social norms, values, and ethical principles. The reliability and consistency of an agent are important trust factors. A more advanced understanding of the trustworthiness of decision-making agents distinguishes trust in competence from trust in willingness. Trust in competence is the ability to execute as expected and make technically sound choices. Trust in willingness is the perception that the decision maker operates with the intention to act in accordance with the interests of the observer, i.e., according to an ethical, moral, reasonable, societal, or institutional norm. It is a necessary trust component that affects decisions when the system's in-action outcomes are unknown, uncertain, or probabilistic, and it becomes critical when the expectations of the observer cannot be verified.

Following the AI community definitions, competence in AI relies on cognitive faculties like perceiving, reasoning, solving, and learning. In AI, competence is usually defined by raw measures of performance, capabilities, and confidence, which are used every time we speak bluntly about the reliability of machine learners. The notion of trust in willingness often goes beyond the technical system's properties and is set at a higher-level observation or sociotechnical operation level.

10.6.2. Ethical Considerations

While analysts are increasingly excited about the potential utility of human-AI collaboration, they can also face ethical and normative puzzles. It is well known that employing machine learning can produce models with biases that are inherited from the training data. When a human works with a biased model, it can reinforce these biases, leading to more biased conclusions. Beyond simply using machine learning models with known biases, employing streaming or more experimental sources of data without understanding all of the latent biases or behavioral features of people behind these data collection processes can reinforce biases and increase privacy concerns. As with the introduction of any new technology, we therefore caution users to keep these potential negative consequences in mind when considering using supervised learning.

We recognize that although significant benefits can come from using machine learning tools to translate raw data into actionable insights, this must be done with care and attention to the potential negative downstream effects of turning to these tools to drive decision-making processes. Our hope is that analysts become more aware of these potential drawbacks and use that awareness to guide their collaborations with the AI systems we developed. Indeed, this is the main goal of the collaborative system. Despite our developmental and prototyping efforts, however, our tools are by no means perfect, and we expect the path toward making them fail-safe to be a long-term and ongoing endeavor.

10.6.3. Job Displacement Concerns

We agree with those concerns. Particularly high on the list of types of tasks where we believe AI will not replace human analyst creativity in the foreseeable future are innovation and creating genuinely novel large-scale reasons. Such reasoning is difficult because it can lead to the unexpected by definition, implying that it is much harder to identify a reproducible process. As a result, the task of automating novelty in a way that will be accepted and comprehended by the user is technically more difficult than automating routine tasks. Moreover, there are some analytical problems that only humans are interested in solving. As above, augmenting human analysts could allow a single human analyst to take care of problems with a high novelty component—where historical data, from which a model could be trained using known labels, is absent or at best very sparse—but this approach merely represents a different problem, rather than a solution.

Our result with the system addressed a particular class of prediction problems with a human-in-the-loop approach. According to the law given by Listfield, the answer to the question, what happens if an AI technology is used in the real world and something bad happens, is: 'Benefit: automaticity. Cost: liability.' This implies that in the near future the main advantage of HFAI over AI techniques that aim to replace human creativity will come from the reduced liability of the human. Even if the system performs only as well as an AI-only system, experts who agree to have HFAC on the diagnostic team will not be held to fault to the same extent as before. Indeed, by providing intermediate judgments that can be assessed and bound by humans, these systems can also help limit the scope of these technologies more critically and define the limits of what could or should be done in a given task.

10.7. Case Studies of Successful Collaboration

We still lack evidence that the analyst-AI partnership is viable or valuable in operational settings. Most AI systems built in the last decade have not prioritized training analysts or being designed for collaborative use and interoperability with existing social structures or tools. As a result, attempts to operationalize AI have required large-scale reorganization, subject matter experts to place trust in opaque systems, or analysts to perform significantly different work than they had prior to the AI systems' integration. Before demonstrating benefits for some complex tasks, AI-based systems for complex analysis in operations have often been insufficiently effective, and timely use of AI has lagged behind what we could achieve by promoting better analyst-AI collaborations. On the other hand, structured approaches to augmenting analysts have been demonstrated to work across a spectrum of strategies. Our main role as AI researchers is to begin with the analysts we hope to support and the tasks in which these analysts can benefit from assistance. In the following chapter, we describe four use cases of our structured approach in sections drawn from activities with our societal sponsor. These examples collectively demonstrate work toward trusted AI in operations, and an applied strategy to contribute to AI models, tools, and processes that are developed in collaboration with the practitioners, analysts, and organizations that require them. AI exists for analysts who need it, when they need it, and in a form they can use.

