Frameworks for Collective Intelligence: A Systematic Literature Review

Article *in* ACM Computing Surveys · February 2020

DOI: 10.1145/3368986

3 authors:

Shweta Suran
Tallinn University of Technology

Vishwajeet Pattanaik
Indian Institute of Science

Dirk Draheim
Tallinn University of Technology

Frameworks for Collective Intelligence: A Systematic Literature Review

SHWETA SURAN, Tallinn University of Technology, Estonia
VISHWAJEET PATTANAIK, Tallinn University of Technology, Estonia
DIRK DRAHEIM, Tallinn University of Technology, Estonia

Over the last few years, Collective Intelligence (CI) platforms have become a vital resource for learning, problem solving, decision making and predictions. This rising interest in the topic has to lead to the development of several models and frameworks available in published literature. Unfortunately, most of these models are built around domain-specific requirements, i.e., they are often based on the intuitions of their domain experts and developers. This has created a gap in our knowledge in the theoretical foundations of CI systems and models, in general. In this paper, we attempt to fill this gap by conducting a systematic review of CI models and frameworks, identified from a collection of 9,418 scholarly articles published since 2000. Eventually, we contribute by aggregating the available knowledge from 12 CI models into one novel framework and present a generic model that describes CI systems irrespective of their domains. We add to the previously available CI models by providing a more granular view of how different components of CI systems interact. We evaluate the proposed model by examining it with respect to six popular, ongoing CI initiatives available on the web.

CCS Concepts: • **Human-centered computing → Collaborative and social computing theory, concepts and paradigms**; **Collaborative and social computing systems and tools**; *Human computer interaction (HCI)*; • **Information systems → Crowdsourcing**; **Collaborative and social computing systems and tools**; • **General and reference → Surveys and overviews**.

Additional Key Words and Phrases: Collective intelligence, Crowdsourcing, Human computer interaction, Web 2.0, Wisdom of crowds, Systematic literature review

ACM Reference Format:
Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2019. Frameworks for Collective Intelligence: A Systematic Literature Review. *ACM Comput. Surv.* 1, 1, Article 1 (January 2019), 35 pages. https://doi.org/10.1145/3368986

1 INTRODUCTION

The concept of 'Collective Intelligence' (CI) (i.e., collaborative problem solving and decision making) has been a keen interest of researchers ever since the 18th century [41, 63]. Since this period, the different applications of CI and its associated concepts have extended throughout a wide spectrum of research domains ranging from sociology, psychology, biology, management, economics to computer science among many others [50]. In our work, we focus on CI in Information and Communications Technology (ICT), and therefore, we adhere to the widely-accepted formal definition of CI in the ICT domain, proposed by Pierre Lévy in 1995 [43]. Levy defined CI as a "form of universally

Authors' addresses: Shweta Suran, shweta@taltech.ee, Tallinn University of Technology, Akadeemia tee 15a, Tallinn, 12616, Estonia; Vishwajeet Pattanaik, vishwajeet.pattanaik@taltech.ee, Tallinn University of Technology, Akadeemia tee 15a, Tallinn, 12616, Estonia; Dirk Draheim, dirk.draheim@taltech.ee, Tallinn University of Technology, Akadeemia tee 15a, Tallinn, 12616, Estonia.

distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills” [43]. Some of the CI platforms of the early period include WikiWikiWeb, Experts-Exchange and Google [50]. Since then, advancements in ICT technologies like Web 2.0 [65, 71], Semantic Web [28, 44] and Crowdsourcing [7, 17] have enabled and drastically eased large-scale collaborations over the Internet; leading to the development of well-known CI platforms like WaterWiki [16, 62], Climate CoLab [34, 51], DDtrac [26, 27], WikiCrimes [19, 68] and Goldcorp [4], which facilitate knowledge sharing, problem-solving and decision making among individual users and groups, through web-based interactions and collaborations.

The success of these systems can be credited to their underlying architectures or frameworks (hereinafter referred to as ‘models’). Unfortunately, most of these models are often defined using system-specific elements, principles, attributes, requirements or their combinations [39]; and are based on specific problems [21]. Since each of these CI systems is designed for a specific problem or use-case, the models proposed for these systems are often presented as completely different entities. However, comparing these models shows that although each new CI system and model expands on our current understanding of CI, nevertheless many of these systems bear a few similarities [48]. Sadly, this abundance of diverse knowledge has not yet lead to the development of a unified CI model [13, 56, 67] that can support the development of new CI systems based on systematic knowledge rather than intuition [39]. Also, many of the existing CI systems are proprietary and are therefore not available in scientific literature. And, systems that are described in scientific literature, focus more towards the theoretical foundations, usability, and future applications of collective intelligence [21], rather than focusing on the implementation [39]. This lack of well-defined and systematic knowledge about the architecture and principles of the underlying CI systems has led to a reproducibility crisis.

In order to achieve comprehensive knowledge of CI systems, it is imperative that we extensively investigate published scientific literature irrespective of the so-called proposed models. We are convinced that although different CI systems are defined in different ways, they must share more than just a few common characteristics. And, identifying these characteristics could help us to achieve a unified formal model for designing CI systems, irrespective of their application. To this end, we contribute by conducting a first of its kind Systematic Literature Review (SLR) of Collective Intelligence models. In this SLR, we extensively investigate the characteristics of 12 CI models, selected from a pool of 219 scientific publications. And, based on the results of our review we develop a novel framework that can be utilized to understand existing CI systems. The proposed framework provides a generic model and a set of requisites that would enable creation of novel CI systems, regardless of their domains. This is achieved by exhaustively combining all attributes of the studied CI models into the proposed framework.

Additionally, to better explain the functioning of CI systems with respect to the proposed framework, we examine the different components of six ongoing CI projects: CAPSELLA, hackAIR, openIDEO, Climate CoLab, WikiCrimes, and Threadless.

In particular, we aim at answering the following research questions:

RQ1: What are the underlying models of existing CI systems? What are the common terminologies used to describe CI models? What are their components? And, how are these components associated to each other?

RQ2: Do any of the available CI models appropriately define all CI systems, irrespective of their applications? Can these models be used to create CI systems for novel challenges?

RQ3: If not, then can we somehow combine the available knowledge of CI models and systems to create a unified model that could define all CI systems?

The paper is structured as follows. In Section 2, we describe the research methodology used for conducting this SLR. Section 3 presents a brief summary of the selected studies, followed by the aggregated list of terminologies used to describe CI systems in Section 4. In Section 5, we present a novel framework for CI systems and evaluate our generic theoretical CI model by means of comparative case studies in Section 6. Finally, Section 7 presents the threats to validity of the SLR and Section 8 concludes by summarizing the key findings of this article and provides insights for future research.

2 RESEARCH METHODOLOGY

To answer the research questions mentioned in Section 1 through a transparent and objective approach, we decided to conduct this review based on Kitchenham's "Guidelines for performing Systematic Literature Reviews in Software Engineering" [37]. A Systematic Literature Review summarizes, critically appraises and identifies valid and applicable evidences in available research by using explicit methods to perform thorough literature search [9, 37, 66]. Based on Kitchenham's guidelines, we perform this SLR in five stages:

(1) Search Strategy
(2) Study Selection
(3) Study Quality Assessment
(4) Data Extraction
(5) Data Synthesis

2.1 Search Strategy

Based on the previously identified research questions, we selected a set of search terms. We then used the combination of these search terms to look for relevant research articles in different academic databases. After this, we applied the inclusion criteria on the identified articles and shortlisted the most relevant articles (which we refer to as 'Primary Studies'). Following Kitchenham's guidelines, we then evaluated the primary studies using the quality assessment criteria. And finally, the selected studies were investigated in the data extraction and synthesis stages of the SLR.

2.1.1 Search Terms. As researchers often use the terms 'Crowdsourcing' and 'Wisdom of Crowds' as synonyms for Collective Intelligence [39, 64], we decided to use all three keywords as the primary search terms. And, for the secondary search terms, we used keywords such as model, framework and others that are commonly used to describe ICT systems. In order to construct the search string we used the following guidelines provided by Kitchenham [37]:

(1) Derive search terms from research questions and from initial literature review.
(2) Identify synonyms for search terms from scientific literature.
(3) Use the Boolean 'AND' and 'OR' to link search terms and their synonyms.

The list of identified primary and secondary search terms, and the resulting search string are presented in Table 1.

2.1.2 Academic Databases. The resulting search string was used to search for pertinent articles in four different databases, namely *(ACM Digital Library*, *IEEE Xplore*, *ScienceDirect (Elsevier)* and *Springer)*. The search was restricted to articles published since the year 2000; because, the first popular Web 2.0 based CI platform 'Goldcorp' and 'Threadless'

Table 1. Search terms identified based on research questions

Primary Search Terms	collective intelligence, wisdom of crowds, crowdsourcing
Secondary Search Terms	model, framework, architecture, requirements, principles, attributes, properties
Search String	("collective intelligence" OR "wisdom of crowds" OR "crowdsourcing") AND ("model" OR "framework" OR "architecture" OR "requirements" OR "principles" OR "attributes" OR "properties")

were launched in the same year [18, 52, 83]. It was only after this period that CI systems became popular and were recognized as a significant area of research in ICT. In order to identify relevant books, technical reports and theses, we also conducted a manual search on Google Scholar.

2.1.3 Search Process. During the search process we found that many of the databases indexed each others' articles, therefore the chances of getting redundant results were high. Thus to avoid duplicate results, we manually selected different options (like: Publication Type, Publisher, etc.) while searching through each database. In total 9,418 articles were identified after removing 430 redundant articles. Table 2 presents the number of relevant articles identified from each academic database.

Table 2. Search results

Year	Database	Total Count
2000 - 2017	ACM Digital Library (proceedings, journals, newsletters, and magazines)	1,289
2001 - 2017	IEEE Xplore (conferences, early access articles, journals, magazine and books)	1,214
2000 - 2017	SpringerLink (from sub-discipline 'Information Systems Applications incl. Internet': chapters, conference papers, articles and books)	3,196
2000 - 2017	ScienceDirect (reviews, research articles and books)	4,096
1997 - 2018 *Manual Search*	Google Scholar (research articles, books, reports and thesis)	53
	Total	9,848
	Total (after screening)	9,418

2.2 Study Selection

To identify the articles relevant to our research questions, we applied a two-phase selection process. During this process two researchers of this review independently analyzed the identified articles and selected the studies, which were most likely related to our research questions.

2.2.1 Selection Phase 1. In this phase we studied the titles and abstracts of the identified articles and assessed them on the basis of the inclusion criteria listed in Table 3. After completion of this phase 216 primary studies (PS) were selected.

We then scanned the reference list of the selected primary studies to identify related articles that we might have missed during our initial search. We found 3 articles which passed our inclusion criteria, and therefore we added these articles to our list of primary studies; making a total of 219 articles (see Table 2 in Supplementary Material).

Table 3. Inclusion criteria for Selection Phase 1

Criteria ID	Inclusion Criteria
IC1	The article explains the theoretical foundations of collective intelligence in computer science
IC2	The article describes the role of collective intelligence in crowdsourcing and open innovation
IC3	The article focuses on architecture/frameworks of CI systems
IC4	The article describes CI systems/applications available on the Web
IC5	The article focuses on knowledge generation and exchange in crowds
IC6	The article is related to at least one aspect of our research questions
IC7	The article should not compare collective intelligence with swarm intelligence and artificial intelligence

2.2.2 Selection Phase 2. In this phase we applied the quality assessment criteria illustrated in Table 4 to the primary studies selected in Selection Phase 1. After completion of this selection phase, 12 primary studies were finally selected. These 12 studies were then used for data extraction and data synthesis. We describe both stages further in the next sections.

Table 4. Quality assessment criteria for Selection Phase 2

Criteria ID	Quality Criteria Check-List
QC1	Are the research objectives clearly defined in the study?
QC2	Does the study propose a new framework or, provide technological details of an existing CI system?
QC3	Is the system architecture/framework/design/experiment clearly defined in the study?
QC4	Is the proposed CI architecture or framework compared to existing CI models or systems?
QC5	Does the study provide insights about the role, importance and behaviour of individuals in the proposed CI system or model?
QC6	Does the study propose novel solutions to crowd management issues in CI?

2.3 Study Quality Assessment

The intention of this phase is to determine the relevance of selected studies while limiting bias in the study selection process. In this phase, all three researchers of this review independently assessed the primary studies by answering the questions presented in Table 4. For each primary study, the researchers answered the questions as ‘Yes’, ‘Partly’, or ‘No’; scoring each criteria as 1, 0.5 and 0 respectively. The individual scores for each question were then added to derive a total score for each primary study. The studies that scored 3 or higher were finally selected for the data synthesis stage. Any conflict of opinion about the process and results of the quality assessment measures were discussed among all three researchers to reach a consensus. The scores of the remaining 12 primary studies that satisfied the quality assessment criteria are presented in Table 5; followed by the title of the studies and the publication type presented in Table 6.

