Chapter

Collective Superintelligence: Enabling Real-Time Conversational Deliberations among Humans and AI Agents at Unprecedented Scale

Louis B. Rosenberg

Abstract

This chapter explores the pursuit of Collective Superintelligence using a novel technology called Conversational Swarm Intelligence (CSI) that enables real-time conversational deliberations among human groups of potentially unlimited size. Traditional methods for capturing the Collective Intelligence (CI) of large groups use data aggregation techniques that are useful for narrow tasks such as numerical estimations and forecasting, but struggle to address the complex, open-ended problems faced by real-world organizations. CSI has overcome these limitations by adopting a dynamic systems approach inspired by the remarkable decision-making abilities of bee swarms, bird flocks, and fish schools. By leveraging the principles of Swarm Intelligence in combination with the power of Generative AI, CSI enables real-time conversational deliberations among networked human groups of potentially any size, empowering large teams to discuss issues, brainstorm ideas, debate alternatives, share knowledge, assess risks, and quickly converge on solutions with significantly amplified collective intelligence. This unlocks a new massively scalable framework for communication and collaboration. It also has the potential to enable large hybrid groups of human experts and AI agents to deliberate together in real-time and solve problems at collective intelligence levels that exceed all individual participants, both human and AI.

Keywords: collective superintelligence, swarm intelligence, conversational deliberation, conversational collective intelligence, generative AI, AI agents

1. Introduction

1.1 Why collective superintelligence (CS)?

The average Fortune 1000 company has over 30,000 employees and has large functional teams with many hundreds of members. Equally large teams exist in government, civil, scientific and defense organizations. And while many companies say that their most valuable asset is the intelligence and creativity of their teams, no technology exists to enable large teams to hold real-time conversations that allow them to leverage their collective knowledge, expertise, insight, sensibility and wisdom in optimized ways.

As will be described below, enabling productive real-time conversational deliberations among groups of *potentially unlimited size* could enable large organizations to efficiently solve complex problems with intelligence, insight, situational awareness, and creativity that significantly exceeds the abilities of all individual members.

In addition, many believe that AI agents will soon play a critical role in the global workforce, bringing their unique informational and analytical strengths to large teams. It is therefore important to consider not just enabling large human groups to hold thoughtful real-time deliberations but to also include AI agents within these deliberations at scale. In fact, research suggests we may be able to enable very large groups of human stakeholders and AI agents to amplify their collective intelligence to levels that significantly exceed all individual participants, both human and AI.

This goal is called Collective Superintelligence (CS) and this paper describes a recent technology called Conversational Swarm Intelligence (CSI) that has made significant progress in this direction. It merges the biological principle of Swarm Intelligence with the flexibility and power of Generative AI to enable productive conversational deliberations by text, voice, or video at *potentially unlimited scale*, and has been shown to significantly amplify the collective intelligence of collaborating teams [1].

1.2 Why scale conversational deliberation?

Collective Intelligence (CI) is the well-known principle that when large human groups combine their knowledge, expertise, and insights, they can outperform the median participant on a range of useful tasks [2]. This "amplification effect" has been shown to scale with group size, especially when members have diverse knowledge, expertise, perspectives, or situational awareness [3, 4].

The concept was famously discovered by Sir Francis Galton in 1906 when he asked 800 farmers to estimate the weight of an ox and found the median value was surprisingly accurate. That was over a century ago and yet today, most CI methods for large groups still aggregate data collected from individuals. While aggregating human insights is useful if you have 800 farmers and an ox whose weight you need to estimate, most organizations need to solve significantly more complex and open-ended problems, usually with interrelated tradeoffs, risk factors, contingencies and alternatives.

In real organizations, Conversational Deliberation (CD) is the fundamental method by which teams interactively share perspectives, brainstorm ideas, debate alternatives, assess risks, prioritize options, and converge on solutions that leverage their combined knowledge, wisdom, expertise, and situational awareness [5]. It has therefore been a longstanding goal of researchers to combine the *deliberative benefits* of real-time conversation with the *collective intelligence* benefits of very large groups.

Unfortunately, authentic real-time conversations do not scale. Research into conversational dynamics suggests that the ideal size for a thoughtful deliberation is only about 4–7 people [6, 7]. At this scale, each individual has a good amount of *airtime* to express their views and low *wait-time* to respond to others. But as group size grows, airtime decreases, wait-time increases, and conversational dynamics rapidly degrade. By 10–12 people, it ceases to be an interactive discussion, devolving into a series of monologs, and above 15–20 people, it becomes a one-to-many presentation [8, 9].

This has made it impossible to leverage the collective intelligence of large groups through natural deliberative conversations. Instead, the primary methods for harnessing the collective intelligence of large groups have involved capturing data from individual members through polls, surveys, or prediction markets and statistically aggregating to produce a combined result [2]. Such aggregation techniques are generally referred to as Wisdom of Crowds (WoC) and have been shown to outperform the median participant in certain tasks. That said, most CI methods have significant limitations:

i. Because traditional CI methods aggregate data instead of enabling true deliberations, they generally have *narrow use-cases* such as numeric estimation, probabilistic forecasting, and multiple-choice selection. In real-world organizations, complex problems often require open-ended brainstorming and debate among stakeholders to find solutions [10].

ii. Because traditional CI systems aggregate preference data, they often fail to reveal the *underlying reasons why* the group favors or rejects the various options under consideration. Often, the *reasons why* contain the most useful insights, especially when surfaced through deliberation.

iii. Because traditional CI systems aggregate data, they generally converge on the *most popular perspective* within a group. This often fails when the “conventional wisdom” is flawed [11]. This problem is common in close-knit teams that develop *organizational conventional wisdom*.

iv. Many CI methods use sequential upvoting methods in online forums or sequential “trades” in online prediction markets. Both methods are highly susceptible to social influence bias (i.e. *herding* or *snowballing*) in which prior upvotes or trades significantly influence future upvotes or trades. This amplifies noise [12, 13] and causes momentum distortions (i.e. bubbles) that dampen intelligence [14].

v. The goal of CI should not be to beat the *median participant*, but to outperform *all participants*. This is rarely achieved by WoC methods that converge on the most popular solution, not the best solution [15].

vi. The future of CI will likely include humans and AI agents [16]. Without group deliberations among large human and AI teams, there is no path for human sensibilities to organically guide hybrid solutions [15].

Overall, aggregation-based CI methods have useful applications but generally do not support the complex, open-ended problems faced by real organizations. As will be described in the next section, CSI aims to address these barriers by combining the *intelligence amplification* of large groups with the *deliberative power* of natural real-time conversations [17, 18]. In addition, CSI supports deliberations among large “hybrid groups” of humans and AI agents, further expanding the effectiveness of conversation-based collective intelligence [16–18]. Together, these benefits suggest that CSI may be a viable pathway for enabling human groups and AI agents to collaborate on complex real-world challenges and collectively exceed the intelligence levels and problem-solving abilities of all members, both human and AI. This objective is called Collective Superintelligence [16].

