iut de BORDEAUX

Informatique

名古屋大学
NAGOYA UNIVERSITY

BUT2 internship report at

ALIFE-CORE

laboratory

Assessment of an artificial life experiment model in the context of the research regarding open-endedness of LLMs.

Student:	Clément BARRIÈRE
Internship supervisor:	Reiji SUZUKI
Responsible teacher:	Bruno MERY
College year:	2024/2025

Abstract

English version

This internship report summarizes a two-month experience at ALIFE-CORE, artificial life laboratory at the Graduate School of Informatics of Nagoya, Japan. As a member of this structure, I participated in rituals and built my research plan.
The primary mission was to assess Evolutionary Ecology of Words, an artificial life experiment model, for its use as an open-endedness driven benchmark for LLMs.
The second objective was to develop more flexible and customizable systems to run experiments, which notably include the implementation of a new and handy way to integrate LLMs into this pythonic model.
The research led to the discovery of a strong language bias and its use as a creativity engine, since multilingual experiments showed interesting alteration of LLMs' behavior. The questioning of the tools and the initial design of the experiment model finally led to a completely new and specialized application, improving the analysis scope of the results.
I had the chance to see this fruitful work integrated into Evolutionary Ecology of Words research advancement, and have been invited to share these results in a national scale conference, assisting Pr. Reiji SUZUKI for the presentation of his experiment model.

French version

Ce rapport de stage résume une expérience de deux mois chez ALIFE-CORE, laboratoire de recherche en vie artificiel à l'Ecole Supérieure d'Informatique de Nagoya, Japon. En tant que membre de cette structure, j'ai eu l'occasion de participer aux différents rituels et de construire mon propre plan de recherche.
La mission principale consistait en l'évaluation de Evolutionary Ecology of Words, un modèle d'expérience de vie artificielle, pour son utilisation en tant que benchmark de non-déterminisme pour les LLMs. Le second objectif était de développer des systèmes plus flexibles et ajustables, ce qui inclut notamment l'implémentation d'un nouveau moyen habile d'intégrer un LLM dans ce modèle d'expérience pythonique.
La recherche a mené à la découverte d'un puissant biais de langage et à son utilisation en tant que moteur de créativité, étant donné les expériences multilangues ont montré des propriétés intéressantes quant à l'altération du comportement des LLMs. La remise en question des outils et de l'architecture initiale du modèle d'expérience ont finalement conduit à une toute nouvelle application spécialisée, améliorant l'amplitude avec laquelle les données sont maintenant analysées.
J'ai eu la chance de voir ce travail fructueux complètement intégré à l'avancée de la recherche sur Evolutionary Ecology of Words, et j'ai été invité à partager ces résultats lors d'une conférence d'ampleur nationale, en tant qu'assistant du Pr. Reiji SUZUKI pour la présentation de son modèle d'expérience.

Table of content

I. The ALIFE-CORE laboratory

Integration into the host structure

This part is either dedicated to the description of the laboratory that welcomed me for this internship, and to my contribution in the activities of this structure.

Main activities

ALIFE-CORE is an artificial life (Alife) and artificial intelligence (AI) laboratory at the Graduate School of Informatics of Nagoya, Japan. The main research areas of this laboratory are the study of:

- the interactions between evolution and apprenticeship,
- the emergence of evolutionary dynamics and personality traits,
- the acoustic interactions for birds,
- the relation between individuals' behavior and societies' structure,
- the open-endedness of artificial models (such as LLMs).

All these phenomena are studied through complex experiments, reproducing environments and constraints to observe targeted dynamics.
The scheme below describes the three pillars of Alife, as well as their respective applications and tendencies:

