This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

Evolution of the Web of Social Machines: A Systematic Review and Research Challenges

Kellyton dos Santos Brito, *Member, IEEE*, Alysson Alves de Lima, Sérgio Endrigo Ferreira, Vanilson de Arruda Burégio, Vinicius Cardoso Garcia, *Member, IEEE*, and Silvio Romero de Lemos Meira

***Abstract*—Social Machines (SM) is the term used to define processes in which the people do the creative work and the machine does the administration. The concept was scarcely studied until 2013, when the series of Workshops on Social Machines was created, and the topic began to receive more attention. However, it is not clear how research has evolved since then. This study aims to investigate and summarize how the research field of SM has evolved since 2013, to outline the state of the art and the practice and identify research opportunities within this field. We performed a systematic literature review analyzing the quantity and quality of publications, the main topics addressed, the current classifications of SMs, the context in which the concepts are used, and the main perceived challenges. We identified and analyzed 56 relevant studies addressing 12 topics, representing the current practical landscape of research regarding SM. Our findings suggest that (i) research interest in SM is increasing, but is still concentrated into two research clusters; (ii) topics are grouped under two main headings: (a) human behavior and (b) software development; (iii) there is still a need for a common taxonomy to define and classify SM; (iv) the main contexts are crowdsourcing and social networks, and the majority of studies are small-scale studies in an academic setup; and (v) more empirical rigor and evidence is needed regarding their use, benefits and challenges, despite some evidence regarding challenges related to user engagement, trust, scalability and a better human-machine collaboration. Finally, a vision of the future of SMs, with the integration of web of people, artificial intelligence and things, is also presented and discussed.**

***Index Terms*— Human-Machine Interaction, Social Computing, Social Machines, Systematic Review, Web 2.0**

I. Introduction

The first use of the term Social Machines is attributed to Tim Berners-Lee [1]. In 1999, he defined it as "processes in which the people do the creative work and the machine does the administration." With certain exceptions, the idea was not thoroughly studied until the beginning of 2010, when two clusters of researchers began new initiatives to better explore the concept. One of these initiatives created the series of Workshops on Social Machines, aiming to bring researchers together to discuss and develop a research agenda.

Five years after the creation of the series of workshops, while there has been a notable intensification of research in this area, it remains unclear as to whether or how research has evolved. Questions regarding results, subareas, benefits and challenges are still very hard to answer; and to the best of our knowledge, we have been unable to find a study that summarizes or even discusses these questions.

The aim of this work is to contribute to close this gap by investigating and summarizing the state of the art and practice of Social Machines research field and identify key research challenges and opportunities. A systematic literature review was performed, starting with all the papers published in the Workshop of Social Machines from 2013 to 2017 and including, in a three-level snowballing strategy, studies that referred to them. We then filtered and analyzed the relevant studies in order to answer questions regarding the quantity and quality of publications, the main topics addressed, the context in which the concepts are used, current classifications of SM and the main perceived challenges.

The analysis was performed on 56 studies. Our findings suggest that research on SM is attracting attention from academics in several areas, mainly focused on 15 topics ranging from social software to software engineering. We also identified two main clusters of researchers, one focused more on human-machine interaction, and another focused on software aspects, such as relationship-aware systems. However, this field of research is still in its infancy, and

Manuscript received December 31, 2018; revised June 01, 2019; accepted December 04, 2019. Date of publication XXXX. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fei-Yue Wang.

This research was partially funded by INES 2.0, FACEPE grants PRONEX APQ 0388-1.03/14 and APQ-0399-1.03/17, and CNPq grant 465614/2014-0.

Kellyton dos Santos Brito is with the Informatics Centre, Universidade Federal de Pernambuco and Department of Computing, Universidade Federal Rural de Pernambuco, 52171-900 Recife/PE, Brazil (e-mail: kellyton@kellyton.com.br).

Alysson Alves de Lima, Sérgio Endrigo Ferreira, Vinicius Cardoso Garcia, and Silvio Romero de Lemos Meira are with the Informatics Centre, Universidade Federal de Pernambuco, 50740-560 Recife/PE, Brazil (e-mail: aal4@cin.ufpe.br, sebbf@cin.ufpe.br, vcg@cin.ufpe.br, silvio@meira.com).

Vanilson de Arruda Burégio is with Department of Computing, Universidade Federal Rural de Pernambuco, 52171-900 Recife/PE, Brazil (e-mail: vanilson.buregio@ufrpe.br).

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

some challenges were identified, such as the lack of a common taxonomy and classification, and the need for more empirical evidence regarding their use, benefits and challenges.

The remainder of this paper is organized as follows: Section II presents the background and previous studies related to this work, followed by Section III, which presents the review method and procedure employed in this study. Section IV provides an overall summary of the results and quality assessment, while Section V discusses the answers to the predefined research questions and highlights the findings. Section VI presents a discussion about the Future of Social Machines, and Section VII reviews the limitations of the study. Finally, Section VIII, concludes and summarize the outcomes.

II. Background and Previous Studies

In his book, Tim Berners-Lee [1] first defined Social Machines as "processes in which the people do the creative work and the machine does the administration."

Until 2010, very little additional research had been conducted based on this concept. We highlight the review published by Roush [2] in 2005, discussing the current technology shift and the concept of continuous computing. Roush also pointed out that "computing means connecting". Additionally, in the early stages of SM research, Meira [3] published a seminal paper bringing a new interpretation of SM in a particular setting:

"A network of programmable machines that are connected to each other and also connect people and institutions in a web of computing, communication and control that needed a much more abstract description and formalization than its external behavior in the form of a public (web) interface and number of APIs on top of the de facto standard internet protocols."

Meira's interpretation was not only focused on social relationships between people and machines, as defined by Berners-Lee, but mainly between machines. Furthermore, his work also proposed an algebra that was able to describe these networks.

In 2012, the "SOCIAM: The Theory and Practice of Social Machines" project was created: an effort funded by the Engineering and Physical Sciences Research Council (EPSRC) involving three leading UK universities: University of Oxford, University of Southampton and University of Edinburgh. The aim, according to the SOCIAM website (sociam.org), was "to produce the first major interdisciplinary research insights into the realm of social machines." Their premises were strongly based on Berners-Lee's initial definition of SM, as stated by Shadbolt [4]:

"*Social Machines can be characterised as assemblies of manually executed and machine-driven (as in 'automatised') services and the interaction of such services.*"

On the other side, in a partnership between Brazilian and the United Arab Emirate (UAE) researchers, the concepts presented by Meira were also developed, focusing on software engineering aspects and software architecture for the development of SM, as stated by Buregio [5]:

We characterize the "Social Machines" paradigm as a result of the convergence of three different visions: (i) Social Software, (ii) People as Computational Units, and (iii) Software as Sociable Entities.

In the first Workshop on Social Machines, which took place at the International World Wide Web Conference held in 2013, these and many other researchers had the opportunity to discuss and share their results in a specific, specialized forum. It may be argued that the first Workshop on Social Machines launched a new area of research; nevertheless, due to the short time span of research, it is still difficult to find studies that summarize the main aspects and results of these initiatives. One initiative in this direction is presented in [6]. Despite the book's objective is not the summarization and analysis of this research field, a strong and detailed correlation between Social Machines and Artificial Intelligence is presented. Another initiative is a very recent book, published after first submission of this work, that summarize the conclusions of SOCIAM project [7]. Although it provides a detailed discussion of Social Machines, the book focuses entirely on the SOCIAM project and neglect many relevant related studies, such as Meira's and Buregio's contributions.

III. Research Method

The method chosen for this research was a systematic literature review, which has proven to be an effective manner with which to identify, evaluate, interpret and compare studies that are relevant to a particular question or area [8]. Following the guidelines defined by [8], the method used in this research is defined below.

A. Research Questions

To define the research questions of this study, we returned to the main objective:

To investigate and summarize how the research field of Social Machines has evolved since 2013, to outline the state of the art and practice and identify research opportunities in this field.

Then, the following research questions were derived:

- RQ1: How many papers focusing on the research and practice of Social Machines were published between 2013 and 2017?

This question aims to verify whether the area is receiving attention from researchers and whether new studies are being conducted.

- RQ2: Which individuals and organizations are most active in SM-based research?

The objective of this question is to investigate if the initially-identified two groups of researchers are still active,

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

if other researchers and organizations are also investigating the area and how this research has spread around the world.

• RQ3. Which topics of Social Machines are being addressed?

Considering that SM is a broad area, covering both social and technical aspects, this question aims to identify which specific topics are being addressed.

• RQ4: What are the current classifications of Social Machines?

This question aims to identify how the characterization of SM evolved after the two initial classifications proposed by Shadbold and Buregio in 2013, whether they are being used in the practice and whether other classifications have replaced them.

• RQ5. In which contexts are the concepts of SM being used?

Aiming to verify whether SM concepts are being used in a particular context or application domains, this question is intended to ascertain as to whether the concepts of SM are best suited to or receiving particular attention from any particular context.

• RQ6. What are the main challenges of developing Social Machines?

This question aims to identify the main perceived challenges, and the knowledge gaps and opportunities in this new area.

B. *Research Team*

A team of six researchers have developed this study, three of whom, Kellyton Brito, Alysson Lima and Sergio Ferreira, are undergoing Ph.D. studies and composed the reviewer group. Vinicius Garcia, Vanilson Buregio and Silvio Meira are full-time lecturers and experts in the research and practice of SM and composed the supervisor group. It is important to note that, as presented in Section II, Silvio Meira and Vanilson Buregio may be considered pioneers in the research of SM.

C. *Decision Procedure*

All team members were involved in defining the scope, research objective, research questions and protocol, as well as discussing the findings. The review process was implemented by the reviewer group, under the direction of the supervisor group.