2. From biological swarms to conversational swarms

Before reviewing the methods of Conversational Swarm Intelligence (CSI), it is useful to cover the biological principle of Swarm Intelligence (SI) that inspired the methods and describe a predecessor technology for leveraging SI among large networked human groups, Artificial Swarm Intelligence (ASI).

2.1 Swarm intelligence (SI)

For over a century, scientists have observed that natural species can amplify their group intelligence by forming real-time systems among members. This phenomenon is commonly referred to as Swarm Intelligence (SI) and it enables a wide range of social organisms, from schooling fish and swarming bees to flocking birds, to solve problems together that are beyond the intellectual capacity of the individual members [19]. These natural systems do not aggregate data like many large human groups but instead form real-time systems in which members "deliberate" by interactively pulling for various alternatives until the group finds a solution that maximizes their collective support.

Honeybees, for example, can collectively solve complex problems such as choosing the optimal location for a new home site from among many candidate sites. They do this by vibrating their bodies in support of various options until the group converges on the solution that maximizes their collective support [20]. This vibration-based deliberation among bees is called a "waggle dance" (because it looks like the bees are dancing) and has been shown to engender an optimized decision-making process that greatly exceeds the mental capacity of the individual bees that comprise the system [21].

Even more surprising, the deliberative systems formed by honeybees follow a similar decision-making process to neurological brains [22]. Both brains and swarms contain large numbers of *excitable processing units* (i.e., neurons and bees) that work in parallel to (a) integrate noisy data, (b) weigh competing alternatives and (c) converge on preferred solutions. In both brains and swarms, outcomes emerge through real-time competition among sub-groups of excitable units. When one sub-group exceeds a threshold level of support, a decision emerges [23, 24].

The similarities between the role of *neurons* in brains and the role of *members* in swarms inspired early researchers to refer to Swarm Intelligence as a "*brain of brains*" and to explore swarm-based models as a path to Collective Superintelligence [25]. This led to Artificial Swarm Intelligence (ASI) in 2014, which was the first swarm-based technology for amplifying collective intelligence [22].

2.2 Artificial swarm intelligence (ASI)

Artificial Swarm Intelligence (ASI) was developed in 2014 to enable networked human groups of potentially any size to deliberate as real-time systems modeled on the dynamics of bee swarms [25]. The first studies of ASI were conducted using an online platform called UNU that provided each user with a graphical magnet to vary the direction and magnitude of their sentiments during real-time deliberations. Intelligent algorithms monitor the motion of each participant's magnet in real-time and uses the dynamics to infer relative conviction within the group. As participants adjust their magnets in response to each other, swarming algorithms update the speed and direction of a collectively controlled puck until the group targets a selected outcome together [26].

Using UNU and successor platforms like SWARM, researchers have studied ASI extensively over the last decade for a wide range of applications ranging from group decision-making and financial forecasting to medical diagnosis. For example, researchers at Oxford University and Unanimous AI published a study in 2017 showing that groups of financial traders, when deliberating as real-time swarms, could amplify their forecasting accuracy by 36% ($p < 0.001$) [26].

Similarly, researchers at Stanford and Unanimous AI published an NSF-funded study showing that small groups of doctors could diagnose chest x-rays as real-time swarms and reduced their diagnostic errors by 33% compared to traditional methods [27, 28]. Studies at California Polytechnic showed "swarming" groups could significantly increase their "social sensitivity" and "social perceptiveness" in subjective judgment tasks [29, 30]. Other studies show significant benefits when groups use ASI to prioritize goals, forecast events, and find socially optimal solutions to ethical challenges or political charged issues [31–33].

Overall, enabling human groups to deliberate as real-time swarms has many advantages over aggregation-based CI methods. That said, ASI also has limitations. For example, graphical swarms have proven most effective for narrow tasks such as numerical estimations, probabilistic forecasting, multiple-choice selection and itemized prioritization. The desire to expand beyond this narrow range of uses has motivated ASI researchers to explore new methods that combine the interactive benefits of *swarming* with the open-ended flexibility of conversational deliberation.

2.3 Conversational swarm intelligence (CSI)

It has been a longstanding goal in the field of collective intelligence to combine the *deliberative power* of natural real-time conversation with the *intelligence amplification* of very large human groups. As described in Section 1.2 above, real-time conversations are inherently unscalable because each member's airtime drops and wait-time increases with each additional participant. A technological solution was therefore required.

In 2023, researchers published the first studies showing that productive real-time conversations among networked human groups could be conducted at potentially unlimited scales and that scaling-up deliberations can significantly amplify collective intelligence [1, 17, 18]. The resulting technology is called Conversational Swarm Intelligence (CSI). It leverages prior learnings from research into ASI technology (based on the dynamics of bee swarms) and expands the conceptual models using the dynamics of fish schools [1, 10, 34].

As background, fish schools are often called "Super Organisms" because they can quickly make critical decisions despite each member having *incomplete information* and most members having limited *situational awareness*. To illustrate this, **Figure 1** below shows a large school facing a theoretical life-or-death situation in which three predators approach from three directions. As portrayed, no individual has *sufficient information* or wide enough *situational awareness* to find an optimal solution.

In fact, most members have no direct awareness of any threats, and only three subsets of members (circled below) are aware of a single threat. This is analogous to human organizations that often face complex real-world problems in which subgroups have limited information, specialized expertise, and narrow situational awareness.

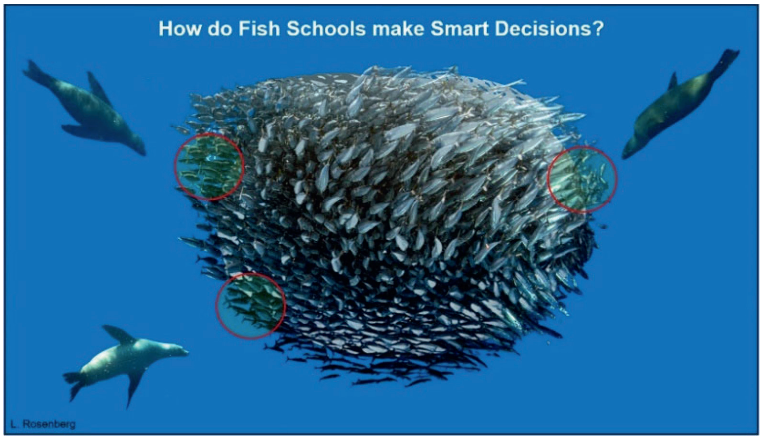

Figure 1.
A large fish school faces a complex life-or-death problem.

2.3.1 So how do fish schools solve this?

Schooling fish evolved the ability to "deliberate" among small groups of nearby fish using a specialized organ called a *lateral line*. This organ detects pressure and vibration changes in the water around them to monitor the *directional intentions* of their neighbors. Because each small group of neighbors overlaps other small groups of neighbors, the local deliberations can rapidly propagate information throughout the full population as shown in **Figure 2**. This allows thousands of fish to quickly deliberate in unison and converge on optimized decisions [35].