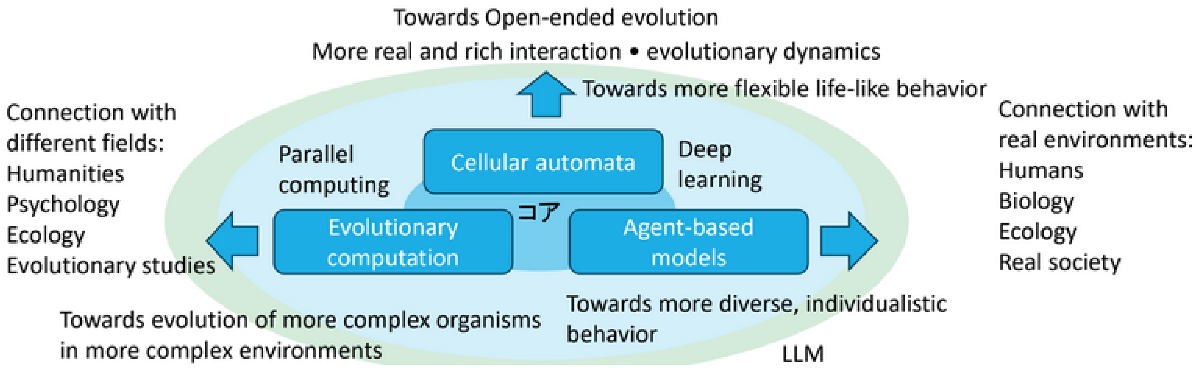


Fig.1 - core concepts of artificial life research

Personal contribution

During this internship, I took part in the research regarding the open-endedness of artificial models. This aspect can be designated by the ability to produce a non-deterministic outcome from an input. It is a core concept in Alife and AI research, which explores systems capable of continuous innovation and increasing complexity. Understanding and applying the novel and creative generative capabilities of LLMs is an important theme. Not only for engineering applications of evolutionary systems, but also for open-ended evolution.

Without going into details (cf. The internship mission), my role in the organization was to assess a prototype artificial life model for use as an open-endedness driven LLM benchmark, and to highlight improvement possibilities. Over the weeks, I could strengthen my analysis and summarization skills, taking my colleagues' work presentations as examples. And I hope my findings will help valuing the promising experiment model I worked on, regarding open-endedness research.

Twice a week, I had a little meeting in my supervisor's office to discuss the advancements I made and potential interrogations and advice that could be shared. As I was only there for two months, the supervision was consequently sustained. While I analyzed my results and explored the possibilities, my supervisor pointed out the relevant paths to follow so as not to get lost.

Structure's organization

This part is dedicated to the working environment I have been evolving in. The following description is factual, and aims to illustrate what it is to be a member of the ALIFE-CORE laboratory. A more subjective viewpoint will be described further (cf. Quality of life at work).

The research staff

This laboratory is led by Pr. Reiji SUZUKI, PhD graduate from the Division of Human Informatics and professor at the Department of Complex Systems Science 4 at Nagoya's Graduate School of Informatics. Which is why the laboratory is also known as the Suzuki unit.
More generally, the Suzuki unit is made up of 22 members, including 1 professor, 1 doctor, 6 PhD students, 6 M2 students, 6 M1 students, and 2 B4 students (last undergraduate bachelor level in Japan).

Given that the members have various research topics and are at different levels of study, exchanging ideas will have been very helpful in gathering different points of view on blocking issues.
To mention examples of research led at ALIFE-CORE:

- Mr. FURUKAWA Tatsuki (PhD student) is currently working on a new type of artificial creature, combining Lenia's model and the biological hierarchy principle.
- Ms. KATO Mahiro (doctor) is currently working on the ability (or not) of LLMs to demonstrate creativity, through the prism of literature.
- Mr. HARA Takumi (M1 student) is currently leading research on gene-based weights of LLMs and fine-tuning for specific tasks, without gradient calculation.

Daily life at the laboratory

The Graduate School of Informatics provides premises for research laboratories, ALIFE-CORE is located on the fifth floor of this building. The head professor has their own office and the rest of the research staff is dispatched into 4 coworking rooms (cf. annexe 1). Each member has their own desktop computer, equipped with last gen CPUs and RTX Nvidia GPUs, ensuring enough computation power to run experiments. The premises are open from 8am to 9pm, but each member has a pass card and the key to their working room.

Each Tuesday and Thursday, all the staff participates in a meeting, and about two members share their research progress. The laboratory also provides several online chat channels such as Teams and Rocket Chat, which facilitates mutual aid.

