ACBE: Artificial Cultural Belief Evolution
Évolution culturelle artificielle de connaissances et de croyances

Principal investigator: Jérôme Euzenat (INRIA)

PRME AAPG2025, E.02: Intelligence artificielle et sciences des données, 54 months

Abstract

Social aspects of human beings have induced their remarkable cognitive capabilities. Cultural evolution is considered to have taken over biological evolution in their development. Yet, current work in artificial intelligence, besides one-to-one dialogues, does not take this into account. We aim at bringing cultural evolution to machines, like artificial intelligence aims to do with intelligence. Hence ACBE is a project of artificial cultural evolution. Although this applies globally to all cultural artifacts, ACBE focuses on knowledge and beliefs considered as a main driver of agent behaviour. ACBE will provide (a) models of cultural evolution games and agents, (b) a simulator for running experiments according to these models, (c) a framework for formally describing experiments to ensure their reproducibility, (d) experiments demonstrating the benefits of the approach, and (e) artificial cultural evolution awareness-raising initiatives. Numerous experiments simulating cultural evolution have been carried out, but they are designed in a one-shot fashion preventing to reuse and compare them. In particular, they do not isolate explicitly the evolutionary mechanisms of variation, selection and transmission. This prevents them from being easily identified and controlled in experiments and thus from assessing their influence on knowledge quality and diversity. ACBE will design a general model for games and agents involved in cultural evolution experiments (a). It will make explicit these parameters. Such a model should be prone to computer simulation and theoretical analysis. It will be implemented in a simulator (b) which allows agents to play different games and games to be played by heterogeneous agents in the same experiment. There exist no such environment yet. Experiments will be described formally (c) so that they can be easily modified, reproduced and their results compared. This will improve experiment reproducibility and permit searching and globally analysing an experiment repository. ACBE will perform experiments (d) demonstrating the benefits of the proposed approach. These will involve different agent populations and generations, controlling the evolutionary mechanisms intertwined in agents behaviour and involving different games and heterogenous agents. In particular, we aim at assessing the necessity and relative influence of each evolutionary mechanisms across different games. We also plan to establish the capability of heterogeneous agent populations, including some based on generative artificial intelligence, to evolve shared knowledge and beliefs. Finally, we will promote the proposed approach towards cultural evolution research communities and particularly those from social sciences and the humanities (e). The long-term prospect of ACBE is to understand how artificial systems can find their place in our societies.

I Proposal's context, positioning and objectives

a Objectives and research hypotheses

Artificial cultural evolution: cultural evolution for machines Cultural evolution considers that the mechanisms applying to the selection of species (variation, transmission, and selection) also play a role in the development of human culture [19]. Culture is understood here as information shared by a population that governs its behaviour, e.g. customs, norms, languages, laws, religions, sciences. It is considered to have taken over biological evolution in human development driving the success of homo sapiens [10, 21].

When artificial systems become involved in our social life, cultural evolution should play the same role. Indeed, for the society to accept artificial systems, a common culture will have to develop. It will also be necessary for groups of heterogeneous artificial systems to act together.

Current work on machine learning (now: artificial intelligence), focusses on the individual aspect of intelligence. It attempts at creating single systems able to solve, and to learn to solve, many problems. However, it can be argued that their social activity made human beings' cognitive capabilities remarkable. Besides one-to-one dialogues, current work in artificial intelligence mostly ignore this aspect. Although AI systems have different capabilities and, like human beings, do not always provide the same answer to the same question, it has been barely considered to make them debate together. However, if such systems are deployed together in our society, they should learn how to behave with people and with each other. That is, to build and evolve a common culture.

For that purpose, it is necessary to understand the laws governing cultural evolution. These may apply to artificial systems as well as other beings. Bringing cultural evolution capabilities to artificial systems aims at providing them with properties of *interoperability* (not wired in, but collectively acquired), *acceptance* (especially by human beings), and *resilience* (the ability to survive drastic changes in their environment). These are long-term properties and not features to be implemented once-for-all in one system.

ACBE aims at bringing cultural evolution to machines, like artificial intelligence aims to do for intelligence. The goal of 'artificial' is *not* to model specific phenomena with realistic agents, but to test the potential of various artificial ones. Hence ACBE is a project of artificial cultural evolution [25]. ACBE ambitions to provide tools to help better understanding cultural evolution mechanisms. For this, we need a general model that (*i*) exposes the mechanisms of cultural evolution, (*ii*) can be actionned to run experiments, and (*iii*) can be theoretically analysed.