2.3.2 Can humans form a swarm intelligence conversationally?

Inspired by fish schools, Conversational Swarm Intelligence works by dividing large, networked groups into overlapping subgroups, each containing about 4–7 members (an ideal size for real-time conversations). Of course, people do not

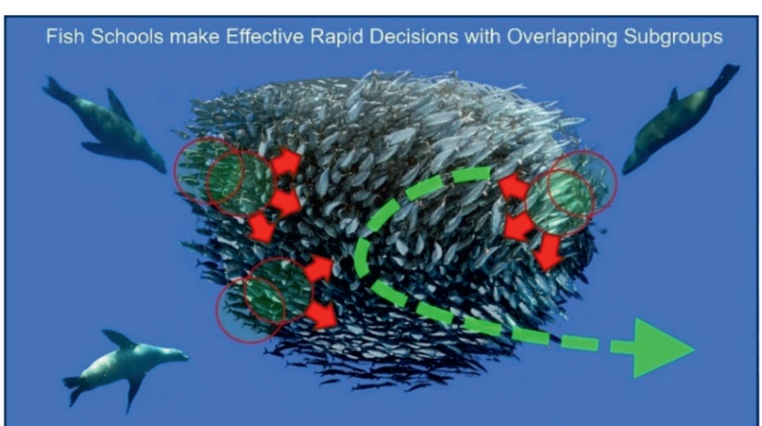

Figure 2.
A collective solution quickly propagates through the fish school.

DOI: http://dx.doi.org/10.5772/intechopen.1010201

possess the fish-like ability to deliberate in overlapping subgroups. In fact, we are unable to focus on multiple conversations at once. This is called the "cocktail party problem" because at cocktail parties many small groups form in close proximity. If you are engaged in a small group conversation at a party and shift your attention to the discussion within a neighboring group, you are likely to struggle to follow either conversation [36].

To address this, CSI uses Large Language Models (LLMs) to deploy an innovative AI agent called a "*Conversational Surrogate*" [1, 10, 16–18, 37, 38]. After CSI divides a large human group into a set of small subgroups, it automatically inserts a Surrogate Agent into each one. This agent is designed to: (a) observe the local deliberation in its subgroup, (b) distill key insights in real time, and (c) pass those insights to other subgroups where that group's local *Surrogate Agent* will express the insights to its members as *natural dialog*. This weaves all subgroups together into a single unified conversation (of potentially any size) in which the full population can discuss issues, brainstorm ideas, evaluate options, and efficiently converge on collective solutions.

Figure 3 below shows an example CSI architecture in which a 100-person group is divided into 14 parallel subgroups, each subgroup having seven or eight human members and one AI agent. This creates a single large conversation in which ideas, insights, assessments, and rationales rapidly propagate until the group converges on one or more solutions that maximize the population's collective confidence or conviction.

It is important to note that Surrogate Agents do not bring outside information into the conversation. Instead, they only pass information between local groups by representing the knowledge, views, or insights of one or more human participants. In addition, the Surrogate Agents are designed to modulate the wording and emphasis of their language to convey the strength of local sentiments as they propagate content across groups.

Also, unlike fish schools, for which groups overlap by virtue of proximity, CSI is often configured to enable all subgroups to *overlap simultaneously*. This takes nature's basic concept of a "swarm" and turns it into a "hyperswarm" that is significantly more efficient than the biology that inspired it [39, 40]. Thus, CSI enables very large groups to hold real-time conversations that quickly leverage their collective knowledge, wisdom, insights, and expertise, and can rapidly converge on optimized solutions.

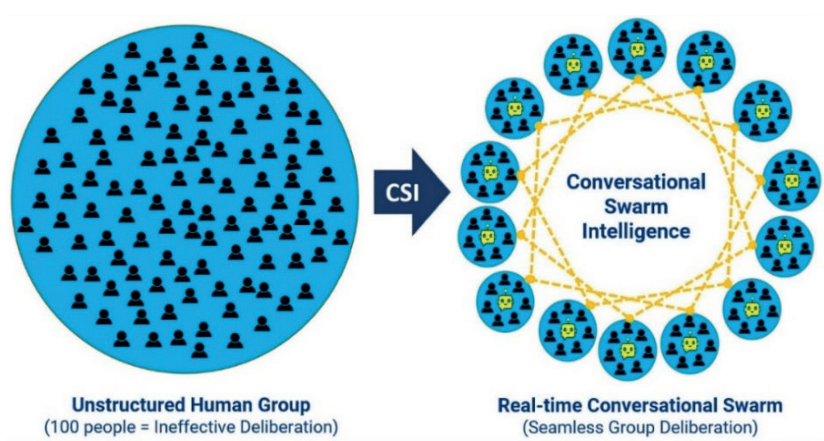

Figure 3.
CSI enables large, networked human groups to hold real-time conversations using LLM-powered Surrogate Agents *that enable all of the groups to overlap in real-time.*

In addition, CSI has inherent benefits that amplify collective intelligence by diffusing the influence of *loudmouth bias*, *authority bias*, and *first-talker bias*, which can significantly distort traditional deliberations. For example, because CSI divides a large group into overlapping subgroups, a single strong personality or authority figure (i.e., a loudmouth) can only have an oversized impact within their local subgroup. For an idea, insight, or perspective expressed by a loudmouth to propagate, it must stand on its own deliberative merits when expressed in other subgroups by a Surrogate Agent.

Similarly, the *first ideas expressed* during a group conversation (large or small) can have an oversized impact on the direction of the deliberation and the ideas that surface. CSI solves this. For example, **Figure 3** above shows a CSI structure with 14 parallel conversations. This means 14 different "first ideas" will be expressed in real-time, each of which must not only gain support within its local group, but it must also compete with the ideas emerging from other groups to propagate throughout the swarm.

Thus, CSI not only allows conversations to scale, it drives *smarter outcomes* by enabling many ideas to surface and percolate in parallel, with the most compelling insights spreading fastest, while the weakest points gradually fade away.

2.4 Validation studies of CSI

A number of studies have been performed to assess the ability of CSI systems to enable effective real-time conversations among large, networked groups, and to significantly amplify collective intelligence. These studies use an online cloud-based CSI platform called Thinkscape® that is accessible to users through standard web-browsers, requires no training for participants, and enables real-time conversations among up to 400 users at once. The platform currently supports text-chat and voice-to-text dialog, and can share media such as images, videos, spreadsheets, or webpages, to all participants.

2.4.1 Groupwise estimation study

In a 2023 study, researchers replicated a classic CI experiment in which large groups were asked to estimate the number of gumballs in a jar [1]. Groups of approximately 240 participants were tested in three scenarios: (i) as individuals filling out a survey, (ii) as aggregated estimates across surveys, and (iii) as a "conversational swarm" using a CSI platform (Thinkscape) that partitioned the 240 participants into 47 subgroups of five or six human members and a single Surrogate Agent. The results showed CSI groups had an average estimation error of 12% which was significantly more accurate than the average individual (55% error) and the traditional CI aggregation method (25% error).