Important activities in a systematic study may lead to conflicts requiring decisions regarding study selection, quality assessment and data extraction. It is thus recommended that such activities be performed by at least two researchers. In order to address these situations, and to diminish threats to validity, for this study we defined a decision and consensus procedure, as described below.

During the review process performed by the reviewer group, the decision procedure began with all three researchers individually performing all activities related to the study selection, quality assessment and data extraction. After individual evaluation, the results were integrated into an Agreement/Disagreement table and a meeting was held. Afterwards, all the results with at least one disagreement were discussed by the members until a final consensus was reached.

D. *Search Process*

This research considered the creation of the series of Workshops on Social Machines was a landmark for initiating the systematic research within this area. Thus, the search strategy was a three-level reverse snowballing search based on articles published in the workshops from 2013 to 2017 and articles that referred to them, as follows:

- Level 1: All articles published in the workshops;
- Level 2: All articles referencing Level 1 papers;
- Level 3: All articles referencing Level 2 papers.

For Level 1 articles, a manual search was performed regarding the proceedings of the workshops. For Level 2 and Level 3, a manual search was performed using Google Scholar and ACM databases, which have options to list articles referencing specific articles. In order to verify consistency, an open search of the term "social machine" was also performed in electronic libraries (ACM Digital Library and IEEE Xplore) and samples of the results were compared with results of the snowball sampling. All search processes were performed in October 2017.

E. *Study Selection*

Study selection was performed by applying inclusion and exclusion criteria.

Inclusion criteria consisted of:

I.1 - Articles published between 2013 and 2017;

I.2 - Articles written in English;

I.3 - Articles published in peer-reviewed journals, conference or workshop proceedings;

I.4 - Articles published as full papers;

I.5 - Articles in which the study or the use of SM are directly addressed;

Exclusion criteria consisted of:

E.1 - Publications dated before 2013 or not written in English;

E.2 - Short or demonstration publications, theses, technical reports and book chapters;

E.3 - Publications in which SM are only briefly cited or not directly addressed.

The criteria used to verify short articles (I.4 and E.2), were (i) articles that are clearly mentioned as one of these categories; and (ii) articles with no more than 3 full pages, excluding references.

F. *Quality Assessment*

One initial difficulty regarding the quality assessment is that there is no established manner with which to define study "quality". In this study we have used the premise suggested by [9], in which quality relates to the extent to which the study minimizes bias and maximizes internal and

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

external validity. Thus, we focused the quality assessment on the rigor of the study. Hence, we proposed the following quality assessment questions:

- QA1: Is the research problem clearly specified?
- QA2: Are the research questions clearly identified?
- QA3: Are the findings/results clearly reported?
- QA4: Are bias and threats to validity clearly discussed?

The scoring procedure was: Yes, if the study clearly answered the question; Partially, if the answers were implicit or could be inferred by the reader; or No, if the study did not address the question. The quality score was Yes = 1, Partially = 0.5 and No = 0 for each question, and the overall quality of a publication was calculated by adding all the quality scores received.

G. Data Extraction and Synthesis

In accordance with previously defined research questions, this study used a data extraction schema to collect relevant data from primary studies. Collected data included publication metadata and evaluation data. Metadata was used to draw a general picture of research on SM, and to answer RQ1 and RQ2. Metadata included authors, affiliation, publication title and year as well as the venue type and name. Evaluation data was used to support the research of remaining research questions, such as whether the study was related to SM classifications, if it presented experimental results, and its context (place, scale and domain). In addition, data already gathered on quality assessment (whether the study discusses its results and validity) was included to support answers for RQ6. The data extraction schema is listed in Table I.

TABLE I. DATA EXTRACTION SCHEMA

ID	Attribute	Data Extraction Question	Resesearch Question
1	Author	Who is/are the authors(s)?	RQ1, RQ2
2	Affiliation	What is/are the author's affiliation(s)?	RQ1, RQ2
3	Publication Title	What is the title of publication?	RQ1, RQ2
4	Publication Year	In which year was the evaluation work published?	RQ1, RQ2
5	Venue Type	What is the venue type in which the study was reported (J=Journal, C=Conference, WS=Workshop)	RQ1, RQ2
6	Venue Name	What is the publication's venue? (Name and acronym)	RQ1, RQ2
7	Keywords	What is/are the publication's keywords?	RQ3
8	Main Topic	What are the main topics addressed in the study?	RQ3
9	Classification	Does the main theme of the study classify or categorize social machines (Y = yes, N = no)?	RQ4
10	Experimental Results	Does the study present experimental results (Y=yes, N=no)?	RQ5, RQ6
11	Context - Place	Context place (A = Academy, I = Industry) where the study was performed	RQ5
12	Context - Scale	Context scale (S = Small, L = Large) where the study was performed	RQ5
13	Context - Domain	Context domain where the study was performed	RQ5
14	Outcome Discussion	Does the study discuss the outcomes, such as main advantages and disadvantages, or benefits and challenges?	RQ6
15	Study Validity	Does the study discuss its limitations and validity?	RQ6

IV. REVIEW RESULTS

In this section, we provide an overall summary of the results and quality assessment. The findings and answers to the predefined research questions are discussed in the next section.

Following the search sequence, we gathered 252 studies using snowball sampling. After applying the first set of inclusion and exclusion criteria (I.1 - I.4 and E.1 - E.2), 175 potential studies was selected. After reading the abstract, introduction and conclusions of each paper and applying the last criteria (I.5 and E.3), which considers whether the study or use of SM are directly addressed by the paper, the final list of 56 selected studies was generated. A partial summary of data collected from the 56 studies is presented in Table II. Table III presents the quality scores for each assessment question.

It is important to highlight that in the last filter we not only considered studies clearly referring to the term "Social Machines", but also studies that referred to other concepts which, depending on the use, could be considered SM. Examples include social computing, when considering a notion of hybrid man-machine intelligence [10]; hybrid-computing elements, referring to systems as compositions of human-based computing elements and machine-based computing elements [11]; and works related to crowdsourcing applications [12], [13], which considers crowdsourcing as one of the most common examples of SM.

From the final list of 56 studies, 8 (14%) were published in academic journals, 22 (39%) published in proceedings and presented at conferences, and 26 (46%) were published in proceedings and presented at workshops. The high number of workshops was expected due to the prevalence of the SOCIAM workshop.

A total of 5 studies was found to be directly related with the classification or categorization of SM and were used as input for the analysis of RQ4. Moreover, 31 papers presented experimental evidence, and were used as input for RQ5 and RQ6.

Regarding quality assessment, we may highlight some limitations of the existing studies on SM. From the 56 selected papers, only 5 studies presented any kind of discussion on limitations or validity, such as conclusion validity, internal or external validity. Therefore, even if conclusions were presented, almost none of the studies could claim that (a) conclusions were related to the use of SM paradigms and no other variables (conclusion validity); (b) the experimental design was able to support the conclusions (internal validity); or (c) the results could be generalized to other situations (external validity). Moreover, only 31 studies (55% of the total) presented any kind of empirical evaluation supporting its results. In addition, 30% of the papers did not provide a clear discussion on the obtained results.

These limitations impaired the results of this review, particularly RQ6: "*What are the main challenges of developing Social Machines?*". This could not be fully answered due to the lack of empirical data and discussions for drawing conclusions. One of the first conclusions of this

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

study therefore, is that a more scientific rigor in studies on SM is required in order to better understand the area.

TABLE II. SUMMARY OF DATA COLLECTED FROM THE 56 PRIMARY STUDIES

Study Ref.	Year	Venue Acronym	Venue Type	Classification	Experimental	Context - Place	Context - Scale	Context - Domains
PS01	2015	EDOCW	C	N	Y	I	S	Enterprises
PS02	2014	CollaborateCom	C	N	Y	A	L	Educational, Social Network
PS03	2016	FORECAST	WS	N	N	A	-	-
PS04	2015	IW3C2	WS	N	Y	A	L	Web Observatory
PS05	2016	ICWE	C	N	N	A	-	-
PS06	2015	CARE/MFSC@AAMAS	C	N	Y	A	L	Urban Mobility
PS07	2013	IW3C2	WS	N	Y	I	L	Crowdsourcing
PS08	2015	Internet Computing	J	N	N	I	-	Enterprises
PS09	2016	PACIS	C	N	N	I	-	Customer Support
PS10	2015	WebSci	C	N	Y	A	S	CrowdSourcing, Social Network
PS11	2015	WebSci	C	N	Y	I	S	Enterprises
PS12	2015	WETICE	C	N	Y	I	L	Enterprises
PS13	2014	IEEE Intell. Syst.	J	N	N	A	-	-
PS14	2014	IW3C2	WS	N	Y	A	S	Creative Media
PS15	2013	IW3C2	WS	N	N	A	-	Crime Open Data
PS16	2016	IJRITCC	J	N	Y	A	S	Quizzes
PS17	2015	J. Web Eng	J	N	Y	A	S	Search, Politician, Urban Mobility, Movies, Images
PS18	2015	WebSci	C	N	N	A	-	-
PS19	2016	ISWC	C	N	N	A	-	-
PS20	2016	IW3C2	WS	N	N	A	-	-
PS21	2015	Comput. Netw	J	N	N	A	-	-
PS22	2016	ICGSEW	C	N	Y	I	S	Crowdsourcing
PS23	2016	WoT	C	N	Y	A	S	Social Network, IoT
PS24	2013	IW3C2	WS	N	Y	A	S	Social Network, Crowdsourcing
PS25	2014	IW3C2	WS	N	N	A	-	Social Network
PS26	2017	COGNITIVE	C	N	N	A	-	-
PS27	2013	EDOCW	C	N	N	A	-	-
PS28	2014	AAMAS	C	N	N	A	-	Social Computing
PS29	2015	WebSci	C	N	N	A	-	Iot
PS30	2015	IW3C2	WS	N	Y	A	-	Urban Mobility
PS31	2014	IW3C2	WS	N	Y	A	S	-
PS32	2016	CSCWD	C	N	Y	A	S	Crowdsourcing
PS33	2016	IW3C2	WS	N	Y	A	S	-
PS34	2013	IW3C2	WS	N	N	A	-	-
PS35	2015	IW3C2	WS	Y	Y	A	S	-
PS36	2016	AMECSE	C	N	Y	A	S	Crowdsourcing
PS37	2017	SAC	C	N	Y	A	L	Urban Mobility
PS38	2016	WebSci	C	Y	N	A	-	-
PS39	2016	Phenom Cogn Sci	J	N	N	A	-	-
PS40	2013	IW3C2	WS	Y	N	A	-	-
PS41	2015	IW3C2	WS	N	N	I	-	-
PS42	2014	IW3C2	WS	N	Y	A	L	Social Network
PS43	2013	IW3C2	WS	N	N	A	-	Health
PS44	2013	IW3C2	WS	N	Y	A	S	-
PS45	2016	IW3C2	WS	N	N	A	-	-
PS46	2013	IW3C2	WS	Y	Y	A	S	-
PS47	2016	IW3C2	WS	N	Y	A	S	Social Network
PS48	2015	IW3C2	WS	N	Y	I	L	Government
PS49	2017	CICM	C	N	Y	A	S	Forums
PS50	2017	ACM Comput. Surv	J	N	Y	A	S	-
PS51	2016	IW3C2	WS	N	N	I	-	-
PS52	2016	ARES	C	N	Y	I	S	Software teams
PS53	2016	IW3C2	WS	Y	Y	A	S	-
PS54	2013	Found. Trends Web Sci.	J	N	N	A	-	-
PS55	2015	IW3C2	WS	N	Y	A	S	Crowdsourcing
PS56	2014	IW3C2	WS	N	N	A	-	CrowdSourcing, Social Network

