3-2024

Superminds at work: The promise of human-AI collaboration

Thomas W. MALONE

Superminds at Work:

The Promise of Human-AI Collaboration

Shaping a collaborative future.

Massachusetts Institute of Technology (MIT) Center for Collective Intelligence Director Professor Thomas W. Malone's scholarship offers deep insights into the promise afforded by the synergies between human intelligence and technology. According to Professor Malone, the boundaries between human intellect and technological prowess are becoming increasingly blurred, but this may not be a bad thing for humankind. In Asian Management Insights' inaugural Pulse Point interview, we get to learn more about the concept of 'collective intelligence', which explores how a partnership between humans and Artificial Intelligence (AI) can be catalysed to make ground-breaking advancements in addressing the wicked problems of our time. At the heart of his arguments is the idea of 'superminds'–entities comprising individuals and computers–that can work together in intelligent ways to make this human-computer partnership possible. Professor Malone also introduces the use of the 'Supermind Design' methodology as a systemic approach to designing such collective intelligence systems.

What do you think of the current debate about the risk of AI displacing or even replacing human intelligence? How do you address this persistent perception that AI and humans cannot work well together? How do these concerns relate to your idea of 'superminds'?

I think we need to spend much less time thinking about people *or* computers, and much more time thinking about people *and* computers. Similarly, *less* time ought to be expended on thinking about how many jobs computers are going to take away from people and *more* time could be devoted to thinking about what people and computers can do together that was never possible before. In other words, we need to ask ourselves: How can people and computers be connected, so that–together–they act more intelligently than any person, group, or computer has ever done before? One way people often talk about this is to say we should 'put humans in the loop'. But I think it's better to start with the human groups that have done almost everything we humans have ever accomplished, and then add computers to those groups. In other words, we need to move from putting *humans in the loop* to putting *computers in the group*.

Using AI and other technologies, these computers can do the things they do better than people–like arithmetic and certain kinds of pattern recognition–and people can do the rest. Perhaps even more importantly, computers can also be used to provide *hyperconnectivity*, i.e., connecting people to other people–and often to computers–at a much larger scale and in rich new ways that were never possible before. Think, for instance, how Internet-based platforms like Wikipedia, Google, and Facebook allow vast numbers of people to create and share information all over the world in a multitude of ways that were never possible when similar kinds of information had to be shared by moving papers around the planet!

A good way to understand this concept is to think about what I call '*superminds*', which I define as *groups of individuals acting together in ways that seem intelligent*.[1] By this broad definition, superminds are all around us, all the time: companies, governments, labour markets, scientific communities, the editors of Wikipedia, and even the global economy. All these examples are groups of people (and often computers) acting together in ways that–at least sometimes–seem intelligent. These superminds can take on various organisational forms for collective decision-making and problem-solving, such as hierarchies (e.g., firms), democracies (e.g., governments and clubs), markets, and communities (e.g., scientific communities and neighbourhoods).

Often, computers can make these superminds smarter. Think of ChatGPT for instance. The system's AI algorithms can conduct intelligent conversations about an amazing range of topics. But this wouldn't have been remotely possible without the vast amount of human-created content on the Internet that was used to train ChatGPT.

Of course, it's also possible for computers to make superminds more stupid, like when fake news influences voters in a democracy. But I think if we use them wisely, computers can help us create much more intelligent human-computer superminds in business, government, local communities, and many other parts of society. And the concept of superminds gives us an evocative new way of viewing AI, not as a *rival* to human intelligence, but as an increasingly valuable *partner* in all the different kinds of groups that make up our human societies.

How does the supermind concept relate to the 'collective intelligence' idea you had also proposed?