2.4.2 Collective IQ study

In a 2024 study conducted by researchers at Carnegie Mellon University and Unanimous AI, groups of 35 participants were given standardized IQ tests using the Thinkscape CSI platform. The results showed that groups using CSI could hold thoughtful deliberative conversations to collectively efficiently answer the IQ questions. The test groups using CSI scored an average IQ in the *97th percentile* (IQ = 128). As shown in **Figure 4**, this score significantly outperformed the median individual

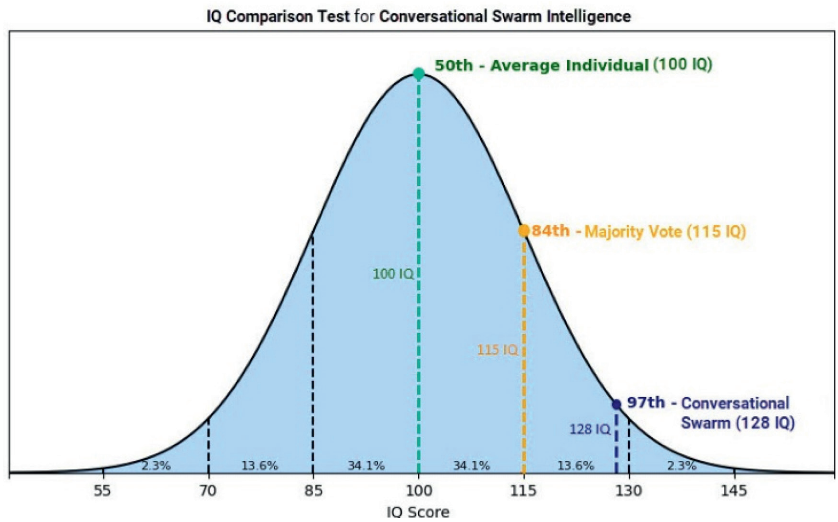

Figure 4.
Results of collective IQ Test study comparing CSI to traditional methods.

in each group (IQ = 100) and significantly outperformed a traditional aggregation-based CI method (IQ = 115) [10].

2.4.3 Large-scale brainstorming study

In another study conducted by researchers at Unanimous AI and Carnegie Mellon, groups of 75 participants were given a standardized brainstorming test known as an Alternative Use Task (AUT). Specifically, participants were asked to collaboratively brainstorm ideas for unconventional uses of common objects with the goal of allowing a surplus of those objects to be sold by a fictional company. Groups were tested in two scenarios: (a) by holding a real-time conversation in a single large chat room similar to Microsoft Teams or Google Chat or (b) by holding a real-time conversation using the Thinkscape CSI platform described above.

All participants were required to brainstorm using each of two test scenarios. Participants were then asked to compare their experiences. The results showed that participants greatly preferred brainstorming using CSI, with a significant majority reporting that the brainstorming process was *more productive*, *more collaborative*, and *surfaced better solutions* when using CSI. In addition, a significant majority of brainstorming participants reported *feeling more heard* and *feeling more ownership* when using CSI [41].

3. Hybrid conversational swarm intelligence (HyCSI)

3.1 Enabling hybrid conversational groups of humans and AI agents

Thus far we have reviewed how CSI enables large human groups to hold productive conversations at scale and can amplify collective intelligence. The next step toward unlocking the potential of Collective Superintelligence is to bring additional AI agents into the CSI architecture. The goal is not to replace human stakeholders, but for AI to serve a supporting role by providing each subgroup with relevant *factual information*, *analytical insights*, and *logical arguments* for consideration during their deliberations.

To achieve these goals, a new type of AI agent was developed for inclusion in the CSI model. Referred to as a "*Contributor Agent,*" the functionality is aimed at observing each local deliberation in real-time and assessing if and when factual, analytical, or logical information should be expressed that (a) has not yet been discussed in that subgroup, and (b) is likely to be relevant and helpful to the ongoing discussion, either by supporting or disputing an idea or alternative that is currently being debated.

Figure 5 shows an example Hybrid CSI (HyCSI) architecture in which each of the parallel subgroups contains six or seven human members and is supported by *one Surrogate Agent* and *one Contributor Agent*. Although only one is shown, CSI can use multiple Contributor Agents to support each local subgroup. This can be valuable if each of the supporting Contributor Agents provides information, expertise, or insight from a different perspective. In addition, CSI can be structured as a hierarchy of increasingly larger subgroups (i.e., swarms of swarms) for mass scaling purposes [42].

It is important to stress the benefit of deploying an independent Contributor Agent in each of the parallel local subgroups in the Hybrid CSI structure. Referring back to **Figure 5** below, the example shows 14 parallel subgroups, each supported by a separate Contributor Agent that provides distinct information to its subgroup that is unique in content, timing, and expression as compared to the Contributor Agents in other groups. This provides *factual diversity* across subgroups, enabling the population to consider a wide range of details in parallel, with the most impactful factual content likely gaining support by spreading between subgroups while the least valuable content fades away.

This ensures that contributions by AI agents, like the contributions of human members, propagate based on *deliberative value*. This leverages the informational power of AI agents but maintains human control over the deliberative process.

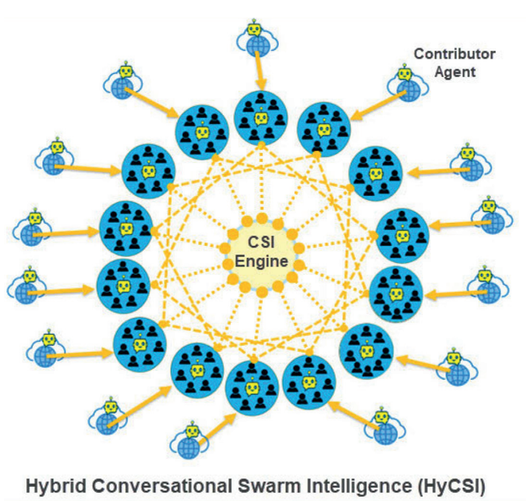

Figure 5.
A hybrid CSI structure is shown with human members and contributor AI agents.

3.2 Validation of hybrid CSI

In 2024, researchers published the first studies in which hybrid groups of human stakeholders and Contributor Agents deliberated together in real-time. As described below, the studies evaluated the usefulness of Contributor Agents designed to provide factual information during groupwise decision-making tasks.

3.2.1 Hybrid decision-making study

In a 2024 study of Hybrid Conversational Swarm Intelligence, networked groups of twenty-give sports fans were tasked with deliberating to collaboratively field Fantasy Baseball teams for competition in a public daily fantasy contest [16]. While fantasy sports may seem frivolous, it is a good model for organizational decision-making because it requires *subject-matter expertise,* probabilistic *forecasting skills*, and the ability to *strategically allocate funds*. That's because each 25-person group was given a fixed payroll to spend on multiple players.

The experiment was run weekly for 10 consecutive weeks and scored using standard fantasy baseball rules. Using the online CSI platform called Thinkscape, the 25 person groups were automatically split into five subgroups, each containing five human members, one Surrogate Agent, and one Contributor Agent. The contributor agents were specifically designed to provide *factual information* about Major League Baseball, including relevant statistics about various players and teams [16].

