Designing Social Machines for Tackling Online Disinformation

Antonia Wild
University of St. Gallen
St. Gallen, Switzerland
antonia.wild@student.unisg.ch

Andrei Ciortea
University of St. Gallen
St. Gallen, Switzerland
Inria, Université Côte d'Azur, CNRS
Sophia Antipolis, France
andrei.ciortea@unisg.ch

Simon Mayer
University of St. Gallen
and ETH Zürich
St. Gallen, Switzerland
simon.mayer@unisg.ch

ABSTRACT

Traditional news outlets as carriers and distributors of information have been challenged by online social networks with regards to their gate-keeping function. We believe that only a combined effort of people and machines will be able to curb so-called “fake news” at scale in a decentralized Web. In this position paper, we propose an approach to design social machines that coordinate human- and machine-driven credibility assessment of information on a decentralized Web. To this end, we defined a fact-checking process that draws upon ongoing efforts for tackling disinformation on the Web, and we formalized this process as a multi-agent organisation for curating W3C Web Annotations. We present the current state of our prototypical implementation in the form of a browser plugin that builds on the Hypothesis annotation platform and the JaCaMo multi-agent platform. Our social machines would span across the Web to enable collaboration in form of public discourse, thereby increasing the transparency and accountability of information.

KEYWORDS

Disinformation, Social Machines, Web Architecture, Linked Data, Multi-Agent Systems, Decentralization

ACM Reference Format:

Antonia Wild, Andrei Ciortea, and Simon Mayer. 2020. Designing Social Machines for Tackling Online Disinformation. In *Companion Proceedings of the Web Conference 2020 (WWW '20 Companion)*, April 20–24, 2020, Taipei, Taiwan. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3366424.3385770>

1 INTRODUCTION

Online social networks have challenged traditional news outlets in their gate-keeping function [12]. This allows for a more diverse, immediate, and unfiltered access to information, but at the same time leaves users with the difficult task of assessing the credibility of information, easing the spread of disinformation.¹ Already in his talk at WWW'94², Sir Tim Berners-Lee was raising the challenge of ensuring the quality of information in a system as open as the

Web – and envisioned the use of *annotations* as a suitable mechanism.³ Such *Web Annotations*, now a W3C Recommendation [16], construct a metadata layer on top of existing resources and without requiring their modification. As such, they can be regarded as a connective fabric that allows users to address statements *within* Web pages⁴ and to bind these statements to metadata – in the domain of online disinformation, for instance, regarding their provenance, truth assessments by others, or opposing views from other sources.

Academics and practitioners have tested human- and machine-based approaches that make use of annotations to tackle online disinformation (see Sec. 2.1.2). However, to-date no solution seems sufficient to reliably and efficiently address the issue. We believe that only a combined effort of both humans and machines will be able to curb so-called “fake news” at Web scale. We refer to such collaborative efforts of humans and machines as *social machines*, following the definition of [19].⁵ We pursue the following research question: *how to design social machines that coordinate human and autonomous agents to perform a transparent and guided credibility analysis of information on the open, decentralized Web – at Web scale and in a timely manner?*

In this position paper, we take a first step towards addressing this question. We defined a fact-checking process that draws upon existing efforts of the *W3C Credible Web Community Group*⁶ and *EUFACHECK*⁷. Then, building upon results from research on *autonomous agents* and *multi-agent systems (MAS)* [23], we conceptualized this fact-checking process as a *multi-agent organization* for curating W3C Web Annotations. To validate our approach, we are currently in the process of implementing a prototypical social machine using *Hypothesis*⁸, an annotation platform that conforms

³<https://videos.cern.ch/record/2671957>, accessed 12.01.2020

⁴<https://bit.ly/2U3IKgU>, accessed 12.01.2020

⁵This term was coined by Berners-Lee, who defines a social machine as a “process in which people do the creative work and the machine does the administration” [2]. In our interpretation, both human and machine agents are regarded as *first-class citizens* of social machines. This view is thus closer to the definition given by Smart and Shadbolt in [19], where social machines are defined as “Web-based socio-technical systems in which the human and technological elements play the role of participant machinery with respect to the mechanistic realization of system-level processes”. This emancipated notion of participatory machine-agents – called hereafter *autonomous agents* – is central to our approach.

