ABSTRACT

Evidence from a variety of domains suggests that the “wisdom of crowds”, the aggregated insights and judgments of large groups of non-experts, can generate results that are equal or better to those of experts. Traditionally, aggregating the insights of many has been a tedious and often infeasible task. Even the group decision support systems of a few years ago were not ready for this challenge. They would usually only facilitate small or medium sized groups and not easily scale. Recent developments in ICTs, however, have made the access to large groups of diverse individuals significantly more feasible, suggesting the feasibility of “group wisdom support systems” (GWSSs). Within this article we determine key requirements for a GWSS, based on the characteristics to be fulfilled for collective intelligence to successfully emerge. We then discuss specific system requirements for two types of systems, a Type 1 numeric belief aggregation system, and a Type 2 knowledge aggregation system. Type 1, represented for instance by prediction markets is relatively simpler, while Type 2, represented for instance by wiki-like mechanisms, is considerably more complex (not necessarily in the software, but in the processes that make the system work). Our article concludes with design suggestions.

Keywords: Group wisdom support system, Group decision support system, Knowledge management, Wisdom of crowds, Collective intelligence, Wiki, Prediction Market

INTRODUCTION

Using software to aggregate the ideas and opinions of a small or even mid-size group of people has long been organizational practice, e.g., via the use of decision support software, knowledge management systems, and related tools. Nevertheless, organizations continue to face significant challenges in effectively collecting and disseminating knowledge and information. For instance, organizations continue struggling to:

- counteract the notorious shortage in formalized expertise;
- cope with increased uncertainty through disruptive trends, such as the “Long Tail” [1] and “Wikinomics” [11] business paradigms, and the ever increasing variety in products, services and customer target groups.

How can organizations deal with these challenges? Recent insights in the ability of large collections of individuals (“crowds”) outside of organizational contexts to generate knowledge and insights, suggest an extension of the group decision support system (GDSS) concept. “Group wisdom support systems” GWSS, as we will call them in this article, offer an opportunity to harness a mostly untapped potential of collective intelligence. This type of intelligence manifests itself in numerous ways. One surprising manifestation (outside of the organizational context) is sports event predictions. Yahoo!, for instance, holds annual contests where experts and the public guess the outcome of NFL football games in weekly "Pick'em" rounds. Every week, each game participant guesses the outcomes of all games, beginning in the first week, until the end of the season. The Yahoo crowd's pick is an aggregation of the entire populations' guesses. Expert guesses are elicited and reported individually. Experts have an advantage over the crowd in that they can review the crowd's guess before placing their own. Nevertheless, the experts have a difficult time in performing as well as the crowd. For the entire 2005 season, the crowd received the absolute best result (beating all experts), while in 2007; it tied with the one best expert. (No results are available for 2006).

Crowd belief collection and aggregation is only feasible with information and communication technology (ICT). Numerous difficulties arise that cannot be solved otherwise. A system such as Pick'em, for instance, must be able to first attract a large number of participants. It must provide them an interface to report their beliefs. This data must be effectively stored and then mined, so as to generate meaningful data aggregates, which can then be
presented to the audience (users). While this brief description already highlights several system needs (e.g., data capture and warehousing and data mining), several unique system demands arise from the requirements of collective intelligence stressed by Surowiecki [10]. These need to be satisfied as well so as to generate both efficient and accurate results. We will look at the requirements of collective intelligence, and at the corresponding needed GWSS features in the following sections.

**BACKGROUND AND THEORY**

**Harnessing Wisdom of Crowds in the Enterprise Context**

In his book “The Wisdom of Crowds” Surowiecki states that companies have not yet recognized the value of bottom-up methods (of the kind of collective intelligence he describes) to solve difficult cognition problems. Such problems include deciding among potential new products to building new factories, forecasting demand, or contemplating mergers [10, p. 267]. The term wisdom is meant to contrast the available information, knowledge or intelligence against the expertise, sophistication or intelligence of a single person that is considered an expert. The term crowd signifies that decisions are based on the collective intelligence of many, as characterized above. In summary, the key point is, that “if you put together a big enough and diverse enough group of people … that group’s decisions will, over time, be ‘intellectually (superior) to the isolated individual,’ no matter how smart or well-informed he is.” [10, p. xviii]. The suggested reasons for the superior performance of crowds are better cognition, coordination of behaviors, and cooperative behaviors. The mechanisms Surowiecki describes rely heavily on information and communication technologies to aggregate the opinions of many into common, shared insights.

