



HAL
open science

Interoperable AI for Self-Organisation

Stefan Sarkadi, Fabien Gandon

► **To cite this version:**

Stefan Sarkadi, Fabien Gandon. Interoperable AI for Self-Organisation. SaSSO 2023 Workshop - Sustainable and Scalable Self-Organisation Workshop (ASCOS 2023), Sep 2023, Toronto, Canada. 10.1109/ACSOS-C58168.2023.00043 . hal-04384816

HAL Id: hal-04384816

<https://inria.hal.science/hal-04384816v1>

Submitted on 10 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Interoperable AI for Self-Organisation

Stefan Sarkadi

Dept. of Informatics
King's College London
London, UK
stefan.sarkadi@kcl.ac.uk

Fabien Gandon

Inria, University Cote d'Azur, CNRS, I3S
Sophia-Antipolis, France
fabien.gandon@inria.fr

Abstract—Knowledge or information sharing is one of the foundational principles shared by the Web and any well functioning decentralised multi-agent system (MAS). Another is interoperability. Engineers of MAS follow design strategies for ensuring interoperability that range from using a common ontology to using translation mechanisms. But what kind of sustainable communication strategy would AI agents themselves use if given the chance to self-organise? In this short paper we summarise our previous findings from [6], where we used an evolutionary agent-based modelling approach to answer this question.

Index Terms—Interoperability, Communication, Self-organising Multi-Agent Systems, Artificial Languages

I. SUSTAINABLE KNOWLEDGE-SHARING ON THE WEB

Decentralised systems are built on the principle of agent-agent communication and interoperation [2]. An important issue is the sustainability of different communication strategies. Indeed, looking back at the history of AI, all approaches come with associated costs in terms of resources and maintenance, from classical knowledge representation methods such as ontologies [4] to latest machine learning ones [7].

If we were to deploy multi-agent systems on the Web, one option would be to specify some ontology that they use to meaningfully communicate with each other. Regarding ontology design and implementation, sustainability directly relates to the strategies adopted in the multi-agent system using the ontologies for its cooperation and communication. It also depends on characteristics of the MAS itself e.g. in an open MAS where any new agent can join the communication strategies may be different from other more controlled systems.

Alternative agent-oriented communication protocols can differ in terms of software maintenance of agents and in terms of costs incurred to store and share knowledge. At the two extremes, the burden could be put either on the agents to align their individual ontologies when they need to interact, or, at the opposite end, on a shared ontology provider to ensure a *lingua franca*. Each extreme has its costs and benefits and there is a continuum of hybrid options in-between. For instance, letting agents have their own ontologies means they will pay the cost of aligning them when they have to communicate, whereas maintaining a shared ontology comes with the price of achieving consensus, commitment and versioning every time it needs to be changed. But how sustainable would these different strategies be from an economy of knowledge perspective?

In terms of economy of knowledge sharing, the Web can be seen as an open MAS, where knowledge is mostly public but

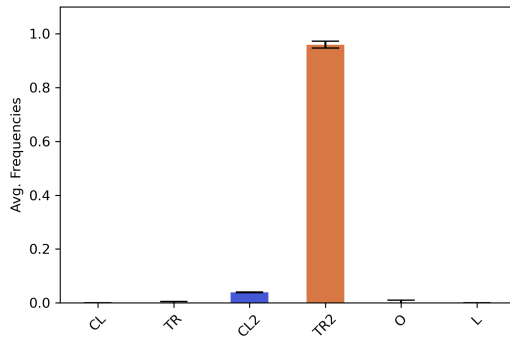
can also be private or restricted to a group or even a couple of users or agents (e.g. B2B interactions, intranets, etc.). We can also say that Web agents that contribute to Wikipedia are building a public good. The over-arching goals of Web agents would be to share as much knowledge as necessary to ensure they achieve their goals without disclosing their entire knowledge base (there might be some knowledge that needs to remain private). From an evolutionary perspective, this search for interoperability makes Web agents compete with each other, and assuming that the knowledge shared is relevant, the agents with the fittest strategies are able to share the most knowledge in the most efficient way possible.

We have known for decades that the diversity of resources in modern infospheres such as the Web calls for artificial ecosystems with a diversity of interacting agents ranging from reactive to deliberative paradigms, maintaining the information ecology [3] and supporting their hybrid communities in sustainable ways [1]. However, improper ethical advancement of AI could lead to unsustainable long-term side effects in knowledge of information sharing systems. On one hand we have the deployment of power-hungry protocols such as Blockchains. On the other hand, there is the risk of self-interested agents that wish to maximise the output of their knowledge sharing goals and cause two main negative outcomes on the *Infosphere* [5]. The first is caused by *exploitation* of resources, while the second is caused by the *pollution* of resources. For example, a Web agent might exploit the system it finds itself in (the Web in this case) by producing redundant information, or by doing so in conjunction with the exploitation of communication channels with other agents.