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

TABLE III. QUALITY ASSESSMENT SCORES

Study Ref.	QA1	QA2	QA3	QA4	Score
PS01	1	0	0	0	1
PS02	1	0	1	1	3
PS03	1	0	0	0	1
PS04	0	0	1	0	1
PS05	0,5	0	0	0	0,5
PS06	0,5	0,5	1	0	2
PS07	1	1	1	0	3
PS08	1	0,5	0	0	1,5
PS09	1	0,5	0	0	1,5
PS10	0,5	0,5	1	0	2
PS11	0,5	0,5	1	1	3
PS12	1	1	0	0	2
PS13	1	1	0	0	2
PS14	1	0,5	1	0	2,5
PS15	1	1	1	0	3
PS16	0,5	0,5	1	0	2
PS17	1	0,5	1	0	2,5
PS18	1	0,5	0	0	1,5
PS19	1	0,5	1	0	2,5
PS20	1	0,5	0	0	1,5
PS21	1	1	1	0	3
PS22	1	0,5	1	0	2,5
PS23	1	0,5	1	0	2,5
PS24	1	0	1	0	2
PS25	1	0,5	1	0	2,5
PS26	1	1	1	0	3
PS27	1	1	0	0	2
PS28	1	0,5	0	0	1,5
PS29	0,5	1	1	0	2,5
PS30	0	1	0,5	0	1,5
PS31	0,5	1	1	0	2,5
PS32	1	1	1	1	4
PS33	1	1	1	0	3
PS34	0,5	0,5	0	0	1
PS35	1	0,5	1	0	2,5
PS36	1	0,5	1	0	2,5
PS37	1	0	1	0	2
PS38	1	1	1	0	3
PS39	1	1	0	0	2
PS40	1	1	1	0	3
PS41	1	1	0	0	2
PS42	1	0,5	1	1	3,5
PS43	1	1	1	0	3
PS44	1	1	1	0	3
PS45	1	0,5	0	0	1,5
PS46	1	0,5	1	0	2,5
PS47	0,5	0	0,5	0	1
PS48	1	0,5	0	0	1,5
PS49	0,5	0,5	1	0	2
PS50	1	0	1	0,5	2,5
PS51	0,5	1	0,5	0	2
PS52	1	0,5	0	0	1,5
PS53	0,5	0,5	1	0	2
PS54	1	1	1	0	3
PS55	0,5	0,5	1	0	2
PS56	0	0,5	0,5	0	1

V. DISCUSSION OF RESEARCH QUESTIONS

In this section, we address the research questions presented in Section III.

A. RQ1: How many papers focusing on the research and practice of Social Machines were published between 2013 and 2017?

The distribution of publishing years is shown in Fig. 1. The distribution demonstrates that, despite a small decrease in 2014 in comparison with 2013, the number of publications began to increase over the following years. One of the reasons for the decrease in 2014 may have been that researchers who published in the first year of the workshops were still developing new results and thus did not publish in 2014. In fact, only three authors, who are co-authors of the same papers in the period, published both in 2013 and 2014. On the other hand, in 2015 and 2016 the number of publications increased, and authors of the 2013 publications appeared once more, presenting new results from their research.

It should be noted that as the search process was performed in October 2017, many papers published in 2017 were not retrieved by this study and were therefore not considered in this analysis.

B. RQ2: Which individuals and organizations are most active in SM-based research?

The 56 selected studies were conducted by a total of 125 authors and co-authors from 18 different countries. Thirty authors were involved in more than one study, and 10 authors were involved in 4 or more studies. A list with authors of three or more publications is presented in Table IV.

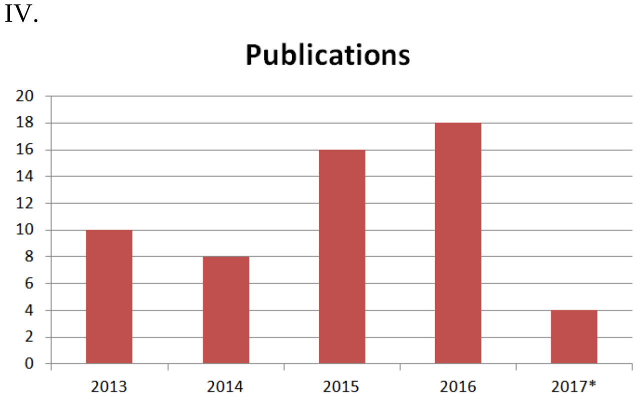


Fig. 1. Study distribution over the publication years

There is no surprise regarding the two most active authors. Indeed, they not only presented the two first classification frameworks of SM in 2013 but have continued to conduct their research on this topic. We also detected groups of researchers who have frequently co-authored studies. There are clearly two researcher and institutional clusters: the

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

SOCIAM project, in the UK, headed by Shadbolt, and a partnership between Brazilian and UAE universities, headed by Buregio.

This data is reinforced by the analysis of the institutions. Authors were distributed amongst 43 institutions, 12 of which appeared in more than one paper. From these, 7 are from one of two identified clusters. On the other hand, new research on SM was also performed in Cairo University (Egypt), the International Institute of Information Technology (India) and the University of Lyon (France). The list showing institutions with one or more papers is presented in Table V.

TABLE IV. AUTHORS OF 3 OR MORE PUBLICATIONS

Author	Papers	Author	Papers
Nigel Shadbolt	8	Daniel A. Smith	3
Vanilson Burégio	7	David De Roure	3
Dave Murray-Rust	5	Eman S. Nasr	3
Ramine Tinati	5	Kieron O'Hara	3
Silvio Meira	5	Mervat Gheith	3
Wendy Hall	5	Michael Rovatsos	3
Leslie Carr	4	Pip Willcox	3
Max Van Kleek	4	Susan Halford	3
Nelson Rosa	4	Tarek Ali	3
Ségolène Tarte	4	Thanassis Tiropanis	3
Clare Hooper	3	Zakaria Maamar	3

These 56 studies focused on SM scattered amongst 27 different venues. As expected, the Workshop on Social Machines (formally identified as the "International World Wide Web Conference Committee - IW3C2"), was the main venue, with 25 publications. In second place, the ACM Web Science Conference (WebSci) appeared with 5 publications. In fact, since 2015 the conference has included "Social machines, collective intelligence, and collaborative production" as possible topics for submission, encouraging additional proposals in the field of SM. In addition to these venues, two articles were published in EDOCW, and all other venues published only one article. This data indicates that the Workshop on Social Machines is still the main forum for researchers in this area, and, although the topic is receiving attention from other web-related conferences, it is still dependent on the workshop.

Considering the mentioned dependence on the SOCIAM Workshop, it should be mentioned that it was not held in 2018, which thereby indicates that the last edition was in 2017. Thus, a further question needs to be investigated over the coming years as to whether or not research in the area will continue and grow in other venues.

C. *RQ3. Which topics of social machines are being addressed?*

Topic identification was performed according to grounded theory data analysis presented by Strauss and Corbin [14]. The first step in encoding the data is the *open coding* process, when researchers first assign all meaningful quotations to a major higher-level category of information. Next, researchers organize these categories further through *axial coding*, and categories emerge from related sub-categories. Finally, *selective coding* is performed, in order to classify all categories around a core category.

In this study, this process was performed twice: (i) by researcher qualitative open coding; and (ii) using study keywords as the input of open coding step.

Qualitative Coding and Analysis

In this analysis, the open coding process was performed by researchers, who read each study abstract, introduction and conclusion, and assigned an open quotation to them. It was suggested that 2 or 3 quotes should be assigned to each study.

In the first round of open coding 110 different quotations were assigned to the 56 studies. By aligning synonyms, the number was reduced to 98. Then, after axial coding similar categories were grouped together, and topics were grouped into 12 different categories, as presented in Table VI.