In order to distinguish the systems enabling crowd wisdom from conventional group decision support systems and to structure their design, we have to clarify the mechanisms and requirements of harnessing the “wisdom of crowds” in the interest of corporate strategy and decision making. Hence we begin with a clarification of the meaning of crowd.

The terms group and crowd are defined broadly by Surowiecki. The characteristics of a “crowd” can implicitly be concluded from the conditions he lists for group wisdom to occur. They are [10, p. 27]:

- diversity of opinion,
- independence, and
- decentralization.

**Diversity of opinion** refers to the availability of multiple viewpoints (ideally many), within the group, as each further viewpoint may help to explain the phenomenon better. **Independence** means that peoples’ opinions are not determined by the opinions of those around them, which is typically the case in homogeneous groups or teams. **Decentralization** points out that a number of people take part that can draw on their local and specific knowledge, in contrast to a single expert, and decision-power fully residing in a central decision authority. The latter also implies that there is some mechanism for transforming individual judgments into a single collective one. The existence of this mechanism is a fourth requirement for crowd wisdom, namely an aggregation mechanism that combines individual opinions or knowledge contributions into a collective decision or collective intelligence. Between diversity, independence, and decentralization on one hand, and aggregation on the other hand, exists a tension that needs to be managed, since heterogeneity of the crowd is desirable, yet leads to more aggregation difficulty and vice versa. In the enterprise context, we can differentiate three crowd categories with their different properties: enterprise internal groups, cross-boundary groups integrating e.g. business partners, customers and suppliers, and fully open, public crowds. Enterprise internal groups are limited to the size of the organization, and thus are likely smaller, and more homogeneous than other types of crowds. For example, when HP sought to capture crowd wisdom to forecast printer sales (using prediction markets), the number of participants was between twenty and thirty. Such relatively small crowds have to be better “controlled” so as to assure diversity, independence, and decentralization. Cross-boundary groups and fully open groups will be larger and thus more heterogeneous; however, participation and information aggregation will be a challenge.

With respect to the information or knowledge processed, we can identify two types of situations, dealing with highly structured (typically quantitative), or with semi/low structured (typically non-quantitative) information / knowledge. Many situations call for numeric insights, e.g., sales and revenue forecasts. In such situations, the aggregation mechanisms are well understood, automatable, and efficient. Other situations require the aggregation and integration of non-quantitative information and knowledge into meaningful aggregates. For example, creating a corporate policy, or help-desk guidelines belong to this task type. For such situations, there are no automatable procedures that assure the quality of the collectively created work product. Each crowd member might actively add their insights, but simply
adding more content to the knowledge base may not improve overall quality or usability of the collectively created knowledge asset, and in contrast, may lead to duplication, inconsistencies, and “information entropy”.

By crossing the different crowd types and information types, we obtain a 2x3 grid of crowd wisdom scenarios, as depicted in Figure 1. The figure identifies the different task situations and techniques presently used to support them.

### Degree of Openness

<table>
<thead>
<tr>
<th>Public Crowds</th>
<th>Cross-boundary Crowds</th>
<th>Enterprise Internal Crowds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting the popularity of emerging technologies (prediction aggregator)</td>
<td>Prioritizing future product features (web-polls with lead-user communities)</td>
<td>Predicting future sales revenue (prediction market)</td>
</tr>
<tr>
<td>Product flaw identification (blogs)</td>
<td>Customer-company support group (discussion forum, wiki)</td>
<td>Product innovation (“innovation jam sessions”)</td>
</tr>
</tbody>
</table>

Figure 1: Crowd Wisdom Grid

As outlined in Figure 1, enterprises already use technology solutions for most situations, albeit providing limited support. In highly structured information aggregation situations (we will call them from now on for simplicity Type 1), the information system deals largely with numerical information, and operates as a "belief aggregator". Examples of such systems include prediction markets (e.g., Newsfutures.com), or prediction aggregators (e.g., Predictify.com); also web-polling for votes falls into this category. Elliott Masie (Masie.com), for example, uses this approach to tap into the wisdom of the community of learning and knowledge professionals.