The results showed that 25-person groups using CSI amplified their performance to the *73rd percentile* (among members who averaged in the 50th percentile). More importantly, when asked to assess the usefulness of the Contributor Agents (called *Infobots* in the study), *87% of participants* supported the statement*: "Our decisions were stronger because of information provided by the Infobot."*

4. Enabling video deliberations at unlimited scale

In the pursuit of Collective Superintelligence, the natural next step for CSI platforms is to expand from real-time text and voice to real-time videoconferencing. In addition, innovative methods are needed for enabling asynchronous engagement as it becomes logistically more difficult to gather groups simultaneously at larger and larger scale.

4.1 Can CSI enable videoconferencing at massive scales?

To enable productive video conversations at very large scale, a next-generation CSI architecture has been developed in which large groups are divided into parallel video-conferencing subgroups, each of which is provided with a Surrogate Agent that appears as an embodied participant (i.e., as an animated avatar.) As shown in **Figure 6** below, the animated avatar performs the surrogate functions described in Section 2.3 above by receiving insights extracted from other local groups and expressing those insights as natural dialog within its own group. This weaves all of the local deliberations together into a unified conversation, enabling the full population to collectively discuss issues, debate alternatives, brainstorm ideas, prioritize options, and converge on groupwise solutions.

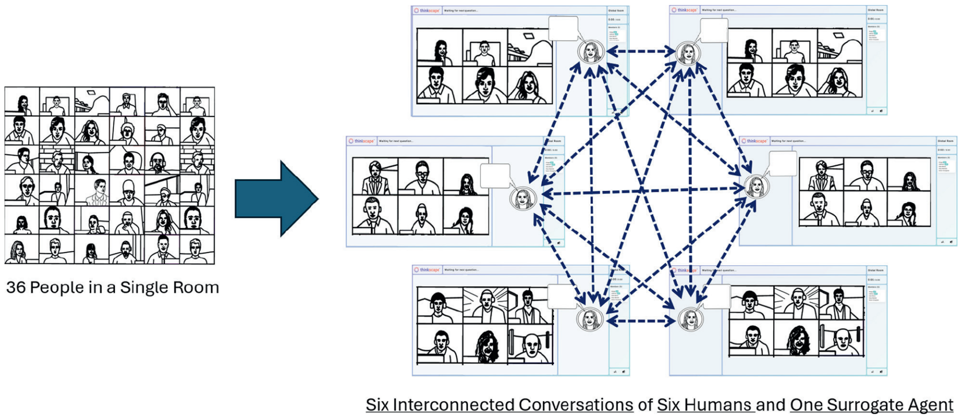

Figure 6.
Example CSI architecture for real-time video-conferencing among 36 people.

While the above figure shows a real-time video conference among 36 individuals, the same structure can be used for hundreds or thousands of participants. Also, hierarchical structures for CSI have been developed (i.e. swarms of swarms) that enable scaling at much larger group sizes. Also, one or more Contributor Agents can be added to each local room as additional avatars to bring relevant informational content into the group dialogs as discussed in Section 3.1 above. In this way, very large groups of human participants and AI agents can deliberate in large-scale video conversations in which they discuss issues, brainstorm ideas, debate alternatives and converge on solutions that maximize their collective conviction.

4.2 Can CSI support asynchronous participation?

By definition, groupwise conversational deliberation is a synchronous social experience in which the participants have the opportunity to react to points made by colleagues as they emerge, enabling groups to build upon each other's ideas, debate supporting or opposing arguments in real-time, and consider a wide variety of opinions, perspectives and rationales as they debate issues. That said, synchrony imposes significant logistical constraints upon teams as it requires large groups to coordinate specific times for deliberative activities. This can be challenging for groups that span diverse geographies, workflows, and time-zones.

For these reasons, asynchronous tools like polls, surveys, forums, and prediction markets have logistical strengths vs. synchronous methods. Unfortunately, those logistical strengths often come with significant limitations that hinder collective intelligence and hamper authentic cross-pollination of insights. As described in Section 1.2 above, when ideas and arguments are contributed sequentially in forums, prediction markets, and other tools that engage participants over time, *temporal bias* and *social influence bias* can significantly amplify noise and distort outcomes. This is because ideas expressed early in an asynchronous process can get overweighted consideration, not because of their merits but because of their timing. In addition, ideas that receive early support by an asynchronous upvote become more likely to receive additional upvotes.

Such "momentum" effects in forums and markets hinder collective intelligence. In a well-known study conducted at MIT, Hebrew University, and NYU, researchers studied the impact of a random *first upvote* on online forums and found that it increased the likelihood of future positive ratings by 32% and caused herding effects that distorted the final ratings by 25% [12]. We can infer similar temporal herding effects in asynchronous messaging tools.

In addition, asynchronous messaging platforms like Slack, Discord, and Microsoft Teams have been found to overwhelm collaborating members due to the frequent notifications and interruptions, and the pressure of growing backlogs of messages, leading to stress and reduced productivity [43–46]. In addition, the growing backlog of sequential messages in platforms like Slack can cause their own temporal biasing effects and can drive uneven participation that distorts outcomes [47].

This begs the question – *Can we combine the deliberative benefits of real-time conversation with the logistical benefits of asynchronous engagement?* One solution that leverages the innovations of CSI is to enable *later participants* to engage with Surrogate Agents that represent the perspectives of earlier participants [48]. To enable this without temporal biasing or social biasing, the CSI architecture can be structured to enable *sequential batches* of members who participate synchronously and introduce time-based Surrogate Agents that represent views and opinions surfaced in prior batches. This is shown schematically in **Figure 7** below.

The example above shows a first networked human group of 100 participants holding a real-time conversation using a CSI structure that divides them into 14 subgroups interconnected by Surrogate Agents to form a unified deliberation. Then, at a future moment in time, a second batch of 100 participants engage in a synchronous conversation, also shown as 14 subgroups woven together by Surrogate Agents. In addition, each subgroup in the second batch has a second Surrogate Agent that represents ideas, views, and perspectives that were surfaced in the first batch. As shown, any number of additional batches can be conducted over time with Surrogate Agents added that represent all prior batches. In this way, the conversation is woven together over time but is done at sufficient parallel scale to avoid momentum effects (i.e. herding) that distorts forums, markets, and other traditional asynchronous collaboration methods.

In addition, it should be noted that at each stage in the process, the CSI structure can also include Contributor Agents that bring informational content into the deliberations, thereby enabling groups of human members and AI agents to collaborate conversationally in real-time batches, combining their collective insights asynchronously over time.