Agent-based simulation for artificial cultural knowledge and belief evolution Artificial cultural evolution experiments consider artificial *agents* performing tasks, also called *games*, using their *culture*. Feedback from these tasks leads agents to *adapt* their culture. The result of these individual processes carried out by a large *population* of agents over a long time is cultural *evolution*. For instance, agents may use ontologies to make decision on how to proceed with objects of their environment. Their task is to agree on a common decision. When this fails, one agent will change its ontology to comply. In the long run, agents always agree, their ontologies describe better the environment, but they remain diverse [5]. Hence, artificial cultural evolution deals with micro-mechanisms considering how a single agent behaves and macro-mechanisms considering the wider context of agent societies or populations.

ACBE concentrates on knowledge and beliefs (hereby used interchangeably) as a key component to drive agents' behaviour. Knowledge is only one aspect of culture. However, its relation to behaviour is easy to understand and it has been well studied in artificial intelligence and other fields. Thus, it makes a good choice to investigate cultural evolution. This work will be based on symbolic knowledge and beliefs, but a large part of it should apply to subsymbolic systems.

Agents will be able to represent their knowledge and beliefs about their social and external environments. They will use it to behave in these environments. We base our work on a clear separation between knowledge and the behaviour that it induces. This separation and explicit representation has positive methodological consequences. One of these is that it is possible to apply measures on one or the other. Diversity, efficiency, distance, or correctness can be measured on both knowledge and behaviour. Separating knowledge from behaviour gives the opportunity to mix, in the same society, agents with different, e.g. symbolic and subsymbolic, knowledge representations and human beings.

There are various ways to study cultural evolution, but it has mostly been investigated through one-shot case studies. These may be wide-ranging observational work or agent-based simulations, but the models concern only one type of result to observe. In particular, they usually do not consider settings in which different types of agents are involved (except for coevolution experiments) and different games are played by these agents. Although this results in valuable models, it provides little information on the interference and complementarity between agent and game types.

Moreover, agents made for playing specific games lead to independent experimental settings making experiments difficult to compare. ACBE aims at improving the usability of artificial cultural evolution experiments by defining formal experiment descriptions that facilitate the reproducibility and repurposability of experiment designs as well as their queryability and comparability.

Characterising the evolutionary mechanisms of cultural evolution Biological evolution is based on variation, transmission and selection. These mechanisms are precise enough to describe evolution and general enough to cover more than biological evolution. However, in cultural evolution, they are far more widely applied and are tightly interdependent. Hence, covering them requires special attention.

Variation One key ingredient to evolution is the capability to generate variants. In the genetic model, variation occurs during gene transmission and depends on cross-over, mutations, and errors. In the case of agents, these variants are the possible knowledge representations adopted by agents. Variation occurs during transmission as well as in other occasions: when revising their beliefs after realising that they are erroneous, when choosing how to perform an action. Variant generation does not have to be random or simply frequentist. It is structured by the space of possible knowledge representations.

Transmission In genetics, transmission is inheritance at birth. In cultural evolution, transmission is not inheritance any more. One can distinguish between vertical and oblique (inter-generational) transmission and horizontal (intra-generational) transmission [19]. These three models may rely on different modalities (teaching, imitating). Moreover, transmission may occur during direct (communicating, teaching) or indirect (observing) transmission events. Such modalities depend on choices from the transmitter (what to transmit, with which modality, to whom) and from the receiver (from whom to learn and how to integrate it in one's knowledge). Only some of these aspects, such as transmission biases, have been extensively studied [20].

Selection In biological evolution, selection is often considered as the result of the fitness of individuals. Either unfit (or non viable) elements are simply discarded, or transmission is performed differentially with respect to fitness. In cultural evolution, agents predominantly select the knowledge that they will preserve or transmit. Contrary to genetic evolution so far, they can change their knowledge if they consider it as unfit or if they find out more promising one. Hence, agents will need criteria to decide how to retain or to discard knowledge. Such criteria, like fitness measures, may come from the outside of the agent, e.g. the observed failure of a task, or from the inside, e.g. its inconsistency with other knowledge.

As can be seen, the evolutionary mechanisms are well present in cultural evolution, but they are intrinsically intertwined. ACBE aims at contributing to understand how.