10.7.1. Finance Sector

Robo-advisors are beginning to appear in banking and other financial institutions. Robo-advisors are sophisticated software programs that combine mathematical rules and historical market data to select securities, generate trading orders, and achieve other financial objectives associated with managing clients’ investments. Like other AI products and services, some robo-advisors are tailored primarily for analysts and portfolio managers, while others are intended more for final wealth management by customers. As with other AI programs and services in the finance sector, robo-advisors are new and not yet widely deployed. They are also likely to bump up against the regulatory preference for inclusion of direct human intuition and judgment in certain aspects of investment management.

The financial models used by investment banking technology firms are increasingly facilitating investment decisions and augmenting other skills associated with investment banking, such as predicting mergers, acquisitions, and spinoffs. The predictive abilities of these AI products appear to relate to time-sensitive big-data events; forecasts should be made based on open market policies. Sports, energy, and health indicators are also built into the predictive models. Business intelligence might be faster, better, and cheaper with artificial intelligence, but it also seems to be somewhat different. Forty-eight percent of fund managers think there will be a lot more man-computer partnerships in finance in the future, while seventy-nine percent of the fund managers think AI is just another quantitative tool.

10.7.2. Healthcare Sector

Overall, healthcare requests tend to be simpler tasks compared to finance use cases. We observe that the types of use cases the healthcare sector requests include requests surrounding five key areas: patient care, diseases, biological concepts, medicines, and medical procedures. Healthcare requests may come from both individuals with health questions and medical professionals, and because they have different requirements, the model working on healthcare requests needs to handle the different questions effectively. For instance, demands for blood glucose levels generally look for a numeric outcome for a patient to track their health. Answering this demand accurately may be lifesaving for the patient.

The model supporting healthcare is designed to consider whether an answer resolves the asker's information needs, and whether the answer is phrased with an appropriate balance of precision versus safety and empathy, and adheres to the core healthcare principles. These healthcare principles consider data privacy, dignity, and respect for both the requestor and theoretical entities represented in the data that the model may encounter through its AI-driven information recommendation. The visualizations provide insight into how the healthcare system responds to stressful or potentially life-threatening information requests, which they take great care in addressing through both model design and internal company policies.

10.7.3. Marketing Sector

The marketing sector, where analysts employ a combination of qualitative and quantitative data to understand consumer behavior, make marketing campaigns, and calculate return on investment, is ideal for leveraging AI capabilities for the exponential data growth and automated data analysis capabilities. With the knowledge process between a domain expert and an AI system less clear than in other sectors, companies are less informed about how to reap the benefits from using AI for marketing. We visit three projects in the marketing sector where companies use machine learning for analyzing clickstream data to help media and e-commerce companies understand customer behavior. The objective is to infer changes in preference states and collectively measure media experiences and behaviors to improve customer experience. Neural networks-based modeling offers new opportunities to advance our understanding and characterization of very rich clickstream data.

Two projects aimed to optimize marketing campaigns for e-commerce companies by targeting individuals with the highest probability of making a purchase with the lowest expense through machine learning-based personalization. This paper studies the specification and estimation of such flexible empirical models of individual choice, with the application of a classic use case in web surfing for internet retailers. Dynamic panel data models of purchase behavior or other economic decisions have been used in a wide variety of economic applications. Marketing is concerned with the interest in individual demographic, economic, and experiential features in predicting both web surfing behavior and purchase behavior. A considerable body of research is concerned with model parameters that vary across individuals and transactional data. In this work, analysts are interested in examining the role that machine learning plays in ensuring accurate individual predictions.