⁶The *W3C CredWeb CG* is an interdisciplinary group committed towards providing technologies that help end-users assess the credibility of information on the Web, without limiting their choices (<https://www.w3.org/community/credibility/>).

⁷*EUFACHECK* (<https://eufactcheck.eu/>) is a fact-checking project of the *European Journalism Training Association* that is directed towards European political news coverage, adhering to the code of principles of the *International Fact-Checking Network* (<https://www.poynter.org/ifcn-fact-checkers-code-of-principles/>).

⁸<https://github.com/hypothesis/client/>

¹See [21] for spread and [3] for consumption of true and false news online.

²The First International Conference on the World-Wide Web, CERN, 25 - 27 May, 1994.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '20 Companion, April 20–24, 2020, Taipei, Taiwan

© 2020 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-7024-0/20/04.

<https://doi.org/10.1145/3366424.3385770>

to the W3C Web Annotation Recommendation, and *JaCaMo*, a platform for the development of MAS⁹.

The rest of this paper is structured as follows. Sec. 2 discusses background and related work on tackling online disinformation, social machines, and MAS. Sec. 3 presents our approach, including detailed reasoning for its encompassing features. We present the current state of our prototypical implementation in Sec. 4 and discuss its limitations.

2 BACKGROUND AND RELATED WORK

We give an overview of existing fact-checking approaches in Sec. 2.1. We argue that human- and machine-driven fact-checking efforts *by themselves* are not sufficient to cope with disinformation at scale and with accuracy. Consequently, in Sec. 2.2 we present the concepts of *social machines* and MAS that allow to design and implement *hybrid* fact-checking campaigns.

2.1 Combating Disinformation Online

To help users assess the credibility of information at the point of consumption, manifold human- and machine-driven approaches have been developed and tested in both academia and practice. Efforts to analyze the content and its metadata can be mapped according to the dimensions of *automation* – i.e., to which degree software drives the credibility assessment, and *abstraction* – i.e., to which degree the analysis abstracts from the content itself and considers external information provided in an open-world scenario.

2.1.1 Human-driven Fact-checking. At a low degree of automation, organizations and groups of experts (notably journalists) have formed manual fact-checking initiatives.¹⁰ Manual fact-checking enables a high degree of abstraction and to a certain extent ensures accountability (through disclosure of the fact-checking organization) and transparency (through disclosure of the fact-checking process), but it cannot scale up to the size of the Web.¹¹ As a consequence, the majority of fact-checking sites has limited coverage regarding topics, languages, and geography. In addition, ideally claims would be verified *before* going viral. This would require both effective identification and prioritization of claims with a high likelihood of virality, as well as a very fast response time from manual fact-checkers.

2.1.2 Automated Fact-checking. Fact-checking approaches with a higher automation degree mostly rely on *natural language processing* (NLP) to analyze information [1]. They then evaluate the truthfulness by using information that is contained *within* the content itself and its immediate metadata (closed-world approach) or by incorporating outside knowledge [20] (open-world approach; high level of abstraction).¹² Despite the progress, automated fact-checking still lacks accuracy, and the underlying models are often trained with topic-, language- or culture-specific content [8], which undermines the adaptability of these systems. Additionally, current

solutions are not yet able to grasp the intent of the text, e.g., irony or metaphors [22], and are vulnerable to bias [18]. Lastly, current automated approaches do not provide an end-to-end solution for reliably determining the truthfulness of an article as they often focus on one aspect of credibility (e.g., the language style or the publisher’s credibility) and thus demand for integration within a larger system that assesses credibility along all its dimensions.

