CrowdSmart

Integrating human co-creation and generative AI (ChatGPT)

Building trust and learning relevance

Relevance of ChatGPT results for co-creation and innovation

We are at the dawn of a new era with Generative AI yielding applications like ChatGPT, Mid-Journey, and many more to follow. The potential impact is immense but there is also significant danger. Public discourse, already reeling from the political impact of fake news, now faces a new challenge: Do large language models create an even more significant threat to the amplification of misinformation threatening all our institutions? How do we know what is true?

As a pioneer in AI[1], and an early adopter of pre-trained transformer models[2], I will share a path forward that is currently delivering practical applications of integrated human collective intelligence and artificial intelligence agents.

Text presented as knowledge requires vetting to be trusted. Large language models promise massive value by putting vast collections of knowledge at our fingertips. This article makes the case that to trust and apply it, we need an additional form of AI based on human curation and contextualization. Let's ask ChatGPT.

Question to ChatGPT: Do Large Language models provide a way to validate the source of knowledge?

ChatGPT: Large language models by themselves do not provide a way to validate the source of knowledge. These models are trained on vast amounts of text data and can generate text that is similar to the text they were trained on, but they do not have the ability to verify the accuracy or source of the information they provide. It's important for people to use critical thinking skills and verify information from multiple sources in order to determine its accuracy and credibility.

Question to ChatGPT: Is there an AI technology that helps humans apply critical thinking to large language model results?

ChatGPT: There is currently no AI technology that specifically helps humans apply critical thinking to large language model results.

While large language models create excitement about their generative potential, we must be extremely cautious about the value of what they generate. Data lineage is foundational to trust in information systems. These models generate from fragments of text without reference to data lineage. Anecdotes are exciting. That is not the same as data provenance.[2]

"L*arge language models drop the provenance of fragments used to generate outputs. By definition, it cannot produce scientific knowledge"*

John Seely Brown[3]

In the remainder of this paper, we will explore how an AI facilitator integrates knowledge models of humans, and intelligent agents (in this case ChatGPT). We will demonstrate a way to productively engage ChatGPT in collaborative co-creation. Finally, we will explore how curation with collective human intelligence creates trusted, contextualized AI-assisted co-creation.

The collective intelligence of humans and machines

In parallel to the development of large language models, a third wave of AI focuses on integrating the collective intelligence of humans and machines with roots to how humans build trusted knowledge: the scientific method.

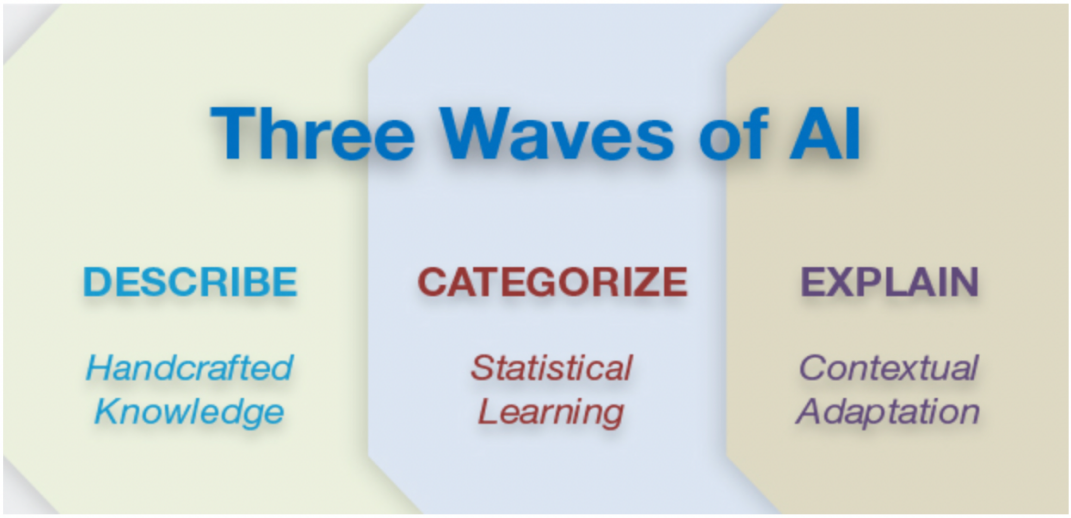
Collective intelligence is believed to underly the remarkable success of human society.[4]

Engaging human collective intelligence is fundamental to building trust in knowledge. The Enlightenment established the scientific method and logic as the process for knowledge aggregation and defending human rights against tyranny. It underpins critical thinking principles, including validating evidence, logical reasoning, and deliberation. In the Structure of Scientific Revolutions, Thomas Kuhn states: *"Revolutions should be described not in terms of group experience but in terms of the varied experiences of individual group members. Indeed, that variety itself turns out to play an essential role in the evolution of scientific knowledge."*[5] Debate, deliberation, and discussions guide the adoption of scientific knowledge.

Constructive deliberations are based on evidence and cite the reason for a prediction or belief about an outcome. In Thomas Bayes's essay on the "Doctrine of Chances," he shows how evidence strengthens or challenges our confidence in what we know. If new evidence affirms our beliefs about an outcome, our confidence in knowledge increases. We challenge the evidence or change our thoughts if it challenges our assumptions.