TABLE V. INSTITUTIONS WITH MORE THAN 1 PUBLISHED PAPER

Institution	Country	Papers
University of Southampton	UK	19
University of Edinburgh	UK	11
University of Oxford	UK	8
Federal University of Pernambuco	Brazil	5
Federal Rural University of Pernambuco	Brazil	4
National Institute of Science and Technology for Software Engineering	Brazil	4
Cairo University	Egypt	3
Independent Researcher	Egypt	3
Zayed University	United Arab Emirates	3
International Institute of Information Technology	India	2
Recife Center for Advanced Studies and Systems	Brazil	2
University Lyon	France	2

Most present categories include software engineering and development (37 studies), and collaboration: crowdsourcing and collaboration (between humans) and human-machine collaboration. Also, other technical categories were also identified, such as data-centric approaches, artificial intelligence, internet of things and smart cities and enterprises. Other non-technical categories, related with human behavior and concerns were also verified, such as questions regarding security and privacy, individuals and web observatories. Finally, studies regarding theory and definitions, as well as the agenda and future of SM was also established.

Lastly, through the selective coding, analyzing and grouping these categories, three main areas emerged in the field of social machines: (i) **human behavior**, including collaboration amongst humans and amongst humans and machines, concerns about human behavior and their security and privacy, and web observatories; (ii) **SM software development**, including software engineering, data-centric approaches, artificial intelligence and smart things, cities and enterprises; and (iii) **theory and future**, including the study of the foundations and future of SMs.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

TABLE VI: IDENTIFICATION OF TOPICS BASED ON QUALITATIVE ANALYSIS, INCLUDING THE NUMBER OF STUDIES IN EACH CATEGORY

Topics	Emerged Categories	Selective Coding	
Collaborative Software Development, Collaborative Systems, Crowdsourcing, Crowdsourced Development, Crowdsourcing Applications, Crowdsourcing Process, Online Communities	Crowdsourcing and Collaboration (12)	Human Behaviour (46)	Social Machines
Socio Technical Networks, People as Social Machines, Human-Based and Machine-Based Computing Elements, Human Machine Agenda, Human Machine Cooperation, Computational Social Systems, People as Computational Units, Human Machine Collaboration, Human-Machine Systems, Social Computation	Human-machine cooperation (18)		
Privacy, Identity in Social Machines, Trust, Trust and Reputation, User Personal Data, Web Use Constraints, Identity	Security and privacy (9)		
Impact on Human Lives, Individual, Human Behavior, Human Habits	Individuals (4)		
Web Observatories, Social Machines Observation	Web observatories (3)		
Social Machines Architecture, Social Machines Development, Social Machines Choreography and Coordination, Social Service Component, Software Api, Software Composition, Software Development, Software Development Cycle, Software Engineering, Software Model, System Design, System Design/Architecture, Systems Development Life Cycle, Collective Adaptative Systems, Development Model, Collaborative Software Development, Social Machines Coordination, Social Machines Interaction, Self-Adaptative Software, Social Machines Certification/Monitoring, Social Machines Assessment, Multi Agent Systems, QoS, Requirements Elicitation, Social Orchestration, Modelling, Distributed Computing	Software Engineering and Development (37)	Software Development (56)	
Semantic Web, Open Data, Government Data, Linked Data, Data Curation, Data-Driven Systems, Real-Time Data Processing	Data-centric (9)		
Machine Intelligence, Machine Learning, Artificial Intelligence	Artificial Intelligence (4)		
Iot, Web of Things	Internet of Things (3)		
Smart Cities, Smart Enterprise, Social Enterprise	Smart Cities and Enterprises (3)		
Social machines characterization/classification, Theory and Definitions	Theory and Definitions (12)	Theory and Future (15)	
The Future of The Web, Philosophy, Social Web Agenda	Agenda and Future (3)		

Keyword Analysis

The direct keyword analysis was performed to check the results of open coding and verify whether any important topic had escaped consideration. In this analysis, we extracted 155 different keywords from the 56 studies. Of these, only 26 appeared more than once, the main highlights being social machines with 29 occurrences, crowdsourcing (6 occurrences), web science (5), web observatory/observatories (4) and socio-technical systems (4), followed by software development-focused topics, such as software engineering (3), design (3), distribution and social computing model (3), and computing element (3).

By performing axial coding without considering the generic keyword Social Machine/Machines, we classified the addressed keywords into 13 categories, as presented in Table VII. Within these categories, some people-centered topics may be noted, such as (i) crowd engagement and online communities, (ii) social computing model, where people participate actively in the software activity; (iii) transparency and participation and (iv) cities and citizens, focusing on the engagement of citizens in government and city activities; (v) identity, privacy and trust, mainly focused on the security of personal data sharing; and (vi) other socially-related topics. On the other hand, software related topics may also be encountered, such as (vii) software engineering, in a broader approach; (viii) agent-based software; (ix) internet of things; and (x) artificial intelligence. Finally, topics regarding (xi) workflows and information flows and (xii) web science, may also be identified.

By performing selective coding, we may identify two core groups. The first consists of categories (i) – (vi) and is related to people activities as part of social machine systems, and deal with its related concerns, such as people engagement and security concerns. The second is related to the development of software that, in many cases, supports the activities of the first group, and consists of categories (vii) – (xii). The union of these categories forms the core category of Social Machines.

This result is very compliant with the open coding results, through the identification of two main subareas: one related to (i) human behavior and people-involved computing; and another related to technical aspects of (ii) SM software development. In addition, some identified subcategories are the same (such as software engineering) or very similar, such as *crowdsourcing and collaboration* and *crowd engagement and online communities*. Thus, based on these similarities and considering that the first classification presented in Table VI is more complete, including the theory and future category, we consider this topic identification the most appropriate.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

TABLE VII. IDENTIFICATION OF TOPICS BASED ON KEYWORDS

Keywords	Emerged Categories	Selective Coding	
Business Process Crowdsourcing, Co-Creation, Collaborative Filtering, Collective Intelligence, Community, Crowd Work, Crowdsensing, Crowdsourced Software Requirements, Crowdsourcing, Crowdsourcing Software, Mass Collaboration, Online Communities, Online/Offline Community	Crowd engagement and Online Communities (23)	People involved computation (89)	Social Machines
Social Computing Model, Collective Adaptive Systems, Computing Element, Distributed Cognition, Human Computation, Human Factors, Human-Based Computation, Human-Machine Networks, Q&A, Social Computation, Social Computing, Sociotechnical Systems, Socio-Technical Systems, User-Generated Content	Social Computing model (23)		
Democracy, Digital Rights, Open Data, Open Government, Participation, Transparency, Web Observatories, Web Observatory	Transparency and Participation (10)		
Cities, Citizen Engagement, Smart City, Urban Interactions, Urban Planning, Urbanization, Wayfaring	Cities and Citizens (7)		
Data Security, Identity, Individual, Personal Apis, Personal Data, Privacy, Pro-Human Web, Trust	Identity, Privacy and Trust (11)		
Social Action, Social Artifact, Social Customer Support System, Social Enterprise, Social Feature, Social Intelligence, Social Media, Social Network(S), Social Software, Social Theory, Social Web, Sociality	Other Social Related (15)		
Anonymized Requirements Elicitation, Architecture, Crowdsourced Software Requirements, Design, Dynamic Requirements Documentation, Software Architectures, Software Engineering, Web Engineering	Software Engineering (12)	Software development aspects (50)	
Agent Communication, Autonomous Agents, Web Agents	Agent based software (3)		
Automation, Internet Of Things, Web Of Things	Internet of things (4)		
Machine Intelligence, Machine Learning	Artificial Intelligence (2)		
Data Growth, Database, Linked Data	Data centric (4)		
Information Cascades, Information Retrieval, Information Systems, Information Theory, Task Management, Workflow, Workflow Composition, Workflow Optimisation, Workflow Orchestration	Workflows and Information Flows (11)		
Web, Web 2.0, Web 2.0 Application, Web Agents, Web Engineering, Web Of Things, Web Platforms, Web Science, Web-Oriented Systems	Web Science (14)		

D. RQ4: What are the current classifications of Social Machines?

To answer this question, data extraction identified which papers were related to classifying social machines, and five papers were selected.

Two seminal papers were already expected, the works of Shadbolt [4] and Buregio [5], both from 2013.

Shadbolt aimed to define an initial classificatory framework based on areas of (a) contributions (tasks, purpose and context of participation), (b) participants (participants and roles) and (c) motivation (motivation and incentives), and constructs related to each area. Examples of constructs include:

• Tasks, purposes and context of participation: activities involving creative production of content/subjective appraisal of content/solving definable computation tasks; tasks are domain specific, and physical location of participation is relevant to the service.

• Participants and roles: participant autonomy; participant anonymity, generality of audience.

• Motivation and incentives: participants are intrinsically motivated (to gain/share knowledge, to be social, for the benefit of society as a whole, etc.), or are motivated by extrinsic reward (payment or status). The list of consolidated constructs of SM, according to Shadbolt, is presented in Fig. 2.

Even though Shadbolt's classification also includes solving computation tasks or sets of problems, his work is strongly based on the human-machine interaction.

Tasks, purpose and context of participation
Activities involve creative production of content
Activities involve subjective appraisal of content
Activities involve solving (a definable) computation task or set of problems
Tasks are domain-specific
The machine owner derives benefit from participation
Activities and tasks are pre-defined or participant-defined
Variation in types of contributions and tasks
Participants' participate to fulfil needs of a role
work-related/professional role
home/family related
social role
leisure/entertainment role
Participation is done via:
mobile devices
Web browsers
apps
sensors/sensing (location sensing and wearable devices)
Participation is done in a mobile context
Physical location is relevant to the service
Participants and roles
Generality of audience
Participant autonomy
Participant anonymity
Extent of hierarchical organisation of roles
Clear separation of roles among participants
Motivation and incentives
Participants are intrinsically motivated:
to gain/share knowledge
to "get something done"
to "be for fun/entertainment"
to "be social"
for the benefit of a specific group of people who need help
for the benefit of society as a whole
Participation is motivated by extrinsic reward (payment, status)

Fig. 2. Consolidated constructs of SM[4]

On the other hand, Buregio proposed a broader classification framework, characterizing SM as a convergence of three visions: (i) social software; (ii) people as computational units;

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

and (iii) software as sociable entities. By using Buregio's framework, all social machines identified by Shadbolt could be classified as social software (such as Facebook) or people as computational units (such as reCaptcha). In addition, Buregio focused on software as sociable entities, such as some types of web services and, perhaps the most innovative contribution of his work, relationship-aware systems. This classification is presented in Fig. 3.