In summary, the *manual* checking of online information cannot be performed in a timely manner [6] (but is mostly accountable and transparent), while *automated* approaches lack the facilities to reliably detect disinformation (but can potentially work at scale). Systems encompassing collaborative efforts of human and autonomous agents might be able to unify the desirable properties of each of these approaches¹³.

2.2 Social Machines on the Web

Social machines could enable such large-scale collaborations among humans and autonomous agents on the Web. Over the past couple of years, the *World Wide Web Consortium* (W3C) – and in particular the *W3C Web Annotation Working Group* and the *W3C Social Web Working Group*¹⁴ – finalized standards that weave social features into the very fabric of the Web. Particularly, the W3C Web Annotation Protocol [16] builds upon the *Linked Data Platform*¹⁵, which is part of *SOLID*¹⁶, thus strengthening the decentralized features of the Web. These standards unlock new opportunities for the development and deployment of social machines on the Web.

Some researchers have identified multi-agent systems (MAS) as a suitable means to conceptualize and engineer social machines [5, 13, 14]. In distributed artificial intelligence, MAS are systems conceptualized in terms of agents situated and interacting in a shared environment, where an *autonomous agent* is commonly defined as “a computer system, situated in some environment, that is capable of flexible autonomous action in order to meet its design objectives” [11, 23].¹⁷ Drawing upon MAS research, in [13] the authors propose to design social machines for crowd-sourcing software development by complementing the *social compute unit* – a model for *ad-hoc* human worker teams – with interaction protocols expressed in *lightweight social calculus*. Interaction protocols are also central to [5]: the authors define *social protocols* as behavioral standards based on the expectations of (human and non-human) participants towards interactions, which are then reflected in a computational model and can be enacted in a decentralized manner. *Decentrality* was also identified in [14] as one of the three key challenges for making social machines easily implementable on the Web – such that many people, rather than only a handful of powerful platforms, can benefit from them. This early research presents encouraging results, yet the use of MAS to conceptualize social machines is still insufficiently investigated.

⁹Due to space constraints, we refer interested readers to [4] for more details on the JaCaMo meta-model and multi-agent platform.

¹⁰These include Pulitzer-price winning *PolitiFact*, *Snopes*, and *FactCheckEU*.

¹¹For instance, the *Washington Post* alone publishes on average 500 stories and videos per day (<https://www.theatlantic.com/technology/archive/2016/05/how-many-stories-do-newspapers-publish-per-day/483845/>, accessed 27.01.2020)

¹²See [6] for a detailed overview of efforts in the field of content validation.

¹³First ideas how to create such systems have been envisioned in [15], which presents an ontology for a hybrid fact check infrastructure that could be embedded into the Web, based on the W3C Web Annotation Recommendation.

¹⁴<https://www.w3.org/TR/social-web-protocols/>

¹⁵<http://www.w3.org/TR/2015/REC-ldp-20150226/>

¹⁶<http://solid.mit.edu>

¹⁷*Autonomy* is central to this definition and refers to the agent’s ability to operate on its own, without the need of direct intervention from humans or other agents.

Figure 1: Disinformation Tackler: six processes (center) that contain several questions each (right) guide users through the inspection of online information along four credibility indicators (left).

3 CONTRIBUTION

We introduce an approach that uses multi-agent organizations to conceptualize social machines for tackling online disinformation. Our objective is to tackle disinformation in a decentralized Web *at scale* and in a *transparent* manner. At the same time, we aim to *avoid undue censorship* and help agents make informed decisions, allowing the system to *adapt to changes* (in the underlying user community and information). Finally, we maintain *openness* in terms of the information being examined and the participating agents to benefit from new (technological) developments.

To this end, we designed a multi-faceted fact-checking process that considers several credibility indicators (see Sec. 3.1). Based thereon, we formalized this process as a multi-agent organization that coordinates the individual fact-checking efforts of humans and autonomous agents (see Sec. 3.2). To separate domain knowledge from operational processes – and to support the re-usability of our concepts and approach – we defined the *Disinformation Tackler Ontology* (see Sec. 3.3), which extends the *W3C Web Annotation Vocabulary* [17].