Table 5. Quality score of Selected Studies

Primary Study ID	QC1	QC2	QC3	QC4	QC5	QC6	Total Score	Selected Study ID
PS1	1	1	0.5	0	0.5	1	4	S1
PS19	1	0	0.5	0	1	1	3.5	S2
PS48	1	1	0.5	0.5	0.5	0	3.5	S3
PS73	1	1	1	1	1	0	5	S4
PS108	1	1	0.5	0	0.5	0	3	S5
PS138	1	0	1	1	0.5	0	3.5	S6
PS154	1	0	1	1	0	0	3	S7
PS155	1	0	1	1	0	0	3	S8
PS156	1	1	1	1	0	0	4	S9
PS173	1	1	1	1	1	0	5	S10
PS174	1	0.5	0.5	0	0.5	0.5	3	S11
PS204	1	1	1	1	0	1	5	S12

Table 6. List of final selected studies

Study ID	Study Title	Publication Type
S1	"Intelligent Collectives: Theory, Applications and Research Challenges" [58]	Journal Article
S2	"Leadership and the Wisdom of Crowds: How to Tap into the Collective Intelligence of an Organization" [53]	Journal Article
S3	"Modelling the Index of Collective Intelligence in Online Community Projects" [73]	Conference Paper
S4	"The Role of Collective Intelligence in Crowdsourcing Innovation" [67]	PhD Thesis
S5	"Collective Intelligence Model: How to Describe Collective Intelligence" [21]	Conference Paper
S6	"Collective Intelligence Systems: Classification and Modeling" [48]	Journal Article
S7	"Designing for Collective Intelligence" [27]	Journal Article
S8	"On Model Design for Simulation of Collective Intelligence" [70]	Journal Article
S9	"A Resource Allocation Framework for Collective Intelligence System Engineering" [81]	Conference Paper
S10	"Harnessing Crowds: Mapping the Genome of Collective Intelligence" [52]	Journal Article
S11	"Leveraging the Power of Collective Intelligence through IT-enabled Global Collaboration" [32]	Journal Article
S12	"Collective Intelligence: a Keystone in Knowledge Management" [3]	Journal Article

2.4 Data Extraction and Synthesis

The intention of data extraction stage is to identify the main contributions of the selected studies, and to present a summary of the work. Table 7 presents the data items extracted from the 12 selected studies and Section 3 presents a summary of the same. The contributions i.e., models and elements of the Selected Studies are presented in Table 16.

The goal of the data synthesis stage is to collate and summarize the contributions of the selected studies. In Section 4 we first catalogue the definition types and classifications of the studied CI models; we then identify all unique and synonymous *characteristics*, *levels*, *requirements*, *properties* and *building blocks* and classify them into 24 distinct attributes

Table 7. Data extracted from Selected Studies

Extracted Data Item	Description
Study Title	see Table 6
Author(s)	see Table 2 in Supplementary Material
Year	see Table 2 in Supplementary Material
Publication Title	see Table 2 in Supplementary Material
Publication Type	see Table 6
Source/Publisher	see Table 2 in Supplementary Material
Summary	see Section 3

(presented in Table 16). Finally, based on these findings we then answer the first two research questions (RQ1 and RQ2) in Section 4 and the final research question (RQ3) in Section 5.

3 SUMMARY OF SELECTED STUDIES

3.1 S1 (Van Du Nguyen et al. 2018)

The aim of this study is to define the criteria necessary for a collective to be intelligent. To do so, the study [58] presents a novel general CI framework based on crucial attributes of a collective.

Influenced by Bonabeau's [4] concept of Decisions 2.0, which is defined as "a new era of decision-making in which the traditional decision making process is supported using the wisdom of crowds through collaboration and collective intelligence" [58], Nguyen et al. state that "collective intelligence is considered as the power of Decisions 2.0" [58]. Based on this premise, the study proposes a CI framework based on characteristics vital for a intelligent collective, as proposed by Surowiecki [77]. According to Nguyen et al., a collective must fulfill four criteria (presented in Table 8) [58] to be intelligent. And, based on these characteristics the authors propose a general framework of CI (namely, *Collective*, *Aggregation Methods* and *Collective Performance Measures*) [58] for wisdom of crowds.

Table 8. Criteria for collective to be intelligent (as presented by Nguyen et al. in S1) [58]

Criteria	Description
Diversity	Individuals must belong to diverse backgrounds, knowledge bases, etc.
Independence	Freedom for individuals to act according to their choice, without others influence
Decentralization	Facilitate individualism and assure diversity in individuals
Aggregation	Appropriate methods to integrate individual solutions [58]

3.1.1 Diversity. A collective must be diverse, as a heterogeneous group of individuals can provide new knowledge and diverse viewpoints to any given problem. Nguyen et al. further categorize diversity as: "diversity in the composition of collective members" [58] and "diversity of individual predictions in a collective" [58]. To explain diversity, the authors used an example of weather forecasting; where accurate weather predictions are a difficult task even if relying on experts. The authors claim that such prediction problems could be solved more easily if multiple individuals were allowed to add extra information and provide different perspectives to solve the problem [58].

3.1.2 Independence. The individuals in a CI system must be allowed to provide their own inputs and their decisions should not be influenced by others [58, 77]. This is important, because information cascades can diminish the intelligence of the collective [1].

3.1.3 Decentralization. This criteria helps individuals act independently, while avoiding others' influence; and thus ensures diversity [58]. To explain this, the authors used the example of Linux, where solutions to specific problems are selected from a pool of solutions submitted by independent programmers from around the world.

3.1.4 Aggregation. This criteria provides the appropriate mechanism to integrate the opinions and solutions provided by the individuals [58]. Examples of such new aggregation methods include prediction markets [82] and social tagging [86].

3.2 S2 (Kurt Matzler et al. 2016)

The aim of this study is to present activities necessary to promote collective intelligence within organizations. The proposed activities are based on the work of Surowiecki [77] and are explained using case studies and real world examples [53].

In this study Matzler et al. argue that although platforms such as wikis, blogs, prediction markets etc. might be enough to harness the wisdom of the crowd from end users; however, such platforms are inadequate to support collective intelligence within organizations. The authors propose that in order to harness the power of collective intelligence within organizations, it is imperative that managers follow the following steps: "create cognitive diversity" [53], "promote independence" [53], "access decentralized knowledge" [53] and "effectively aggregate knowledge" [53].

3.2.1 Cognitive Diversity. To explain cognitive diversity Matzler et al. refer to the work done by Scott E. Page [61]. Page states that cognitive diversity can be explained as combination of "diverse perspectives" [53], "diverse interpretations" [53], "diverse heuristics" [53] and "diverse predictive methods" [53]. Matzler et al. explain the relevance of these attributes in organizations by the means of two case studies; namely "How diversity can drive innovation" [30] and "The CEO's role in business model reinvention" (a case study from Infosys Technologies Limited) [24].

3.2.2 Promote Independence. Matzler et al. emphasize the importance of this step, by explaining how lack of independence or peer pressure may force employees to convey incorrect or sugar-coated information to their managers; which may lead to biased decisions [53]. The authors suggest that managers should create an atmosphere of open dialogue where all employees can share their honest opinions and ideas; the authors recommend techniques like the PreMortem exercise [38] to create such an independent environment.

3.2.3 Access Decentralized Knowledge. In regard to this step Matzler et al. state how, in the past, knowledge was organized hierarchically in organizations; where as now, due to globalization, decentralization and data ubiquity, knowledge within organizations is not limited to the organizations themselves [53]. In other words, when looking for novel solutions and ideas, organizations now rely heavily on participants via online contests, social media platforms, blogs and wikis [74]. Matzler et al. argue that organizations could boost their internal collective intelligence by allowing their employees to tap into this decentralized knowledge aggregated from the social web [53]. Employees could then use this knowledge to come up with ideas and solutions to support the organization's growth, while being aligned with the organization's vision and mission.

3.2.4 Effectively Aggregate Knowledge. The final step for promoting collective intelligence within organizations is to effectively aggregate dispersed knowledge. In this study, Matzler et al. briefly discuss techniques (such as averaging individual opinions) that could be utilized to aggregate knowledge from different sources [53]. The authors further describe this step using the examples of predictive markets and peer review systems, which have been found to be effective knowledge aggregation techniques [29]. Lastly, Matzler et al. discuss the effectiveness of Wikipedia's peer review system [53], by comparing the accuracy of its knowledge base to that of Britannica, as investigated by Jim Giles [22].

3.3 S3 (Aelita Skarzauskiene et al. 2015)

The aim of this study is to propose measures to quantify the minimum potential required by community projects, necessary to transform them into CI systems. The authors do so by investigating the trends in engagement and participation in online communities in Lithuania. Skarzauskiene et al. conduct both qualitative and quantitative research by extensively interviewing 20 individuals and by conducting a public opinion survey with 1022 Lithuanian participants between the age of 15 and 74 [73]. Finally, the authors propose three levels/measures a community project must fulfill in order to be considered as a CI system [73].

Before conducting qualitative and quantitative research, Skarzauskiene et al. briefly analyzed several CI frameworks proposed by researchers. Based on the literature, Skarzauskiene et al. proposed a conceptual framework for assessing the potential of CI [73]. The authors define the proposed conceptual framework in three levels, presented in Table 9 [73]. Using the proposed levels, combined with results of qualitative and quantitative analysis, the authors calculate a CI Potential Index; which they claim could assist developers and initiators of community projects by helping to assess the CI potential of such projects [73].

Table 9. Levels for assessing CI potential (as presented by Skarzauskiene et al. in S3) [73]

Level	Description
Capacity Level	Describes possible user actions, both as individual and as a member of the community [48]. It also includes massive participant interactions [47] that promote knowledge creation and innovation [3].
Emergence Level	Describes a system state [48] that supports self-organizing, "emergent" behavior, "swarm effect" [47] and mechanic development [3].
Social Maturity Level	Describes the clarity of system goals [3] and community/individual objectives [48].

3.4 S4 (Juho Salminen 2015)

The aim of this doctoral thesis is to explore the role of collective intelligence in crowdsourcing innovations [67]. Salminen's work is motivated by the fuzzy nature of collective intelligence, that has lead to different interpretations of the concept including 'wisdom of crowds' [77] and 'swarm intelligence' [5]. In his work, Salminen attempts to defuzzify the notion of collective intelligence, by investigating its emergence as a complex-adaptive system [67].

To do so, the author conducted a systematic literature review of published case studies discussing three CI platforms (OpenIDEO, Quirky and Threadless). He then observed user behaviour on each of these platforms for over a month, and gathered relevant data including web clips, diary entries and statistics. Salminen also conducted a literature review of available CI frameworks, based on which he proposed a new theoretical framework for CI. Finally, he evaluated the proposed framework based on his observations from the previously analyzed CI platforms. Salminen defines the proposed framework through three levels of abstraction [67]:

- *Micro*: “enabling factors of collective intelligence”
- *Emergence*: “from local to global”
- *Macro*: “output of the system and wisdom of crowds” [67]

Table 10 presents the elements of Salminen’s proposed theoretical framework based on themes from literature. Apart from the proposed theoretical CI framework, Salminen also highlights the crucial issue of biased feedback. When observing the previously mentioned CI platforms, the author found that participants would often create multiple accounts to vote up their own ideas and would demoralize their competitors by providing negative/incorrect feedback and down-votes. Salminen states that to prevent such issues, researchers must create measures to evaluate the accuracy of crowd decisions [67].

Table 10. Themes and elements of the CI theoretical framework (as presented by Salminen in S4) [67]

Level	Theme	Elements of the theoretical framework	References
Micro	Humans as social animals	Human capabilities for interaction *	
	Intelligence		
	Personal interaction capabilities		[12, 35, 85]
	Trust		[2]
	Motivation		
	Attention		
	Communities [67]		[8, 10]
Emergence	Complex adaptive systems	Agents, activities, feedback, emergence	[60]
	Self-organization	Agents, activities, feedback	[36]
	Emergence	Emergence	[14]
	Swarm intelligence	–	
	Stigmergy	Agents, activities, feedback, distributed memory	[78]
	Distributed memory [67]	Distributed memory	[6]
Macro	Decision making	Output	
	Wisdom of crowds	Output	[77]
	Aggregation	–	
	Bias	–	
	Diversity	–	[31]
	Independence [67]	–	

* Same for all themes in micro level.