When the information or knowledge is relatively low in structure (from now on Type 2), the available technology needs to aggregate crowd contributions, so as to generate an integrated knowledge asset. This type of situation is at present supported by several Web 2.0 technologies. For example, members of the public write blogs, which are either linked and organized by other members of the public, or aggregated at some level through aggregators such as Technorati (Technorati.com), Digg (Digg.com), or Del.icio.us. A less individualistic mechanism is the joint participation in wikis (e.g., Wetpaint.com), where participants can actually re-write and re-organize others content so as to improve integration and overall knowledge quality. Another collaborative approach is what IBM calls “Jamming”. Recently the IBM InnovationJam (https://www.collaborationjam.com/get_started2006/index.shtml) invited ideas from several thousand IBMers in more than 160 countries. While this has been an enterprise internal approach, IBM recently has also invited outsiders, creating cross-boundary jams.

### Design Science Perspective

While individual solutions to support crowds exist, the objective of this article is to identify more general principles for the design of group wisdom support systems. GWSS are an emerging kind of system, designing and building of which is original work. A research method that strives to provide relevant results for practice while complying to rules of rigorous research is design science. Design science is defined as fundamentally a problem-solving paradigm. Hevner et al. [5] describe design science research’s performance in information systems via a concise conceptual framework and clear guidelines for understanding, executing, and evaluating the research. In the design science paradigm, knowledge and understanding of a problem domain and its solution are achieved in building and applying the designed artifacts, here a GWSS. According to the guidelines for effective design science, research requires the creation of an innovative, purposeful artifact (Guideline 1) for a specified problem domain (Guideline 2). Because the artifact is purposeful, it
must yield utility for a specified problem’s owner or holder. Hence, thorough evaluation of the artifact is crucial (Guideline 3). Novelty is also crucial since the artifact must be innovative, solving a currently unsolved problem, or solving a known problem in a more effective or efficient manner (Guideline 4). This therefore distinguishes design science research from the practice of design. The artifact itself must be rigorously defined, formally represented, coherent, and internally consistent (Guideline 5). The process through which it is created, and often the artifact itself, incorporates or enables a search process whereby a problem space is constructed and a mechanism posed or enacted to find an effective solution (Guideline 6). Finally, the results of the design science research must be communicated effectively (Guideline 7), both to a technical audience (researchers who will extend them and practitioners who will implement them), and to a managerial audience (researchers who will study them in context, and practitioners who will decide whether the artifacts should be implemented within their organizations). While the limitations of this article may only enable us to sketch a solution corresponding to these guidelines, any future research in this direction should certainly benefit from adherence to these guidelines.

Insights from GDSS Research

For more than two decades, group decision making processes have been supported through information technology. Group decision support systems differed vastly from earlier “information systems”. While such systems delivered an outcome, an optimized solution, or a set of processed transactions, GDSS focused on process. The underlying concept was to support a staged set of activities, first helping the group to generate many ideas through free thinking and association (commenting on the ideas of others), followed by idea categorization or other form of aggregation, and then selection of the best option among those generated [4; 9].

The underlying paradigm for supporting the process was simple, yet elegant. Group research had identified techniques for collective idea generation and solution finding that would lead to gains, but also losses. For example, anonymity resulted in more and better ideas, but also social loafing. Hence, a “support system” would need to be designed to amplify process gains and reduce process losses.

Systems to enable crowds to generate group wisdom need to adopt the same paradigm, as, once again, the system cannot automatically generate the right outcomes. As such, group wisdom support systems will need to target potential process gains and losses of large, heterogeneous groups, then amplify the gains, and reduce the losses. Gains and losses will arise in the tension between group heterogeneity, which is needed to generate a broad range of insights on one hand, and coordination, collaboration, and aggregation, which are needed to successfully shape the individual insights into a collective output.

From G-D-SS to G-W-SS

Group decision support systems are defined as “interactive computer-based environments that support concerted and coordinated team effort toward completion of joint tasks.” [8, p. 165]. Here the group has a sense of a common identity, and interacts intensely to seek consensus. In this type of groups, the groupthink-effect [6] indirectly works towards conformity in judgment and loss of idea diversity.