The first generation of AI captured the experience and evidence of expert knowledge in hand-crafted systems (expert systems).[6] The technologies used, logic and frame-based knowledge representations systems, explained how the system came to an answer or prediction. Explanations revealed the facts and assumptions underlying reasoning. Truth maintenance was a thing! When assumptions or underlying facts changed, the system logic reflected that in explanations. Systems could create multiple world views allowing exploration of multiple reasoning paths. Model-based reasoning demonstrated it was possible to build effective systems in generating alternative outcomes based on alternative models.[7] Model-based thinking is important in linking the science of collective intelligence to next-generation AI.[8] Scott Page's book "The Model Thinker" explores the power of multiple-model thinking in prediction accuracy and decision-making.[9]

Current AI technologies learn models by correlating patterns learned from historical data. BlackBox AI, often called, learns statistical patterns from data and does not lend itself to explanation. Blackbox AI delivers answers without explanation, as ChatGPT so eloquently explained in the opening paragraphs of this article.[10]



A DARPA Perspective on Artificial Intelligence (click for presentation)

A new generation of AI, the third generation, brings humans, context, and explanation into the picture, integrating and extending aspects of the first two generations. The third generation creates the path to trusted knowledge through explanation and context. Explanatory AI can answer questions like: "How did you come to that conclusion?" or "Why did you give that answer to the question?". Judea Pearl makes the case in his book "The Book of Why" that true intelligence explores and imagines based on causal models.[11] We don't know how to build AI systems that automatically embed the kind of knowledge that 4-year-old humans do—explore the implications of "what if"—exploring counterfactuals. Counterfactuals explore future possibilities, like what might happen if kangaroos had no tails?

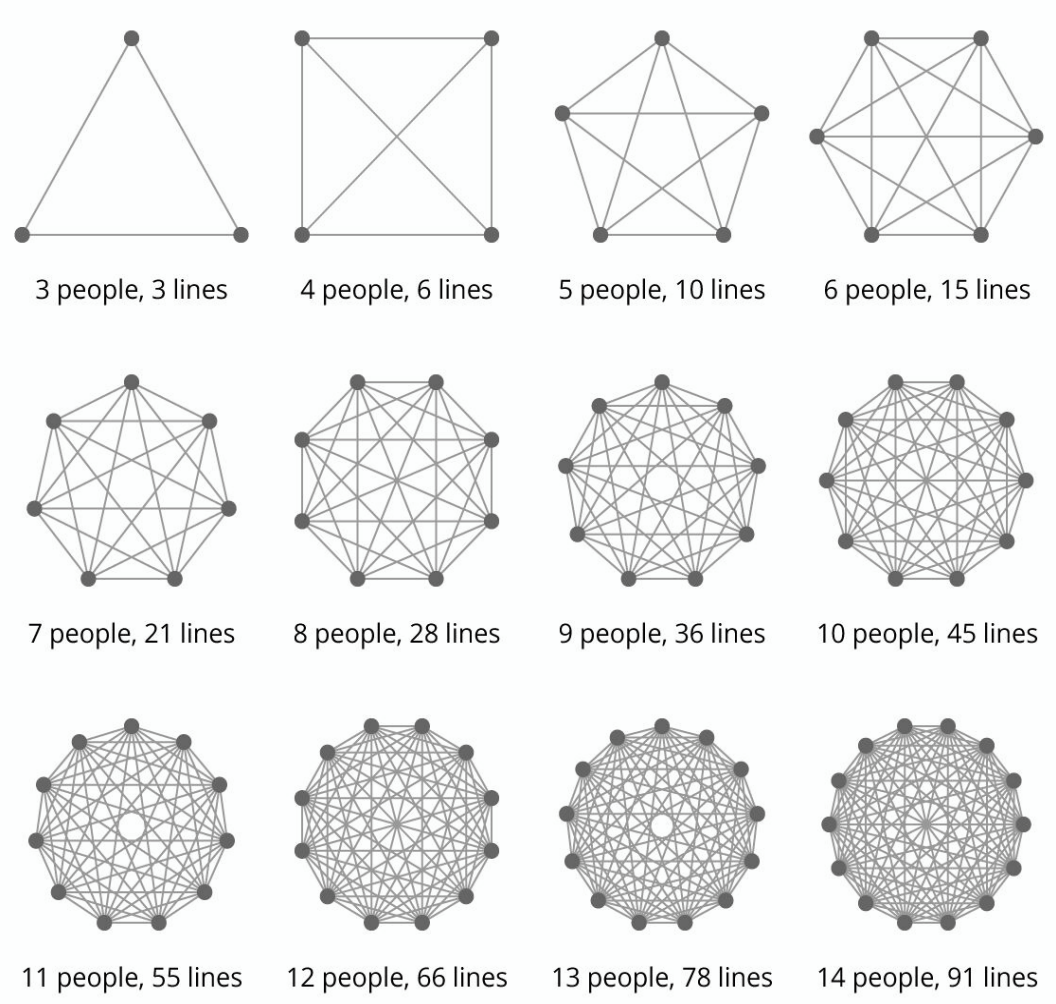
The task is to design an AI that explores the power of why. Imagine an AI that brings humans and intelligence agents (AIs) together to *explore the power of counterfactuals in generating alternative future forecasts.* Causal models are essential for any task where we use AI systems in decision-making, forecasting, or co-creation. The fastest path to that goal requires actively capturing collective human reasoning in co-creation and problem-solving tasks.