3.5 S5 (Sandro Georgi et al. 2012)

The aim of this study is to build a comprehensive model based on available literature while recognizing the characteristics that describe collective intelligence [21].

Georgi et al. draw attention to a very important issue in the field of collective intelligence, i.e., that research about the topic in general is very limited, as most available research is application- and type-specific. The authors state that although numerous scientific articles and reports have been published about CI platforms, frameworks and models, only little research has been done on “how to describe collective intelligence in general” [21]. To fill this gap the authors first studied the existing scientific literature and choose three models of collective intelligence, namely “the collective intelligence genome” by Thomas W. Malone et al. [52], “mitigating biases in decision tasks” by Eric Bonabeau [4] and "the collective intelligent system" by Ioanna Lykourentzou et al. [49]. Combining these three models, the authors propose five novel characteristics and argue that these can appropriately describe collective intelligence. Table 11 presents these characteristics and their descriptions as stated by Georgi et al. [21].

Table 11. Characteristics that define collective intelligence (as presented by Georgi et al. in S5) [21]

Characteristics	Description
Objective of task	Can be described as the outcome that the collective intelligence intends to achieve. These objectives can be categorized as 'create' (creation of knowledge or ideas or physical objects) and 'decide' (correctness or best or most suitable, respectively).
Size of contribution	Represents the amount or volume of contribution, and can vary depending upon the complexity of the problem and form/structure of the collective intelligence.
Form of input	Can be presented in form of rules or data/information (pictures, text, data-sets etc.); and can be categorized as: instructions, challenge descriptions or raw material.
Form of output	Can be of two types: knowledge (i.e., intangible) or products (i.e., tangible).
Stakeholder	Defines stakeholders of a CI system based on their roles. 'initiators' are those whose objective is to reach a desired goal. 'contributors' do the actual work and use their intelligence to provide solutions. Finally, 'beneficiaries' are those who profit from the outcomes of such systems. [21]

3.6 S6 (Ioanna Lykourentzou et al. 2011)

This study aims to design a modelling process that can identify the common characteristics of CI systems. Additionally, the process helps to identify challenges that prevent the construction of a generic CI system [48].

Lykourentzou et al. claim that their work is the first attempt in classifying the common shared characteristics of CI systems. The authors state that although all CI systems may seem to be substantially different from each other, they all seem to share quite a few characteristics. After analyzing published literature on collective intelligence, Lykourentzou et al. proposed that all CI systems could be categorized as either 'active' or 'passive' systems. Additionally, 'active' CI systems could further be classified into 'collaborative', 'competitive' or 'hybrid' systems [48]. The authors suggest that in 'passive' CI systems, groups of users would exhibit behavior of swarms, irrespective of whether the system requires such a behaviour or not. Whereas in 'active' systems, crowd behavior is created and coordinated by the system itself [48].

Lykourentzou et al. further state that based on this classification, CI systems have several common attributes (described in Table 12) [48]. The authors also highlight issues of 'critical mass', 'task and workload allocation' and 'motivation' that should be considered when designing CI systems. Finally, Lykourentzou et al. model three types of CI systems (*Collaborative*: Wikipedia and open source software development communities; *Competitive*: Innocentive, BootB, DesginBay, DARPA Network Challenge; *Passive*: vehicular ad-hoc networks) using the previously identified attributes [48].

3.7 S7 (Dawn G. Gregg 2010)

The aim of this study is to demonstrate the requirements for designing CI applications. Gregg states that a CI application harnesses the knowledge of its users by facilitating human interaction and decision making; and therefore, new CI applications must center around the importance and use of user defined data [27]. Inspired by the work of Tim O'Reilly [79], Gregg proposes seven key requirements for CI applications (described in Table13) [27].

Table 12. Common characteristics that define a CI system (as presented by Lykourentzou et al. in S6) [48]

Characteristics	Description
Set of possible individual actions	Set of actions that an individual is allowed to perform when contributing (in some form or another) within the system
System state	Set of minimal variables that completely define the system
Community and individual objectives	List of goals that a community or an individual intends to achieve by using the system
Expected user action function	Effort expected from users, necessary to achieve individual/community goals
Future system state function	Expected state of the system after some time, given the system's current state and user actions
Objective function	Measures the extent to which individual/community goals of the system have been achieved [48, 81]

Table 13. Requirements for CI applications (as presented by Gregg in S7) [27]

Requirements	Description
Task-specific representation	CI applications should support task-specific views depending upon the application domain
Data is the key	The effectiveness of CI applications is directly proportional to it's data quality and quantity, and therefore should facilitate data collection and sharing among it's users
Users add value	CI applications should help users to improve the usefulness of data, by providing mechanisms that enable user-oriented addition, modification or enhancement of data
Facilitate data aggregation	Keeping the importance of data in mind, CI applications should be designed with necessary feature that enable data aggregation throughout the duration of systems' use
Facilitate data access	CI applications should offer services and mechanism that facilitate reuse of data outside the application
Facilitate access for all devices	CI applications should provide services that are usable not just with PCs and internet servers, but also portable devices like PDAs, smart-phones, tablets
The perpetual beta	New features must be added to CI applications from time to time, depending upon the community needs and requirements [27]

To illustrate how these requirements could be used to design CI applications, the author developed the 'DDtrac' application for children with special needs. The application was intended to support decision making in special education and therapy. DDtrac is a web-based CI application and has two main objectives: first, the application facilitates communication between therapists and teachers so that they could share information about the needs of the children; second, the application allows data collection and provides tools for data analysis to understand a child's progress and to determine adjustments necessary for a better development of the child. The application was deployed for a duration of 18 months with one autistic student and his teachers and therapists. After the conclusion of the trial, all participants reported that the application successfully achieved both its core objectives and helped to improve the academic performance of the student [27].

3.8 S8 (Martijn C. Schut 2010)

This study aims to provide systematic guidelines and instructions for development of CI models, irrespective of the developer's domain. To come up with these guidelines the author first conducted a number of research studies and identified key contributions which distinguish CI systems from other ICT systems. Based on the literature, Schut compiled a list of properties of CI systems [70] presented in Table 14. After this, the author investigated several strands of research like: complex adaptive systems, swarm intelligence and others; that are often described as being synonymous or at least associated to collective intelligence [70]. Based on the findings, Schut finally proposed a "systematic approach for designing CI system models" [70] and illustrated the proposed methodology using two case studies namely the 'Chinese Whispering Room' and the 'Braitenberg collectivae' [70].

Table 14. Properties of CI systems (as presented by Schut in S8) [70]

Types	Properties
Enabling CI properties	These properties enable the emergence of collective intelligence in a system. • *Adaptivity* refers to the capability of a system to change its behaviour or structure depending upon the environment. • To understand system behaviour, it is important to understand both individual actions and *interactions* among individuals as a whole. These interactions enable the flow of information within systems. • Individual or system behaviour can be described fundamentally using *rules*. Such rules implicitly represent the relationship between system inputs and outputs. [70]
Defining CI properties	If these properties exist in a system, it can be considered a CI system. • *Global-Local* are levels which distinguish between aggregation at system and individual level, respectively. This distinction is important for understanding adaptivity and emergence. Adaptivity can occur at local and/or global level, whereas emergence is achieved by going from local to global. • Complex systems must have elements of *randomness* in order to behave as self-organized critical systems. • *Emergence* is defined as the principle that "the whole is greater than the sum of its parts" [14]; and occurs when moving from "the lowest abstraction level (individual) to the highest abstraction level (system)" [70]. • *Redundancy* means that the system should allow it's users to access/visualize available knowledge and information at different locations within the system's user interface. • Redundant data can make the system *robust*, as data that are lost due to malfunctions could still be recovered from other sources. [70]

The CI system modelling approach proposed by Schut is divided into three phases, i.e., 'system design', 'model design' and 'models' (which are further categorized into 'generic', 'system' and 'computer' models) [70]. The components of the 'system design' phase are inspired by examples from self-organization, multi-agent systems and swarm intelligence;

whereas the components of the 'model design' phase are influenced by the work of John P. van Gigch [80] on system modeling and of meta-modeling.

3.9 S9 (Dimitrios J. Vergados et al. 2010)

The aim of this study is to present a framework that can foster the emergence of collective intelligence in web community based platforms. Based on published research, Vergados et al. describe a generic CI system as having three main components, i.e., 'human community', 'machine intelligence' and 'system information' [81].

Vergados et al. argue that although the proposed CI framework may lead to the development of completely different CI systems, all systems would share a number of common characteristics [49]. The authors describe these characteristics as follows [81]:

- **System attributes** (same as described in Table 12)
  - *Set of possible individual actions*
  - *System state*
  - *Community and Individual objectives* [48, 81]
- **Functions** (same as described in Table 12)
  - *Expected community member action functions*
  - *Future system state functions*
  - *Objective functions* [48, 81]
- **Other elements**
  - *Resource allocation algorithms*: These algorithms define the required user actions (depending upon the system state), necessary to reach user/system goals and to maximize the usefulness of the system.
  - *Critical mass*: This indicates "the minimum number of users necessary for the system to function effectively" [81].
  - *Motivation*: A vital factor, important to improve the quantity and quality of user participation in a CI system. [81]

Finally, Vergados et al. evaluate the proposed framework by means of simulation where they analyze how the quality of Wikipedia articles could be improved, if the system was based on the proposed concepts. The authors compare the performance of their approach against the current approach used in Wikipedia by using the mathematical functions of the proposed framework. The authors claim that, based on their framework, a CI-enabled Wikipedia community could significantly improve the quality of articles, while reducing the time required for these articles to reach satisfactory quality [81].

3.10 S10 (Thomas W. Malone et al. 2009)

This study aims to propose a new framework that explains the underlying model of CI systems. To do so, Malone et al. examined 250 web-enabled CI systems; and based on their findings, identified the building blocks or 'genes' (analogy adopted from biology) of collective intelligence [52]. The authors then classified these building blocks, using two pairs of fundamental questions [52], i.e.,

- "Who is performing the task? Why are they doing it?"
- "What is being done? How is it being done?" [52]

The answers to these questions with respect to *staffing*, *incentive*, *goal* and *structure/process* were then proposed as the 'genes' of CI systems [52]. Malone et al. state that different CI systems could be modelled using the combination and recombination of these building blocks. A brief overview of these 'genes' is presented in Table 15 [52].

Table 15. Building blocks of collective intelligence (as presented by Malone et al. in S10) [52]

	Genes	Description
Who?	Hierarchy	In this gene, tasks are assigned to individuals or groups by someone in authority (similar to traditional hierarchical organizations).
	Crowd	In this gene, individuals within the group can indulge in activities if they choose to do so; and there is no authoritative figure. [52]
Why?	Money	Financial gain can be a big motivator for individuals in markets and organizations.
	Love	In many situations, emotional states such as love, affection, passion or simply interest could be a great motivator for participants.
	Glory	Recognition by competitors, colleagues or general public is another important motivator. [52]
What?	Create	In this gene, participants create something like: a T-shirt design, a piece of code or an innovative solution to a given problem.
	Decide	In this gene, participants evaluate and select items from a set of options; primarily, submitted by other participants. [52]
How?	*Create*	
	Collection	This gene occurs when participants create solutions independently. A sub-type of this gene is the *contest gene*, which occurs when one or many contributions are recognized as best and are rewarded.
	Collaboration	This gene occurs when participants create solutions as a group, and the proposed solutions are interrelated or interdependent.
	Decide	
	Group	This gene occurs when "members of a crowd make a decision that applies to the crowd as a whole" [52]. Important variants of this gene include: *voting*, *consensus*, *averaging*, and *prediction markets*.
	Individual	This gene occurs when members of the crowd make their own independent decisions, that might be influenced by other members but are not necessarily identical. Two important variants of this gene are: *markets* and *social media*. [52]

To explain these genes further, Malone et al. examined four web-enabled CI systems: *Linux*, *Wikipedia*, *InnoCentive* and *Threadless*. Finally, the authors claim that the 'sequences of genes' of each of these systems could be combined into *genomes* that could help us to understand these CI systems better [52].

3.11 S11 (Luca Iandoli 2009)

The aim of this study is to provide a model for management of collective intelligence, and to raise issues that must be considered when designing CI systems. Iandoli argues that although there are several open issues in collective intelligence, all of these issues could be organized into two macro-areas i.e., "management of collective intelligence" [32] and "design of collaborative tools" [32].

Iandoli states that online/virtual communities could be viewed as organizations and, therefore, could also be modeled as such. Based on this hypothesis, Iandoli et al. proposed five characteristics of online/virtual communities ‘when modelled as organizations’ [33]:

(1) “Clear goals and objectives” [32, 33] coherent with a predefined mission.
(2) “A large number of participants” [32, 33] who can offer their time and efforts to achieve the system goals (by knowledge sharing, creation and consensus activities); in return for incentives.
(3) “A set of processes” [32, 33] that allow participants to develop, submit or evaluate new ideas, artifacts and decisions by collaborating with others.
(4) “Rules” [32, 33] that govern how participant interact with the system and each another.
(5) “Participant roles and responsibilities” [32, 33].