With group wisdom support systems, we seek to employ a different approach. First of all, the “group” does not have to be a group in the traditional sense [12], where the participants have gone through stages of group development (i.e., forming, storming, norming). On the contrary, diversity of thought (with participants maintaining private theories about the phenomenon and different information) are sought as beneficial. Large group size is therefore preferable, as it allows more diversity of thought. Furthermore contributions by each individual are usually small, and “light weight”, requiring little effort from any single participant. As a result, there is less need to manage traditional group losses (compare [9]), although more opportunity for coordination and aggregation losses. As a result, the process we find in GDSS is likely to be replaced by a more streamlined process of opinion capture and aggregation. At the same time, more emphasis must be placed on meta-processes such as identifying candidates for participation, contacting them, collecting their opinions, and later maintaining the group wisdom within the system. In contrast, GDSS often have little “memory” after the decision has been taken. Table 1 summarizes these characteristic differences.

<table>
<thead>
<tr>
<th></th>
<th>GDSS</th>
<th>GWSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>n x … n x 10</td>
<td>n x 10... n x 100,000</td>
</tr>
</tbody>
</table>
Group Wisdom Support Systems: Aggregating The Insights Of Many Through Information Technology

<table>
<thead>
<tr>
<th>Group composition</th>
<th>Homogeneity desired (performing group, [12])</th>
<th>Heterogeneity desired (crowd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information / knowledge sources for aggregation</td>
<td>ideas, preferences</td>
<td>beliefs (facts), “light knowledge”</td>
</tr>
<tr>
<td>System support</td>
<td>process</td>
<td>process, meta-process, generation of process meta-information</td>
</tr>
<tr>
<td>Group process</td>
<td>brainstorm, categorize, vote (or similar)</td>
<td>(\text{Type 1: Numeric belief aggregation (&quot;belief aggregator&quot;)} ) (\text{Type 2: Knowledge Aggregation (&quot;wiki&quot;) \rightarrow brainstorm, rewrite (&quot;Delphi&quot; [2]) vote} )</td>
</tr>
<tr>
<td>Information volume provided by individual participants</td>
<td>Considerable (multiple ideas and comments, votes)</td>
<td>(\text{Type 1: Little (typically votes)} ) (\text{Type 2: Broad range from almost none to very considerable} )</td>
</tr>
<tr>
<td>Collective Memory</td>
<td>Available, but typically unused</td>
<td>Historical database an important component</td>
</tr>
</tbody>
</table>

Table 1. Characteristic Differences between GDSS and GWSS

Participants. While GDSS traditionally handled small or medium size groups in synchronous or near-synchronous settings, group wisdom support systems need to scale to much larger groups. IBM’s jam sessions now involve tens of thousands of IBMers, while large open systems, such as Wikipedia engage over 100,000 contributors.

Group composition. While GDSS favor performing groups (largely to avoid process losses), GWSS favor heterogeneous groups, to broaden the range of opinions and insights.

Information / Knowledge Sources. Knowledge sources are relatively similar for both types of systems, yet GWSS focusing less on preferences and more on beliefs, and on individual contributions that can be very small and by themselves almost meaningless.

System support. Both types of systems are process systems, managing the collaborative processes. While GDSS generally also rely on a facilitator for monitoring and managing activities, GWSS need more system support and thus more information about the process (meta-information). Furthermore, they need to support processes outside of the group collaboration, such as identification of potential collaborators (meta-process).

Group process. GDSS, while differing in their design, usually include a divergent idea generation (brainstorming element), followed by convergent elements such as idea categorization and voting. GWSS may have similar components, but in particular need support for coordination and aggregation.

Information volume. GDSS participants generally produce a significant amount of information in a relatively short amount of time. Crowd members oftentimes contribute little information (e.g., the majority of Wikipedia contributors has only contributed once), and individual contributions maybe by themselves almost meaningless, yet meaningful in the aggregate.

Collective memory. GDSS keep internal databases where results of sessions are stored. Typically, however, the process information (the broad idea pool) is not processed further, and is not easily available past the interactive GDSS session. For GWSS we would expect the maintenance of a historic database, tracking past contributors and contributions, as for instance done in Wikipedia, or in various prediction markets. Such historical databases can then be mined for further knowledge and meta-information.