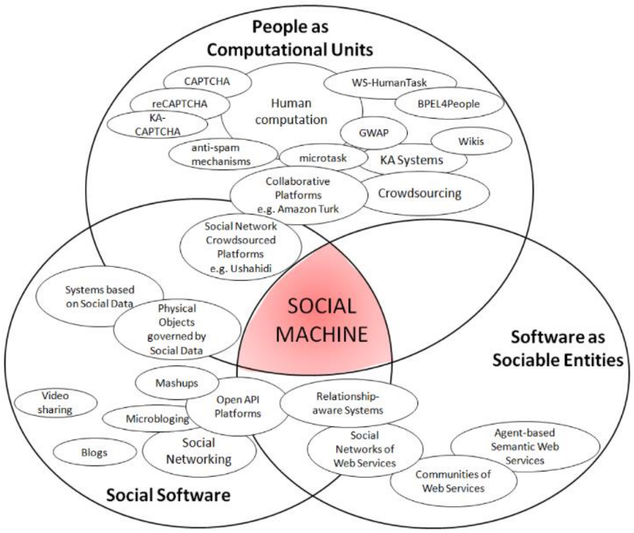


Fig. 3. Classification of social machines by Buregio[5]

In addition to these seminal studies, three others were identified. In 2015, Vass & Munson [15] presented one approach to describe and analyze SM based on developing principles of social constructionist theory adapted for web science. They focus on the concepts of agency, which may be human or non-human, and reflexivity, composed of levels of recognition and responsivity.

Based on qualitative observations and their experiences building SM, in 2016 Hooper et al. [16] presented an approach based on scope and change. Hence, they defined four analytical dimensions: (i) geographical scope, (ii) temporal scope, (iii) social scope, and (iv) changing: functionality, users, data and context. Also in 2016, Halcrow et al. [17] proposed a further classification, a model of an online/offline community to analyze the sociality of SM, called the SPENCE model. The model is constituted of 5 facets: (a) Settlement, (b) Proximity, (c) Exchange, (d) Network, (e) Channels, and (f) Entrepreneurship, and each facet has one or more concepts and sub-concepts, as presented in Fig. 4. For example, the settlement facet has external and internal concepts, proximity may be geographical or psychological, and exchange may occur through communication, information-seeking or diffusion.

By analyzing these classification frameworks, certain conclusions may be drawn. First, to date, there is no common classification framework with which to describe and analyze social machines. Identified studies presented classifications starting from different viewpoints, as presented in this section, and it is even difficult to compare proposed classifications. Similarly, they did not present case studies or practical data on the same social machines, thereby not allowing a comparison of the presented models.

In addition, despite different starting points, some classification overlaps may be identified. For example, in spite of using different definitions, three studies, [4], [16], [17], defined a geographical characteristic of SM. In addition, using different words and definitions, three studies, [15]–[17], presented some characteristics very close to what Buregio [5] had already defined as relationship-aware systems.

Based on these results, we are able to indicate two future studies that could lead to improvements in the area of characterizing SM: (i) one that applies all of these classification models to the same set of social machines, in order to evaluate and compare their quality and usefulness; and (ii) one that identifies their similarities and proposes a common framework for understanding, defining and classifying social machines.

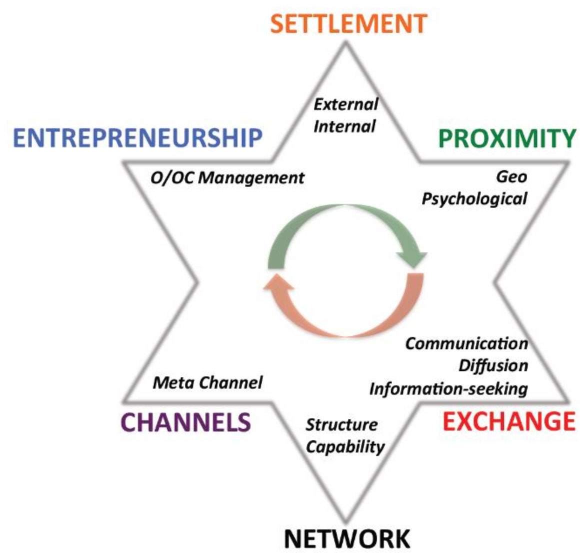


Fig. 4. The SPENCE model [17]

E. RQ5. In which contexts are the concepts of SM being used?

In accordance with suggestions by [8], this study has considered context to signify (a) place: academia or industry; (b) scale: small or large; and (c) application domain.

Identifying the context of studies was a challenging task. First, many studies were purely theoretical, and experimental evidence could be found in only 31 studies (55% of the total). In addition, even in these papers, the context was very often not clearly stated.

Regarding place, we considered as academia all studies that did not clearly state a relationship to an industrial environment. Considering this criterion, the majority (80%) of the 56 studies were performed within an academic context, while only 20% were performed in an industrial setup.

For scale analysis, only the 26 studies with experimental evidence were taken into consideration. From these, scale information could be found in 21 studies: 4 were performed in large scale projects (more than 100 estimated subjects), while

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

the other 17 were small-scale projects. These numbers signify that only 7% of the studies provided empirical evidence based on large-scale projects, thereby constituting a very small percentage of the total.

Finally, we planned to perform the same process previously presented and used in the analysis of Q3, open coding and axial coding, for domain identification. Thus, by performing an open coding step only 19 domains were identified in the 56 studies, even considering that some studies were applied in many domains. We therefore performed a simplified qualitative analysis, and have presented all the identified domains in Table VIII.

Most identified domains pointed to crowdsourcing (10 studies) and social networks (9 studies). This may be explained because many researchers do not aim to build new social machines, but rather focus on analyzing those that already exist, mainly Wikipedia and Galaxyzoo/Zooniverse (crowdsourcing), and Facebook and Twitter (social networks). However, some new proposals may be found, such as studies related to building enterprise SMs, urban mobility (including ridesharing) and studies related to the internet of things, thereby demonstrating that these areas, focused not only on the web but also in “offline” topics, are receiving some attention from researchers.

F. RQ6. What are the main challenges of developing Social Machines?

As presented in the quality assessment (Section IV), there is a lack of empirical data to assess the challenges (and benefits) of developing social machines: only 55% of the studies presented any empirical data, and only five (9%) discussed the limitations or threats to validity. In addition, even when empirical data was present, in most cases the discussion was not directly related to the development of SM, but to an associated area such as crowdsourcing. Furthermore, in 30% of the papers, a clear discussion of the results was missing. We therefore consider that the first challenge of research in the area of social machines is to increase the number of empirically based studies, with a clear discussion on its results and limitations.

TABLE VIII. IDENTIFIED DOMAINS

Domain	Studies
crowdsourcing	10
social network (s)	9
enterprises	4
urban mobility	4
iot	2
creative media	1
crime	1
customer support	1
education	1
forums	1

Domain	Studies
government	1
health	1
images	1
movies	1
politics	1
quizzes	1
search	1
software teams	1
web observatory	1

To identify the main challenges of developing SM, we considered the five studies that discussed their limitations, where we may usually find the problems encountered, challenges and proposals for future work. Three of them presented results directly related to the development of SMs.

The authors of [18] focused on user engagement of SM ecosystems. They concluded that user values and experiences are essential factors for the success or failure of a SM. The study also indicated the importance of presentation and the granularity of geographical and time data, as well as the existence of mechanisms to facilitate trust and fun within a SM.

Trust was also presented as a challenge in conjunction with scalability in [19]. Moreover, they highlighted that a better characterization of the human-machine network is still needed, in agreement with our discussion in RQ4, regarding research on human-machine cooperation. Finally, [13] performed some experiments focused on testing their tool in "diverse and complex decisions of crowdsourcing processes". Despite presenting some lessons learned, although many of them were not directly related to the development of social machines, one of the main findings is that DSSs (decision support systems) may improve the user performance in designing crowdsourcing (seen as social machines) processes.

In addition to this discussion, we would reinforce that the main conclusion is that the very first identifiable challenge of SM research is the need for more empirical evidence, preferably capable of generalizations, and to analyze and assess the real impact of the initiatives of SM.

Moreover, by analyzing other research questions, more challenges were identified:

As presented in Subsection B of Section V, most publications have appeared in the Workshop on Social Machines, held in conjunction with the World Wide Web conference. However, it seems that the last year of this workshop was in 2017, verification is required as to whether the research and publishing of studies related to SM will spread to other venues.

As presented in Subsection D of Section 5, there is still a lack of a common definition and classification of social machines. A cross-study was not encountered that used or compared current classifications, nor a study that identified their similarities and proposed a common framework in order to understand, define and classify social machines.

As presented in Subsection E of Section 5, the concept has mainly been studied in the domains of crowdsourcing and social networks. Although there are studies regarding other topics, such as enterprises, urban mobility and IOT, it is still very incipient, and more research focused on domains other than crowd and social networks are needed.

VI. THE FUTURE OF SOCIAL MACHINES

The results indicate that SM research is still in its infancy. In this section, a discussion about its future is presented.