Iandoli further argues that even if virtual communities are modelled as organizations, such communities would still face major governance issues; because of the many differences between virtual communities and real organizations. Three of these issues [32] are:

- *Attention governance*: This involves, reducing the possibility of premature, incomplete or biased decisions; caused due to the lack of correct and unbiased knowledge or due to peer pressure.
- *Participation governance*: The system must facilitate and support participation of large number of individuals from diverse backgrounds. Participants must be provided with suitable incentives to keep them inspired and motivated to share their information and knowledge, and to help achieve the system objectives in an unbiased fashion.
- *Community governance*: Appropriate rules must be established to enable smooth and stable interactions among participants and communities; the system should be organized hierarchically and individuals should be given clearly defined roles, responsibilities and incentives [32].

Finally, Iandoli states that even if all the above mentioned issues are resolved, there would still be two challenges i.e., “designing proper visualization tools” [32] and “designing trust and reputation appraisal systems” [32] that would have to be dealt with; irrespective of the technologies used when designing such collaborative platforms [32].

3.12 S12 (Andre Boder 2006)

This study aims to establish a new model for collective intelligence in organizations. The model is inspired by Nonaka’s work on “The Knowledge-Creating Company” [59] and provides insights that enable transformation from tacit to explicit knowledge within and among organizations; from a collective intelligence perspective [3].

Pertaining to literature on knowledge management in organizations, Boder argues that the process of how organizational elements (such as individuals, their expertise, formal and informal networks, methods of communication and implicit cultural norms) interact to enable knowledge creation, represents a form of collective intelligence. Based of this argument, Boder presents the building blocks of organizational collective intelligence [3]:

- Block A (*Development of competencies*) i.e., The first block “is the development of competencies” [3]. Although difficult to realize, organizations should aim to develop complementary competencies. This could possibly be achieved by human resource managers, who should identify individuals with different competencies gained from different situations; and once such individuals are identified, knowledge managers should bring them together so that their competencies complement each other. Doing so, organizations could take advantage of individual competencies and therefore create new knowledge.

- Block B (*Goal development*) The second block "is the development of a common representation of the goals" [3]. Although each group or department within the organization could have its own goals and objective, these goals and their representations should be aligned with the organizations overall objectives and should be coherent.
- Block C (*Mechanic development*) The third block "is the development and alignment of processes into mechanics of interactions between entities involved" [3] i.e., *organizations*. The formal and informal norms of the organization must be stated explicitly; additionally, employees should respect each others expectations and should trust each others competencies, because such a culture would enable smooth articulation when dealing with new problems or challenges. [3]

To illustrate how these building blocks could be used in the process of building collective intelligence, Boder breaks down these actions into six groups; and describes six generic tools that could be utilized to apply these actions [3]. He then uses these tools and actions to describe three scenarios: "the value chain" [3], "co-integration of key competences to achieve a critical medical mission" [3] and "innovative problem-solving" [3]. Finally, the author concludes by stating that organizations must create novelty to survive and evolve. And this is only possible, if organizations build collective intelligence by combining the know-how of their employees and integrate organizational knowledge with partner organizations by "coordinating their respective value chains" [3].

4 DATA SYNTHESIS

In this section, we look at the different definitions and classifications of elements that describe a CI model, as proposed by the studies discussed in Section 3. Looking at all these elements, it is clear that different authors define CI systems using different terminologies such as *characteristics*, *levels*, *requirements*, *properties* and *building blocks*; however, a deeper examination of these models proves that each of these definition types propose similar (if not same) concepts. Similarly, many of the selected studies explain collective intelligence from different perspectives (such as: CI in organizations, CI as self-organizing systems and others); however, the characteristics of CI presented in these studies are very much alike. Table 16 presents the list of all characteristics proposed in the selected studies and classifies them into 24 *unique* attributes (described in Section 5) according to their definitions (described in Section 3). It is important to note here, that some of the selected studies have proposed combinations of characteristics from previous research; and therefore, are presented as combinations of attributes in Table 16.

Table 16. Terminologies used (in S1 - S12) to describe CI systems and attribute ID(s) of their respective classifications

Study ID	Definition Type	Classification	Sub-classification	Attribute ID(s)
S1	Characteristics	Diversity		A1
		Independence		A2
		Decentralization		A1, A2
		Aggregation [58]		A16
S2	Steps	Cognitive diversity		A1
		Promote independence		A2

		Access decentralized knowledge		A17
		Effectively aggregate knowledge [53]		A16
S3	Levels	Capacity level	Set of possible individual actions	A18
			Massive participant interaction	A19
			Competencies development	A7
		Emergence level	System state	A20
			Self-organizing	A8
			Emergent behaviour	A9
			Mechanic development	A10
		Social maturity level	Community & Individual objectives	A11
			Goal development [73]	A12
S4	Levels	Micro-level	Humans as social animals	A13
			Personal interaction capabilities	A19
			Trust	A10
			Motivation	A3
			Attention	A19, A3
			Communities	A4
		Level of emergence	Complex adaptive systems	A8, A9
			Self-organization	A8
			Emergence	A9
			Swarm intelligence	A8, A9
			Stigmergy	A8, A9
			Distributed memory	A17
		Macro-level	Decision making	A13
			Wisdom of crowd	A13
			Aggregation	A16
			Diversity	A1
			Independence [67]	A2
S5	Characteristics	Objective of a task		A12
		Size of contribution		A5
		Form of input		A21
		Form of output		A21
		Stakeholder [21]		A4
S6	Characteristics	Set of possible individual actions		A18
		System state		A20

		Community & Individual objectives		A11
		Critical mass		A5
		Task & workload allocation		A14
		Motivation [48]		A3
S7	Requirements	Task-specific representation		A22
		Data is the key		A23
		Users add value		A6
		Facilitate data aggregation		A16
		Facilitate data access		A17
		Facilitate access for all devices		A17
		The perpetual beta [27]		A8, A9
S8	Properties	Enabling CI properties	Adaptivity	A8, A9
			Interactions	A19
			Rules	A10, A21
		Defining CI properties	Global-local	A16
			Randomness	A8
			Emergence	A9
			Redundancy	A22
			Robustness [70]	A24
S9	Characteristics	System attributes	Community & individual objectives	A11
			Set of possible individual actions	A18
			System state	A20
		Other elements	Resource allocation algorithms	A14
			Critical mass	A5
			Motivation [81]	A3
S10	Genes	Staffing	Crowd	A4
			Hierarchy	A4
		Incentive	Extrinsic motivation	A3
			Intrinsic motivation	A3
		Goal	Create	A13
			Decide	A13
		Structure/Process	Collection	A15
			Collaboration	A15
			Group Decision	A15
			Individual Decision [52]	A15

S11	Characteristics	Clear goals coherent with mission	A12
		Large number of motivated participants	A5, A3
		A set of processes	A15
		Rules	A10, A21
		Roles & Responsibilities [32]	A18
S12	Building Blocks	Competencies development	A7
		Goal development	A12
		Mechanic development [3]	A10

Based on the findings of the data extraction and data synthesis stages, we now answer the first two research questions.

4.1 Research Question (RQ1)

What are the underlying models of existing CI systems? What are the common terminologies used to describe CI models? What are their components? And, how are these components associated to each other?

4.1.1 RQ1.1. What are the underlying models of CI systems?

Literature shows that collective intelligence is a multidisciplinary field, drawing concepts and techniques from a number of different disciplines including computer science [23], organizations [25], social media [69], complexity sciences [70] and psychology [84]; therefore, different scholars have described collective intelligence from different perspectives. However, over the years only three definitions of collective intelligence have been widely adopted in ICT; two of which were proposed in this decade. The first formal definition of collective intelligence (in ICT) was proposed by Pierre Lévy (1997) [43], followed by Jerome C. Glenn (2013) [23] and Thomas W. Malone (2015) [50]. Although each of the definitions describes collective intelligence in its own distinct way; nevertheless, when examined together, the definitions express CI as having three main components, i.e., *individuals* (with data/information/knowledge); *coordination and collaboration activities* (according to a predefined set of rules); and *means/platform for real-time communication* (viz. hardware/software). When combined, these components enable intelligent behaviour in groups or crowds.

Table 17 is the result of segregating all the characteristics defined in Section 3 in terms of the just discussed three main components of CI systems.

4.1.2 RQ1.2. What are the common terminologies used to describe CI models?

As suggested in the selected studies, CI models have been described using terminologies such as *characteristics* (S1, S5, S6, S9 and S11), *steps* (S2), *levels* (S3 and S4), *requirements* (S7), *properties* (S8), *genes* (S10) and *building blocks* (S12). And, each of these terminologies are further segregated into different classification and sub-classifications as described in Section 3. However, as mentioned in the previous sections, the terminologies used in these models describe similar concepts; and therefore can be classified into unique attributes as presented in Table 16.

4.1.3 RQ1.3. What are the components of CI models? And, how are these components associated to each other?

Typically the components of ICT systems are classified as data, hardware, software, information, procedures and people. However, since the selected studies describe CI models by the means of their characteristics; these characteristics

Table 17. Unique attributes of CI (from S1 - S12) segregated according to the components of collective intelligence

Component	Characteristics	Attr. ID	Study ID(s)
individuals	diversity	A1	S1, S2, S4
(with data, information, knowledge)	independence	A2	S1, S2, S4
	motivation	A3	S4, S6, S9, S10, S11
	crowd	A4	S4, S5, S10
	critical mass	A5	S5, S6, S9, S11
	users add value	A6	S7
coordination & collaboration activities	competencies development	A7	S3, S12
(according to a predefined set of rules)	self-organization	A8	S3, S4, S8
	emergence	A9	S3, S4, S8
	trust & respect	A10	S3, S4, S8, S11, S12
	community & individual objectives	A11	S3, S6, S9
	clear goals & objectives	A12	S3, S5, S11, S12
	wisdom of crowd	A13	S4, S10
	task & workload allocation	A14	S6, S9
	set of processes	A15	S10, S11
means for real-time communication	aggregate knowledge	A16	S1, S2, S4, S7, S8
(viz. hardware/software)	access to decentralized knowledge	A17	S2, S4, S7
	roles & responsibilities	A18	S3, S6, S9, S11
	massive interactions	A19	S3, S4, S8
	system state	A20	S3, S6, S9
	predefined input/output types	A21	S5, S8, S11
	task specific representation	A22	S7, S8
	data is key	A23	S7
	robust	A24	S8

can be interpreted as the components of CI models. Based on the definitions of collective intelligence [23, 43, 50], we can segregate these characteristics/attributes and their relationship into the three main components of collective intelligence as described in Section 4.1.1 and presented in Table 17.

4.2 Research Question (RQ2)

Do any of the available CI models appropriately define all CI systems, irrespective of their applications? Can these models be used to create CI systems for novel challenges?

Comparing all characteristics of the studied CI models (see Table 16) with the components of CI systems (described in Section 4.1.1 and presented in Table 17), we see that none of the studied models have all 24 unique attributes, and therefore cannot define all CI systems completely. However, the existing models provide insights that can assist in planning when designing a CI system; and point out challenges that would have to be solved in order to achieve a robust and adaptive CI system. Most authors themselves state that their proposed CI models only describe collective intelligence in specific domains (S2, S3, S4, S9, S11 and S12); and suggest that further research and investigation is required to gain a better understanding of generic CI systems (S1, S3, S6, S8, S10 and S11). Therefore, although particular CI models can be used to define CI systems for specific domains; the same models might not be as useful when designing CI systems from other disciplines.

Furthermore, since the proposed models are evaluated using either quantitative/qualitative interviews (S3) or, case studies (S4, S8) or, examples from scenarios (S12) or, simulations (S9) or, applications/systems build based on the models (S6, S7, S10); it is not possible to identify a single model that can be used (in its current state) to design CI systems for novel challenges. For now, the most generic CI model available in literature is the one proposed by Malone et al. (S10) [21]; however, this highly cited and accepted model needs to developed further for deeper and more accurate understanding of collective intelligence [21, 52, 76].

5 A NOVEL FRAMEWORK FOR COLLECTIVE INTELLIGENCE

Using the findings from the data extraction phase of the SLR, we now attempt to contribute to the available CI models by proposing a unified framework for collective intelligence; combining the 24 unique attributes (see Table 17) of CI models identified from studies S1 - S12. The purpose of the proposed framework is to answer the final research question (RQ3) and provide additional insights and explanations that can help us better understand CI systems in general. In order to evaluate the proposed 'generic' CI model, we will compare the model to multiple CI systems, each designed for a different objective and belonging to different disciplines (see Section 6).