ACBE investigates the novel topic of artificial cultural knowledge and belief evolution. It is challenging by the questions it aims to answer, the current lack of precise delineation of concepts, and the interrelated solutions to develop.

ACBE has for objectives to:

Define a general **model** of artificial knowledge and belief evolution, involving games and agents, offering the capability to
- have agents playing several games and games played by several types of agents,
- be used for both experimental simulations and theoretical analyses.

Implement an integrated **simulator** based on this model allowing to
- straightforwardly experiment in a parameterised and extensible way,
- identify and make actionable evolutionary mechanisms as key parameters.

Support **open science** practices and in particular
- semantic experiment descriptions for quality insurance and reproducibility,
- findability, queryability and comparability of experiment descriptions.

Produce and reproduce outstanding **experiments** in cultural knowledge evolution
- involving experiments over multiple populations and generations,
- taking into account evolutionary mechanisms distinctively,
- involving multi-games and multi-agent types, including LLMs and humans,

Promote the approach and tools through open workshops and tutorials to **disseminate** towards
- different audiences, with a specific aims at reaching social sciences and the humanities,
- different culture types beyond knowledge and belief representation.

b Position of the project as it relates to the state of the art

Simulating cultural evolution

Artificial evolution has been considered in evolutionary computation [13]. However, this work closely adheres to the genetic metaphor which is too strict for cultural evolution. This is the case even when attempts at cultural evolutionary computation are made [22]. These are essentially tools for optimisation, though evolution has no aim, and not that of optimising.

Cultural evolution may be studied through observation of phenomena, population dynamic models (macro), physical small-scale experiments or agent-based simulation (micro). Cultural evolution experiments have been performed for long through multi-agent simulations: a society of agents adapts its culture through a precisely defined protocol [4]. Agents perform repeatedly a specific task under changing conditions, called game, and their evolution is monitored. This protocol aims to discover experimentally the states that agents may reach and the properties of these states. This has been successfully and convincingly applied to the evolution of natural languages [18, 28]. Agents play language games and adjust their vocabulary and grammar as soon as they are not able to communicate properly, i.e. if they misuse a term or do not behave in the expected way. Such experiments showed their capacity to model various settings in a systematic framework and to provide convincing explanations for linguistic phenomena. For instance, specific experiments have shown how agents can agree on a colour coding system or a grammatical case system.

Recently, artificial systems and especially large language models (LLM) have been considered for simulating cultural evolution [30]. However, most work focuses on their impact on human cultural evolution [6] or their adoption of human biases [2], not on their own cultural evolution.

A last type of simulation method is the one based on opinion propagation. It has been adapted to belief propagation [24] and we have started working on joint opinion/beliefs propagation [17]. So far, this has been based on simplistic mechanisms unrelated to cultural evolution. We plan to address this by allowing agents to use their knowledge when propagating either opinions or beliefs.

Cultural evolution experiments are usually designed from scratch and do not benefit from reuse. They are either implemented in a one-shot ad hoc way to test one feature at once [4, 28], or in all encompassing experiments that require setting a whole new environment each time [5]. Accordingly, there does not exist dedicated simulators exposing typical cultural evolution mechanisms. Our own framework, *Lazy lavender*, suffers from the same issue: it only offers agents able to play a single type of game.

Simulation platforms such as NetLogo [26], Gama [14] or R [1] do not offer direct cultural evolution mechanisms. This is one of the goals of ACBE to offer a dedicated cultural evolution simulator that (*i*) separates agents from games, and (*ii*) exposes evolutionary traits. Such a modular experiment design should contribute reusing, comparing and reproducing experiments.

Experiment reproducibility and open science

Historically, experiments were and still are described within papers expected to provide the necessary information to reproduce them. Laboratory notebooks record experiment information to ensure that paper descriptions correspond to what has been performed. Nowadays, electronic notebooks incorporating various tools are often used. Although such notebooks may be very precise and contain a lot of information, they are not always structured. Moreover, they are usually designed as independent artefacts, making them difficult to analyse globally as their descriptions differ.

We have developed our own methodology for describing experiments before performing them, and publishing repositories containing descriptions, results and analysis on our sake.re server. Although this contributes to reproducibility, this does not go as far as providing formal experiment descriptions.