The science of collective intelligence demonstrates that a cognitively diverse group of informed participants outperform individual experts in prediction tasks. Many readers are likely familiar with the works of Phillip Tetlock[12], Scott Page[9], and Tom Malone[13]. Their published work establishes the power of collective intelligence in forecasting. Prior work in collective intelligence did not focus on capturing the collective reasoning process driving the prediction.

The collective reasoning process of humans engaged in co-creation tasks is foundational to a trusted AI. We can build a knowledge acquisition system from this foundation that

integrates collective human intelligence with artificial intelligence systems. We will show this integration leads to a system that can unlock the generative potential of humans and machines.

Capturing how our opinions and ideas align in an outcome-driven discussion is not an easy problem. We all have experienced the difficult problem of group brainstorming and deliberation. For that reason, most avoid the problem by keeping deep deliberation sessions to small groups (e.g., seven or fewer). Balancing between idea exploration and prioritization is hard, and the complexity grows with the size and diversity of the group. The relationship between lines of communication and complexity is illustrated in the figure below:

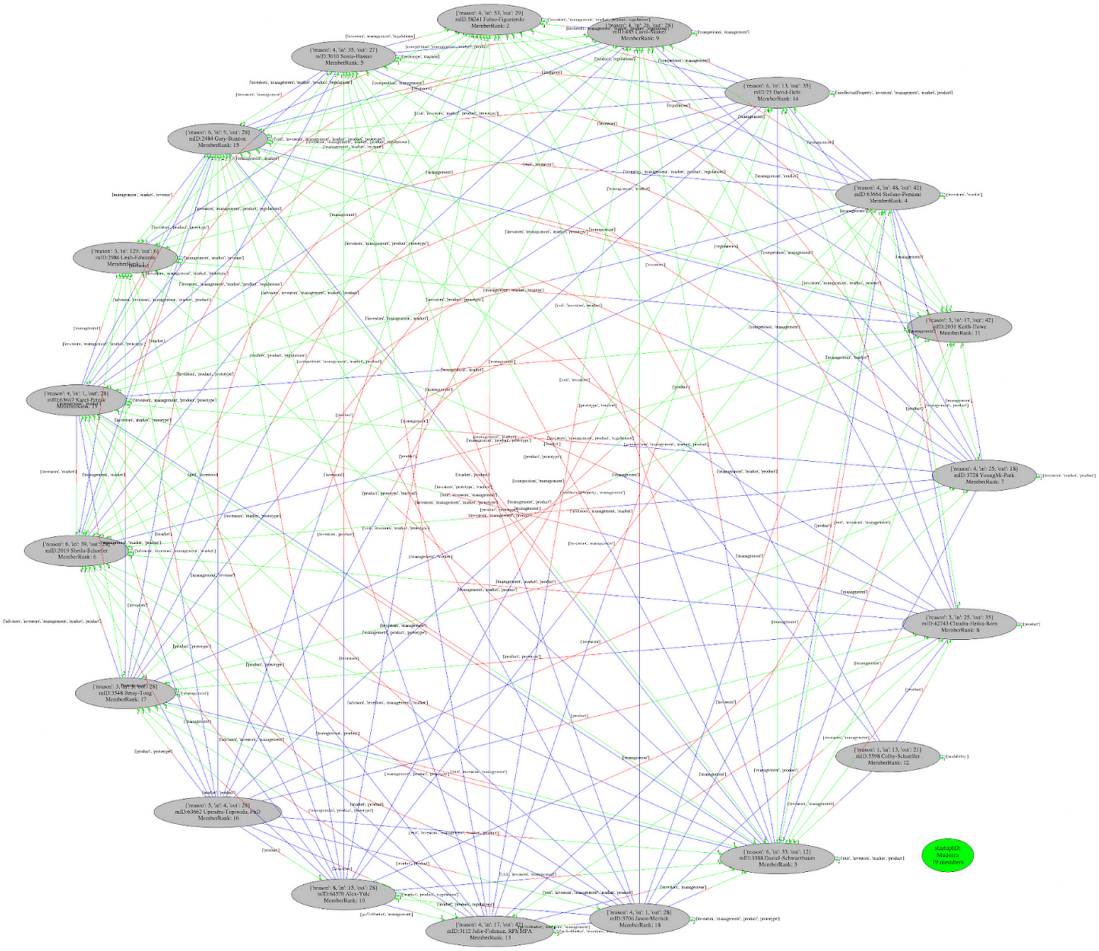


The complexity of the problem of balancing new id grows exponentially with the size and engagement level of the group, making it a problem for AI.