RQ3. Can we somehow combine the available knowledge of CI models and systems to create a unified model that could define all CI systems?

We combine the knowledge of the CI models studied in this SLR, and propose a novel framework that describes CI systems in a fine-grained manner. We do so, by comprehensively classifying all components of the studied CI models into 24 unique attributes (see Table 16), and then categorizing them into three sections:

- a 'generic' model that defines all CI systems
- additional requisites for CI systems
- CI as a complex adaptive system

While taking inspiration from the building blocks for CI proposed by Malone et al. [52], combined with the findings from Section 4.1.1, we propose a model that describes CI systems by the means of *staff*, *process*, *goal* and *motivation*. Designed as an extension to Malone's concept of building blocks, the proposed generic model segregates the originally proposed *genes* into more fine-grained *types*; introduces a new classification, namely *interactions*; and suggests vital *properties* for the *staff* and *goal* building blocks of the generic model. Finally, remaining attributes that could not be accommodated into the building blocks, are aggregated into the additional requisites category.

5.1 A Generic Model for CI Systems

As mentioned in Section 3, Malone's genome model for collective intelligence [52] is based on two pairs of questions: *"Who is performing the task? Why they are doing it?"* and *"What is being done? How is it being done?"* [52]. Based on these questions, the authors proposed the analogy of *genes* categorized as staffing, incentives, goal and process. Each of these categories, were subdivided into individual *genes* which, when combined, created the genome of CI systems. Drawing from the literature, we decided to move away from the concept of genes, and rather examine the proposed *genes* as *types*. Doing so we realized, that the available *genes* could be segregated into new types and sub-types. And, while some of the *genes* could be better understood as *interactions* between *types*, others could be explained as necessary *properties* inherent to these new *types*.

5.1.1 Who is performing the task? The *staff* in CI are the actors who perform different tasks within the system (as suggested in S10). As literature suggests, these actors or individuals must interact with each other based on certain rules depending upon the structure (hierarchical/non-hierarchical) of the system. And, when viewed as a collective, the *staff* of a CI system must exhibit a specific set of properties for the system to function effectively.

- *Types*: The actors or individuals in a collective are the first component of a CI system, and therefore play a vital role in describing how the system functions (A6). Typically, these actors (A4) can be segregated on the basis of their roles and responsibilities (A18) within the system. Drawing insights from S4, S5 and S10; we determine that actors in a CI system can be classified as follows:
  - *Passive actors* or beneficiaries are individuals who aim to gain from the outputs produced by the CI system, but do not wish to contribute in the problem solving process. These beneficiaries could either be stakeholders who are financially motivated, or end-users who simply want to exploit the knowledge produced by the system (but do not wish to actively contribute). Examples of stakeholders in CI systems could be seen in the following projects: Threadless, InnoCentive [15] and GoldCorp [83]. Here, the host organizations crowdsource their problems (designing of T-shirts, research & development, and identifying ideal mining locations, respectively) to the general public; with the intention of using the produced knowledge or artifacts for their own advantage.
  - *Active actors* or contributors are individuals who are involved in CI processes (defined in 'How'); such actors use their knowledge and expertise and help to create innovative solutions to the given problem. Such contributors can be further divided into two categories, namely *crowd* and *hierarchy*.
    - *Crowd* in a CI system comprises actors who actively contribute new knowledge, information or artifacts to the system. Such actors are allowed to carry out a predefined set of actions, based on concrete sets of rules and regulations; however, there is no authoritative figure that has direct control over the actors' individual actions. Examples of crowd in CI systems can be seen in the following projects: Climate CoLab [34], WikiCrimes [19, 68] and WeKnowIt [42] where users contribute data and information about the weather, crimes and disasters; and also help verify the authenticity of the accumulated knowledge. Whereas, in projects such as Threadless [15], members of the crowd contribute by creating new artifacts and deciding on the best.
    - *Hierarchy* in a CI system comprises administrators and experts who are responsible for allocating tasks to the crowd. While the administrators monitor crowd behaviour in the system and make sure that the community and individual goals of the collective are achieved, the experts analyze and verify the contributions of the crowd. Additionally, in some cases the experts also help in identifying the best contributions or solutions. An ideal example of such a hierarchy can be seen in WikiCrimes: institutional agents, monitor agents, reputation agents and others are responsible for different administrative activities within the system [19, 68].
- *Properties*: To ensure that a collective exhibits intelligent behaviour, the collective of actors in a system must have a few crucial properties. According to S1, S2 and S4, a CI system must promote diversity and independence among its actors, as this can enable the creation of novel solutions exploiting knowledge from individuals familiar with multiple domains and with different experiences. Also, these actors should be allowed to act independently, as this can help to get rid of peer pressure, and therefore to reduce user-generated bias. Finally, to enable an effective collective intelligence, a collective must have critical mass or a minimum number of actors as suggested in S5, S6, S9 and S11.

  - *Diversity* (A1) in CI systems refers to the heterogeneous nature of actors, who belong to different age groups, genders and educational, financial and cultural backgrounds. This is important as, such diverse actors can provide diverse pieces of knowledge, perspectives, interpretations and experiences; and this could lead to the creation of innovative solutions and better decisions. An example of the advantages of diversity in actors can be seen in InnoCentive [15]: organizations with small R&D groups crowdsource their problems to acquire new and innovative ideas.
  - *Independence* (A2) means that opinions of one actor should not be influenced by the opinions of others. Independence among actors is vital, as it can help avoiding information cascades where users pass information that they assume to be true (without appropriate evidence or knowledge); and therefore make irrational choices and decisions [1, 46, 55].
  - *Critical mass* (A5) in collectives is defined as minimum number of actors who must participate in system processes for the system to function effectively. Although studies suggest that critical mass is an imperative property that enables effective creation and constant exchange of diverse knowledge and information; however, the concept needs to be investigated further as critical mass in different CI systems can often depend upon the system goals and objectives.
- *Interactions*: Interactions in CI systems can either exist between two or more actors, or among actors and the contributions of others. Such interactions can be categorized as follows:
  - *Trust and respect* (A10) are two preconditions for cooperation. When dealing with new problems or challenges, actors in a collective must treat each other with respect and should trust each others' abilities and competencies; as doing so can enable smooth and efficient flow of knowledge and information within the system.
  - *SECI* "Socialization, Externalization, Combination and Internalization" [59] (A7) are the four components of Nonaka's model for knowledge creation in organizations [59]. Using these knowledge dimensions, organizations can convert their employees' tacit knowledge into explicit organizational knowledge and back. Since the SECI model was originally designed to promote sustainable innovation in organization, these concepts can also be utilized in CI systems to enable competency development in actors (as suggested in S12 and [11]).

  Finally, as suggested in S3, S4 and S8, a CI system must support such interactions in massive volumes (A19).

5.1.2 Why they are doing it? *Motivation* (A3) in CI systems is essential to maintain user engagement and encourage participation. Depending upon on the objectives of a system, users in a CI system could be motivated by their desire to gain knowledge (as in Wikipedia [45]), by money and glory (as in Threadless [8] and InnoCentive [57]) or by social cause (as in hackAIR [54]). According to Malone et al. (S10), money, love and glory can be considered high-level motivations for people participating in CI systems [52]; whereas, Vergados et al. (S9) categorize motivation as tangible, intrinsic and self-fulfilling [81]. Combining the recommendations from S4, S6, S9, S10 and S11, we categorize motivation as *intrinsic* and *extrinsic*.

- *Intrinsic* motivations such as social cause, interest, passion and self-fulfillment encourage actors in a collective to collaborate and contribute for the betterment of the community or its individuals. An example of such motivation can be seen in DDtrac: school teachers and therapists collaborate to understand a child's needs and determine necessary adjustments in teaching techniques for better development of children with special needs [26, 27].
- *Extrinsic* motivations are factors external to CI tasks that encourage actors to contribute in hopes of getting rewards. Such motivations can be either *tangible* like money and trophies or, *intangible* like fame and glory. In CI projects like Threadless [8], InnoCentive [57] and Goldcorp [83] participants are offered cash rewards and

prizes for submitting ideas and designs; whereas, in WikiCrimes [19] participants gain reputation based on the reliability their contributions.

5.1.3 What is being accomplished? Unlike Malone's gene model (S10) that attempts to answer the question of 'What is being done?' (from an organizational perspective), we decided to focus on the question of 'What is being accomplished?' for our proposed model. Based on the literature, we found that our question is a better fit, as it could appropriately define the different types of objectives/goals (of CI systems) presented as characteristics in several selected studies. In general, these *goals* can be defined as 'observable and measurable desired results bound to one or more objectives, that have to be achieved by committed actors within a finite time-frame'. Since, collective intelligence initiatives are typically motivated by *community* or *individual* objectives (A11) as suggested by S3, S6, S9; we segregate CI goals into the two aforementioned *types*. These types can be seen again in Threadless: individuals with niche in T-shirt designing participate in competitions to present their contributions to the community, learn from others' feedback and earn money; whereas, the community's goal is to bring new T-shirt designs to the marketplace by choosing and popularizing trending designs [8]. Additionally, drawing from the contributions of S3, S5, S11 and S12 the requisite properties of these CI system goals could be categorized as *well-defined* and *objective* (A12).

5.1.4 How is it being done? Malone et al. categorized the processes in CI systems as combinations of dependent-independent and create-decide activities, where the create-decide activities answered the question "What is being done?" [52]. In our proposed model, however, we describe CI *processes* (A15) as *types* of activities and *interactions*. As literature suggests, the activities can be either *create*, where actors come up with new ideas or design new artifacts; or it can be *decide*, where actors express their likes or dislikes for a particular subject or artifact. Since both of these activities can be either be done by individual actors or groups of actors, these activities could also be viewed as *dependent* or *independent* interactions. To add more granularity to process types, create activities can be further classified into *contest* (S10) and *voluntary*. As the names suggest, *contest* create activities are carried out in competitive environments and are extrinsically motivated, whereas *voluntary* create activities are intrinsically motivated. It is the combination of these three types (decide, contest and voluntary) and interactions (dependent and independent) that enables intelligence in collectives (A13).

- *Collection* (i.e., *create* plus *independent*): In such activities or processes actors participate as individuals and their contribution to the system is a result of their independent work. An example of *collection through contest* can again be seen in Threadless: individuals compete for cash rewards by creating and submitting new T-shirt designs [8]. Whereas, in WikiCrimes, actors contribute through *voluntary collection* by reporting criminal activities they witness in their local vicinity [19].
- *Collaboration* (i.e., *create* plus *dependent*): Such activities are carried out by groups of actors or communities where multiple individuals work together as a single entity and create new ideas or products.
  As an instance for *voluntary collaboration* we can again look at DDtrac: therapists and teachers work together to maximize the learning outcomes of students with special needs [27]. Similarly, in hackAIR, volunteers from NGOs conduct workshops to build citizen interest in the hackAIR platform and educate them on how they could become a part of the project's community and help to gather air quality data from their local vicinity [40]. Whereas, *collaboration in contests* is seen in openIDEO, where multiple participants work as a team and propose solutions to societal challenges, in hopes of getting financial rewards [67].

- *Individual decision* (i.e., *decide* plus *independent*): Such decisions are made by individuals acting as independent entities and can be different for different actors. However in some cases these decisions may be influenced by the information provided by other actors. For instance, in Threadless the members of the community independently vote for T-shirt designs submitted by the participants. Unfortunately, as suggested by Salminen in S4, in some cases participants tend to creates multiple accounts with the intension to down-vote their competitors; thereby influencing other members and generating biased community feedback [67].
- *Group decision* (i.e., *decide* plus *dependent*): In such activities decisions are made by multiple individuals as a group or a community, and the outcome of the decisions impacts the community as a whole. For instance, such consensus can be seen in Threadless: the employees of the organization review the t-shirt designs chosen by the community and finally decide which designs to produce and award [8].

5.1.5 Input and Output. The final component of a CI system is the *flow of information* or, form of input/output (A21); and can be explained as interactions between the 'who' and the 'how' of the system. The flow of information starts from the *actors* who are responsible for providing inputs like individual knowledge and experiences, data from sensors, or end-user opinions and feedback from social media platforms. The collected inputs are then processed using different activities in 'how'; and the results of these activities are then presented back to the *actors* who now take new decisions or produce new artifacts based on this new found knowledge. Since this flow of information between the *actor*s and the *process*es of the CI system is so vital, we decided to add it to our generic CI model.