10.8. Future Directions in Human-AI Collaboration

In this article, we advanced the discussion on title and abstract: Dramatic improvements in and adoption of AI technology can be facilitated by making analysts the prominent cognitive center in intelligence work and leveraging automation to help with time-consuming, error-prone, and low-level work. To this end, we formulated the types of augmentation: an AI sitting at the helpdesk and an AI standing beside the wall, as well as the settings and implications of the human-AI collaboration in intelligence work. In formulating this framework, we discussed the roles of AI and humans, the interaction flows across these roles, the computing and human science challenges in achieving intelligence augmentation, and the ways for human-AI collaboration to unfold in reality. Finally, we provided three burning research directions in intelligence augmentation based on factors within or around organizations and developed an illustrative scenario about different roles and interaction flows across humans and AI in the intelligence profession.

In the future, we endeavor to address the following burning research questions: We hope that this article can inspire much interest in these compelling research directions in human-AI collaboration and advance the realization of intelligence augmentation so as to confidently make the best use of valuable analysts in intelligence work while accommodating the drawbacks of technology. We believe that elevating the focus of AI in the value chain of the knowledge production process is key to fostering a mutual reliance that will make analysts at the center and AI at the edge a feasible and successful arrangement for the future and that situating the relationship within a workflow informs the role of automation and the form that collaboration should take.

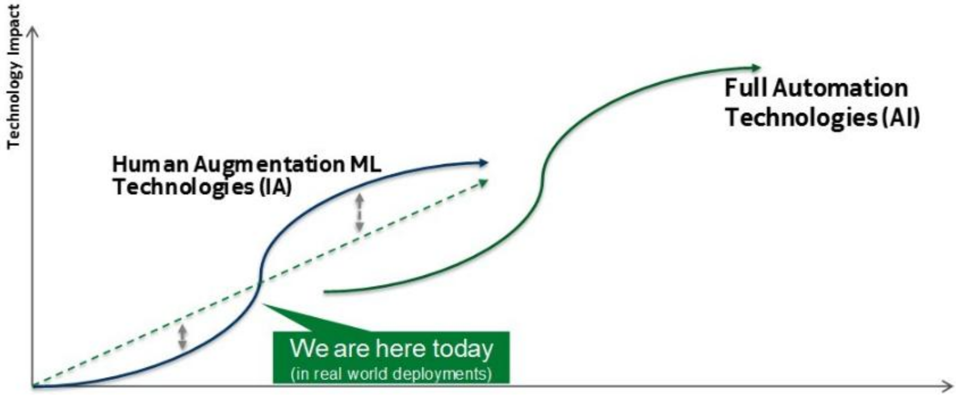

Fig 10 . 3 : Artificial Intelligence — Human Augmentation

10.8.1. Emerging Technologies

Emerging technologies have the potential to revolutionize the impact of AI-enabled tools, which augment analysts and intelligence personnel by increasing the speed, accuracy, and effectiveness of their work. For example, a brain-machine interface is being developed that will allow individuals to control computers using electrical activity in the brain. Virtual reality (VR) and augmented reality (AR) would be implemented into many more systems and products. The Internet of Things (IoT) has the potential to generate substantial data that can be analyzed to gather insights and provide recommendations for numerous applications. It can also be used to collect structured and unstructured data that can be used to train AI models.

Analysts and intelligence personnel are individuals who could benefit from the assistance of AI to synthesize and automate routine aspects of their work. They are some of the most prominent groups who stand to benefit from automation in the near future. Consequently, tools that effectively act as aides to these individuals and allow them to do their work, including the potential to perform novel functions that they were otherwise unable to perform, stand to benefit these individuals significantly. These tools, however, can require a large amount of finesse from developers to produce effectively useful results, otherwise known as an AI-hard problem. Thus, analysts and intelligence personnel are a 'gold standard' for AI automation, because if this effort is successful in helping this 'most difficult' user to automate, then these similar systems have a higher likelihood of generalizing to other less sophisticated users.