In addition, if we ensure that "seed population" (i.e. first batch) is large enough, the subsequent batches can be relatively small and still avoid temporal biasing and social biasing impacts. In fact, the subsequent batches could be scaled down to single

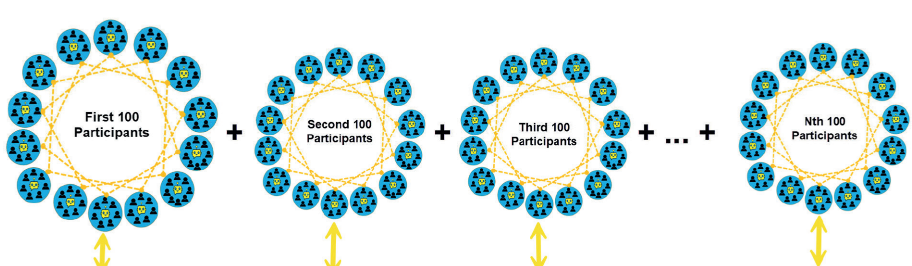

Figure 7.
Example CSI architecture semi-synchronous deliberations at unlimited scale.

deliberative groups (e.g., 4–7 people) or could be scaled down further to a single individual engaging with one or more Surrogate Agents. These options for asynchronous engagement over time are illustrated in **Figure 8** below.

To avoid the traditional problems of asynchronous engagement a sufficiently large seed population is recommended in the first batch [49]. That said, once a large enough deliberation has occurred, single individuals can engage directly with a single Surrogate Agent that represents the views, opinions, and or perspectives surfaced in prior deliberations. Such a *Pluribus Agent* is a personified embodiment of the collective intelligence at that moment in time. A single individual can therefore interview this emergent collective intelligence, or even argue with it, by holding a real-time interactive conversation with an animated *Pluribus Avatar* that represents the prior deliberations.

In addition, if the population of prior participants is large enough, and if the prior deliberations included Contributor Agents that bring the informational and analytical power of AI into the conversations, the Pluribus Avatar could be a real-time representation of a Collective Superintelligence that can outperform all members, both human and AI, when solving complex real-world problems that require human oversight [40].

4.3 Why enable hybrid collective superintelligence?

As outlined above, the goal of Collective Superintelligence is to enable large human groups to collaboratively solve complex open-ended problems at intelligence levels that exceed all participating members. This would offer significant value for organizations, enabling teams of potentially any size to leverage their collective knowledge, wisdom, insights and expertise to quickly address critical issues. Such abilities would be particularly useful for tackling multifaceted problems that require input from a diverse range of disciplines and specialties and a broad spectrum of situational awareness.

In addition, there is a larger moral imperative for the pursuit of Collective Superintelligence stemming from the rapid advancement of AI systems. Many researchers expect AI agents will soon be widely deployed within organizations that can solve expert-level problems at speed, accuracy, and intelligence levels that surpass the abilities of most human employees [50–52]. Whether or not these AI agents achieve the definition of Artificial General Intelligence (AGI), when AI systems do reach a perceived state of *"Cognitive Supremacy"* with respect to human stakeholders, we could see a growing tendency for organizational managers to defer important decisions to automated agents under the assumption that AI systems have deeper

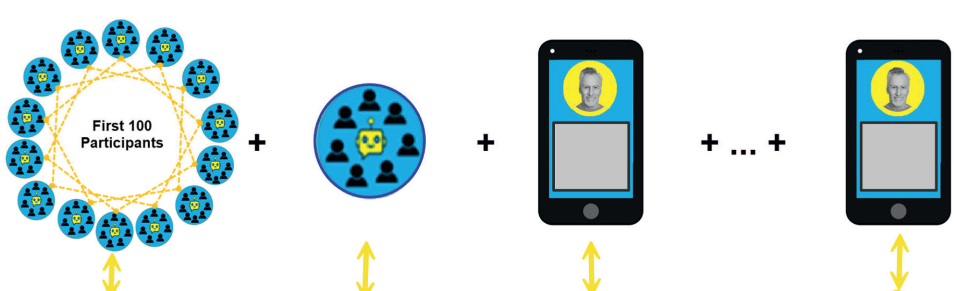

Figure 8.
Example CSI architecture semi-synchronous deliberations at various subsequent scales.

knowledge and stronger analytical skills [40]. This would be a dangerous direction for governance at all levels of management, as many critical decisions require not just sufficient knowledge and analytical prowess, but human values, morals, interests and sensibilities [53].

Thus, by enabling large human groups to form real-time deliberative systems in collaboration with AI agents, we may be able to foster a hybrid Collective Superintelligence that can solve complex organizational problems at intelligence levels that exceed all participants, both human and AI. In addition, such systems could help ensure that human values and societal interests remain inherent to the deliberative process.

4.4 What are the risks of hybrid collective superintelligence?

While CSI has potential to enable human deliberations at unlimited scale and to significantly amplify collective intelligence, we must consider the potential risks of enabling Surrogate Agents and Contributor Agents to participate within groupwise conversations. A significant risk of all human-AI interactions is the current tendency of participants to perceive information provided by AI agents as more *authoritative* or *accurate* than content from human members [54, 55]. This is sometimes referred to as "AI authority bias" and it can distort how humans reach decisions, form subjective judgments, or make forward-looking predictions.

To counter this effect within CSI and HyCSI systems, effort should be made to ensure that participants do not view the AI agents as *authority figures* within the groupwise deliberation. This means not giving the AI agents a "facilitator" or "moderator" role but instead designing the agents to contribute content using similar conversational tone, phrasing, and confidence as human participants [42].

In addition, it is important to remind human participants that the contributions made by Surrogate Agents in each local dialog were not formulated independently by an AI system but were derived based on real-time human discussions in other subgroups. In this way, the human participants are more likely to view the AI contributions with the same skepticism they would have if the comment had come directly from human members of the group. These reminders have been implemented in CSI systems in two ways:

First, the AI agents are generally designed to provide an "attribution indicator" next to their comments. These indicators inform participants of the subgroup(s) from which their comments were derived. In addition, CSI systems provide human participants with a graphical representation of the CSI network and the AI messages that are passed among subgroups. Such graphical reminders help ensure the human participants do not forget that the Surrogate Agents are sharing content between groups and not offering their own AI-generated opinions or perspectives. An example of such a graphical representation is shown below in **Figure 9**.

The figure above depicts a group of 102 participants broken into 21 small subgroups and uses animated "AI" icons to represent the real-time passing of ideas, arguments, counter arguments, and other content between subgroups. In addition, "connection traces" are left behind by the AI icons as they pass between groups, showing the growing network of connections during deliberations. These visual indicators remind the human participants that their group's conversation is deeply connected to the network of other groups and that the agents are acting as surrogates that merely share content between and among participants.

Beyond countering authority bias, additional protections should be put in place to ensure AI agents are not deliberately influencing or manipulating participants.

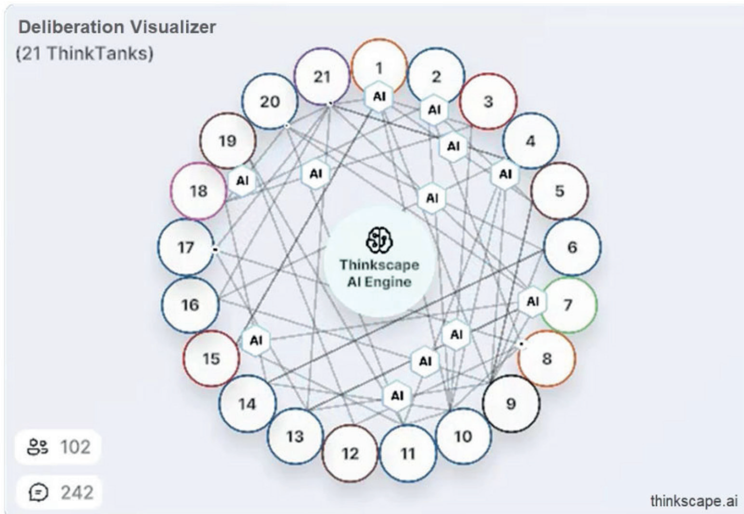

Figure 9.
Snapshot of animated "Deliberation Visualizer" in Thinkscape.ai CSI platform.