The aggregation of the aforementioned components is illustrated as the proposed 'generic' model for CI systems, in Figure 1.

5.2 Additional Requisites

Although, any CI system can be described as the combination of the above mentioned components, there are a few additional requisites that must exist in a CI system, for the system to work effectively.

- *System state* (A20). This can be expressed as the minimum set of variables that completely define a CI system. As discussed in S3, S6 and S9 the system state can include challenges/issues raised by the members of the community, the identified solutions, activities of the users and the system resources. Since our proposed model defines CI systems as the combination of different processes, actors, motivations and goals; unique combinations of the same can be used to express the system state of a collective intelligence system.
- *Data is the key* (A23). "Collective intelligence draws on user-generated content and sharing of information, knowledge and ideas" [69] and, therefore, data or information/knowledge provided by members of the collective is a vital component of a CI system. For a CI system to be able to reach its goals, the system must allow its users to collect, manipulate and share large volumes of data; this can enable robust innovations and decisions.
- *Aggregate knowledge* (A16). Since the effectiveness of a collective intelligence relies primarily on user generated data/information, CI systems must have mechanisms and processes that aggregate this data/information. These aggregation processes are important as information provided by the community can often come from a variety of sources and could be incorrect or biased [20]. Aggregating the information, however, could help resolving conflicting information and could therefore allow for better innovations and reliable decisions. Additionally, systems should also provide mechanisms that allow users to aggregate their knowledge by means of social tagging (for information retrieval), collaboration (for exchange of vocabularies) and task specific representation.

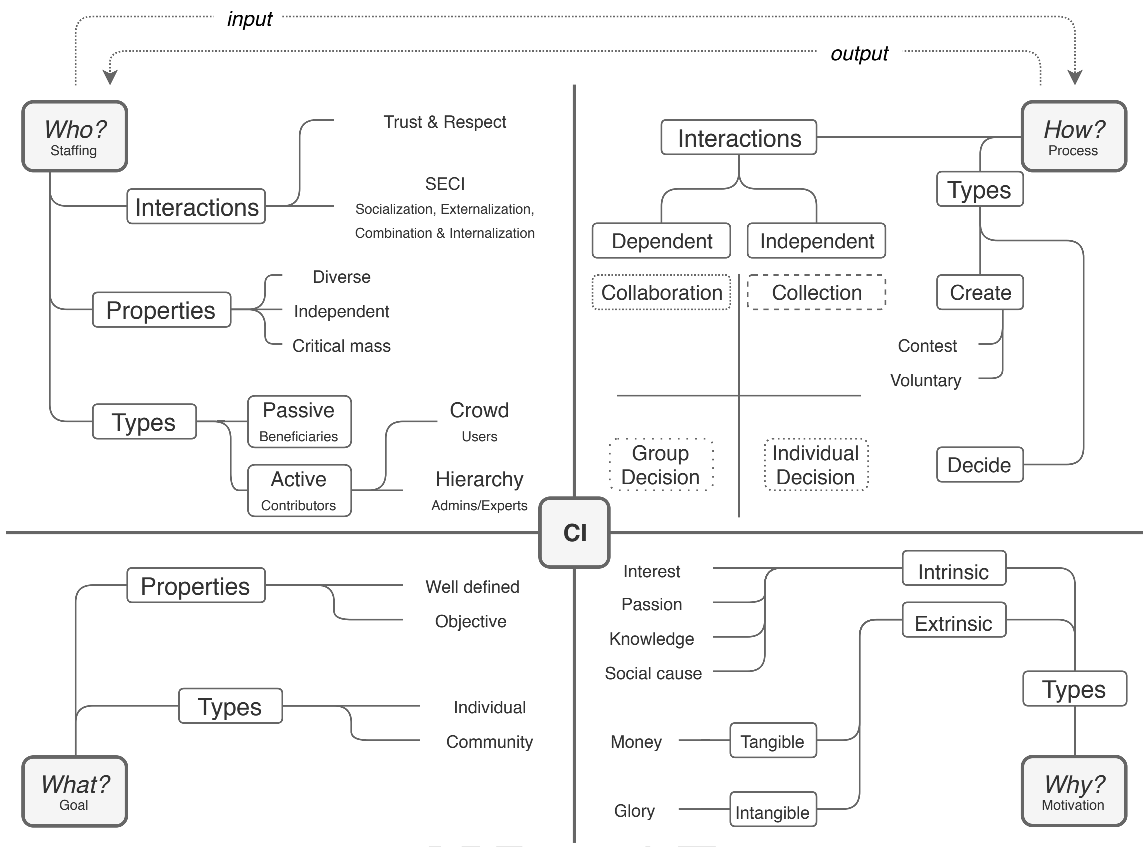

Fig. 1. Generic Model for Collective Intelligence systems

- *Access to decentralized knowledge* (A17). Thanks to growing number of internet users, more and more people are able to communicate, collaborate and share information on the web. Keeping user interest in mind, it is important for CI systems to allow users from different parts of the world to participate and gain from the knowledge or artifacts generated by the system. To do so, the system must facilitate access across multiple devices like PCs, laptops, smart phones, servers and others. Furthermore, CI systems should support open data and open innovation practices, and should allow data access to users even outside the system.
- *Task & workload allocation* (A14). Another important aspect that should be kept in mind when designing CI systems, are methods for coordination and resource allocation. When designing CI systems, the expected tasks of different actor types should be predefined; and based on these tasks, the rules and extent of interactions among actors and actor's access to the aggregated knowledge must be outlined. For instance, participants should be allowed to add new solutions and view solutions submitted by others; however, they should not be allowed to make changes to others' contributions without the contributor's consent. On the other hand, system administrators should have complete access to the data/information/knowledge produced within the system.
- *Task specific representation* (A22). To support knowledge creation and enable fluid information exchange among actors from asynchronous groups, CI systems should provide task specific representations like tables, charts,

histograms, plots and knowledge graphs. Additionally, depending on the task or problem, the system should allow its users to visualize the same knowledge/information in different forms.

- *Robust* (A24). Finally, since CI systems are designed as complex systems with multiple components, actors, users and resources, it is important for such systems to be able to handle redundant and erroneous inputs. In addition to this, the system should also have appropriate mechanisms for data/information/knowledge backup and recovery, in case of a system crash or malfunction.

5.3 CI as Complex Adaptive System

CI systems are complex by nature [70] and should be able to adapt to their environments, making such platforms complex adaptive systems (as suggested in S4 and S8). However, for a system to be complex adaptive, the system must exhibit adaptivity, self-organization and emergence [60, 67, 70, 75].

Adaptivity means that the system or its components should allow constant changes over the period of its existence, depending upon the needs of its collective [70]. System developers should regularly update and evolve the platform by bringing in new technologies and services, based on user feedback and requirements; under the condition that these requirements are aligned with the system goals and objectives.

Self-organizing (A8) [47, 67] means the systems should be able to organize and re-organize its internal structure without the need of an external control [36, 72]. This behaviour could be facilitated by allowing creation of communities, where each member of the community would have a reputation that they could gain by providing useful contributions (in form of insights, knowledge or artifacts) and through up-votes/stars given to them by other members of their community. Such a reputation model can help create a structure within these communities, and therefore further interactions between such communities can lead to self-organizing behaviour within the system.

Emergence (A9) in a system occurs when simple interactions among low-level system components give rise to new and unexpected patterns or properties, disparate from the properties of the system as a whole (based on the definition of emergence proposed by Damper [14]). In adaptive and self-organizing systems, regular modifications to the system and ever-changing user behaviour may lead to the creation of unforeseen patterns, properties or outcomes; thereby exhibiting emergent behaviour.

6 COMPARATIVE CASE STUDIES

In this section we evaluate the proposed generic model from Section 5.1 by examining six CI platforms with respect to the aforementioned model. The CI platforms were chosen on the basis of the following criteria: the platforms should belong to different disciplines/domains, the systems should be available for use (during the time of study), the platforms should have been published/discussed in scientific literature, the deliverables of the platforms should be available online and lastly, the platforms should be recent or ongoing.

Based on these criteria we identified six CI platforms (see Table 1 in Supplementary Material). To analyze the platforms, we created user profiles on each of the platforms and observed system processes for create and decide activities; over the duration of six months, i.e., starting January 2018 to the end of June 2018. During this period, we interacted with the system as passive users. We created projects/ideas to analyze the creation process, however, we never submitted the projects/ideas for evaluation. We observed submissions for other participants, and feedback from active users; and analyzed how the system communities and hierarchies work synchronously to come up with new contributions and innovative ideas. Additionally, we studied the available technical reports, scientific publications, FAQs and other useful resources for each of the platforms. Aggregating our observations, we found that different aspects each of the six CI platforms could be described using our proposed generic model. Tables 18, 19 and 20 present the ‘What’, ‘Who’ and ‘Why’ - ‘How’ for each of the platforms, respectively.

6.1 What?

Table 18. List of studied CI platforms and their ‘What’

CI Platform	Year	What*	Domain
CAPSELLA	2016 - 2018	*IG*: Learn about new ICT technologies that can help improve agronomic practices. *CG*: Develop new ICT solutions, software’s and applications; and promote start-ups that can provide such solutions for agrifood business and farmers.	Agrobiodiversity
hackAIR	2016 - 2018	*IG*: Learn about the concentration of air pollutants (especially particulate matter) in cities and its effect on the health of local residents. *CG*: Provides citizens with real-time information about air pollution levels in their local vicinity and enable conversations for possible improvements in air quality.	Air pollution
openIDEO	Ongoing since 2010	*IG*: Demonstrate their skills and expertise to solve complex challenges; and learn from others’ work. *CG*: Tackle global challenges by developing innovative solutions using human-centric collaboration activities.	Innovation platform
Climate CoLab	Ongoing since 2009	*IG*: Participate in initiatives to help reach global climate goals. *CG*: Collaborate with other communities and experts; and help design/choose solutions to help identify sustainable growth initiatives.	Climate change
WikiCrimes	Ongoing since 2008	*IG*: Report criminal incidents. And keep track of crime rates in local vicinity. *CG*: Assist governing bodies in validating reports of crimes provided by individuals. Help maintain a public record of all criminal activities.	Crime monitoring
Threadless	Ongoing since 2000	*IG*: Showcase their artistic ability by creating new T-shirt design. *CG*: Express community interest and select best T-shirt designs. Bring new and treading T-shirts designs to the market place.	Apparel design

* *IG*: Individual goals. *CG*: Community goals.

The goals of the six CI platforms can be summarized as follows:

The *CAPSELLA* project is designed to enable creation of new ICT solutions for farmers and agricultural experts. The platform focuses on ICT contributions that facilitate collection and exchange of data and experiences from individuals working in agriculture and bio-diversity.

hackAIR is designed as a platform where citizens can collect and access information about air quality in different parts of the world. The system empowers citizens by providing openly available DIY sensor designs, tool-kits and tutorials; thereby enabling citizens to be a part of the data collection process.

Similar to hackAIR, the *openIDEO* and *Climate CoLab* platforms deal with climate change and other environmental/societal challenges. However, both of these platforms are designed to enable creation of new and innovative solutions by means of collaboration. While the contributions in Climate CoLab are focused towards global climate change goals; the contributions in openIDEO are focused more towards open innovation practices for societal change.

The *WikiCrimes* platform allows residents to anonymously report criminal activities in their local vicinity. This is especially useful in countries where citizens are not willing to contact the law enforcement agencies due to fear or lack of trust. The platform also allows its users to track the frequency and scale of criminal activities in different areas, thereby helping users in making better decisions when visiting specific locations.

Finally, the *Threadless* platform is meant for e-commerce and focuses on retail of apparels. The platform enables artists and designers to showcase their talent by sharing their T-shirt designs with the community. The best designs are then made available for sale on the Threadless marketplace; thereby providing artists and designers with a means of income.

We further elaborate the goals of these CI platforms, as *individual* and *community goals* and domains in Table 18.

6.2 Who?

Table 19 presents the different actors of the analyzed CI platforms, segregated into three categories namely *active (crowd)*, *active (hierarchy)* and *passive/beneficiaries* based on our proposed generic model. The table also indicates whether the platforms provide open data for future research or not.