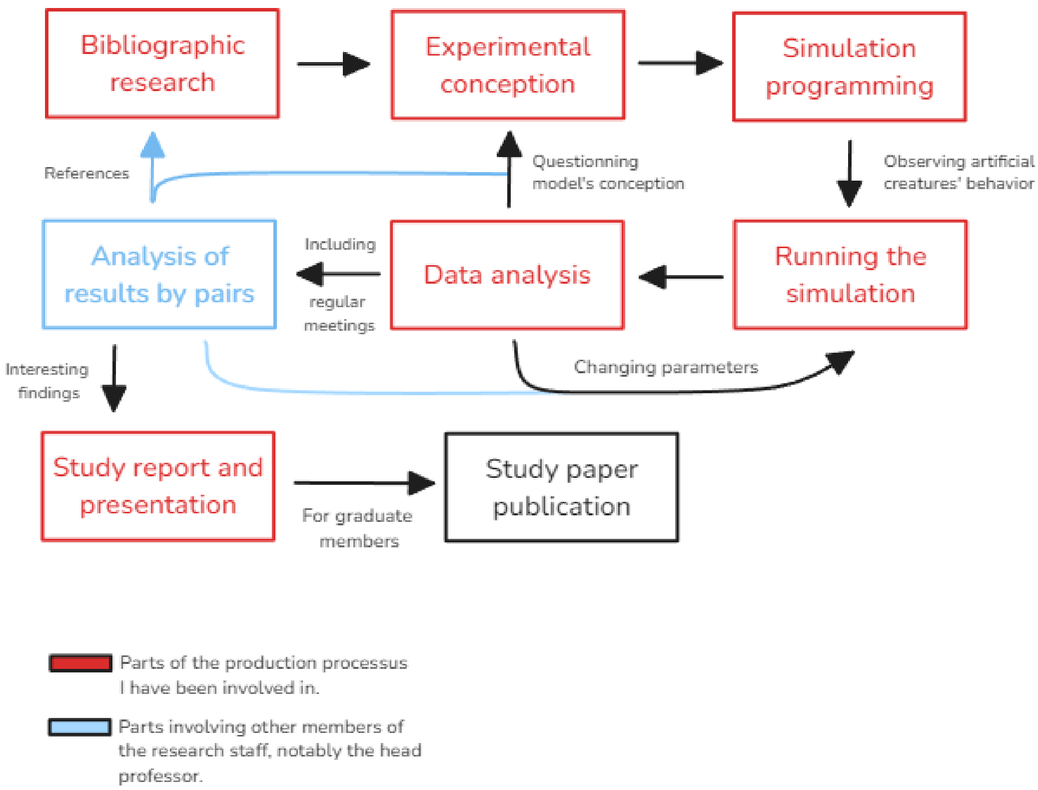


Fig.2 - production process at the laboratory, personal contribution and other actors

The internship mission

This part is an overview of what I was involved in during this internship. Particularly focused on my professional experience and the management of my R&D project.

Discovery of the research profession

During this internship, Pr. SUZUKI offered me the chance to elaborate my own research plan. He suggested I work on a new simulation model called Evolutionary

Ecology of Words (abbreviated to WEM), which will be detailed in a next part (cf. Evolutionary Ecology of Words). In a nutshell, this model involves ecosystems of words and their semantic derivation using the rich linguistic expression and inference ability of LLMs.

In March 2025, Pr. Reiji SUZUKI and his colleague Pr. Takaya ARITA, published a paper (cf. Webography) upon the early stages of this new experiment model, which is rooted in game theory. I therefore dedicated the first two weeks to resume the model conception and previous analysis, down to the raw data used in the original experiment. This kind of documentation work would have lasted until the very end of this internship, throughout very interesting papers referenced by Pr. SUZUKI.

The range of possible angles from which to study WEM is wide. The major part of my work would have been focused on a case study: the impact of languages on experiments' results. On the premise that a language is imbued with the culture it comes from, I really liked the idea that changing prompt language could open up different, more or less specific, semantic fields. Experiments following this hypothesis turned out to be very fruitful, highlighting a strong language bias, and exploiting it as an engine of creativity for LLMs; while discovering other unique properties linked to the different languages tested. This advance demanded the elaboration of new metrics and therefore also allowed questioning the visualization tools.