In my mind, saying something is a "supermind" is just shorthand for "collectively intelligent system". And the phrase "collective intelligence" also suggests that we might be able to measure the intelligence of superminds the way we measure the intelligence of individual humans. In fact, my colleagues and I did exactly that. We created an 'IQ (Intelligence Quotient) test' for groups using the same statistical techniques that psychologists use to create such tests for individuals. We found that, just as for individuals, there is a single statistical factor for groups that predicts how well a group will perform on a wide range of tasks. We called this factor 'collective intelligence'. To our surprise, we also found that the collective intelligence of a group was only weakly correlated with the average individual intelligence of the group members. And it was significantly influenced by three other factors: the average social perceptiveness of the group members, the extent to which the group's conversations were not dominated by a few members, and the percentage of women in the group.[2] And the last factor was mostly accounted for statistically by the first factor. In other words, it was well-known before we conducted our research that women on average score higher than men on this measure of social perceptiveness, and this may well explain why groups with more women did better. In addition, we found that social perceptiveness was a significant predictor of collective intelligence not only in face-to-face groups but also in online groups, and not only in laboratory groups but also in classrooms and online games. In short, to create collectively intelligent superminds, we don't just need smart people, we also need people who are good at working with other people.

Given the sensitive nature of the human-computer relationship, how do you create such AI systems?

One way of thinking about this is to consider the five cognitive processes that are needed by any intelligent system, whether it is a person, computer, or group. Before taking any action, you need to *decide* first what action to take. In order to do that, you usually need to *create* a few options. You can also usually *decide and create* something better if you can sense the world around you and remember the past. And if you're really smart, you can *learn* to do all these things better over time. In my book *Superminds*, I give examples of how computers can help human groups do all these things more intelligently.

For groups, perhaps the most important cognitive process is decision-making, and in the book, I describe five basic ways groups can make decisions. In Hierarchies, group decisions are made by delegating them to individuals in the group, and technology makes this possible in very new ways. For instance, in Google, all the operational decisions taken to generate the results for a user's search query are made by computer algorithms, and the key role of humans is to manage, train, and maintain these algorithms. In Democracies, group decisions are made by voting, and computers make new kinds of democratic decision-making possible. For example, in 'liquid democracies', humans can vote directly on any issue they want to, but for issues they don't want to pay attention to, they can delegate the action of voting to human proxies or to computers.

In Markets, group decisions are just the combination of many buyer-seller agreements. For instance, computers have long been able to carry out automatic trading in financial markets that also include humans, and computers can now do online retailing with very little human intervention. In

The concept of superminds gives us an evocative new way of viewing AI, not as a *rival* to human intelligence, but as an increasingly valuable *partner* in all the different kinds of groups that make up our human societies.

Communities, group decisions are made by a kind of informal consensus based on group norms, and online communities like Wikipedia show how groups of people and bots can jointly edit the same documents with very good results. But social media filtering algorithms illustrate an unfortunate possibility that can arise from the interaction among communities for online discussions and markets for online advertising. In many cases, these filtering algorithms are designed to optimise advertising revenues, not the experience of their users, and this often leads to undesirable social consequences, such as people becoming addicted to their news feeds.

For the first four types of superminds, there is some degree of cooperation among the group members. But there is also a fifth kind of supermind, which I call Ecosystems. Here, the group decisions are made by the law of the jungle—people with the most power get what they want—and the survival of the fittest. Just as individuals often compete for power in a group, superminds also often compete with one another for power. For example, countries compete with one another in economic markets, cultural spheres, and military wars. Within a given society, too, there are often conflicts for power among corporations, governments, voters, and communities. And there are many ways that computers can change the balance of power among these different types of superminds–some are probably good (like the rapid spread of innovations in society), and some are probably bad (like using killer robots in wars).

I believe this is a powerful framework for thinking about a lot of what goes on in the world, since almost all the human groups we know of belong to one of these five categories. In most situations, several of these types of superminds (sometimes all of them) are in play at the same time. This is also a helpful way to analyse and come up with ideas for how computers, especially AI, can help make superminds smarter. For instance, collective intelligence can be incorporated into corporate strategic planning by using crowdsourcing, prediction markets, and data analytics to get a wider range of stakeholders involved in the planning process.

Can you give us some examples of how to design such systems?