There are two benefits to more formal experiment descriptions:

Reproducibility Describing experiments in a formal way should allow to replay them, to reproduce them and to repurpose them [31]. Beyond rerunning the actual experiment, running different statistical analyses of the data should be possible. Further formalising elements, such as experimental hypotheses, opens the opportunity to check automatically wether they are satisfied or not.

Meta-analysis Formal experiment description would also enable us to consider a base of experiments from which to find and query them [9]. This would allow us to compare experimental conditions and results across experiments. Ideally, it will be possible to generate a meta-analysis on a specific topic from the base.

This would contribute to make experiments FAIR, i.e. Findable, Interoperable, Accessible and Reproducible [32]. Currently, there is no notebook repository that can be queried on specific attributes. Dataset search engines, such as DataCite or Google data set search, which often contain experiment results, are searcheable on very poor metadata (besides dates and authors).

Recently, leaderboards like Kaggle have helped standardising some aspects of experiments which allow one to compare them. However, they focus on one type of problem to solve and rather constrain the 'agents' that solve them. Their descriptions lack the necessary flexibility.

Concerning computational experiments, of particular interest is the COMSES Computational Model Library [23], which tracks mostly agent-based models in life and social sciences. It aims at reproducing computational simulations. This is very useful but the descriptions are not sufficient for answering a question such as:

> *Are there agreement experiments with more than three benevolent agents which failed to show knowledge improvement.*

Such a query contains a lot of implicit information not available to a property-value-based query engine: types of objects (benevolent agents), types of experiment (agreement experiments), measurements (knowledge improvement). Other queries of interest would be:

> *Are there experiments whose results contradict those of this one?*

In order to achieve this, formal experiment description can rely on semantic web technologies.

Concerning the objects of scientific statements, many resources express these with semantic web technologies. Methodological aspects of research have not been left aside. The Open Research Knowledge Graph project (ORKG) [3, 16] ambitions to represent all that can be found in a research paper: this is wider than experiments, but experiments have reproducibility requirements which lead their descriptions to go beyond what is available in papers. Some work has proposed ontologies describing experiments [27], hypotheses [11] or claims [7]. On the experimental side: the researchobject project has provided a way to describe protocols [15] as well as SMART protocols [12]. Attempts are also made to describe evaluation methods [29].

The state of the art allows for expressing formally broad aspects of scientific knowledge. ACBE will build on it to specify formal experiment descriptions applying to cultural evolution experiments that improves reproducibility and meta-analysis.

c Methodology

Overall organisation

The project is organised in work lines. These are called work lines instead of work packages to insist on the idea that they are carried out during the whole project and beyond. Because most of the outcome of the project are research results, they also do not end at the end of the tasks: it is expected that the project results will be published through conference and journal papers as well as PhD dissertations. However, they are split in specific tasks with well-identified ends at which they must deliver a specific product.

Description of work

The project is organised along 5 work lines of distinct nature which correspond to the 5 objectives:

- WL1 Model cultural knowledge and belief evolution experiments
- WL2 Implement a simulator based on the designed model
- WL3 Support open science through reproducibility and meta-analysis
- WL4 Experiment cultural evolution, based on the features of the simulator
- WL5 Disseminate project results to cultural evolution researchers

WL1: Models of artificial cultural knowledge and belief evolution

The first work line will develop general **models** of artificial knowledge and belief evolution, involving games (T1.1) and agents (T1.2) that take part to cultural evolution experiments. An initial model will be provided to support the implementation of the simulator (WL2) and experiments (WL4). It will be constantly updated upon experiments and will reach a second version (T1.3).

There is currently no such general model of cultural evolution games. Some attempts have recently been made albeit focussing on one single game [8]. One important challenge of this work is that the designed models should be sufficiently precise so that (a) they can be processed by the simulator, and (b) they can be theoretically exploited in order to prove games or setting properties. Such models will offer the capability for agents to play several games and for games to be played by several types of agents.

Task 1.1: Game model Duration: **12m** ($T_0 - T_0 + 12$)
Goal: Defining the models of games used in cultural knowledge evolution
Output: Game and environment model (D1.1, report, M12)

Task 1.2: Agent model Duration: **12m** ($T_0 - T_0 + 12$)
Goal: Defining the models of agents used in cultural knowledge evolution
Output: Agent and population model (D1.2, report, M12)

Task 1.3: Extended models Duration: **12m** ($T_0 + 18 - T_0 + 30$)
Goal: Extending the agent and game models proposed previously
Output: Revised game and agent models (D1.3, report, M30)

WL2: Simulator

An integrated **simulator** will be developed based on the models developed in WL1. It will enable to identify key parameters of cultural evolution and to straightforwardly experiment in a parametric and extensible way. The first iteration of the simulator (T2.1) will process games with agents as defined in WL1. The second iteration (T2.2) will support agents playing several games at once and different types of agents playing. This will also allow for specific experiment scheduling, such as synchronous playing, which is often used in some types of simulations.