As it turns out, the adaptive learning methods used to bring order to the web provide a foundation for creating a system that enables the third generation of AI based on explanation and context. Research in modeling the scientific citation process underpins Google's foundational PageRank algorithm. For mathematicians and computer scientists readers, solving the problem of ordering information on the web became an eigenvalue problem.[14]

Citing the source of information is fundamental to scientific publications, scholarship, and critical thinking. Coupling AI modeling with citation analysis at the individual level provides the key to a system for human knowledge acquisition at scale.

The figure below shows a group of people (nodes in the network) sharing knowledge. (I will show later how we construct this web.) As mentioned, the nodes are people, and the links represent an individual citing another individual's reasons for their score in predicting an outcome of a decision. For example, if the discussion was "How likely is [specific innovative idea] going to succeed in the market?". Collective reasoning learns the reasons for scores—the knowledge sharing that creates the web. The labeling of the links denotes the content of what was shared.



On a plane trip back from Tokyo some years ago, I was playing with this structure and noticed that it could be turned into the same problem used in citation analysis and PageRank. It was an eigenvalue problem, and the solution gave a ranking of knowledge influence for that specific group of collaborators in a particular task. We called the result InfluenceRank. For any group of humans, no matter how large, engaged in a focused knowledge-sharing task, the peer-reviewed InfluenceRank becomes a mathematical calculation. InfluenceRank learns by a single-blind process who are the thought leaders in an outcome-focused process (e.g. should we invest in a specific innovative product). In a similar process, RelevanceRank learns what are the ideas and themes that represent the collective mind of the group.

RelevanceRank and InfluenceRank bring a scalable mechanism for the common experience of a group "resonating" with an idea. RelevanceRank represents the prioritized statements the group resonated with within a collective reasoning engagement. In physics, the eigenvalues of a vibrational system depict the resonance modes, as in striking a tuning fork. In collective reasoning, RelevanceRank identifies the group's resonance modes—*it learns what ideas resonate with the participating group, no matter how large.*

Collective reasoning, the name given to the process, delivers a scalable means to build group alignment on why the contributors believe a project will succeed or fail. The cognitive diversity and expertise of the group determine the quality of knowledge generated.

The mechanism for building the web of knowledge is easy to use. It requires simple well-formed questions. Questions are the drivers of decision deliberations, predictions, and knowledge-building exercises. "Do you believe [new product] will fit market needs and grow organically?" The prediction accuracy measures the collective intelligence of the participating group.

Predictive assessments with Likert or quantitative scales, as demonstrated in the question: "How likely is [some innovative product] to succeed in the market?" are the framework of predictive models. For simplicity, assume a scale of 1 to 10.

Collective reasoning applies design thinking methodologies to understanding the why behind the accuracy in the predictive score and, as such, may be an amplifier to the predictive accuracy of the group. When people deliberate and brainstorm, they often generate new ideas. Collective reasoning enlivens that process. In collective reasoning, we focus on reasons for scores. Score values are not shared. In addition, the *identity of the knowledge source is masked in the deliberation process, echoing the practice in the scientific community of double-blind reviews.*

Identity masking reduces bias and opens the door to creating new areas of alignment—finding common ground for groups of any size.

The AI, in this case, is a super facilitator. We call it a Collective Reasoning Facilitator (CRF) because it facilitates a collective deliberation and reasoning process necessary for group alignment on a decision or prediction. It does this through intelligent sampling that creates a short, readable list of 7 items. It is intelligent because it adapts its sampling strategy based on the responses and interactions of the group. The sampling algorithm balances discovery and prioritization by paying attention to emerging results. The list is curated for each participant and evolves as the group deliberates. The system uses large language models to embed the collective reasoning process. The system creates a language model of the collective mind. It uses the model of the collective mind to sample statements that stimulate divergent thinking.

The participants respond to a sample list of their peer's reasoning (identity masked). *A "single-blind" process reduces bias by focusing on the ideas and reasons rather than the position or personal identity of the contributor.* As mentioned above, the sample list is uniquely constructed for each participant based on their reason and the current state of the "collective mind." They prioritize the items they believe are most relevant to the deliberation process. If they don't see anything that resonates, they can request a new list. If they get a serendipitous moment seeing others' reasoning, they can enter a new idea at the moment of inspiration. At each step, the system updates the model and uses the updated model to create a new sample list.

Intelligent agents, like ChatGPT, function simply as *one of contributors to the co-creation process. Misinformation is unlikely to survive a single-blind process if the number and diversity of contributors are managed properly.* Since it is impossible for the contributors to know if the idea came from a human or AI contributor, the system provides a kind of "Turing Filter." If evidence or reasoning is accepted and rises in relevance to the group, it is accepted on equal ground—the identity of the contributor is not relevant—relevant knowledge is all that matters.

> Turing Filter: If an idea generated by an AI Agent rises in peer-reviewd relevance in a collective reasoning process with human reviewers we make the conjecture that it has passed the Turing Test.

In summary, the Collective Reasoning Facilitator prompts and guides participants (human or intelligent agents) through a templated process of deliberation and prediction, exploring the why behind our imaginations, judgments, and predictions, to create a persistent knowledge model of our reasoning about some future outcome.