To help design such collective intelligence systems, my colleagues and I have developed a methodology called 'Supermind Design'.[3] The methodology fosters innovative thinking about how to design groups for solving specific problems. It does this by systematically asking questions about factors such as different types of group decision-making and technology uses. The methodology has been used in organisations such as Takeda Pharmaceuticals and Deloitte Consulting, and trialled in Singapore by the National Robotics Programme.

We call these systematic questions 'moves', and the first set of moves is also used as part of many other creativity techniques. It comprises *Zoom In* to focus on the parts and types of a problem, *Zoom Out* to concentrate on the bigger picture of which the problem is a part or type, and *Analogise* to look for analogies to the problem.

The second set of moves is focused specifically on designing superminds, that is, groups of people and, often, computers. Those moves include *Groupify*: What are the different kinds of groups (hierarchies, democracies, etc.) that could help solve a problem? The next move is *Cognify*: How can the different cognitive processes be used in solving a problem? And the third is *Technify*: How can different kinds of technologies be used to help solve a problem?

We have also recently developed an AI tool called Supermind Ideator to help people design superminds. For example, if you want to come up with innovative ideas about how to reorganise the structure of a company or use a new AI capability, you could use the Supermind Ideator tool to help you. What the Ideator tool essentially does is applying the Supermind Design moves automatically. It is built on top of OpenAI's GPT-3.5 and GPT-4, and uses generative AI techniques to provide people with ideas that can inspire them to think of things they would never have thought of before.

We are currently using Supermind Design and Supermind Ideator in our work with Singapore's Changi Airport. For example, we're exploring opportunities for improvement in the aeroplane turnaround process, which is what happens between the time the plane arrives at the gate and the time it pushes back to go to the runway. There are many processes that happen then–passengers getting off the aircraft while others prepare to board it, refuelling, as well as unloading and loading of baggage, food, etc.–and these have to happen fast. Hence, we want to use our approach of analysing processes to try to develop more innovative processes that can be more efficient and flexible.

We also want to look at security screening and think about what can be done to improve the process. AI can do some part of it, with humans doing the rest, so there are a lot of interesting possibilities.

While it is, of course, important to design technology, we believe that the successful use of AI technologies will also depend on being able to devise the processes in which those technologies are embedded. And we think it's possible to develop a body of scientific knowledge about process design that can be very useful for doing this.

The advent of superminds sounds promising, but what could be some of the perils? How would the role of humans change in such partnerships?

I've already talked about some of the potential risks of human-computer superminds, like fake news, addictive social media algorithms, and killer robots. And I don't think it is guaranteed that using computer technology will make things better. But, in general, I'm optimistic about the long-term potential of the ways we use computers. I believe there are huge opportunities for very good things we can do with these technologies.

One important perspective to emphasise is that no matter how much computers can accomplish, we will probably still want humans to play the role of setting goals for computers and using human values to judge the results. In other words, people will often be managers of computers and not just of other people.

In the future, as the world becomes more and more interconnected and AI software becomes more capable, thinking of people and computers as parts of a cohesive global supermind will prove invaluable. And perhaps our future as a species will hinge on how effectively we can harness this global collective intelligence, to make choices that are not only smart, but also wise. AMI

Dr Thomas W. Malone
is the Patrick J. McGovern Professor of Management, and Director of the MIT Center for Collective Intelligence at the MIT Sloan School of Management

Endnotes

1 See Thomas W. Malone, "Superminds: The Surprising Power of People and Computers Thinking Together", Little Brown, 2018.

2 The study was conducted by Professor Anita Williams Woolley of Carnegie-Mellon, Professor Christopher F. Chabris of Geisinger Health System, Professors Alexander Pentland and Thomas W. Malone of MIT, and Assistant Professor Nada Hashmi of Babson College. For details, refer to "What Makes Teams Smart", MIT Sloan Management Review, October 4, 2010.

3 MIT Center for Collective Intelligence, "Supermind Design Primer", June 2021.