Table 19. Comparative of CI platforms - 'Who'

CI Platform	Who			Open Data
	Active (Crowd)	Active (Hierarchy)	Passive / Beneficiaries	(*Yes/No*)
CAPSELLA	Farmers, food & seed communities	Agro-ecology, agri-food, ICT experts	Farmer communities, technology providers, other organizations	*Yes*
hackAIR	Citizens, open source communities	Environmental/ health/ educational organizations, scientific communities	Enterprises, local governments	*Yes*
OpenIDEO	Participants, innovators, alliances	Experts, challenge sponsors, advisory board	Participants (who only wish to participate in workshops)	*No*
Climate CoLab	Participants, community members	Fellows, judges	Government bodies, business organizations, civil society, individual citizens, consumers	*No*
WikiCrimes	Citizens	Agents, news media, government agencies	Citizens, government agencies	*No*
Threadless	Designers, consumers	Organization (Threadless)	Consumers	*No*

6.3 Why and How?

Table 20 illustrates, how each of the analyzed CI platforms motivates different kinds of actors using different sets of intrinsic and extrinsic motivators; and how different kinds of actors carry out different create and decide activities based on their roles within the system.

Table 20. Comparative of CI platforms - 'Why' and 'How'

CI Platform	Why			How*		
	Intrinsic	Extrinsic		Create		Decide
	Interest (I)/ *Passion* (P)/ *Knowledge* (K)/ *Social cause* (S)	*Tangible*	*Intangible*	*Contest* (CL/CB)	*Voluntary* (CL/CB)	(ID/GD)
CAPSELLA	IKS	Money	-	CL	Both	GD
hackAIR	IPKS	hackAIR sensors	Points, badges	CL	Both	GD
OpenIDEO	IPKS	Money	Glory	CL	Both	Both
Climate CoLab	IPKS	Money	Points	Both	CL	Both
WikiCrimes	IKS	-	Reputation	-	Both	Both
Threadless	IPK	Money	Design Quotient	CL	CL	Both

* CL: *Collection.* CB: *Collaboration.* ID: *Individual decision.* GD: *Group decision.*

After mapping our observations (from each of the platforms) to our generic model, we found some interesting relationships between *actor* types, their *motivations* and their *activities*:

- *Decide* activities are typically *intrinsically* motivated.
- *Contest (create)* activities by individuals of the *active (crowd)* are always *extrinsically* motivated. Whereas, *voluntary (create and decide)* contributions by individuals of the *active (crowd)* are always *intrinsically* motivated.
- *Voluntary (create)* contributions can be of two types: as data or information contributed by *crowd*, as in CAPSELLA, hackAIR and WikiCrimes, or as feedback and suggestions given by *crowd* and members of *hierarchy* to help improve participants' contributions like in OpenIDEO, Climate CoLab and Threadless.

7 THREATS TO VALIDITY

The primary threats to validity of this Systematic Literature Review include bias in search strategy, bias in selection process and inaccuracies in data extraction.

The selection of studies relied on the search strategy which included the selection of search terms and literature resources, and the search process. The search terms were selected based on both the research questions and an initial literature review; followed by a three-step process to construct the search string as described in Section 1. We then choose four prominent academic databases of computer science and used the formulated search string to identify relevant literature. Table 2 presents the number and types of research articles identified from each of the academic databases. To avoid bias in our search strategy and to identify relevant technical reports, books and thesis, we conducted manual search on Google Scholar.

To avoid bias in the study selection process, we first reviewed the titles and abstracts of the identified studies and then selected only those studies which fulfilled the inclusion criteria. We then studied these selected articles and manually checked their references to make sure that we did not miss any relevant articles during the search process. Finally, the selected studies were then evaluated based on the quality assessment criteria. As a result of the study selection phase, we were able to identify the most relevant studies with respect to our research questions.

To eliminate inaccuracies in data extraction, each primary study was independently studied by all researchers and any disparities in findings were resolved through discussions. During the process, we found two pairs of studies, i.e., S1, S2 and S6, S9, which shared a couple of similarities. The first pair (S1, S2) described the characteristics of CI systems using similar classifications, while the second pair (S6, S9) was written by the same authors. By consensus, we decided to keep both pairs in our selected studies; as S1 and S2 described CI systems from different perspectives whereas, S6 and S9 provided different contributions.

8 CONCLUSION

The objective of this article was to analyze different collective intelligence models described in the scientific literature and to identify a generic model that could be utilized to design new CI platforms. To this end, we conducted a Systematic Literature Review, in which we identified 9,418 articles on collective intelligence models. Out of these articles, we selected 12 studies based on an exhaustive selection process. We then critically analyzed these selected studied and found that none of the models provided a generic view of CI systems, as each of the models were designed based on specific perspectives. And, the models that could potentially be used to design domain independent CI systems lacked granularity and needed to be researched further. So, to fill this research gap, we aggregated the components of the CI models described in the selected studies and proposed a unified framework for understanding CI systems. The proposed framework describes CI systems in three parts. First, a generic model, which describes CI systems as a combination of goals, staff, motivation and processes; which are further described as types, interactions and properties. Second, a list of requisites necessary for CI systems to work effectively. And third, guidelines that could enable complex adaptive behaviour in CI platforms.

To evaluate if the proposed model could define CI systems from different domains, we selected a set of ongoing CI projects and observed user activities within the platform, over a duration of six months. After this, we systematically organized our observations and segregated them according to the different components of our proposed generic model. We found that our model successfully described the components of each of the CI platforms and revealed some interesting relations between the types of actors, their activities and motivations. The evaluation of the proposed model also gave us the opportunity to present our unified CI framework by means of examples (i.e., 6 ongoing CI initiatives). It was imperative that we describe the components of these CI platforms in terms of the proposed CI model, so that both researchers and system designers/developers in the field could utilize our novel model to design and develop new CI systems. The 24 unique attributes that describe the proposed framework could provide initial insights to system designers and developers; and could be beneficial during the requirement elicitation process when developing new CI systems. We recognize that we need to further examine the proposed framework by comparing it to a larger set of CI platforms, as doing so would help us gain a deeper understanding about how the proposed framework could be used to design new CI systems. Additionally, we would like to evaluate the proposed framework by conducting qualitative interviews with domain experts and researchers working on upcoming CI initiatives. And finally, we would also like to investigate different trust and reputation models that could be utilized to reduce user bias within CI platforms, thereby enhancing user experience and enabling smooth exchange of knowledge and information within communities.

REFERENCES

[1] Lisa R. Anderson and Charles A. Holt. 1997. Information Cascades in the Laboratory. *The American Economic Review* 87, 5 (1997), 847–862. http://www.jstor.org/stable/2951328

[2] Kirsimarja Blomqvist. 1997. The many faces of trust. *Scandinavian Journal of Management* 13, 3 (1997), 271 – 286. https://doi.org/10.1016/S0956-5221(97)84644-1

[3] Andre Boder. 2006. Collective intelligence: a keystone in knowledge management. *Journal of Knowledge Management* 10, 1 (2006), 81–93. https://doi.org/10.1108/13673270610650120

[4] Eric Bonabeau. 2009. Decisions 2.0: The Power of Collective Intelligence. *MIT Sloan Management Review* 50, 2 (Winter 2009), 45–52.

[5] E Bonabeau and C Meyer. 2001. Swarm intelligence - A whole new way to think about business. *Harvard Business Review* 79, 5 (May 2001), 106+.

[6] T. Bosse, C.M. Jonker, M.C. Schut, and J. Treur. 2006. Collective Representational Content for Shared Extended Mind. *Cognitive Systems Research* 7, 2-3 (2006), 151–174. https://doi.org/10.1016/j.cogsys.2005.11.007

[7] Daren C. Brabham. 2008. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence* 14, 1 (2008), 75–90. https://doi.org/10.1177/1354856507084420

[8] Daren C. Brabham. 2010. Moving the crowd at Threadless. *Information, Communication & Society* 13, 8 (2010), 1122–1145. https://doi.org/10.1080/13691181003624090

[9] David Budgen and Pearl Brereton. 2006. Performing Systematic Literature Reviews in Software Engineering. In *Proceedings of the 28th International Conference on Software Engineering (ICSE '06)*. ACM, New York, NY, USA, 1051–1052. https://doi.org/10.1145/1134285.1134500

[10] Romina Cachia, Ramón Compañó, and Olivier Da Costa. 2007. Grasping the potential of online social networks for foresight. *Technological Forecasting and Social Change* 74, 8 (2007), 1179 – 1203. https://doi.org/10.1016/j.techfore.2007.05.006

[11] Kyung Jin Cha, Yang Sok Kim, Byeonghwa Park, and Choong Kwon Lee. 2015. Knowledge Management Technologies for Collaborative Intelligence: A Study of Case Company in Korea. *International Journal of Distributed Sensor Networks* 11, 9 (2015), 368273. https://doi.org/10.1155/2015/368273

[12] Cary Cherniss. 2010. Emotional Intelligence: Toward Clarification of a Concept. *Industrial and Organizational Psychology* 3, 2 (2010), 110–126. https://doi.org/10.1111/j.1754-9434.2010.01231.x

[13] Prerna Chikersal, Maria Tomprou, Young Ji Kim, Anita Williams Woolley, and Laura Dabbish. 2017. Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 873–888. https://doi.org/10.1145/2998181.2998250

[14] R. I. Damper. 2000. Editorial for the Special Issue on 'Emergent Properties of Complex Systems': Emergence and levels of abstraction. *International Journal of Systems Science* 31, 7 (2000), 811–818. https://doi.org/10.1080/002077200406543

[15] Miguel de Castro Neto and Ana Espírtio Santo. 2012. Emerging Collective Intelligence Business Models. In *MCIS 2012 Proceedings*. Mediterranean Conference on Information Systems. https://aisel.aisnet.org/mcis2012/14

[16] Tony Diggle. 2013. Water: how collective intelligence initiatives can address this challenge. *Foresight* 15, 5 (2013), 342–353. https://doi.org/10.1108/FS-05-2012-0032

[17] Anhai Doan, Raghu Ramakrishnan, and Alon Y. Halevy. 2011. Crowdsourcing systems on the World-Wide Web. *Commun. ACM* 54, 4 (2011), 86. https://doi.org/10.1145/1924421.1924442

[18] Ketry Gorete Farias dos Passos and Edna Lúcia da Silva. 2012. Effect of collective intelligence in organizations. *Transinformação* 24, 2 (May 2012), 127–136. https://doi.org/10.1590/s0103-37862012000200005

[19] Vasco Furtado, Leonardo Ayres, Marcos de Oliveira, Eurico Vasconcelos, Carlos Caminha, Johnatas D'Orleans, and Mairon Belchior. 2010. Collective intelligence in law enforcement - The WikiCrimes system. *Information Sciences* 180, 1 (2010), 4 – 17. https://doi.org/10.1016/j.ins.2009.08.004

[20] Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han. 2015. Truth Discovery and Crowdsourcing Aggregation: A Unified Perspective. *PVLDB* 8, 12 (2015), 2048–2049. https://doi.org/10.14778/2824032.2824136

[21] Sandro Georgi and Reinhard Jung. 2012. Collective intelligence model: How to describe collective intelligence. In *Advances in Intelligent and Soft Computing*. Vol. 113. Springer, 53–64. https://doi.org/10.1007/978-3-642-25321-8_5

[22] Jim Giles. 2005. Internet encyclopaedias go head to head. *Nature* 438, 7070 (December 2005), 900–901. https://doi.org/10.1038/438900a

[23] Jerome C. Glenn. 2013. Collective Intelligence and an Application by The Millennium Project. *World Futures Review* 5, 3 (2013), 235–243. https://doi.org/10.1177/1946756713497331

[24] Vijay Govindarajan and Chris Trimble. 2011. The CEO's role in business model reinvention. *Harvard Business Review* 89, 1-2 (2011), 108–14.

[25] Antonietta Grasso and Gregorio Convertino. 2012. Collective Intelligence in Organizations: Tools and Studies. *Computer Supported Cooperative Work (CSCW)* 21, 4 (01 Oct 2012), 357–369. https://doi.org/10.1007/s10606-012-9165-3

[26] Dawn Gregg. 2009. Developing a collective intelligence application for special education. *Decision Support Systems* 47, 4 (2009), 455 – 465. https://doi.org/10.1016/j.dss.2009.04.012

[27] Dawn G. Gregg. 2010. Designing for Collective Intelligence. *Commun. ACM* 53, 4 (April 2010), 134–138. https://doi.org/10.1145/1721654.1721691

[28] Tom Gruber. 2008. Collective knowledge systems: Where the Social Web meets the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web* 6, 1 (2008), 4 – 13. https://doi.org/10.1016/j.websem.2007.11.011

[29] Friedrich August Hayek. 1945. The use of knowledge in society. *The American Economic Review* 35, 4 (1945), 519–530.

[30] Sylvia Ann Hewlett, Melinda Marshall, and Laura Sherbin. 2013. How diversity can drive innovation. *Harvard business review* 91, 12 (2013), 30–30.