Co-creation, decisions, and predictions in practice

Large language models offer a tremendous ability to leverage the power of accumulated human knowledge as represented in the vast store of digitized text that emerged due to the internet. For it to be of practical value to make reliable predictions and decisions, we need trust and contextualization. Trust grows with repeated tested reliability of predictions. Explanations of that link reasoning to results further increase confidence. We will now look at three examples of why this matters.

Forecasting the outcome of a decision with context-specific knowledge models

Early on, we at CrowdSmart tested and trained the methodology and system on a hard problem: predicting the fundability of early-stage startups. Given the lack of data, the context is an excellent environment for assessing the predictive power of collective human intelligence with applications to decisions to fund innovations, new products, or services. For more detail, see my paper "AI-guided co-creation."

For four years, beginning in 2016, we created predictive models by engaging a group of angel investors, experts, and founders in creating probabilistic models of their collective reasoning to predict the likelihood of getting follow-on funding (predicting the ability to survive and extend their funding runway). A best practice model for early-stage investing guided the process.

Initially, we did not use intelligent agents primarily because it was early in the development of transformer models. Early attempts at language generation failed to be sufficiently coherent to pass through a "Turing Filter" of human participants.

The top-scoring company of approximately 150 companies reviewed had developed technology for a neural network chip running on very low power. We created a decision forecasting model for early-stage investing based on research done in seed investing. More on the process is described in the attached note.[15]

The collective reasoning process for the deep learning chip company resulted in a score of 92%. Less than 100 days later, they raised a series A at $16M pre, and their market cap now is estimated to be in the several hundred million. This strongly indicates the predictive power of collective intelligence alluded to earlier. This was not a one-off case. A majority of the companies survived the tough climate of the last 3 years. The takeaway is simply this: collective human intelligence worked and was predictively accurate. (For the entire data set, we were able to show that the system was able to predict startup survivability with >80% accuracy while simultaneously reducing bias).

The collective reasoning process produces a probabilistic graphical model (a Bayesian Belief Network—BBN).[16] The model provides the user with a prediction and an explanation of why. The sources of knowledge used are referenced and verifiable. *The knowledge model produced approximates a causal model for a decision or prediction.*

This new type of knowledge model derived from human and machine intelligence can be queried or simulated for "what if" analysis:

If you ask the knowledge model: "What is the market demand?" it returns:

'Love the combination of 10K X improvement of computing power with low electric power needed. Paired with the ability to have local Machine Learning algorithms embedded makes this a very interesting platform for many applications as well as brand new markets. Compelling because of 4 major things. 1. Chip is not dependent on the Cloud, thus can be inserted into almost any device. 2. Low power consumption, thus does not generate heat, keeping overall device small while device battery-life is not impacted & potentially excludes current competitive chips. 3. Market identification & initial interest seems real, with little competition within target market. 4. Extensively experienced leadership team.'

"What are the risks?" returns: *'The risk is if the initial chip design does not meet expectations and a re-design is required.'*

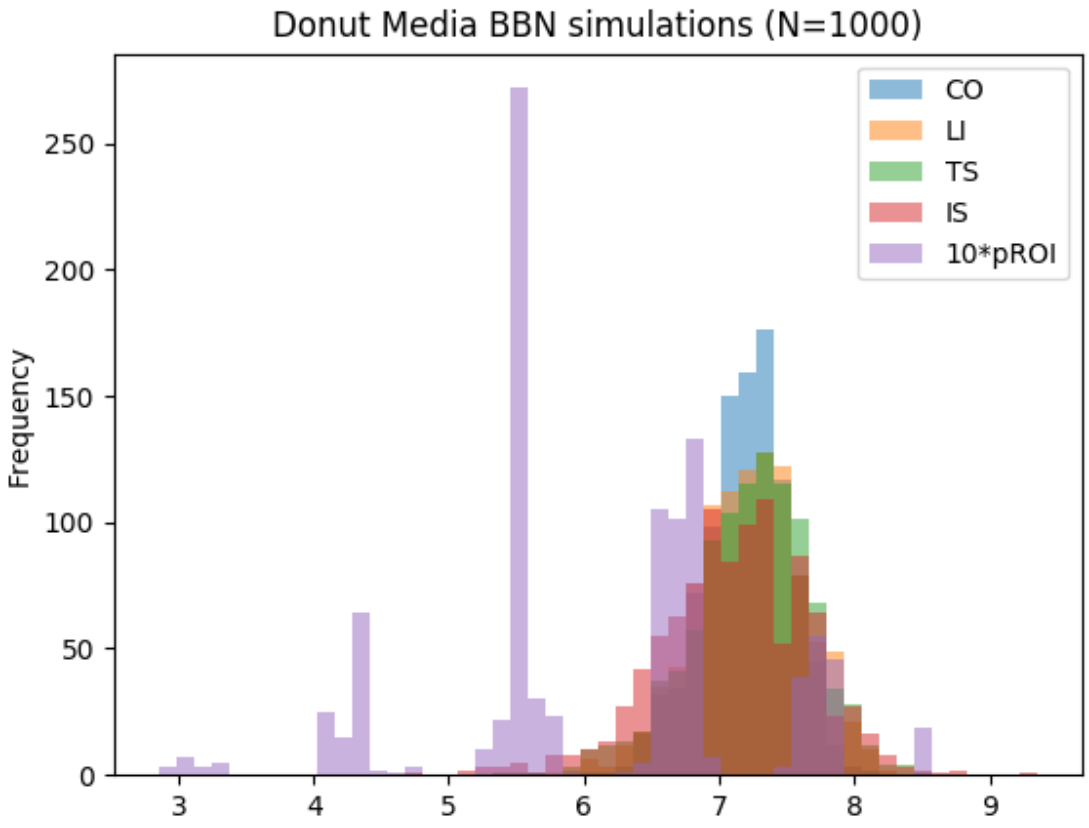
The above is a simple illustration of embedding contextualized knowledge into a large language model that now provides specific answers to questions. The exact wording of the answer is linked to the person who said it. While the collective reasoning process is single-blind, the CrowdSmart system *keeps track of the authorship of the reason given*. *The system also ranks contributors by the influence of their reasons on the outcome under a metric called "InfluenceRank."* In this case, the top influencer had extensive experience in semiconductor fabrication.