A. Common Definitions

As Shadbolt [7] observes, “*Social machines are ... in one sense, as old as the hills, and in another, as novel as the technologies that underlie them—the World Wide Web, social networking, smartphones, and so on.*” Indeed, researchers use the term in a variety of ways. In some cases, they refer to any kind of human–computer interaction system or social network. Thus, as presented and discussed in Section V, studies aimed at defining a common SM taxonomy are necessary because clear definitions may enhance its development by reducing misunderstanding.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

B. The Human Cloud, Collective Intelligence, Security, and Privacy

SMs and the social web are on the top of the era of people as a service—when people, or “the crowd”, are requested by machines to solve tasks, issues, or problems that require cognitive intelligence. This phenomenon is also called “the human cloud” [20][21], a metaphor from a machine’s point of view: machines and software in the digital world access the cloud (of humans) as needed to perform difficult tasks.

This scenario promises to amplify computers’ capabilities and lead to various new questions, such as (i) how to persuade people to provide their time and effort as a service; (ii) how to manage and merge people, tasks, and knowledge; and (iii) how to deal with people data and inherent tradeoffs, such as privacy, accountability, security, publicity, and copyright. Indeed, who is the owner and responsible party for systems based on collective intelligence—and how trustworthy are these systems?

C. Artificial Intelligence and Intelligent Social Systems

Advances in artificial intelligence (AI) and machine learning models and algorithms, allied with the advent of big data, facilitate the processing and learning of information on an unprecedented scale.

However, AI requires human knowledge in many forms, such as problem modeling and providing training data. As noted by [22], one limiting factor of automated processing solutions is the availability of sufficiently well-structured annotated training data sets. However, one of the key advantages of the human cloud is the provision of such data sets. Thus, as [6] observes, AI needs SMs.

The integration of human (collective) intelligence and AI is also quite promising: if humans provide annotated data for AI learning systems, these systems can process the data, learn from them, and apply new knowledge to solve problems more efficiently. Research regarding this continuous integration is therefore quite promising.

D. Social Internet of Things (SIoT)

In parallel with current SM theory and practice, researchers have investigated the potential for integrating social networking concepts into internet of things (IoT) solutions. The social behavior of such “things,” including main concepts, architecture and network characterization [23], thing relation modeling [24], and the evolution of things into social objects [25], has also begun to attract interest.

In this scenario, the social Iot (SIoT) and SMs may be seen as complementary, and studies regarding the evolution of “things” from mere sensors and actuators to devices with intelligence (artificial, collective, or a mix thereof) and social behavior is also promising.

E. Software Architectures and Reengineering the Web

Proper software architecture is necessary to support this new web of people integrated with AI and the IoT. The classification presented on Table VI shows that most studies are related to software engineering and development, especially the architecture of SM systems and their scalability, security, and interoperability requirements.

In this sense, SMs may be considered a high-level abstraction of self-contained sociable software-or-human computational units. In this vision, they would be implemented using current architectural styles and technologies, such as microservices, and can inherit their concepts and capabilities.

On the other hand, SMs can also lead to the definition of a new software architecture style. In this sense, the majority of the field remains unexplored in terms of definition and adoption, as well as requirements, constraints, and capabilities.

We can envision the future of SMs as an integration of the web of people, the web of AI, and the web of things, all supported by proper software architectures and technologies capable of integrating them in an efficient, scalable, and secure way. Figure 5 presents this vision.

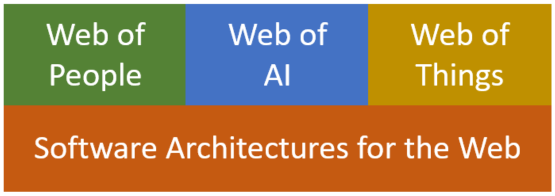


Fig. 5. The web of people, AI, and things

In the near future, when communicating with someone digitally, we will not be able to tell if we are communicating with another person, an AI system, an object (thing), or a combination of them.

VII. Limitations of This Study

Despite the rigor with which this study was conducted, it is possible that it may have been affected by the threats of validity, particularly with regard to finding all the relevant studies, assessing their quality and extracting data.

Given the increasing number of studies in the area of SM, there is no guarantee that all the relevant studies were identified. Even by applying the snowballing strategy on Google Scholar and ACM databases, some papers may have escaped inclusion. To reduce this issue, an open search of the term "social machine" was performed in electronic libraries (ACM Digital Library and IEEE Xplore) and samples of those results were compared with already-collected papers.

Additionally, study collection was performed in October 2017, signifying that many studies published in 2017 were not included.

Quality assessment and data extraction were individually performed by each of the three reviewers. Disagreements were discussed in a consensus meeting. In the last instance, members of the supervisor group were called upon for final decisions. Although this procedure increased our confidence in the reliability of this study, we nonetheless found that quality assessment and data extraction may have been compromised by the way most of the studies were reported. The report organization of some studies made it difficult to locate the required information in the extraction process. Furthermore, many papers did not present sufficient information, and, in many cases, information had to be inferred from the text.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

Therefore, despite the effort to reach a consensus during data extraction and quality assessment, there may have been some inaccuracies in the inferred data.

VIII. Concluding Remarks and Future Works

This study analyzed 252 articles, of which 56 were focused on Social Machines, investigating and summarizing how this new research field had evolved since 2013. Amongst these studies, 5 presented classification schemas, 31 presented some empirical evaluation, 11 were related to an industrial context and 4 were applied in a large-scale context.

This study demonstrates that the number of publications focused on SM is increasing, mainly in two research clusters located in the UK and Brazil-UAE respectively. Publications also concentrated on the SOCIAM Workshop, but have started to spread to other conferences, such as WebSci. Because the last edition of the workshop took place in 2017, it is worth investigating whether research in this area will actively grow in another venues.

Second, identified research subareas may be characterized by three groups: (a) human behavior, using social machines; (b) software development, developing them; and (c) theory and future. Human behavior focuses mainly on crowdsourcing and human collaboration, human-machine cooperation, security and privacy, individual concerns and web observatories. Software development focuses on software engineering and the development of social machines, data-centric approaches, the use of artificial intelligence, and the use of SM concepts on internet of things and smart cities and enterprises. Finally, theory and future studies theoretical aspects.

Third, while there are some characterizations and classifications of SMs, there is still a lack with regard to common definitions, classification and especially their use. Thus, there is a need for further studies based on (and using) current characterizations, and especially that identify their similarities and propose a common framework to understand, define and classify social machines.

With regard to context, this study identified that the main contexts in which SM are used are mostly social networks and crowdsourcing, but some may lead to new research, such as enterprise systems, internet of things and urban mobility. However, most studies are small in scale and performed in academic settings, and more large-scale studies, performed in an industrial setup are needed in order to gain a better understanding of the area. Also, as main challenges not already summarized, we would highlight the lack of empirical evidence, concerns about user engagement, trust, scalability, and a better human-machine collaboration.

Finally, the study presented the authors point of view of the future of SM as an integration of the web of people, the web of AI, and the web of things, all supported by proper software architectures and technologies capable of integrating them in an efficient, scalable, and secure way.

The results from this review have contributed to the research field of SM by providing the academic community with a better understanding of the research landscape of SM and illustrate some of the gaps in the area that opens opportunities for future research. In this sense, our future work will unfold in two directions: (i) to perform a study defining a common classification of social machines, based on the studies identified in this review; and (ii) to define, discuss and implement a software engineering agenda for the implementation of SM software systems.

This literature review may also be extended in certain ways. A search extension may be performed to expand the search strategy and number of sources, thereby performing a broader study; a temporal update would be performed without modifications to the protocol, to expand the timeframe and compare results over the time; and finally, both approaches would be combined.

Appendix - List of Primary Studies

[PS01] Z. Maamar, V. Buregio, N. Faci, D. Benslimane, and Q. Z. Sheng, "'Controlling' Web 2.0 Applications in the Workplace," in 19th International Enterprise Distributed Object Computing Conference, 2015.

[PS02] D. Murray-Rust, O. Scekic, H.-L. Truong, D. Robertson, and S. Dustdar, "A Collaboration Model for Community-Based Software Development with Social Machines," in Proceedings of the 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2014.

[PS03] A. Aldini, "A Formal Framework for Modeling Trust and Reputation in Collective Adaptive Systems," in Workshop on FORmal methods for the quantitative Evaluation of Collective Adaptive SysTems, 2016.

[PS04] R. Tinati, X. Wang, I. Brown, T. Tiropanis, and W. Hall, "A Streaming Real-Time Web Observatory Architecture for Monitoring the Health of Social Machines," in Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion, 2015, pp. 1149–1154.

[PS05] M. Luczak-Roesch, R. Tinati, S. Aljaloud, W. Hall, and N. Shadbolt, "A universal socio-technical computing machine," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, vol. 9671, pp. 559–562.

[PS06] M. Rovatsos, D. Diochnos, and M. Craciun, "Agent Protocols for Social Computation," in International Conference on Advances in Social Computing and Multiagent Systems - 6th International Workshop on Collaborative Agents Research and Development, and Second International Workshop on Multiagent Foundations of Social Computing, 2015.

[PS07] R. Tinati, T. Tiropanis, and L. Carr, "An approach for using Wikipedia to measure the flow of trends across countries," Proc. 22nd Int. Conf. World Wide Web companion, no. May, pp. 1373–1378, 2013.

[PS08] V. Buregio, Z. Maamar, and S. Meira, "An architecture and guiding framework for the social enterprise," IEEE Internet Computing, vol. 19, no. 1. pp. 64–68, 2015.

[PS09] J. Lin, L. C. Lin, and S. Huang, "An Architecture for Social Customer Support System," in Proceedings of 20th Pacific Asia Conference on Information Systems, 2016.