The ability for the same agent to enter different games using the same kind of knowledge and beliefs is a point that will require a specific attention and will constitute a significant advance on the state of the art. The fact that different agents cannot enter the same set of games does not raise problem and will offer the opportunity for new experiments.

Task 2.1: Implementation of the cultural knowledge evolution simulator Duration: **12m** ($T_0 + 6 - T_0 + 18$)
Goal: Developping a cultural knowledge evolution simulator
Output: ACBE Simulator (D2.1, software, M18)

Task 2.2: Simulator version 2 Duration: **12m** ($T_0 + 24 - T_0 + 36$)
Goal: Providing an improved and more stable version of the simulator
Output: ACBE Simulator v2 (D2.2, report, M36)

WL3: Reproducibility and open science

In order to support **open and reproducible experimental science**, this work line will provide a way to describe experiments, their hypotheses and their results so that they can be easily repurposed and rerun (T3.1). Such precise descriptions are key to the independent reproduction of experiments. In a second time, (T3.2), these descriptions will be extended and exploited in order to be able to retrieve, select and compare a large base of experiments and their results.

The difficulty of this work is to provide formal descriptions that allow experiments and global analyses to be processed. A balance has to be found between the depth of descriptions and the capability to address a wide range of experiments.

Task 3.1: Support for reproducibility Duration: **12m (**T_0+12-T_0+24**)**
Goal: Defining procedures and resources to warrant that simulations can be reproduced
Output: Support for experiment description and reproducibility (D3.1, report, M24)

Task 3.2: Support for meta-analysis Duration: **12m (**T_0+36-T_0+48**)**
Goal: Providing metadata descriptions for finding and querying experiment records
Output: Support for experiment meta-analysis (D3.2, report, M48)

WL4: Experiments

The aim of this work line is to demonstrate that the promoted approach can be used to obtain significant results in experimental artificial cultural evolution. Based upon the simulator provided in WL2, cultural knowledge and belief evolution **experiments** will be carried out. This will also offer a first library of games and agents to be reused in further experiments. A first round of experiment (T4.1) will aim at reproducing existing experiments within the framework. This should provide a fast feed-back on its implementation. The second round of experiments (T4.2) will more specifically show the benefits of clearly identifying evolution mechanisms within agent models. Finally, the third experiments (T4.2) will provide evidence of the capacity of the system to support the integration of different kinds of agents simultaneously playing different games.

The goal of this work line is not simply to test the proposed environment: it aims at providing valuable research results in cultural knowledge and belief evolution, and specifically to its core with respect to evolutionary mechanisms. Beyond reports, deliverables are expected to be experiment records including experiment descriptions, results and analysis, and that will be published in papers.

Task 4.1: Experiments with populations and generations Duration: **12m (**T_0+12-T_0+24**)**
Goal: Showing the capability of the simulator to deal with multi-populations and multi-generation experiments
Output: Cultural evolution experiments with populations and generations (D4.1, report, M24)

Task 4.2: Experiments with distinct evolutionary mechanisms Duration: **12m (**T_0+24-T_0+36**)**
Goal: Showing the benefits of identifying evolutionary mechanisms in experiments
Output: Cultural evolution experiments with distinct evolutionary mechanisms (D4.2, report, M36)

Task 4.3: Multi-game and multi-type experiments Duration: **12m (**T_0+42-T_0+54**)**
Goal: Performing experiments in which heterogeneous agents play simultaneously together several games
Output: Multi-game and multi-type cultural evolution experiments (D4.3, report, M54)

WL5: Dissemination and management

This work line deals with the management of the project and the dissemination of the outputs of ACBE. It organises the project (T5.1) so that it delivers the best results, and takes care of their dissemination (T5.2). It will promote the proposed approach and platform through open workshops and tutorials aiming to **disseminate** towards different audiences, with a specific goal to reach social sciences and the humanities, and opens to different culture types beyond knowledge and belief representation.