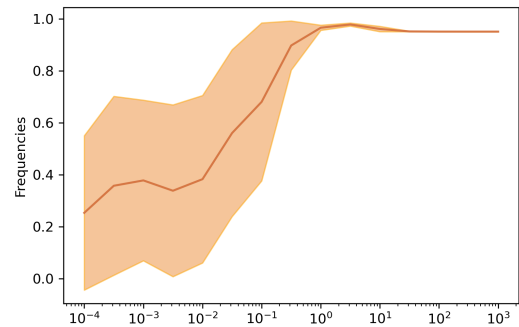
Information ecosystems are a multidimensional world where the importance of agents, their execution, the information resources, the infrastructure, the communication, ... must be treated from the perspective of sustainability as a complex self-organising system where every component has costs of design, production, running and maintenance that their owners may want to keep as low as possible.

II. METHOD & RESULTS

In “Interoperable AI: Evolutionary Race Towards Sustainable Knowledge Sharing” [6], we described a model simulation to study the sustainability of different strategies for cooperation in decentralized systems focusing on knowledge sharing among agents in a MAS. We assumed that the cost



(a) Average frequencies for voluntary interactions.



(b) Imitation effect for TR2.

of knowledge sharing is greatly impacted by the strategies for ontology-based communication being used by the agents.

We also assumed that Web agents want to be more sustainable and that they might compete with each other in finding the most sustainable means to share knowledge. One way to make self-interested agents more ‘sustainable’ would be to allow them to optimise between the strategies that they choose to share knowledge (and in some cases they would inevitably share) and the cost of doing so in various contexts. If these agents were able to choose between different knowledge sharing strategies, and able to learn from their peers the costs and benefits of these strategies, what strategy would they choose?

To test the idea that “the more ambitious goal of a common universal ontology is unlikely to be realized” [8], we have formulated a set of research questions and hypotheses to study the emergence on communication strategies in different knowledge-sharing ecosystems. Our main research question was: Does the choice of a common universal ontology emerge as an evolutionary stable strategy in populations of self-interested knowledge-sharing agents that learn from each other through imitation?

We identified seven strategies that range from using individual ontologies and translation services to using a common ontology. These are: 1) using own language (*O*), 2) translation (*TR*), 3) translation per knowledge (*TR2*), 4) using a common language with exponential cost of maintenance (*CL*), 5) using a common language with a maintenance cost proportional to the number of agents that use it (*CL2*), 6) a hybrid strategy that uses a combination of *TR* and *CL* named (*H*), and 7) a hybrid strategy that uses a combination of *TR2* and *CL2* (*H2*). We implemented these strategies and performed extensive simulations of evolutionary agent-based models. We tested the strategies when interactions were non-voluntary (agents had to participate in knowledge sharing) and voluntary where agents could decide not to participate in knowledge sharing by choosing the additional Loner strategy (*L*).

We found that, under the assumption of cooperation and the possibility of decentralization, translation strategies are more sustainable than shared ontology strategies for knowledge sharing between autonomous agents, and that hybrid strategies did not perform as well as expected. Figure 1a

shows that among the different strategies the translation per knowledge strategy (*TR2*) is the most sustainable one. *TR2* is the communication strategy that Web agents choose to pay a cost for translating the knowledge that they wish to share with other agents, but agents only pay for the translation of knowledge that they do not understand, i.e., for the exact pieces of knowledge for which they do not share exactly the same ontologies. The effect of imitation in Figure 1b shows us that agents with better strategy adoption/learning abilities will choose *TR2* more frequently than other agents.

As a concluding note, there is currently a lack of translation services of this kind. As mentioned in the paper, we use our results to motivate the deployment of translation-based approaches for making knowledge-sharing in decentralised systems more sustainable.

ACKNOWLEDGMENTS

The authors would like to thank the 3IA Cote d’Azur ANR-19-P3IA-0002 and the HyperAgents Project ANR-19-CE23-0030 for their support.

REFERENCES

- [1] Bettina Berendt, Fabien Gandon, Susan Halford, Wendy Hall, Jim Hendler, Katharina E Kinder-Kurlanda, Eirini Ntoutsi, and Steffen Staab. Web Futures: Inclusive, Intelligent, Sustainable The 2020 Manifesto for Web Science. *Dagstuhl Manifestos*, 2021.
- [2] Samuel H Christie, Lalana Kagal, Alessandro Ricci, and Munindar P Singh. Decentralized systems. *IEEE Internet Computing*, 26(6):5–6, 2022.
- [3] Fabien L Gandon. Combining reactive and deliberative agents for complete ecosystems in infospheres. In *IEEE/WIC International Conference on Intelligent Agent Technology, 2003. IAT 2003.*, pages 297–303. IEEE, 2003.
- [4] Asunción Gómez-Pérez, Mariano Fernández-López, and Oscar Corcho. *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer, 2006.
- [5] Gian Maria Greco and Luciano Floridi. The tragedy of the digital commons. *Ethics and Information Technology*, 6(2):73–81, 2004.
- [6] Stefan Sarkadi, Andrea GB Tettamanzi, and Fabien Gandon. Interoperable AI: Evolutionary Race Toward Sustainable Knowledge Sharing. *IEEE Internet Computing*, 26(6):25–32, 2022.
- [7] David Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. Hidden technical debt in machine learning systems. *Advances in neural information processing systems*, 28:2503–2511, 2015.
- [8] Barry Smith. Chapter 11: Ontology. In L. Floridi, editor, *Blackwell Guide to the Philosophy of Computing and Information*, pages 155–166. Blackwell, Oxford, 2003.