[PS10] S. Tarte, P. Willcox, H. Glaser, and D. De Roure, "Archetypal Narratives in Social Machines: Approaching Sociality Through Prosopography," in Proceedings of the ACM Web Science Conference, 2015, p. 24:1--24:10.

[PS11] C. J. Hooper, M. Nind, S. Parsons, A. Power, and A. Collis, "Building a Social Machine: Co-designing a TimeBank for Inclusive Research," in Proceedings of the ACM Web Science Conference on ZZZ - WebSci '15, 2015, pp. 1–9.

[PS12] Z. Maamar, V. Burégio, and M. Sellami, "Collaborative Enterprise Applications based on Business and Social Artifacts," in 24th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2015.

[PS13] M. Strohmaier and C. Wagner, "Computational Social Science for the World Wide Web," IEEE Intell. Syst., vol. 29, no. 5, pp. 84–88, 2014.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

[PS14] A. Guy and E. Klein, "Constructed Identity and Social Machines: A Case Study in Creative Media Production," in International World Wide Web Conference Committee, 2014.

[PS15] M. B. Evans, K. O'Hara, T. Tiropanis, and C. Webber, "Crime Applications and Social Machines: Crowdsourcing Sensitive Data," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 891–896, 2013.

[PS16] T. Ali, M. Gheith, and E. S. Nasr, "CrowdCE: A Collaboration Model for Crowdsourcing Software with Computing Elements," Int. J. Recent Innov. Trends Comput. Commun., vol. 4, no. 2, 2016.

[PS17] A. Bozzon, M. Brambilla, S. Ceri, A. Mauri, and R. Volonterio, "Designing Complex Crowdsourcing Applications Covering Multiple Platforms and Tasks," J. Web Eng., vol. 0, no. 0, pp. 0–0, 2015.

[PS18] M. J. Day, L. Carr, and S. Halford, "Developing the 'Pro-human' Web," in ACM Web Science Conference, 2015.

[PS19] P. Pareti, "Distributed linked data as a framework for human-machine collaboration," in International Semantic Web Conference, 2016.

[PS20] S. A. Applin and M. D. Fischer, "Exploring Cooperation with Social Machines," in International World Wide Web Conference Committee, 2016.

[PS21] G. Demartini, "Hybrid human-machine information systems: Challenges and opportunities," Comput. Networks, vol. 90, pp. 5–13, 2015.

[PS22] T. Ali, M. Gheith, E. S. Nasr, and P. Elbaz, "Hybrid-computing elements: A multi-sourcing model for managing crowdsourcing software," in Proceedings - 11th IEEE International Conference on Global Software Engineering Companion Proceedings, ICGSEW 2016, 2016, pp. 1–6.

[PS23] A. Ciortea, A. Zimmermann, O. Boissier, and A. M. Florea, "Hypermedia-driven Socio-technical Networks for Goal-driven Discovery in the Web of Things," in Proceedings of the Seventh International Workshop on the Web of Things - WoT '16, 2016, pp. 25–30.

[PS24] P. Singh and N. Shadbolt, "Linked Data in Crowdsourcing Purposive Social Network," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 913–918, 2013.

[PS25] D. Murray-Rust and D. Robertson, "LSCitter: Building Social Machines by Augmenting Existing Social Networks with Interaction Models," in International World Wide Web Conference Committee (IW3C2), 2014, pp. 875–880.

[PS26] P. R. Smart, "Machine Intelligence and the SocialWeb: How to Get a Cognitive Upgrade," in International Conference on Advanced Cognitive Technologies and Applications, 2017.

[PS27] V. A. Burégio, S. L. Meira, N. S. Rosa, and V. C. Garcia, "Moving towards 'relationship-aware' applications and services: A social machine-oriented approach," in Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC, 2013, pp. 43–52.

[PS28] M. Rovatsos, "Multiagent systems for social computation," in 13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014, 2014, vol. 2, pp. 1165–1168.

[PS29] D. De Roure, C. Hooper, K. Page, S. Tarte, and P. Willcox, "Observing Social Machines Part 2: How to Observe?," in Proceedings of the ACM Web Science Conference on ZZZ - WebSci '15, 2015.

[PS30] D. Murray-rust, S. Tarte, M. Hartswood, and O. Green, "On Wayfaring in Social Machines," in International World Wide Web Conference Committee, 2015.

[PS31] V. Buregio, L. M. Nascimento, N. S. Rosa, and S. R. de L. Meira, "Personal APIs as an Enabler for Designing and Implementing People as Social Machines," in Proceedings of the 23rd International Conference on World Wide Web, 2014, pp. 867–872.

[PS32] N. H. Thuan, P. Antunes, and D. Johnstone, "Pilot experiments on a designed crowdsourcing decision tool," in Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2016, 2016, pp. 601–606.

[PS33] J. Zhao, R. Binns, M. Van Kleek, and N. Shadbolt, "Privacy languages : Are we there yet to enable user controls?," WWW '16 Companion Proc. 25th Int. Conf. Companion World Wide Web, pp. 799–806, 2016.

[PS34] B. Dalton, "Pseudonymity in Social Machines," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 885–890, 2013.

[PS35] J. Vass and J. Munson, "Revisiting the Three Rs of Social Machines: Reflexivity, Recognition and Responsivity," in Proceedings of the 24th International Conference on World Wide Web, 2015, pp. 1161–1166.

[PS36] T. Ali, E. S. Nasr, and M. Gheith, "Self-management of Distributed Computing Using Hybrid-Computing Elements," in Proceedings of the 2nd Africa and Middle East Conference on Software Engineering - AMECSE '16, 2016, pp. 15–20.

[PS37] M. Rovatsos, D. I. Diochnos, Z. Wen, S. Ceppi, and P. Andreadis, "SmartOrch: An adaptive orchestration system for human-machine collectives," in Proceedings of the ACM Symposium on Applied Computing, 2017, vol. Part F1280.

[PS38] C. Hooper, B. Bailey, H. Glaser, and J. Hendler, "Social Machines in Practice: Solutions, Stakeholders and Scopes," in Proceedings of the 8th ACM Conference on Web Science, 2016, pp. 156–160.

[PS39] S. O. Palermos, "Social machines: a philosophical engineering," Phenomenol. Cogn. Sci., vol. 16, no. 5, pp. 953–978, 2017.

[PS40] V. Buregio, S. Meira, and N. Rosa, "Social Machines: A Unified Paradigm to Describe Social Web-Oriented Systems," in 22nd International World Wide Web Conference (WWW 2013 Companion), 2013.

[PS41] M. Van Kleek, D. A. Smith, D. Murray-Rust, A. Guy, K. O'Hara, L. Dragan, and N. R. Shadbolt, "Social personal data stores: The nuclei of decentralised social machines," WWW 2015 Companion - Proc. 24th Int. Conf. World Wide Web, pp. 1155–1160, 2015.

[PS42] H. Halpin and A. Capocci, "The Berners-Lee hypothesis: power laws and group structure in flickr," in Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion, 2014, pp. 885–890.

[PS43] M. Van Kleek, D. A. Smith, W. Hall, and N. R. Shadbolt, "'The Crowd Keeps Me in Shape': Social Psychology and the Present and Future of Health Social Machines," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 927–932, 2013.

[PS44] R. Tinati, L. Carr, S. Halford, and C. Pope, "The HTP Model: Understanding the Development of Social Machines," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 921–926, 2013.

[PS45] A. Merchant, T. Jha, and N. Singh, "The Use of Trust in Social Machines," in International World Wide Web Conference Committee2, 2016.

[PS46] N. R. Shadbolt, D. A. Smith, E. Simperl, M. Van Kleek, Y. Yang, and W. Hall, "Towards a classification framework for social machines," 22nd International World Wide Web Conference (WWW 2013 Companion). Rio de Janeiro, Brazil, pp. 905–912, 2013.

[PS47] S. Poddar, S. K. Ernala, N. Singh, and R. A. Samvara, "Towards a Ubiquitous Model of an Individual in Social Machines," in International World Wide Web Conference Committee, 2016.

[PS48] V. Burégio, K. Brito, N. Rosa, M. Neto, V. Garcia, and S. Meira, "Towards Government as a Social Machine," in Proceedings of the 24th International Conference on World Wide Web, 2015, pp. 1131–1136.

[PS49] J. Corneli, U. Martin, D. Murray-Rust, and A. Pease, "Towards mathematical AI via a model of the content and process of mathematical question and answer dialogues," in International Conference on Intelligent Computer Mathematics, 2017, vol. 10383 LNAI, pp. 132–146.

[PS50] M. Tsvetkova, T. Yasseri, E. T. Meyer, J. B. Pickering, V. Engen, P. Walland, M. Lüders, A. Følstad, and G. Bravos, "Understanding Human-Machine Networks: A Cross-Disciplinary Survey," ACM Comput. Surv., vol. 50, no. 1, pp. 1–35, 2015.

[PS51] D. Ahlers, P. Driscoll, E. Löfström, J. Krogstie, and A. Wyckmans, "Understanding Smart Cities as Social Machines," Proc. 25th Int. Conf. Companion World Wide Web, pp. 759–764, 2016.

[PS52] I. Alvertis, D. Papaspyros, S. Koussouris, S. Mouzakitis, and D. Askounis, "Using crowdsourced and anonymized Personas in the requirements elicitation and software development phases of software engineering," in International Conference on Availability, Reliability and Security, 2016.

[PS53] C. A. Halcrow, L. Carr, and S. Halford, "Using the SPENCE Model of Online/Offline Community to Analyse Sociality of Social Machines," in Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion, 2016, pp. 769–774.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

[PS54] K. O'Hara, "Web Science: Understanding the Emergence of Macro-Level Features on the World Wide Web," Found. Trends® Web Sci., vol. 4, no. 2–3, pp. 103–267, 2013.