The figure below shows a three-dimensional projection of a context-specific knowledge model embedded in a large language model (the one shown here is BERT based). The figures present a visualization of the collective mind of the participants deliberating on the specific task of evaluating the neural chip company. Each point in the embedded space is labeled by: topic, probability of relevance to the group, author, and a vast number of attributes supporting a rich and comprehensive analysis of the language component of the model.

In addition to a comprehensive language model, a BBN connects language to scoring behavior and provides an active, simulatable mini “expert system” of the decision process.

Context-specific models at scale provide a foundation for generative techniques. Combining BBN models with common contexts provides a foundation for in-depth analysis and generative exploration.

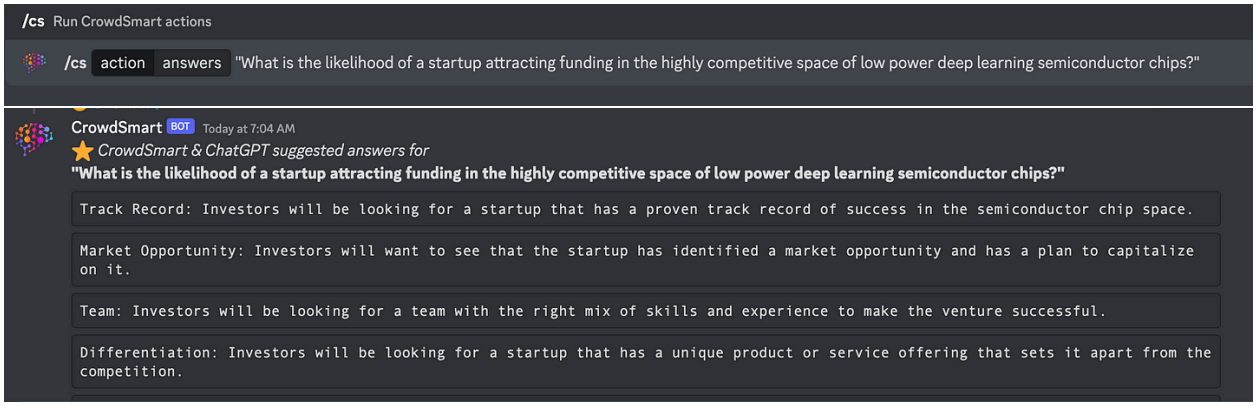
The models are persistent representations of the collective intelligence of the reviews. Presenting text to them will generate possible scoring patterns. For example, the knowledge model for a media company was presented with a transcript from a video conference call discussing the company. The following figure shows the inferred scoring pattern from the text:



The purple shows probabilistic scoring patterns inferred from the knowledge model. The other four items show data distributions that were used in building the model. The video conference call was not a structured conversation and therefore produced noisy data. In spite of the unfocused nature of the conversation, one of the model's distribution patterns fits rather well with the actual scored data. Examples like this show the potential for a long and highly productive relationship between collective knowledge models and generative AI.

Engaging ChatGPT as an intelligent contributor to a collective reasoning process

Intelligent agents like ChatGPT provide a powerful role in considering possible outcomes for a decision or investment. Staying on the topic of deep learning low-power chips, ChatGPT plays a participating role as a generator of answers to specific questions.



These comments are generic and could be improved with Prompt Engineering.[17] Prompt Engineering allows adding context and instructions to the query. Intelligent agents for the collective reasoning process are created by embedding the context and instructions in the decision model template guiding the Collective Reasoning Facilitator.

As an example suppose your team is looking for insightful ways to create a predictive decision model for investing in web3 companies. The template given to the CRF contains a single open-ended question:

> What do you believe are the key features that create a high level of interest for web3 investors? Name the feature (e.g. openness) and why it is relevant.