3.1 Disinformation Tackling

In our system, the tackling of disinformation consists of 46 distinct questions that are assigned to six processes of varying length (see Fig. 1). Each process focuses on a different dimension of evaluating the credibility of an online information resource. These processes guide users through content evaluation along the four credibility indicators defined by the *W3C Credible Web Community Group (CredWeb CG)*: *inspection* – assessing the content itself, *corroboration* – identifying claims and checking for verification by outer sources, *reputation* – assessing the credibility of the information provider, and lastly *transparency* – evaluating the self-declaration of publisher and author [25]. To each dimension we assigned corresponding elements of the disinformation tackling process of EU-FACTCHECK. While EUFACTCHECK served as a starting point for our fact-checking process, we abstracted from its domain-specificity and reliance on experts by aligning with the efforts of the *W3C CredWeb CG*. In our process, users are required to provide evidence for their answers at several instances, e.g., by linking to opposing sources.

Figure 2: Structural and selected functional specification of our MOISE organization

3.2 Organizational Model

As the dissemination and forms of online information evolve, so do the mechanisms of disinformation producers and the tactics of malevolent system users. To cope with these dynamics, our social machine is expected to change, thus needing to be *adaptive*. Further, the system should be *open* with regards to participant agents and the integration of technical solutions. Lastly, we require some sort of normative power within our system to combat malevolent influences and to foster self-regulation in the system. Thus, we are in need for a formal representation of our organization geared towards tackling disinformation. Doing so, we follow Hendler and Berners-Lee, who demand for technology that “allows user communities to construct, share and adapt social machines so that successful models can evolve through trial, use and refinement” [9, 14].

3.2.1 Multi-agent Organizations with MOISE. Organizations in MAS are characterized by a purpose-directed structure of emerging or predefined agent interaction, which is typically designed by organization engineers [4]. To cope with the autonomy of agents (see Sec. 2.2), a normative dimension is typically used to impose constraints on the agents’ behavior, while at the same time allowing them the freedom to reason about the organization and the goals they (ought to) fulfill. We formalize our fact-checking process as a multi-agent organization using the *MOISE model* [10], which defines a multi-agent organization on three dimensions: a *functional*, *structural*, and *normative* dimension. The *functional dimension* defines goals and goal decomposition schemes which coordinate the achievement of goals (see Fig. 2). The *structural dimension* defines roles and groups within an organization. The *normative dimension* assigns goals to roles via norms meant to guide the agents’ behavior [10]. The concept of roles allows for decoupling the fact-checking process from individual persons and autonomous agents, who can thus freely enter or leave these (open) organizations. Furthermore, this allows new goal coordination schemes to be deployed at run time, which enables the dynamic reconfiguration of the fact-checking process. Such reconfiguration could be deployed,

Figure 3: Main concepts of Disinformation Tackler Ontology

for instance, with regards to changes in the underlying content (e.g., the language the information is displayed in).

3.2.2 A MOISE Organization for Tackling Disinformation. Fig. 2 shows the MOISE organization used to formalize our fact-checking process in Sec. 3.1. MOISE organizations are defined declaratively, and we represent the specification of our organization in RDF using the MOISE ontology defined in [24]. Our organization can then be instantiated on a per-needs basis to meet the needs of tackling disinformation at scale in a decentralized Web. Both people and autonomous agents can reason on their specifications (e.g., to decide to join a fact-checking organization or which goals to achieve). The *functional dimension* in Fig. 2 is composed of six schemes which correspond to the six processes of our fact-checking process, where all sub-goals can be achieved in parallel. Our *structural dimension* highlights the different groups that agents within the system can join. Each group in our organization can contain one of three roles (see Fig. 2), where the active user role encompasses our (human or autonomous agent) fact checkers. With regards to the *normative dimension*, the active user role is permitted to perform fact-checking tasks, such as Content evaluation & article corroboration.

3.3 Vocabulary

To enable unification and re-usability of our domain concepts, we conceptualize the *Disinformation Tackler Ontology*. Fig. 3 depicts the core concepts of our ontology.