[PS55] M. Luczak-Rösch, R. Tinati, and N. Shadbolt, "When Resources Collide : Towards a Theory of Coincidence in Information Spaces," in International World Wide Web Conference Committee, 2015.

[PS56] S. Tarte, D. De Roure, and P. Willcox, "Working out the plot: the role of Stories in Social Machines," in International World Wide Web Conference Committee, 2014.

REFERENCES

[1] T. Berners-Lee and M. Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. New York: Harper Collins, 1999.

[2] W. Roush, "Social Machines - Computing means connecting," in *MIT Technology Review, August 2005*, 2005.

[3] S. R. L. Meira *et al.*, "The Emerging Web of Social Machines," in *Computer Software and Applications Conference (COMPSAC)*, 2011, pp. 26–27.

[4] N. R. Shadbolt, D. A. Smith, E. Simperl, M. Van Kleek, Y. Yang, and W. Hall, "Towards a classification framework for social machines," *22nd International World Wide Web Conference (WWW 2013 Companion)*. Rio de Janeiro, Brazil, pp. 905–912, 2013.

[5] V. Buregio, S. Meira, and N. Rosa, "Social Machines: A Unified Paradigm to Describe Social Web-Oriented Systems," in *22nd International World Wide Web Conference (WWW 2013 Companion)*, 2013.

[6] J. Hendler and A. M. Mulvehill, *Social Machines: The Coming Collision of Artificial Intelligence, Social Networking, and Humanity*, 1st ed. Berkeley, CA: Apress, Berkeley, CA, 2016.

[7] N. Shadbolt, K. O'Hara, D. De Roure, and W. Hall, *The Theory and Practice of Social Machines*, 1st ed. Springer, 2019.

[8] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," *Engineering*, vol. 2, p. 1051, 2007.

[9] Centre for Reviews and Dissemination, "Systematic reviews: CRD's guidance for undertaking reviews in health care," *University of York*, 2009. .

[10] M. Rovatsos, "Multiagent systems for social computation," in *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014*, 2014, vol. 2, pp. 1165–1168.

[11] T. Ali, E. S. Nasr, and M. Gheith, "Self-management of Distributed Computing Using Hybrid-Computing Elements," in *Proceedings of the 2nd Africa and Middle East Conference on Software Engineering - AMECSE '16*, 2016, pp. 15–20.

[12] A. Bozzon, M. Brambilla, S. Ceri, A. Mauri, and R. Volonterio, "Designing Complex Crowdsourcing Applications Covering Multiple Platforms and Tasks," *J. Web Eng.*, vol. 0, no. 0, pp. 0–0, 2015.

[13] N. H. Thuan, P. Antunes, and D. Johnstone, "Pilot experiments on a designed crowdsourcing decision tool," in *Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2016*, 2016, pp. 601–606.

[14] J. M. Corbin and A. Strauss, "Grounded theory research: Procedures, canons, and evaluative criteria," *Qual. Sociol.*, vol. 13, no. 1, 1990.

[15] J. Vass and J. Munson, "Revisiting the Three Rs of Social Machines: Reflexivity, Recognition and Responsivity," in *Proceedings of the 24th International Conference on World Wide Web*, 2015, pp. 1161–1166.

[16] C. Hooper, B. Bailey, H. Glaser, and J. Hendler, "Social Machines in Practice: Solutions, Stakeholders and Scopes," in *Proceedings of the 8th ACM Conference on Web Science*, 2016, pp. 156–160.

[17] C. A. Halcrow, L. Carr, and S. Halford, "Using the SPENCE Model of Online/Offline Community to Analyse Sociality of Social Machines," in *Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion*, 2016, pp. 769–774.

[18] C. J. Hooper, M. Nind, S. Parsons, A. Power, and A. Collis, "Building a Social Machine: Co-designing a TimeBank for Inclusive Research," in *Proceedings of the ACM Web Science Conference on ZZZ - WebSci '15*, 2015, pp. 1–9.

[19] M. Tsvetkova *et al.*, "Understanding Human-Machine Networks: A Cross-Disciplinary Survey," *ACM Comput. Surv.*, vol. 50, no. 1, pp. 1–35, 2015.

[20] E. Kaganer, E. Carmel, R. Hirscheim, and T. Olsen, "Managing the human cloud," *MIT Sloan Manag. Rev.*, 2013.

[21] P. R. Smart, "Machine Intelligence and the SocialWeb: How to Get a Cognitive Upgrade," in *International Conference on Advanced Cognitive Technologies and Applications*, 2017.

[22] C. J. Lintott and J. Reed, "Human Computation in citizen science," in *Handbook of Human Computation*, 1st ed., P. Michelucci, Ed. New York: Springer-Verlag New York, 2013.

[23] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The social internet of things (SIoT) - When social networks meet the internet of things: Concept, architecture and network characterization," *Comput. Networks*, vol. 56, pp. 3594–3608, 2012.

[24] A. Li, X. Ye, and H. Ning, "Thing relation modeling in the internet of things," *IEEE Access*, 2017.

[25] L. Atzori, A. Iera, and G. Morabito, "From 'smart objects' to 'social objects': The next evolutionary step of the internet of things," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 97–105, Jan. 2014.

Kellyton dos Santos Brito (M'15) received the B.S. degree in computer science from Universidade Federal da Bahia, Salvador, Brazil, in 2000 and the M.S. degree in computer science from Universidade Federal de Pernambuco, Recife, Brazil, in 2007. He is currently pursuing the Ph.D. degree in computer science at Universidade Federal de Pernambuco.

He is Professor at Universidade Federal Rural de Pernambuco since 2011. Besides, he is a researcher at INES (National Institute of Science and Technology in Software Engineering) and he worked as software engineer and carried out industrial projects in many companies, such as CESAR (Center for Advanced Studies and Systems of Recife), Brazil, from 2005-2008, and Medicware Systems, Brazil, from 2003-2005. His current research interest includes software architecture, social machines, open government data, machine learning and elections forecasting.

Alysson Alves de Lima is a Ph.D student in computer science at the Universidade Federal de Pernambuco, Recife, Brazil, M.S. in informatics from the Universidade Federal da Paraíba, João Pessoa, Brazil (2016), MBA in software engineering from Faculdade Estácio de Sá, João Pessoa, Brazil (2013), and Graduated in computer technology and systems from the Faculdade de Tecnologia da Paraíba, PB, Brazil (2011).

He has solid knowledge in software engineering, software architecture, cloud computing and algorithm complexity. His main topics of interest are software engineering, cloud computing, fog computing, internet of things, artificial intelligence and social machines.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCSS.2019.2961269

Sérgio Endrigo Ferreira was born in Recife, Pernambuco, Brazil in 1982. He received the B.S in system analysis from Universidade Salgado de Oliveira, Recife, Brazil (2007) and M.Sc degree in software engineering from the CESAR (Center for Advanced Studies and Systems of Recife), in 2011. He is currently pursuing the Ph.D. degree in computer science from Universidade Federal de Pernambuco, Recife, Brazil.

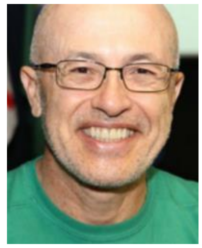
Since 2005, he is a Software Engineer at Auttar Hut Processamento de Dados. His research interests include software engineering, social machines, and microservices.

Vanilson de Arruda Burégio is a Professor in the Computing Department of Universidade Federal Rural de Pernambuco (UFRPE) with Ph.D. (2014), M.Sc. (2006) and B.Sc. (2003) in computer science from Informatics Centre of Universidade Federal de Pernambuco (UFPE), Recife, Brazil.

He is a software engineering expert with more than 19 years of experience in the development of several information systems in private and government institutions as well as national and international companies. His research interests include software architecture of data-intensive applications and initiatives that explore how to use the social machine paradigm in the development of emerging collaborative platforms and social enterprises.

Vinicius Cardoso Garcia (M'08) was born in Rio de Janeiro, Brazil. He has a B.S. degree in computer science from Universidade Salvador, Salvador, Brazil (2000), M.S. in Computer Science from Universidade Federal de São Carlos, São Carlos, Brazil (2005) and a Ph.D. from Universidade Federal de Pernambuco, Recife, Brazil (2010).

He is an Associate Professor at Informatics Centre of Universidade Federal de Pernambuco since 2010. Besides, he is an associate researcher at INES (National Institute of Science and Technology in Software Engineering) and he worked as a software and systems engineer and software productive consultant at CESAR (Center for Advanced Studies and Systems of Recife) from 2005 to 2010, where he carried out several industrial projects focused on various aspects of software engineering. Since 2010 Dr. Vinicius Garcia has been working as a researcher in agreements and partnerships with startups, small and medium-sized companies of Porto Digital (an initiative in Recife, Brazil to foster technological innovation in the northern region of Brazil) in the areas of Cloud Computing, Software-Defined Storages, Site Reliability Engineering and Continuous Software Engineering with a primary focus on improvements for the Brazilian industry and state of the practice improvement of these areas.



Silvio Romero de Lemos Meira received his B.S degree in electronic engineering from Instituto Tecnológico de Aeronáutica, Brazil (1977), M.S degree in computer science from Universidade Federal de Pernambuco (1981) and Ph.D. degree in computer science from University of Kent at Canterbury, England (1985).

He is Emeritus Professor at Informatics Centre of Universidade Federal de Pernambuco, Brazil, Senior Researcher at INES (National Institute of Science and Technology in Software Engineering) and Senior Researcher at Instituto SENAI de Inovação para Tecnologias da Informação e Comunicação, Brazil, and was also Associate Professor at Fundação Getulio Vargas, Brazil. In 2012-2013, he was Fellow at Berkman Center from Harvard University, and was also Faculty Associate until 2015. His research interests include information life cycle, social machines, information systems, social networks, software engineering, creativity, innovation and entrepreneurship.

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.