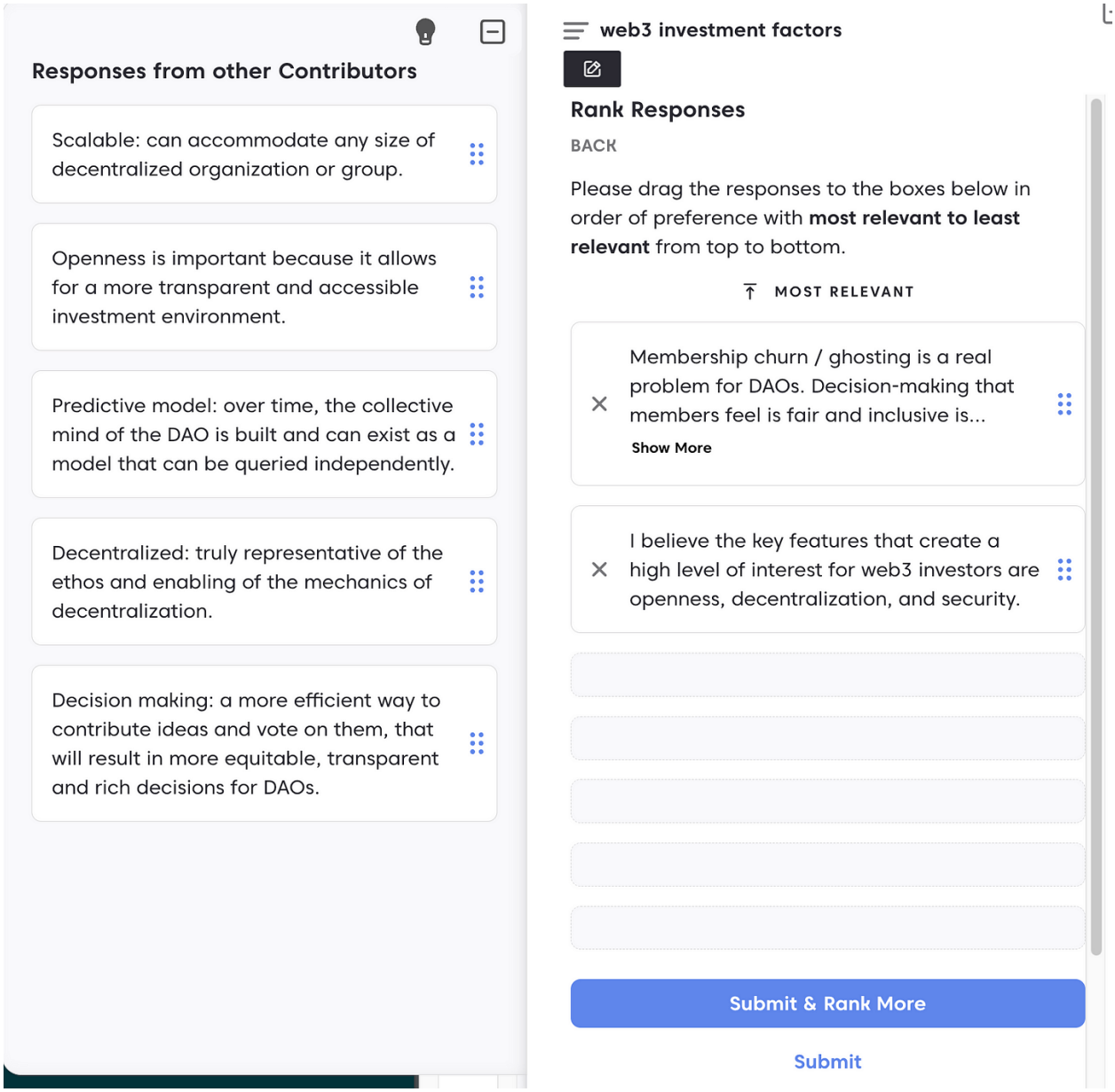
We add a prompt instruction:

> Answer in first person as if you were human. How would you advise me on a web3 investment?

The result blends nicely into a collective reasoning process:

> I believe the key features that create a high level of interest for web3 investors are openness, decentralization, and security. Openness is important because it allows for a more transparent and accessible investment environment. Decentralization is important because it allows for a more distributed and secure investment environment. Finally, security is important because it ensures that investments are safe and secure from malicious actors. All of these features create a high level of interest for web3 investors because they provide a secure and reliable.

ChatGPT's inputs are blended with human-generated inputs in the sampling process:



There you have it! An integration of generative AI with collective human intelligence. The generated BBN for the engagement will likely contain a blending of human-generated ideas with ChatGPT-generated ideas. Whatever passes through the "Turing Filter" is retained in a persistent knowledge model of the collective reasoning process.

Application to socio-political applications

AI's destructive impact on the socio-political front is largely due to its ability to mask misinformation as truth and isolate groups into like-minded contexts. For that reason, it is

not prudent to blindly embrace a new form of BlackBox AI, such as ChatGPT. Human collective intelligence, as represented in the types of collective knowledge models generated by the methods discussed above, provides a powerful trust-building mechanism. The knowledge models generated are explanatory, have citation links to sources, and achieve the "critical thinking" requirement suggested by the inquiry to ChatGPT discussed at the beginning of this paper.

CrowdSmart is working on a civic impact project with a Cincinnati group on applying collective intelligence and the methods discussed above to find ways to reduce gun violence. Asking ChatGPT will only generate from the text trails of what we have historically researched, discussed, and debated (without much progress). Any horizon of hope for the future will come from the inclusive collective intelligence exploring and assessing new paths forward. It will not come from correlations in the text paper trails of our past failed efforts.