[31] Lu Hong and Scott E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101, 46 (2004), 16385–16389. https://doi.org/10.1073/pnas.0403723101

[32] Luca Iandoli. 2009. Leveraging the Power of Collective Intelligence through IT-enabled Global Collaboration. *Journal of Global Information Technology Management* 12, 3 (2009), 1–6. https://doi.org/10.1080/1097198X.2009.10856494

[33] Luca Iandoli, Mark Klein, and Giuseppe Zollo. 2009. Enabling On-Line Deliberation and Collective Decision-Making through Large-Scale Argumentation. *International Journal of Decision Support System Technology* 1, 1 (Jan 2009), 69–92. https://doi.org/10.4018/jdsst.2009010105

[34] Joshua Introne, Robert Laubacher, Gary Olson, and Thomas Malone. 2011. The Climate CoLab: Large scale model-based collaborative planning. In *Proceedings of the 2011 International Conference on Collaboration Technologies and Systems, CTS 2011*. 40–47. https://doi.org/10.1109/CTS.2011.5928663

[35] Caroline Just. 2011. A review of literature on the general factor of personality. *Personality and Individual Differences* 50, 6 (2011), 765 – 771. https://doi.org/10.1016/j.paid.2011.01.008

[36] Stuart A. Kauffman. 1993. *The origins of order: Self-organization and selection in evolution.* Oxford University Press, USA.

[37] Barbara A. Kitchenham and S. Charters. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering.* Technical Report EBSE-2007-01. Keele University.

[38] Gary A Klein. 2004. *The power of intuition: How to use your gut feelings to make better decisions at work.* Crown Business.

[39] A. Kornrumpf and U. Baumöl. 2014. A Design Science Approach to Collective Intelligence Systems. In *2014 47th Hawaii International Conference on System Sciences*. 361–370. https://doi.org/10.1109/HICSS.2014.53

[40] Evangelos Kosmidis et al. 2018. hackAIR: Towards Raising Awareness about Air Quality in Europe by Developing a Collective Online Platform. *ISPRS International Journal of Geo-Information* 7, 5 (2018). https://doi.org/10.3390/ijgi7050187

[41] Hélène Landemore and Jon Elster. 2012. *Collective Wisdom: Principles and Mechanisms.* Cambridge University Press. https://doi.org/10.1017/CBO9780511846427

[42] Vita Lanfranchi and Neil Ireson. 2009. User Requirements for a Collective Intelligence Emergency Response System. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (BCS-HCI '09)*. British Computer Society, Swinton, UK, UK, 198–203. http://dl.acm.org/citation.cfm?id=1671011.1671035

[43] Pierre Levy. 1997. *Collective Intelligence: Mankind's Emerging World in Cyberspace.* Perseus Books, Cambridge, MA, USA.

[44] Pierre Lévy. 2010. From social computing to reflexive collective intelligence: The IEML research program. *Information Sciences* 180, 1 (2010), 71–94. https://doi.org/10.1016/j.ins.2009.08.001

[45] Randall M. Livingstone. 2016. Models for Understanding Collective Intelligence on Wikipedia. *Soc. Sci. Comput. Rev.* 34, 4 (Aug. 2016), 497–508. https://doi.org/10.1177/0894439315591136

[46] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108, 22 (2011), 9020–9025. https://doi.org/10.1073/pnas.1008636108

[47] Shuangling Luo, Haoxiang Xia, Taketoshi Yoshida, and Zhongtuo Wang. 2009. Toward collective intelligence of online communities: A primitive conceptual model. *Journal of Systems Science and Systems Engineering* 18, 2 (01 Jun 2009), 203–221. https://doi.org/10.1007/s11518-009-5095-0

[48] Ioanna Lykourentzou, Dimitrios J. Vergados, Epaminondas Kapetanios, and Vassili Loumos. 2011. Collective Intelligence Systems: Classification and Modeling. *Journal of Emerging Technologies in Web Intelligence* 3, 3 (August 2011). https://doi.org/10.4304/jetwi.3.3.217-226

[49] Ioanna Lykourentzou, Dimitrios J. Vergados, and Vassili Loumos. 2009. Collective Intelligence System Engineering. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES '09)*. ACM, New York, NY, USA, Article 20, 7 pages. https://doi.org/10.1145/1643823.1643848

[50] T.W. Malone and M.S. Bernstein. 2015. *Handbook of Collective Intelligence.* MIT Press. https://books.google.ee/books?id=Px3iCgAAQBAJ

[51] Thomas W. Malone and Mark Klein. 2007. Harnessing Collective Intelligence to Address Global Climate Change. *Innovations: Technology, Governance, Globalization* 2, 3 (2007), 15–26. https://doi.org/10.1162/itgg.2007.2.3.15

[52] Thomas W. Malone, Robert Laubacher, and Chrysanthos N. Dellarocas. 2009. Harnessing Crowds: Mapping the Genome of Collective Intelligence. *SSRN Electronic Journal* (2009). https://doi.org/10.2139/ssrn.1381502

[53] Kurt Matzler, Andreas Strobl, and Franz Bailom. 2016. Leadership and the wisdom of crowds: how to tap into the collective intelligence of an organization. *Strategy & Leadership* 44, 1 (2016), 30–35. https://doi.org/10.1108/SL-06-2015-0049 arXiv:https://doi.org/10.1108/SL-06-2015-0049

[54] Gavin McCrory, Carina Veeckman, and Laurence Claeys. 2017. Citizen Science Is in the Air – Engagement Mechanisms from Technology-Mediated Citizen Science Projects Addressing Air Pollution. In *Internet Science*. Springer International Publishing, Cham, 28–38. https://doi.org/10.1007/978-3-319-70284-1_3

[55] Lev Muchnik, Sinan Aral, and Sean J. Taylor. 2013. Social Influence Bias: A Randomized Experiment. *Science* 341, 6146 (2013), 647–651. https://doi.org/10.1126/science.1240466

[56] J. Musil, A. Musil, D. Weyns, and S. Biffl. 2015. An Architecture Framework for Collective Intelligence Systems. In *2015 12th Working IEEE/IFIP Conference on Software Architecture (WICSA)*. 21–30. https://doi.org/10.1109/WICSA.2015.30

[57] Kimberly Nehls. 2015. Crowdsourcing. *Learning, Media and Technology* 40, 1 (2015), 123–126. https://doi.org/10.1080/17439884.2014.929144

[58] Van Du Nguyen and Ngoc Thanh Nguyen. 2018. Intelligent Collectives: Theory, Applications, and Research Challenges. *Cybernetics and Systems* 49, 5-6 (2018), 261–279. https://doi.org/10.1080/01969722.2017.1418254

[59] Ikujiro Nonaka and Takeuchi Hirotaka. 1995. *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation.* Oxford University Press. https://books.google.ee/books?id=B-qxrPaU1-MC

[60] Julio M Ottino. 2004. Engineering complex systems. *Nature* 427, 6973 (2004), 399. https://doi.org/10.1038/427399a

[61] S.E. Page and Princeton University Press. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies.* Princeton University Press. https://books.google.ee/books?id=FAFVHnJ7uK0C

[62] Chloe Parker. 2015. The IWA WaterWiki: an open-access publishing platform providing free content on all aspects of water, wastewater and the environment. *Environment and Urbanization* 27, 1 (2015), 137–138. https://doi.org/10.1177/0956247814547172

[63] Michael A Peters and Richard Heraud. 2015. Toward a Political Theory of Social Innovation: Collective Intelligence and the Co-creation of Social Goods. 3, 3 (2015), 7–23. https://researchcommons.waikato.ac.nz/handle/10289/9569

[64] Rolf Pfeifer, Jan Henrik Sieg, Thierry Bücheler, and Rudolf Marcel Füchslin. 2010. Crowdsourcing, open innovation and collective intelligence in the scientific method : a research agenda and operational framework. (2010). https://doi.org/10.21256/zhaw-4094

[65] P. Clint Rogers, Stephen W. Liddle, Peter Chan, Aaron Doxey, and Brady Isom. 2007. WEB 2.0 learning platform: Harnessing collective intelligence. 8, 3 (2007), 16–33. http://tojde.anadolu.edu.tr/makale_goster.php?id=348

[66] Frantz Rowe. 2014. What literature review is not: diversity, boundaries and recommendations. *European Journal of Information Systems* 23, 3 (01 May 2014), 241–255. https://doi.org/10.1057/ejis.2014.7

[67] Juho Salminen. 2015. *The role of collective intelligence in crowdsourcing innovation.* PhD dissertation. Lappeenranta University of Technology.

[68] H. Santos, L. Ayres, C. Caminha, and V. Furtado. 2012. Open Government and Citizen Participation in Law Enforcement via Crowd Mapping. *IEEE Intelligent Systems* 27 (07 2012), 63–69. https://doi.org/10.1109/MIS.2012.80

[69] Detlef Schoder, Peter A. Gloor, and Panagiotis Takis Metaxas. 2013. Social Media and Collective Intelligence—Ongoing and Future Research Streams. *KI - Künstliche Intelligenz* 27, 1 (01 Feb 2013), 9–15. https://doi.org/10.1007/s13218-012-0228-x

[70] Martijn C. Schut. 2010. On model design for simulation of collective intelligence. *Information Sciences* 180, 1 (2010), 132 – 155. https://doi.org/10.1016/j.ins.2009.08.006 Special Issue on Collective Intelligence.

[71] Toby Segaran. 2007. *Programming collective intelligence : building smart web 2.0 applications.* O'Reilly, Beijing Sebastapol CA.

[72] V. Singh, G. Singh, and S. Pande. 2013. Emergence, Self-Organization and Collective Intelligence – Modeling the Dynamics of Complex Collectives in Social and Organizational Settings. In *2013 UKSim 15th International Conference on Computer Modelling and Simulation.* 182–189. https://doi.org/10.1109/UKSim.2013.77

[73] Aelita Skarzauskiene and Monika Maciuliene. 2015. Modelling the Index of Collective Intelligence in Online Community Projects. In *International Conference on Cyber Warfare and Security.* Academic Conferences International Limited, 313.

[74] Daniel Stieger, Kurt Matzler, Sayan Chatterjee, and Florian Ladstaetter-Fussenegger. 2012. Democratizing Strategy: How Crowdsourcing Can Be Used for Strategy Dialogues. *California Management Review* 54, 4 (2012), 44–68. https://doi.org/10.1525/cmr.2012.54.4.44

[75] Tim Sullivan. 2011 (accessed March 8, 2019). Embracing Complexity. https://hbr.org/2011/09/embracing-complexity

[76] Shweta Suran, V. Pattanaik, Sadok B. Yahia, and Dirk Draheim. 2019. Exploratory Analysis of Collective Intelligence Projects Developed Within the EU-Horizon 2020 Framework. In *Computational Collective Intelligence.* Springer International Publishing, 285–296.

[77] James Surowiecki. 2005. *The Wisdom of Crowds.* Anchor.

[78] Guy Theraulaz and Eric Bonbeau. 1999. A Brief History of Stigmergy. *Artif. Life* 5, 2 (April 1999), 97–116. https://doi.org/10.1162/106454699568700

[79] O'Reilly Tim. 2007. *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software.* MPRA Paper. University Library of Munich, Germany. https://EconPapers.repec.org/RePEc:pra:mprapa:4578

[80] John P. van Gigch. 1991. *System Design Modeling and Metamodeling.* Springer US. https://doi.org/10.1007/978-1-4899-0676-2

[81] Dimitrios J. Vergados, Ioanna Lykourentzou, and Epaminondas Kapetanios. 2010. A Resource Allocation Framework for Collective Intelligence System Engineering. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES '10).* ACM, New York, NY, USA, 182–188. https://doi.org/10.1145/1936254.1936285

[82] Jennifer H Watkins. 2007. Prediction markets as an aggregation mechanism for collective intelligence. (2007). https://escholarship.org/uc/item/8mg0p0zc

[83] Sean Wise, Robert A. Paton, and Thomas Gegenhuber. 2012. Value co-creation through collective intelligence in the public sector: A review of US and European initiatives. *VINE* 42, 2 (2012), 251–276. https://doi.org/10.1108/03055721211227273

[84] Michael A. Woodley and Edward Bell. 2011. Is collective intelligence (mostly) the General Factor of Personality? A comment on Woolley, Chabris, Pentland, Hashmi and Malone (2010). *Intelligence* 39, 2 (2011), 79 – 81. https://doi.org/10.1016/j.intell.2011.01.004

[85] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* 330, 6004 (2010), 686–688. https://doi.org/10.1126/science.1193147

[86] Koji Zettsu and Yasushi Kiyoki. 2006. Towards Knowledge Management Based on Harnessing Collective Intelligence on the Web. In *Managing Knowledge in a World of Networks.* Springer Berlin Heidelberg, 350–357. https://doi.org/10.1007/11891451_31

Received March 2019; revised October 2019; accepted October 2019