I have been encouraged to share these results with the rest of the laboratory in one of our regular meetings. A real challenge of summarisation, concision, and valuing. Either orally and graphically with the slides, to make it the most understandable, even for those who didn't speak english .
In addition, I had the honor to see my work fully integrated into the WEM research advancement. Pr. SUZUKI and Nagoya University invited me to participate in the 39th annual conference of the Japanese Society for Artificial Intelligence, held in Osaka. There, I assisted the professor to present WEM and share my findings. A golden opportunity to delve deeper into the experience of researcher professional activity.

Improvement of the model's design

This internship will have been twice more formative as I literally led two projects at the same time. Beside the research, Pr. SUZUKI asked me if I could search for a more flexible way to manage LLMs within the simulation, since the actual configuration required to download specific files manually. I so wanted to make it more or less automatic and, in order to be in line with the BUT's program, I refactored the original simulation code into a specialized pythonic application. Even though I was the only developer on the project, I could apply management techniques learned along BUT's courses, and drove the development following OOP concepts.

After two weeks of state-of-the-art existing technologies, I was able to propose a solution matching the need of managing LLMs more easily. This new module now only requires to indicate a model name to load from Hugging Face (most known AI models hub). Once this high-priority need was settled, I started the conception of the overall code I was about to refactor, constructing the UML mockups and then organizing the tasks with the GitHub issues board (cf. annexe 2). Scheduling and imposing deadlines was also important as my time was divided between this development and the research. The project ended in the 8th week, ready to be used as a demonstration tool for the open day of the laboratory, where each member presented their work to visitors.

The improvement of the architecture of the simulation code highly helped saving time over the maintenance. Notably when the research results indicated a need to review certain aspects or implement new features. To mention an example, the elaboration of the emergence score (cf. Development of new metrics) and the associated graph demanded to implement the computation and the log of a mutation history that happened during the experiment.

II. Research & Development

Contribution to the research

This part delves into details of the artificial life model I have been studying, and on the results I was able to produce. Many technical terms are used there, which are defined in the glossary below.

Evolutionary Ecology of Words

Following a WEM based experiment, we can compute several metrics and create graphs from the results (cf. annexe 4). Before all, it is important to know that such an experiment is partly orchestrated thanks to a set of four specific prompts, one for each phases:

- a prefix *context prompt* to indicate the role that the model should bear,
- the generation of initial *existing animal* species names list,
- the judgment between two names to decide *which one is stronger*,
- the generation of *mutation possibilities related to a mutating word*.

As described in Fig.3, and regarding those prompts, these three phases are the articulations that allow any state change within the population of words. Notably, conquest dynamics with *competitions* and new words emergence with *mutations*.

The first available type of visual is a collection of graphs studying the B most present names over the steps. With this one, we can observe the emergence of prey-predator dynamics, food chains, and even epidemic phenomena (cf. annexe 7).

In this case, it translates the alignment of models, since without even meaning to, we can observe the development of nature like ecosystems.

The second type of visual is a semantic graph that uses the Sentence Transformers library to vectorize the names and the UMAP library to draw trajectories. These last materialize the semantic derivation from the initial population of names to the final one, regularly plotting the average vector of all the semantic word vectors within the current population. The more the distance between two plots, the less the two populations have to do with each other. With such a graph, we can identify semantic areas or fields and so observe dynamics of specialisation, exploration or exploitation of words and even concepts that could emerge.

WEM aims to create an environment in which LLMs can maximize their creativity and so test their limits on that aspect. It can notably be measured through the number of novelties that emerged along the experiment. It's not a question of whether the results are realistic or not.

- The LLM generates **A** *existing animal species names.*
- **N** agents are created:
  - it is assigned one of the names in the list,
  - it gets random initial coordonates on the grid.
- For **S** steps:
  1. *Random walk*:
     - each agent take a step forward in a random direction.
  2. *Competition:*
     - if the agent has neighbours, compete with a random one,
     - regarding the words carried by each competitors, the LLM is asked to chose «which one is stronger»,
     - the loser agent gets its word overwritten by the winner's one.
  3. *Mutation:*
     - each agent has a probability **P** to mutate,
     - the LLM generates **B** words related, but more or less different, from the agent current one,
     - the new word of the agent is randomly chosen among these possibilities.