The dissemination task (T5.2) is of specific importance for ACBE as outreach to at least a part of researchers involved in cultural evolution research belongs to our objectives. For that purpose, paper publication about our progress will also be part of it.

Task 5.1: Management and relation with ANR Duration: **54m** ($T_0 - T_0 + 54$)
Goal: Managing the project and coordinating with ANR
Output: Annual report (D5.2, report, M12, M24, M36, M48), Final report (D5.3, report, M54)

Task 5.2: Dissemination Duration: **54m** ($T_0 - T_0 + 54$)
Goal: Designing and implementing dissemination actions
Output: Website (D5.1, report, M3) Workshops on artificial cultural belief evolution (D5.4, workshop, M30) Tutorials on ACBE simulation (D5.5, tutorial, M42) Interdisciplinary workshop on Artificial cultural evolution (D5.6, workshop, M48)

II Impact and benefits of the project

ACBE is not directly applied and rather aims at facilitating scientific research. However, we foresee its potential impact in the long term on research, industry and society.

a Impact on research

The research fields concerned by the outcome of the ACBE project are many. The most direct ones are artificial intelligence and multi-agent simulation of which we are regular contributors.

However, still remaining in computer science, we aim at having an impact on fields such as social network analysis and opinion dynamics, by introducing the effect of beliefs and opinions in these fields, and artificial life, to which our integrative experiments (T4.3) may claim to belong.

But besides computer science, we aim at providing more relevant tools for research in cultural evolution, i.e. in all social sciences and the humanities. This is to that extent that our dissemination effort (T5.2) is primarily oriented. In addition to the specific tools provided by ACBE, the goal of the outreach efforts (web site, workshop and tutorial) is to promote the methodology and approach. Our main goal is to convince social scientists that studying artificial cultural evolution should help their fields progress beyond, and together with, terrain-oriented approaches. Hence, the sought impact is in fostering important research results in these fields.

A side-effect of the project is its contribution to better experiment descriptions that has the potential to impact all experimental sciences.

b Impact on industry

Beyond providing insights in the nature of artificial cultural belief evolution, ACBE has the potential to impact socio-technical system fields such as social robotics which mixes artificial systems with people. For instance, robots designed to be part of our societies will have to acquire knowledge about their physical and social environment. They will have to adapt their knowledge in order to be integrated and well-behaving.

There are some fears that such systems may not abide to human values and beliefs. Another worry is that too rigid systems may be rejected by potential users. Finally, the race towards short-term optimisation may reduce diversity and be detrimental to long-term resilience.

Artificial cultural evolution is definitely an approach to address these points. Evolving such systems with cultural evolution mechanisms may help integrate common values and achieve a smoother integration. Yet we want to understand if the behaviour of artificial systems must be explicitly programmed within such systems (value-based engineering) or if it can be constrained by their very cultural evolution mechanisms (socially negotiated, hence not hard-wired).

More generally, the proposed approach can be used to simulate information propagation and adoption in concrete organisations, e.g. companies, institutions.

In another direction, though the work on reproducibility aims at developing a proof of concept, its outcome may be submitted for consideration to be integrated in future standards, e.g. Research data alliance.

c Impact on society

The former paragraphs already touch on societal issues. What applies to industrial deployment necessarily impacts our societies. More generally, the work on artificial cultural evolution aims at apprehending how artificial systems may evolve a culture, here knowledge and beliefs, together with other members of the society. Understanding, how this can be achieved will help creating better artificial systems and help better developing our societies [10].

A benefit of separating knowledge and beliefs from the context-dependent behaviour is the opportunity to preserve knowledge diversity supporting compatible behaviour. This is an important issue as, in evolution theory, diversity is considered as an important indicator of resilience.

Another social implication of our work is the study of knowledge and belief propagation. Models of opinion dynamics are rather simplistic. Introducing knowledge and beliefs and some evolutionary mechanisms may bring a better understanding of this propagation. This may be applied to counter 'fake news' and 'conspiracy theory' propagation.

III References related to the project

[1] Alberto Acerbi, Alex Mesoudi, and Marco Smolla. *Individual-based models of cultural evolution— a step-by-step guide using R*. Routledge, London (UK), 2022. URL: https://acerbialberto.com/IBM-cultevo/.

[2] Alberto Acerbi and Joseph Stubbersfield. Large language models show human-like content biases in transmission chain experiments. *Proceedings of the national academy of science* 120.44 (2023), e2313790120. DOI: 10.1073/pnas.2313790120.