In this project, we are connecting with impacted neighborhoods, people "in the life," social workers, scientists, and foundations to co-create and assess potential paths forward that are highly likely to work as viewed by all participants.

Building trust in generative AI in scientific research

In the January 26 issue of Nature an editorial piece titled raises a critical issue for scientific research and generative AI "Tools like ChatGPT threaten transparent science.."[18] Referring back to John Seely Brown's quote, the impact of large language models in science presents a real danger of eroding the foundations of knowledge discovery.

Thomas Kuhn made the case that group experience and the variety of individual experiences come together in a collective process that leads to scientific breakthroughs. In any given research group, "tribal knowledge" (the implicit knowledge that emerges from groups working together) plays an important role. Collective reasoning offers a means to make that knowledge explicit.

Collective reasoning and the resulting knowledge models build trust because they follow a process very similar to the scientific method:

1. Propositions (reasons) are linked to a source.
2. Propositions are peer-reviewed in a single-blind process
3. Bayesian learning models the group acceptance of evidence and propositions.

One can imagine a group of scientists leveraging the immense powers of generative AI in collaboration while collectively reasoning together on the results. Generative AI will

continue to play a productive role in scientific research. Likely, curation and contextualization governed by collective reasoning methods in this paper will play a yet-to-be-discovered role.

Collectively more intelligent and trusted

The MIT Center for Collective Intelligence explores how people and computers can be connected so that

— collectively —

they act more intelligently than any person, group, or computer has ever done before.

Progress in generative AI with applications like ChatGPT combined with technology like the Collective Reasoning Facilitator unleashes co-creation possibilities that promise immense business and socio-political benefits.

---

(1) Eric Schmidt recently said: "Tom, you were doing AI before there was AI". While that is not exactly true, I published my first NLP paper in the late 70s.

(2) CrowdSmart's patented knowledge acquisition technology is built on a foundation of transformer models and probabilistic graphical networks.

(3) Data provenance refers to records of the inputs, entities, systems, and processes that influence data of interest, providing a historical record of the data and its origins.

(3) private conversation

(4) Lead sentence in the paper: Human collective intelligence as distributed Bayesian inference, Peter M. Krafft, Julia Zheng, Wei Pan, Nicola´s Della Penna,Yaniv Altshuler, Erez Shmueli,4, Joshua B. Tenenbaum, Alex Pentland, arXiv:1608.01987v1 [cs.CY] 5 Aug 2016

(5) Thomas S. Kuhn's Foreword to Paul Hoyningen-Huene, Reconstructing Scientific Revolutions: Thomas S Kuhn's Philosophy of Science (1993)

(6) Kehler, T. and Flkes, R. 1985. The role of frame-based representation in reasoning. Commun. ACM 28(9): 904–920

(7) Under a research grant from DARPA, our team at IntelliCorp, built a multiple-model reasoning system. "A New Generation of Knowledge System Development Tools."

(8) Reasoning with worlds and truth maintenance in a knowledge-based programming environment (Communications of the ACM volume 31Issue 4April 1988 pp 382–401)

(9) Page, Scott (2018). The Model Thinker: What you need to know to make data work for you". Basic Books.

(10) Not all second-generation AI systems are BlackBox. Building explanatory capabilities for deep learning models is an active area of research and development.

(11) *Pearl, J., & Mackenzie, D. (2019).* The book of why*. Penguin Books.*

(12) Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The art and science of prediction.* Crown Publishers/Random House.

(13) Malone, Thomas, W. (2018) Superminds: The surprising power of people and computers thinking together

(14) Eigen in German means "own." In physics and mathematics, eigenvalues and eigenvectors capture and characterize the behaviors of a system. When you strike a tuning fork, the sound you hear is defined by its "own" vibrational states—its eigenvalues. In our application here, we learn the resonance modes of a group thinking together. We learn what ideas resonate with the participating group.

(15) The community of individual investors and experts participated in an interactive process (identity masked) on various aspects of the startup's potential. The process was guided by a templated model for early-stage investing derived from research in the category. Each engagement included the startup team and a team of ~20 individuals with diverse backgrounds. Accredited investors on the team were allowed to co-invest if interested. The startup team was engaged in interactive Q&A during the deliberation and often evolved their business strategy due to the collaboration. The results of the model were interpreted as a "probability of a return on investment as measured by the ability of the startup to raise follow-on funding. It was a survivability model. We then tracked all companies over a period of at least 2 years after the initial evaluation. For companies with scores >75%, we invested ~$100k (not enough to skew the results) and allowed qualified participants to co-invest.

(16) Bayesian Belief Networks provide a powerful way to create causal links between a belief or fact and an outcome. In our case, we use them to link "reasons for a score" to scores, thus linking reasoning patterns to outcomes.

(17) Prompt engineering is a concept in artificial intelligence, particularly natural language processing (NLP). In prompt engineering, the description of the task is embedded in the input, e.g., as a question instead of it being implicitly given. (source Wikipedia)

(18) Nature Vol 613/Issue 7945

(19) If you are interested in learning how you might get involved in this project, please contact me tom@crowdsmart.ai

---