3.3.1 Weaving W3C Web Annotations into the Web. We rely on *annotations* as our method of choice for Web-based credibility assessment. Since these annotations are independent of the information being annotated and have their own resource identifiers and representations, an entity that *owns* an information resource does not own the dialogue about it. This is crucial for the independent assessment of resources. Standardization efforts of the W3C Web Annotation Working Group also allow annotations to be weaved into the Web in the future.

3.3.2 The Disinformation Tackler Ontology. We defined *Disinformation Tackler* (see Fig. 3), an OWL ontology that allows the uniform representation of annotations following from our fact-checking process. Our ontology is aligned with the W3C Web Annotation Vocabulary [17], which allows for inclusion of information about

annotations on a more technical level (e.g., talking about different information content types such as audio or video). We anchor our ontology via the concepts of *oa:Annotation*¹⁸ and *oa:Motivation*¹⁹. The concept of *oa:Annotation* is central to our ontology. Annotations are created by *dt:Agents* and motivated by different *oa:Motivations* – e.g., a false claim a user has spotted in the text. The concept of *Agents* subsumes both human and non-human autonomous agents, treating both as first-class citizens in our process. A *dt:Agent* can take on different *dt:Roles*, such as *dt:ActiveUser*, which can *dt:annotate* a *dt:InformationResource*, for instance through *dt:highlights* and *dt:tags*.

The Disinformation Tackler Ontology also contains extension points for other ontologies. The *dt:Agent* class is aligned with similar concepts from the FOAF²⁰ and PROV-O ontologies²¹ to include network effects and provenance information into our system in the future. This is of importance, for instance, when assessing whether a user is a bot, based on previous Web-activity or his/her network. Lastly, *dt:Role* serves as an anchor point for the MOISE ontology²².

4 PROTOTYPICAL IMPLEMENTATION

In the following, we present the current state of our prototypical implementation of our social fact-checking machine. Sec. 4.1 presents an architectural overview of our system. We discuss the limitations of our current implementation in Sec. 4.2. While the system has already been conceptualized, the implementation is still ongoing.

4.1 Architectural Overview

We have designed and implemented a browser plugin that allows agents to self-assess the credibility of online information.²³ The resulting annotations, in turn, can be discovered by other agents. Additionally, the implications of the annotations are clustered and aggregated to give agents a summarized overview of the evaluation results. A MOISE-based intermediary keeps track of the organization’s status – tracking level of goal fulfillment, committed agents, etc. – while our back end processes and stores annotations.

The browser plugin is based on *Hypothesis*²⁴ and follows closely the workflow process defined in Sec. 3.1: it allows users to choose which annotations they would like to create throughout the article evaluation process. The annotations are then stored on a back end that conforms to the *W3C Web Annotation Protocol* [16]. To this end, we use the cloud-based service provided by Hypothesis. For the display and aggregation of annotations, the browser plugin retrieves those annotations from the back end, sorts and aggregates them on the front end. The intermediary component sits between the front and back end, managing and keeping track of the organization (see Sec. 3.2) to help coordinate the fact-checking of articles.

¹⁸<https://www.w3.org/TR/annotation-vocab/#annotation>

¹⁹<https://www.w3.org/TR/annotation-vocab/#motivation>

²⁰<http://xmlns.com/foaf/spec/>

²¹<https://www.w3.org/TR/prov-o/>

²²<https://www.emse.fr/~zimmermann/Ontologies/moise.owl>

²³Although currently users can only integrate automated content evaluation by *Factmata* via linking, we could envision additional tasks such as the *Factmata* rating or *NewsGuard* source evaluation (<https://www.newsguardtech.com/>) to be carried out by automated solutions via APIs at a later stage – in fact, *Factmata* is starting to provide its API to clients (<https://factmata.com/api.html>).