**DESIGN PROPOSITIONS FOR GWSS**

Requirements
The fundamental principles of crowd wisdom, as outlined above, define core characteristics of the GWSS, namely (1) enabling the creation of a
heterogeneous pool of beliefs, and (2) facilitating the aggregation of results. In addition, existing implementations of group wisdom suggest another important aspect, namely (3) the creation of meta-knowledge about the group wisdom process.

**Enabling the creation of a heterogeneous pool of beliefs** Surowiecki [10] identified three requirements for collective wisdom to emerge: diversity, decentralization of opinion, and independence. Taken together, all these criteria point to the same issue. Any GWSS software must enable access to/by a group of individuals from a variety of backgrounds, and must enable them to voice their opinion separately and independently from each other. Why so? Access to personal knowledge from a range of sources is important. After all, any crowd member’s “opinion” is expression of the outcome of a “theory” applied to the situation at hand. A broader range of “theories”, in the end, will explain more of the variance in a phenomenon, and thus lead to better results. Independence of ideas is equally important. Although the literature on prediction market research does not explain it, prior research on human judgment and biases [12] suggests that another individual's opinion can serve as an anchor, leading to only marginal adjustments in voicing one's opinion, thus reducing variance of opinions and possibly also biasing judgments. In essence, although crowd in this case potentially could have access to a broad range of knowledge, it does not take advantage of this resource if impendence is not maintained. In other words, independence maintains the breadth of knowledge that has been created by diversity and decentralization. Consequently, to avoid process losses, diversity, decentralization, and independence must be monitored and potential disturbances must be identified.

**Facilitating the aggregation of results.** The aggregation of results is simple when information consists only of highly structured (Type 1) knowledge or information. As such, aggregating the price opinions of many in a prediction market is relatively easily automatable and straightforward. The same is true for probabilities, probability surrogates, or other well structured guesses (e.g., “Who will win the 2008 Presidential elections?”). However, in Type 2 situations, where we need to aggregate relatively unstructured, non-quantitative information or knowledge, the complexity of aggregation increases. Wiki agreement is one such example. Let us consider for example the Wikipedia article on Kowloon (an urban area of Hong Kong). In an earlier version of the article, the origin of the word Kowloon was in question. One Wikipedian raised the uncertain belief that Kowloon may be derived from a story of “nine dragons”. It took a while until another Wikipedian confirmed this notion, and even longer until this belief moved from Wikipedia’s “talk pages” to the actual content pages. Essentially, wikis do not have a formal aggregation mechanism. And as such, aggregation remains a task of considerable difficulty, performed by only the minority of wiki users [7]. Effective and efficient aggregation of Type 2 crowd intelligence thus remains a key requirement for GWSS.

**Creation of Meta-Knowledge about the Group Process.** Coordinating the collaboration of many is a difficult task. While GWSS ignore the “small group problem“, motivating a small group of participants to be active and cooperative, they must deal with large group problems, and must do so with relatively little facilitation or “governance”. Hence, GWSS will need to signal system users information about the process, especially if they affect the process outcome. For example, if an unregistered user is making significant changes to a shared wiki, other members of the collective may want to be alerted to ensure that there is no vandalism. Alternatively, if the crowd participating in a vote is too homogeneous, the system should inform participants. Likewise, if some crowd members demonstrate exceptional capabilities, the system could also signal this.

**Features** The previous section outlines the essential requirements for GWSS, potential feature implementations of which we will discuss in this section.