Often referred to as the AI Manipulation Problem, the danger is that AI agents could be deployed with super-human persuasive abilities [56]. This risk is present in all AI systems that use conversational agents and should be addressed through policy protections. In the absence of policy, the recommended solution is transparency that enables stakeholders to audit the AI contributions so they can ensure that AI agents accurately and objectively represented the views of subgroups [57]. These auditing abilities are currently present in CSI platforms and should be maintained.

5. Conclusions and future work

Conversational Swarm Intelligence (CSI) and its hybrid variant (HyCSI), offer a unique opportunity to enable real-time communication, collaboration, and collective problem-solving at unprecedented scales. This could provide significant value to large organizations that currently have no methods for conducting real-time conversational deliberations among large functional and cross-functional teams.

Future research should expand validation studies to diverse real-world applications, including strategic planning, organizational decision-making, policy deliberations, collaborative forecasting, civic engagement and deliberative democracy. In addition, studies should expand the size of collaborating groups from hundreds to many thousands of simultaneous members and assess the deliberative efficiency, decision-making effectiveness, analytical accuracy, and collective intelligence amplification.

In addition, future research should explore the logistical constraints caused by synchronous deliberations in real-world organizations and study the effectiveness of asynchronous CSI methods that employ Surrogate Agents to pass insights from subsequent individuals or subgroups to future individuals or subgroups as described in Section 4.2 above.

Also, future work should refine the abilities of Contributor Agents to provide domain-specific informational support while ensuring that human stakeholders maintain control over the deliberative process. Future work should also explore the ethical implications of large-scale hybrid deliberations among human members and AI agents, and work to protect human agency and autonomy.

And finally, research should explore the ability of CSI to enable large groups of humans and artificial agents to collaboratively solve problems at intelligence levels that exceed all individual members, both human and AI. This could enable large hybrid groups function as a Collective Superintelligence that inherently keeps human values, morals, interests, and sensibilities inherently in the loop.

Author details

Louis B. Rosenberg
Unanimous AI, Pismo Beach, California, USA

*Address all correspondence to: louis@unanimous.ai

References

[1] Rosenberg L, Willcox G, Schumann H. Towards collective superintelligence, a pilot study. In: 2023 International Conference on Human-Centered Cognitive Systems (HCCS). IEEE; 2023. pp. 1-6

[2] Riedl C, Kim YJ, Gupta P, Malone TW, Woolley AW. Quantifying collective intelligence in human groups. National Academy of Sciences of the United States of America. 2021;**118**(21):e2005737118

[3] Woolley AW, Aggarwal I, Malone TW. Collective intelligence and group performance. Current Directions in Psychological Science. 2015;**24**(6):420-424. DOI: 10.1177/0963721415599543

[4] Malone TW. Superminds: The Surprising Power of People and Computers Thinking Together, Amazon. Little, Brown and Company; 2019

[5] Nelson LM. Collaborative problem solving. In: Instructional-Design Theories and Models. Routledge; 2013. pp. 241-267

[6] Cooney G et al. The many minds problem: Disclosure in dyadic vs. group conversation. Current Opinion in Psychology. 2020;**31**:22-27. Special Issue on Privacy and Disclosure, Online and in Social Interactions edited by L. John, D. Tamir, M. Slepian

[7] Fay N, Garrod S, Carletta J. Group discussion as interactive dialogue or as serial monologue: The influence of group size. Psychological Science. 2000;**11**(6):481-486

[8] Hackman JR, Vidmar N. Effects of size and task type on group performance and member reactions. Sociometry. 1970;**33**(1):37-54. DOI: 10.2307/2786271

[9] Olson GM, Olson JS, Carter MR, Storrosten M. Small group design meetings: An analysis of collaboration. Human–Computer Interaction. 1992;**7**(4):347-374. DOI: 10.1207/s15327051hci0704_1

[10] Rosenberg L, Willcox G, Schumann H, Mani G. Towards collective superintelligence: Amplifying group IQ using conversational swarms. In: Proceedings of the 26th International Conference on Enterprise Information Systems. Vol. 1. ICEIS, SciTePress; 2024. pp. 759-766. DOI: 10.5220/0012687500003690. ISBN 978-989-758-692-7; ISSN 2184-4992

[11] Bénabou R. Groupthink: Collective delusions in organizations and markets. Review of Economic Studies. 2013;**80**(2):429-462

[12] Muchnik L, Aral S, Taylor SJ. Social influence bias: A randomized experiment. Science. 2013;**341**(6146):647-651

[13] Lorenz J, Rauhut H, Schweitzer F, Helbing D. How social influence can undermine the wisdom of crowd effect. Proceedings of the National Academy of Sciences of the United States of America. 2011;**108**(22):9020-9025

[14] Ottaviani M, Sørensen PN. Aggregation of information and beliefs in prediction markets. In: London Conference on Information and Prediction Markets. Finance Research Unit; 2007

[15] Navajas J, Niella T, Garbulsky G, et al. Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. Nature Human Behaviour. 2018;**2**:126-132. DOI: 10.1038/s41562-017-0273-4

[16] Rosenberg L et al. Conversational swarms of humans and AI agents enable hybrid collaborative decision-making. In: 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON 2024). 2024, October; New York, NY, USA.

[17] Rosenberg L, Willcox G, Schumann H. Conversational swarm intelligence (CSI) enables rapid group insights. In: 2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA: IEEE; 2023. pp. 0534-0539. DOI: 10.1109/UEMCON59035.2023.10316130

[18] Rosenberg L, Willcox G, Schumann H, Mani G. Conversational swarm intelligence (CSI) enhances groupwise deliberation. In: 7th International Joint Conference on Advances in Computational Intelligence (IJCACI 2023; Oct 14, 2023). New Delhi;

[19] Krause J, Ruxton GD, Krause S. Swarm intelligence in animals and humans. Trends in Ecology & Evolution. 2010;**25**(1):28-34

[20] Seeley TD, Buhrman SC. Nest-site selection in honeybees: How well do swarms implement the 'best-of-N' decision rule? Behavioral Ecology and Sociobiology. 2001;**49**:416-427

[21] Seeley TD et al. Stop signals provide cross inhibition in collective decision-making by honeybee swarms. Science. 2012;**335**(6064):108-111

[22] Rosenberg L, Willcox G. Artificial swarm intelligence. In: Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys). Vol. 1. Springer International Publishing; 2020. pp. 1054-1070