²⁴<https://web.hypothes.is/>

We designed the intermediary component using JaCaMo [4] – a MAS platform that provides also the reference implementation for MOISE – to define and deploy multi-agent organizations. In our MOISE organization, people are proxied by simple software agents: agents enact roles in the organization on behalf of humans in order to signal the achievement of goals, and they forward to humans all their permissions and obligations within the organization (e.g., goals to be achieved). The intermediary component in our prototype is used to enrich a back end that conforms to the W3C Web Annotation Protocol, but the integration with both the front end and the back end is still ongoing.

4.2 Discussion and Limitations

The current progress with our prototypical implementation demonstrates the feasibility of our approach. Our prototype uses the Hypothesis back end, which conforms to the W3C Web Annotation Protocol. Although W3C Web Annotations have yet to gain support from browser vendors (which would help drive their large-scale adoption) they can already be used with browser plugins such as Hypothesis. However, Hypothesis uses *OAuth 2*, which requires the centralized authentication and authorization of HTTP requests.²⁵ To further support decentralization, we intend to extend this component to implement the SOLID open specifications.

A MOISE organization defines an explicit and objective standard of behavior, but then it is necessary to monitor the behavior of participants against the defined standard. In the MOISE framework we use in our implementation, the monitoring is centralized. However, because we use formal explicit representations of organizations in RDF, we could potentially reuse and instantiate organizations as needed to cope with tackling disinformation at scale. We currently implement our organization using an intermediary component that enriches the Hypothesis back end, but in the future we intend to further decouple these components to have (i) components concerned with running organizations, and (ii) components concerned with storing and managing W3C Web Annotations. Clients could then discover relevant components at run time via hypermedia. To further support decentralization, in future research we also intend to complement our work by investigating additional mechanisms that are *conceptually decentralized*, such as the use of *social protocols* [5].

5 CONCLUSION

To tackle disinformation at scale in a decentralized Web in a transparent, adaptable, and open manner, we designed and implemented a system that enables and coordinates fact-checking efforts of both humans and machines. To this end, we distilled ongoing efforts for tackling online disinformation into an overall fact-checking process that is split into six sub-processes. We used MOISE to conceptualize and formalize this process as a multi-agent organization, and defined the Disinformation Tackler Ontology to allow for unification and reuse of domain concepts.

The proposed contribution opens the door for several interesting problems. For instance, efforts could be directed towards investigating mechanisms to keep user communities *healthy*. Introducing malevolent user detection or discriminatory power at the system level (e.g., for verified fact-checkers) could be promising avenues to

explore. Further, more research is needed in the area of establishing the ground truth of information in crowd-based systems (cf. [7]).

Our system aims to put forward the vision of the Web as a collaborative knowledge base that is not mainly generating value as a dissemination mechanism for content, but rather through the discussions this content sparks. In such a Web, the gate-keeper function regarding content quality is distributed on the shoulders of both humans and machines, allowing the individual to knowledgeably navigate through the online information jungle.