[3] Sören Auer, Viktor Kovtun, Manuel Prinz, Anna Kasprzik, Markus Stocker, and Maria-Esther Vidal. Towards a Knowledge Graph for Science. In: *Proc. 8th International Conference on Web Intelligence, Mining and Semantics (WIMS), Novi Sad (RS)*. ACM press, 2018, 1:1–1:6. DOI: 10.1145/3227609.3227689.

[4] Robert Axelrod. The dissemination of culture: a model with local convergence and global polarization. *Journal of conflict resolution* 41.2 (1997), pp. 203–226. DOI: 10.1177/0022002797041002001.

[5] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. Knowledge Improvement and Diversity under Interaction-Driven Adaptation of Learned Ontologies. In: *Proceedings of the 20th international conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Online*. 2021, pp. 242–250. URL: http://www.ifaamas.org/Proceedings/aamas2021/pdfs/p242.pdf.

[6] Levin Brinkmann, Fabian Baumann, Jean-François Bonnefon, Maxime Derex, Thomas Müller, Anne-Marie Nussberger, Agnieszka Czaplicka, Alberto Acerbi, Thomas Griffiths, Joseph Henrich, Joel Leibo, Richard McElreath, Pierre-Yves Oudeyer, Jonathan Stray, and Iyad Rahwan. Machine culture. *Nature human behaviour* 7 (2023), pp. 1855–1868. DOI: 10.1038/s41562-023-01742-2.

[7] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. Expressing High-Level Scientific Claims with Formal Semantics. In: *Proc. 11th International Conference on Knowledge Capture (K-Cap), Online*. ACM press, 2021, pp. 233–240. ISBN: 978-1-4503-8457-5. DOI: 10.1145/3460210.3493561.

[8] Paul Van Eecke, Katrien Beuls, Jérôme Botoko Ekila, and Roxana Rădulescu. Language games meet multi-agent reinforcement learning: A case study for the naming game. *Journal of language evolution* 7.2 (2022), pp. 213–223. DOI: https://doi.org/10.1093/jole/lzad001.

[9] Jérôme Euzenat. Beyond reproduction, experiments want to be understood. en. In: *Proc. 2nd workshop on Scientific knowledge: representation, discovery, and assessment (SciK), Lyon (FR)*. 2022, pp. 774–778. URL: https://dl.acm.org/doi/pdf/10.1145/3487553.3524676.

[10] Henry Farrell, Alison Gopnik, Cosma Shalizi, and James Evans. Large AI models are cultural and social technologies. *Science* 387.6739 (2025), pp. 1153–1156.

[11] Daniel Garijo, Yolanda Gil, and Varun Ratnakar. The DISK Hypothesis Ontology: Capturing Hypothesis Evolution for Automated Discovery. In: *Proc. Workshops and Tutorials of the 9th International Conference on Knowledge Capture (K-CAP), Austin (TX US)*. Ed. by Ilaria Tiddi, Giuseppe Rizzo, and Óscar Corcho. Vol. 2065. CEUR Workshop Proceedings. CEUR-WS.org, 2017, pp. 40–46. URL: http://ceur-ws.org/Vol-2065/paper08.pdf.

[12] Olga Giraldo, Alexander Garcia, Federico Lopez, and Oscar Corcho. Using semantics for representing experimental protocols. *Journal of Biomedical Semantics* 8 (2017), p. 52. DOI: 10.1186/s13326-017-0160-y.

[13] David Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading (MA US), 1989.

[14] Arnaud Grignard, Patrick Taillandier, Benoit Gaudou, Duc An Vo, Nghi Quang Huynh, and Alexis Drogoul. GAMA 1.6: Advancing the Art of Complex Agent-Based Modeling and Simulation. In: *Proc. 16th international conference on principles and practice of multi-agent systems (PRIMA), Dunedin (NZ)*. Vol. 8291. Lecture notes in computer science. 2013, pp. 117–131. DOI: 10.1007/978-3-642-44927-7_9.

[15] Kristina Hettne, Harish Dharuri, Jun Zhao, Katherine Wolstencroft, Khalid Belhajjame, Stian Soiland-Reyes, Eleni Mina, Mark Thompson, Don Cruickshank, Lourdes Verdes Montenegro, Julian Garrido, David de Roure, Oscar Corcho, Graham Klyne, Reinout van Schouwen, Peter 't Hoen, Sean Bechhofer, Carole Goble, and Marco Roos. Structuring research methods and data with the research object model: genomics workflows as a case study. *Journal of biomedical semantics* 5.1 (2014), p. 41.