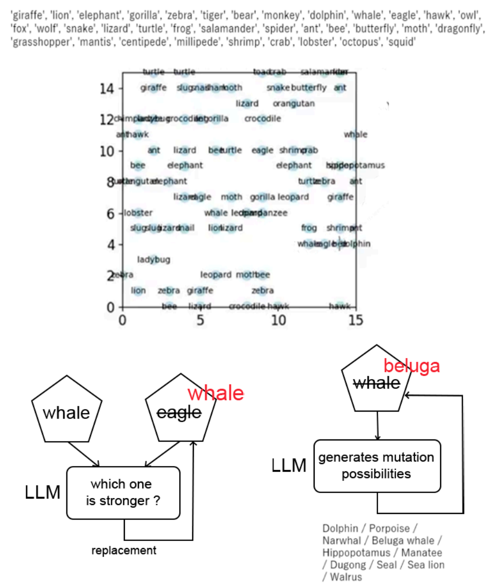


Fig.3 - Evolutionary Ecology of Words model execution

Language bias

My contribution to the research mainly focused on the impact of prompt language on the results. In my opinion, the animal species theme has a really good potential in highlighting the difference between cultural representation of the same name. Legends and even pop culture shape the meaning of the words, in different languages. It took three shots of 10 experiments for me to arrive at my final interpretation of the results, about a week in total of running experiments non-stop.

So I began by selecting several languages, each more or less different from the others in terms of culture and writing style: English, French, Russian and Chinese.

This revealed two things: firstly, when moving from English to another language, the LLM tends to get lost in translation, resulting in the development of multilingual experience; secondly, different translations of the same word, coexisting, had significantly different semantic vectors (cf. annexe 5). This told us that, despite the words themselves, the language used itself had a significant semantic value.

A language bias had been discovered, but whether it was a problem or a useful feature remained to be seen. In the world of IT production, a bias is very often something negative that can falsify results. In this case, we wondered whether the simple translation of a word, while adding to the graphs, could be considered as a real emergence of novelty. So we went back to the basic metrics, more than the semantic analysis: the literal number of emergencies (in new words encountered). This was of course inflated by translations, which is why I developed the emergence score (cf. Development of new metrics), that enabled me to correlate the discovery or consumption of semantic fields with positive or negative trends on the emergence of novelties. These results underscored the potential of multilingual experiments, as zig-zagging between languages helped diversify the species categories within the population, and therefore prevent stagnation in one semantic field. Apart from some particular cases of over-specialized languages, such as Latin, that allow the emergence of specific species categories, such as extinct ones, but tend to clusterize and won’t let the trajectory out easily.

Finally, the different experiments with different languages showed the emergence of characteristic patterns and dynamics, in terms of exploration, exploitation and even robustness of these languages (cf. annexe 6). In other words, the language used to query the LLM influences its generation behavior, which can be used to test different aspects.

Development of a specialized application

This part focuses on the “simulation” module of the application I have been developing to run WEM based experiments. Many technical terms are used there, which are defined in the glossary below.

Independent AI module and optimisation

The original WEM code used the Python llama-cpp library. This library is renowned for its efficiency and optimization. However, the need was not to select the fastest systems, but the most flexible ones, so as to be able to compare several models without revising the whole configuration. I chose Hugging Face's Transformers library, as it provides so-called “Auto” classes which, given a model name, attempt to resolve the necessary configuration and download the required files (e.g. tensor files, tokenizer configuration file, etc...).