REFERENCES

- [1] Alexandre Alaphilippe, Alexis Gizikis, et al. 2019. Automated tackling of disinformation. *Panel for the Future of Science and Technology* (2019).
- [2] Tim Berners-Lee and Mark Fischetti. 2001. *Weaving the Web: The original design and ultimate destiny of the WWW by its inventor*. DIANE Publishing Company.
- [3] Alessandro Bessi, Mauro Coletto, George Davidescu, et al. 2015. Science vs conspiracy: Collective narratives in the age of misinformation. *PloS one* 10, 2 (2015), e0118093.
- [4] Olivier Boissier, Rafael Bordini, et al. 2013. Multi-agent oriented programming with JaCaMo. *Science of Computer Programming* 78, 6 (2013), 747–761.
- [5] Amit Chopra and Munindar Singh. 2016. From social machines to social protocols: Software engineering foundations for sociotechnical systems. In *Proc. of the 25th Int. Conf. on WWW*. Int. WWW Conferences Steering Committee, 903–914.
- [6] Miriam Fernandez and Harith Alani. 2018. Online misinformation: Challenges and future directions. In *Comp. Proc. of TWC2018*. 595–602.
- [7] Simson L. Garfinkel. 2008. Wikipedia and the Meaning of Truth. *Technology Review* 111, 6 (2008), 84–86.
- [8] Lucas Graves. 2018. Understanding the promise and limits of automated fact-checking. *Reuters Institute, University of Oxford* 2 (2018).
- [9] Jim Hendler and Tim Berners-Lee. 2010. From the Semantic Web to social machines: A research challenge for AI on the WWW. *Artificial Intelligence* 174, 2 (2010), 156–161.
- [10] Jomi Hübner, Jaime Sichman, and Olivier Boissier. 2007. Developing organised MAS using the MOISE+ model: programming issues at the system and agent levels. *Int. J. of Agent-Oriented Software Engineering* 1, 3–4 (2007), 370–395.
- [11] Nicholas Jennings and Michael Wooldridge. 1998. Applications of Intelligent Agents. In *Agent Technology: Foundations, Applications, and Markets*, Nicholas Jennings and Michael Wooldridge (Eds.). Springer, Berlin, Heidelberg, 3–28.
- [12] Miriam Metzger, Andrew Flanagin, Keren Eyal, et al. 2003. Credibility for the 21st Century. *Annals of the Int. Communication Assoc.* 27, 1 (2003), 293–335.
- [13] Dave Murray-Rust, Ognjen Sekic, et al. 2014. A collaboration model for community-based software development with social machines. *10th IEEE Int. Conf. on Collaborative Computing* (2014), 84–93.
- [14] Petros Papapanagiotou, Alan Davoust, Dave Murray-Rust, et al. 2018. Social Machines for All. In *Proc. of the 17th Int. Conf. on Autonomous Agents and MAS (AAMAS '18)*. Int. Foundation for Autonomous Agents and MAS, 1208–1212.
- [15] Georg Rehm, Julian Moreno-Schneider, and Peter Bourgonje. 2018. Automatic and Manual Web Annotations in an Infrastructure to handle Fake News and other Online Media Phenomena. *Proc. of the 11th Int. Conf. on Language Resources and Evaluation*.
- [16] Robert Sanderson. 2017. *Web Annotation Protocol*. W3C Recommendation. World Wide Web Consortium (W3C). <https://www.w3.org/TR/annotation-protocol/>
- [17] Robert Sanderson, Paolo Ciccarese, and Benjamin Young. 2017. *Web Annotation Vocabulary*. W3C Recommendation. World Wide Web Consortium (W3C). <https://www.w3.org/TR/annotation-vocab/>
- [18] Jennifer Skeem and Christopher Lowenkamp. 2016. Risk, Race, and Recidivism: Predictive Bias And Disparate Impact. *Criminology* 54, 4 (2016), 680–712.
- [19] Paul Smart and Nigel Shadbolt. 2015. Social machines. In *Encyclopedia of Information Science and Technology, Third Edition*. IGI Global, 6855–6862.
- [20] James Thorne and Andreas Vlachos. 2018. Automated fact checking: Task formulations, methods and future directions. *arXiv preprint arXiv:1806.07687* (2018).
- [21] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.
- [22] Byron C Wallace. 2015. Computational irony: A survey and new perspectives. *Artificial Intelligence Review* 43, 4 (2015), 467–483.
- [23] Gerhard Weiss. 2000. *MAS: a modern approach to distributed artificial intelligence*. MIT press.
- [24] Alexandra-Madalina Zarafin, Antoine Zimmermann, and Olivier Boissier. 2012. Integrating Semantic Web Technologies and MAS: A Semantic Description of Multi-Agent Organizations. In *Proc. of the First Int. Conf. on Agreement Technologies (CEUR WS)*, Vol. 918. 296–297.
- [25] Amy Zhang, Martin Robbins, Ed Bice, et al. 2018. A Structured Response to Misinformation. In *Comp. of The Web Conf. 2018*. ACM Press.

²⁵See Section 1.8 (Interoperability) in RFC 6749: <https://tools.ietf.org/html/rfc6749>