[23] Seeley TD. Honeybee Democracy. Princeton Univ. Press; 2010

[24] Seeley TD, Visscher PK. Choosing a home: How the scouts in a honey bee swarm perceive the completion of their group decision making. Behavioral Ecology and Sociobiology;**54**(5):511-520

[25] Rosenberg LB. Human swarming, a real-time method for parallel distributed intelligence. In: 2015 Swarm/Human Blended Intelligence Workshop (SHBI). IEEE; 2015. pp. 1-7. DOI: 10.1109/SHBI.2015.7321685

[26] Rosenberg L, Pescetelli N, Willcox G. Artificial swarm intelligence amplifies accuracy when predicting financial markets. In: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON). IEEE; 2017. pp. 58-62

[27] Patel BN, Rosenberg L, Willcox G, et al. Human–machine partnership with artificial intelligence for chest radiograph diagnosis. NPJ Digital Medicine. 2019;**2**:111. DOI: 10.1038/s41746-019-0189-7

[28] Rosenberg L, Lungren M, Halabi S, Willcox G, Baltaxe D, Lyons M. Artificial swarm intelligence employed to amplify diagnostic accuracy in radiology. In: 2018 IEEE 9th Annual Information Technology Electronics and Mobile Communication Conference (IEMCON). IEEE; 2018. pp. 1186-1191

[29] Askay D, Metcalf L, Rosenberg L, Willcox D. Enhancing group social perceptiveness through a swarm-based decision-making platform. In: Proceedings of 52nd Hawaii International Conference on System Sciences, HICSS-52. IEEE; 2019

[30] Rosenberg L, Willcox G, Askay D, Metcalf L, Harris E. Amplifying the social intelligence of teams through human swarming. In: 2018 First International Conference on Artificial

Intelligence for Industries (AI4I), Laguna Hills, CA, USA. 2018. pp. 23-26. DOI: 10.1109/AI4I.2018.8665698

[31] Willcox G, Rosenberg L, Burgman M, Marcoci A. Prioritizing policy objectives in polarized groups using artificial swarm intelligence. In: 2020 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA). Victoria, BC, Canada: IEEE; 2020. pp. 1-9. DOI: 10.1109/CogSIMA49017.2020.9216182

[32] Rosenberg L, Baltaxe D, Pescetelli N. Crowds vs swarms, a comparison of intelligence. In: 2016 Swarm/Human Blended Intelligence Workshop (SHBI). Cleveland, OH, USA: IEEE; 2016. pp. 1-4

[33] Rosenberg L, Willcox G. Artificial swarms find social optima. In: 2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Boston, MA, USA: IEEE; 2018. pp. 174-178. DOI: 10.1109/COGSIMA.2018.8423987

[34] Rosenberg L, Willcox G, Schumann H, Mani G. Conversational swarm intelligence amplifies the accuracy of networked groupwise deliberations. In: 2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA: IEEE; 2024. pp. 0086-0091. DOI: 10.1109/CCWC60891.2024.10427807

[35] Parrish JK, Viscido S, Grünbaum D. Self-organized fish schools: An examination of emergent properties. Biological Bulletin. 2002;**202**(3):296-305

[36] Bronkhorst AW. The cocktail party phenomenon: A review on speech intelligibility in multiple-talker conditions. Acta Acustica United with Acustica. 2000;**86**:117-128 [Accessed: November 16, 2020]

[37] Rosenberg L, et al. Conversational swarm intelligence, a pilot study. arXiv.org. 2023. Available from: https://arxiv.org/abs/2309.03220

[38] US Patent 11,949,638. Methods and systems for hyperchat conversations among large networked populations with collective intelligence amplification

[39] Willcox G et al. Hyperswarms: A new architecture for amplifying collective intelligence. In: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE; 2021

[40] Graylin A, Rosenberg L. Our Next Reality. Nicholas Brealey Pub; 2024

[41] Rosenberg L, Schumann H, Dishop C, Willcox G, Woolley A, Mani G. Large-scale group brainstorming using conversational swarm intelligence (CSI) versus traditional chat. arXiv arXiv:2412.14205. 2024

[42] US Patent 12,231,383. Methods and systems for enabling collective superintelligence

[43] Marsh E, Vallejos EP, Spence A. Overloaded by information or worried about missing out on it: A quantitative study of stress, burnout, and mental health implications in the digital workplace. SAGE Open. 2024;**14**(3). DOI: 10.1177/21582440241268830

[44] Mark G, Gudith D, Klocke U. The cost of interrupted work: More speed and stress. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008. pp. 107-110

[45] Thompson, C. Slack is Ruining Work, Wired. Available from: https://www.wired.com/story/slack-ruining-work

[46] Marsh E, Vallejos EP, Spence A. The digital workplace and its dark side: An

integrative review. Computers in Human Behavior. 2022;**128**:107118

[47] Montrief T, Haas MR, Gottlieb M, Siegal D, Chan T. Thinking outside the inbox: Use of slack in clinical groups as a collaborative team communication platform. AEM Education and Training. 2020;**5**(1):121

[48] USPTO US2025001612. Methods and systems for enabling large-scale conversational deliberations among human groups and AI-powered conversational agents. 2023. DOI: 10.13140/RG.2.2.23689.35686

[49] Willcox G et al. Validating a new collective intelligence Technology for Accurate Ranking Using Artificial Swarm Intelligence. In: Proceedings of the Future Technologies Conference. Cham: Springer Nature Switzerland; 2023. pp. 290-304

[50] Jaiswal A, Arun CJ, Varma A. Rebooting employees: Upskilling for artificial intelligence in multinational corporations. In: Artificial Intelligence and International HRM. Routledge; 2023. pp. 114-143

[51] Gabsi AEH. Integrating artificial intelligence in industry 4.0: Insights, challenges, and future prospects–A literature review. Annals of Operations Research. 2024:1-28

[52] Farrow E. Determining the human to AI workforce ratio–exploring future organizational scenarios and the implications for anticipatory workforce planning. Technology in Society. 2022;**68**:101879

[53] Rosenberg L. Keeping up with AI: Why humanity must cultivate a 'hive mind'. Futurism. 2016. Available from: https://futurism.com/keeping-humanity-must-cultivate-hive-mind

[54] Schemmer M, Hemmer P, Kühl N, Benz C, Satzger G. Should I follow AI-based advice? Measuring appropriate reliance in human-AI decision-making. arXiv preprint arXiv:2204.06916. 2022

[55] Rastogi C, Zhang Y, Wei D, Varshney KR, Dhurandhar A, Tomsett R. Deciding fast and slow: The role of cognitive biases in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction. 2022;**6**(CSCW1):1-22

[56] Rosenberg L. The manipulation problem: Conversational AI as a threat to epistemic agency. In: 2023 CHI Workshop on Generative AI and HCI (GenAICHI 2023). ACM; 2023. DOI: 10.48550/arXiv.2306.11748

[57] Rosenberg L. Generative AI as a dangerous new form of media. In: Proceedings of the 17th International Multi-Conference on Society, Cybernetics and Informatics (IMSCI 2023). 2023. DOI: 10.54808/IMSCI2023.01.165