10.8.2. Long-term Implications

The combination of human analysts and their AI helpers will also have longer-term consequences. The first is in the development of a new class of analysts. In theory, contentious analysts should actively resist automation, as it replaces some of the loosely defined functions and expectations of their work. In practice, some may embrace this new way of working, actively deploying and orchestrating a wide range of tools, agents, and devices to cooperatively produce intelligence that neither will be able to easily achieve on their own. To the extent that they do, they will become increasingly capable of rapidly detecting and addressing the problems associated with a rapidly changing intelligence environment. As a result, they will have less need to coordinate a large team of analysts and can focus on developing the unique and valuable skills associated with expert human intelligence collection.

The involvement of AI can also increase the accuracy, validity, and credibility of intelligence reports, improving the reputation of those analysts who are capable of effectively utilizing it. Together, we suspect that these superficial similarities conceal the fact that there are several longer-term implications associated with combining humans and AIs, all of which should increase the demand for good human analysts within the field of intelligence. First and foremost, analysts will have to rapidly develop their ability to effectively, responsibly, and creatively wield a wide range of technologies, making individuals who have an intuitive or educated understanding of these systems even more valuable. This will require them to learn more about the strengths and limitations of these technologies.

10.9. Conclusion

Contemporary algorithmic technology provides fundamentally new opportunities to introduce automation into the information analysis process at scale. Such automation should not, however, be viewed as a "black box" that has complete decision-making authority, displacing human judgment. Instead, automation and algorithmic systems should be leveraged to amplify human judgment in a wide range of tasks. Throughout this work, we outlined several specific human-AI collaborations, as well as other ways in which AI can improve judgment. In doing so, we sought to open up some theoretical questions in understanding how increased levels of automation might affect human performance in judgment and decision problems.

However, the future is opaque, and understanding the effect of automation on human judgment is nonetheless complex, posing significant disciplinary, empirical, and theoretical questions. Some scholars argue for the preservation of human judgment for the intrinsic value of exercising judgment, while others advocate for automation to mitigate systematic errors in assessing complex judgments. We provided signposting milestones along the way and propose a possible road forward for conducting research into these important questions.

10.9.1. Final Thoughts and Reflections on Human-AI Synergy

Recent advances in AI capabilities have generated significant excitement and progress in their application. A growing theme is that AI tools are becoming increasingly useful to enable people in completing a wide range of collaborative tasks. In general, the most effective teams pair humans and AI in a complementary fashion rather than fully automating tasks. In the best of cases, the human-AI synergy can prove to be more accurate, more informed, and, overall, wiser than humans or AI on their own. This chapter focuses on three high-level perspectives on how humans and AI interact in complex tasks.

First, there is a need for better alignment between the psychological, social, and ethical considerations raised by AI and analysts' work practices and contexts. This is important for application areas including automated journalism, expert-sensitive settings in intelligence analysis, and many other areas across AI. We proposed a three-lens approach to think about these considerations. We suggested that AI applications should consider the grand sweep of AI technologies with three basic strategies: AI can serve as a collaborator for humans, helping them achieve more with less; AI offers a set of perceptions and interventions; and AI's collective decision-making tools that support humans care for each other. AI supported by qualitative human analyses can effectively mandate well-being, help people make evidence-based decisions, and judge the ethics of AI systems. Our work towards AI systems emphasizes the importance of continually improving the influence of AI so that AI augments human expertise, improves human understanding of the impacts of AI, complements human work, and designs systems with the power of collective human decision-making.

References

Brynjolfsson, E., & McAfee, A. (2017). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.

Deng, L., & Liu, Y. (2018). Deep Learning in Natural Language Processing: A Survey. IEEE Transactions on Neural Networks and Learning Systems, 29(4), 1–29. https://doi.org/10.1109/TNNLS.2017.2765317

Chui, M., & Manyika, J. (2018). The Potential for Artificial Intelligence to Augment Human Decision Making in the Workplace. McKinsey Quarterly, 1–10. https://doi.org/10.2139/ssrn.3162852

Amershi, S., Weld, D. S., Horvitz, E., & Kamar, E. (2014). Crowdsourcing, Voice, and Human–Computer Collaboration in Machine Learning. ACM Transactions on Computer-Human Interaction, 21(3), 1–35. https://doi.org/10.1145/2641582

Zhang, L., & Zeng, D. (2019). Exploring Human-AI Collaboration: Challenges and Opportunities. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3293663.3293683