**Communication and sampling component.** Enabling the opinion gathering from large groups in an organizational environment requires the availability of a communication component. In simple terms, one might consider this to be a communication client with an associated directory, which allows the knowledge seeker to quickly address a large group with a particular question. For better functionality, the system will contain a group directory with adjustable characteristics and dynamic updating capability. Specifically, the directory should have members listed with an “expertise inventory” identifying their specific subject matter knowledge. The client could have various characteristics, including “unstructured email”, wiki or weblog access, and structured voting, depending on the information requirement. Based on the inputs, the system would be able to draw samples for subsequent aggregation.
**Analysis component.** The analysis component would aggregate the range of inputs. As is common in online polling, such a component should, at least for numerical information, seek input from a user, before disclosing aggregate results, so as to maintain independence as much as possible. Clearly this will be a need for Type 1 intelligence, which is predominantly aggregated through mathematical operations such as summary, average, majority/minority, trend, or probability revision (e.g., Bayes). For Type 2 intelligence aggregation, the task is more complex. We must envision a GWSS with a wiki-like component where aggregation is largely through addition, but where controversial (uncertain) issues can also easily be put to vote, thus breaking Type 2 knowledge into non-controversial issues where addition is mostly sufficient, and controversial issues which are formalized to be settled via Type 1 components. Clearly the solution requires some more complex aspects (e.g., “refactoring” of content to remove ambiguities and such). Nevertheless, the combination of allowing aggregation through addition for non-controversial issues and aggregation through quantitative means for disputed or uncertain issues offers a good starting point for subsequent further enhancements.

**Meta-analysis component.** Existing information systems that aggregate the knowledge of many already compile information beyond that which is the purpose of the information system itself. For example, wikis typically keep records of past changes, authors, volumes of change, and such, thus allowing the determination of most active contributors, and so on. Similarly, prediction markets capture all trades as transactions, and “keep score”, thus accumulating assets for those who frequently predict right. In other words, these types of software extract knowledge from the contribution which can then be used to adapt future contributions. Past Wikipedia contributors with a very active record may later become editors. Past prediction market “winners” accumulate more assets, thus giving them more votes in future markets. Consequently, it appears important that a GWSS records not only information and knowledge about the topic of discovery, but also about the process and its contributors, so as to tweak the process in future towards further efficiency and effectiveness. We would envision further features such as the computation and reporting of “diversity indices”, so as to allow the decision maker to know how broad the input base is. Furthermore, analyzing sources through a social network analysis tool, the system could determine degrees of separation between those inputting knowledge and thus determine, or at least alert to, potential challenges to independence. Expertise inventories should be dynamically updatable, either based on member inputs, or findings from previous rounds of assessment, thus enabling access to a broad, diversified base of personal knowledge.

**DISCUSSION AND CONCLUSIONS**

In this article we offered guidelines for a new category of software, group wisdom support systems. The essential purpose of this software type is the aggregation of the beliefs of many (non-experts) to derive judgments that can beat the predictions of experts and relieve the reliance on this scarce resource. While our proposal of GWSS is by no means complete, it differentiates GWSS from GDSS and highlights changes that must be made to the architecture so as to fulfill the promise of crowd wisdom. Our arguments identify some surprising issues. Specifically, the architecture of GDSS may favor building on others’ ideas in collective brainstorming, yet is counter-productive to the collection of independent thoughts. This surprising finding may be in part an explanation for the fact that GDSS research found nominal groups to perform equally as well or better than actual groups in brainstorming experiments [3; 11]. Our argument also points out that wiki technology, which is currently a favored technology for the aggregation of the knowledge of many, is without formal mechanisms for aggregation. It is thus highly dependent on the effort of individuals to carry out this complex cognitive task without tools to assist their work. We suggest the addition of voting or polling mechanisms to such software so as to formalize knowledge aggregation on complex, undecided issues. In addition, our research suggests the incorporation of meta-knowledge components to improve predictive quality. Again, the implications are potentially controversial. While the aggregation of human preferences typically adopts the notion of “one person one vote”, we realized that when aggregating beliefs (facts believed to be true), the same may not be true. Prediction markets, for instance, give more “votes” (cash resources) to those whose past predictions have been true. Thus, those who predict with higher accuracy subsequently gain more votes. While the prediction market mechanism, where someone can use the cash earnings from predictions in one domain to weigh in heavily on the predictions in another domain may not be ideal, a weighted prediction mechanism may be desirable, and thus the cause for the creation of “meta information”.

There is much more to learn about GWSS. Our
Group Wisdom Support Systems: Aggregating The Insights Of Many Through Information Technology

This article only covers the phenomenon at a very broad level, but it also outlines areas that offer an opportunity for more formal empirical investigation and systems development. We hope that by identifying GWSS as a new class of information systems, it will help other researchers and practitioners to better recognize the shortcomings of existing systems used to support collective intelligence tasks and thus lead to the development of more purposeful GWSS.

REFERENCES