[16] Mohamad Yaser Jaradeh, Allard Oelen, Kheir Eddine Farfar, Manuel Prinz, Jennifer D'Souza, Gábor Kismihók, Markus Stocker, and Sören Auer. Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge. In: *Proc. 10th International Conference on Knowledge Capture (K-Cap), Marina Del Rey (US)*. 2019, pp. 243–246. DOI: 10.1145/3360901.3364435.

[17] Hiro Kataoka, Jérôme Euzenat, and Koji Hasebe. Exchanging and updating opinions and beliefs reinforces echo chambers. In: *Proc. 39th annual conference of the Japanese society for artificial intelligence (JSAI), Osaka (JP)*. 2025.

[18] Simon Kirby, Hannah Cornish, and Kenny Smith. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the national academy of sciences* 105.31 (2008), pp. 10681–10686.

[19] Alex Mesoudi. *Cultural Evolution: How Darwinian theory can explain human culture and synthesize the social sciences*. University of Chicago Press, Chicago (IL US), 2011.

[20] Alex Mesoudi and Andrew Whiten. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical transactions of the royal society B* 363 (2008), pp. 3489–3501. DOI: 10.1098/rstb.2008.0129.

[21] Jürgen Renn. *The evolution of knowledge: rethinking science for the anthropocene*. Princeton university press, Princeton (NJ US), 2020.

[22] Robert Reynolds, ed. *Cultural algorithms: tools to model complex dynamic social systems*. Piscataway (NJ US): IEEE press, 2020. DOI: 10.1002/9781119403111.

[23] Nathan Rollins, Michael Barton, Sean Bergin, Marco Janssen, and Allen Lee. A Computational Model Library for publishing model documentation and code. *Environmental modelling and software* 61.4 (2014), pp. 59–64. DOI: 10.1016/j.envsoft.2014.06.022.

[24] Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sébastien Konieczny, and Pierre Marquis. Belief Revision Games. In: *Proceedings of the 29th AAAI Conference, Austin (TX US)*. 2015, pp. 1590–1596. DOI: 10.1609/aaai.v29i1.9415.

[25] Herbert Simon. *The sciences of the artificial*. Cambridge (MA US): The MIT press, 1969.

[26] Paul Smaldino. *Modeling social behavior: mathematical and agent-based models of social dynamics and cultural evolution*. Princeton university press, Princeton (NJ US), 2023.

[27] Larisa Soldatova and Ross King. An ontology of scientific experiments. *Journal of the Royal Society Interface* 3.11 (2006), pp. 795–803. DOI: 10.1098/rsif.2006.0134.

[28] Luc Steels, ed. *Experiments in cultural language evolution*. John Benjamins, Amsterdam (NL), 2012.

[29] Markus Stocker, Manuel Prinz, Fatemeh Rostami, and Tibor Kempf. Towards research infrastructures that curate scientific information: A use case in life sciences. In: *Proc. 13th International Conference on Data Integration in the Life Sciences (DILS), Hannover (DE)*. Vol. 11371. Lecture Notes in Computer Science. Springer, 2018, pp. 61–74. ISBN: 978-3-030-06015-2. DOI: 10.1007/978-3-030-06016-9_6.

[30] Aron Vallinder and Edward Hughes. Cultural Evolution of Cooperation among LLM Agents. In: *Proceedings of the 24th international conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Detroit (US MI)*. 2025, to appear.

[31] Luisa Werner, Pierre Genevès, Nabil Layaïda, Jérôme Euzenat, and Damien Graux. Reproduce, replicate, reevaluate: the long but safe way to extend machine learning methods. In: *Proc. 38th AAAI Conference on Artificial Intelligence (AAAI), Vancouver (CA)*. 2024, pp. 15850–15858. URL: https://ojs.aaai.org/index.php/AAAI/article/view/29515/30854.

[32] Mark Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip Bourne, Jildau Bouwman, Anthony Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair Gray, Paul Groth, Carole Goble, Jeffrey Grethe, Jaap Heringa, Peter 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott Lusher, Maryann Martone, Albert Mons, Abel Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3 (2016), p. 18. DOI: 10.1038/sdata.2016.18.