The corresponding class in the application therefore acts as a wrapper for this library, allowing us to instantiate a model locally. Consequently, I implemented all the relevant configuration options for such use: local offload, whether or not to use the GPU, or even the quantization of weights. Once the model has been instantiated and configured, this object can be used in a loop to generate responses by giving it a prompt. The usefulness for WEM is also being able to change models during an experiment, or simply to launch several experiments with different models very easily. Moreover, the new LLM handling module is completely independent from the rest of the application, meaning that it can be used to integrate LLMs into any other program without having to adapt its code.

```
self._model: LanguageModelHandler = LanguageModelHandler(
    model_name=self.config["model"]["name"],
    log_event=self._log_event,
)


self._model.configure_model(
    local_offload= self.config["model"]["local_offload"],
    quantization= self.config["model"]["quantization"],
    temperature= self.config["model"]["temperature"],
    use_gpu= self.config["model"]["use_gpu"],
)


response = self._model.generate_response(
    prompt,
    min_new_tokens=n,
    max_new_tokens=self.config["model"]["max_tokens_per_word"]*n,
    rep_penalty=self.config["model"]["rep_penalty"],
)
```

Fig.4 - Instantiation, configuration and use of the new LLM handler system

Loading models can require an enormous amount of RAM, and given that a WEM based experiment involves repeated use of the model, emphasis has also been placed on controlling memory release. They are therefore methods that use the python gc (“garbage collector”) library and take care to force free all variables after use. For good measure, the same strategy has been applied to all the other classes in the application.

Implementation of new features

All experiment, model and workspace parameters are now centralized in a JSON configuration file. Launching an experiment is as simple as modifying this file and running the corresponding executable for the user's OS. In addition, precise event management, using assertions and warnings, ensures easy debugging in the case of a problem. For example, classes such as Simulation or Judge inherit from WemActor and are therefore forced to implement a verification method for parameters required in the config (cf. annexe 2).

The application currently has two modules: ‘simulation’ and ‘makegraph’, respectively responsible for running an experiment and generating visuals. The interfaces between these two modules are the result data files (cf. Available data and visuals); one generates them, the other uses them. These two entities form a complementary, communicating system.

Beyond revising the code architecture, given that an experience takes a long time to complete, I wanted to be able to see in real time what was going on inside it. One of the new features is therefore the ability to choose to log experience activity. A log file describes phases and gives details on the state of variables and operations in progress: call to LLM, current emergence rate, warnings, resolving configuration, etc… . To continue on real time feedback dispositives, the application also implements graphs plotting current data and actualized it at each step.

Improved parameter customization now allows the user to test more abstract parameters, such as prompts, criteria (different than “stronger”, e.g. “the most talkative”), the size of the words that an agent can carry (a sufficiently large size may allow the introduction of phrases rather than single words), etc … . In fine, this facilitates exploration of the model's possibilities.

Data science and visualization tools

This part is dedicated to my experience with data science and the development of adequate visualization tools during this internship.

Available data and visuals

During the experiment, data is logged from generation to generation. There are currently four types of output data:

- The history of the state of the N agents at each generation, in the form of a CSV file with header [gen, id, x, y, word].
- The history of judgments rendered by the LLM, in the form of a JSON file associating two-word tuples with the winning word.
- The history of competitions between agents, in the form of a JSON file associating tuples of two agent ids with the winner’s one.
- The history of mutations, in the form of a JSON file, which associates a source word with its mutation(s), itself associated with the list of possibilities from which it was selected.

These results files are read by several classes such as ExpTrajectoryAnalyzer or ExpEmergenceAnalyzer, which can be used to create specific visuals: static graphs, animations, visual boards; everything to facilitate analysis of the results from as many perspectives as possible. Calls to methods to create visuals have of course multiple parameters: select only certain trials, output format, plot interval of certain data, return of the created visual or use of a given support to draw on, etc … .

As the project progressed, we constantly questioned the relevance of our perspective on analyzing the results. It was also an opportunity for me to further explore the possibilities offered by the matplotlib and pyplot libraries.
At this time, the available visuals are:

- The semantic derivation trajectory graph and animation,
- The convex hull and percentiles contour areas of the plotted vectors,
- The density of the plotted vectors within the semantic space,
- The top B frequency curves and histogram over generations,
- The position of agents on the grid at each generation,
- The emergence scores evolution over generations.

Whether these ones already existed in the initial code, I have been working and understanding all of them since I built the classes integrated to the application.

Development of new metrics

During research and development, especially after the discovery of the language bias (cf. Language bias), we realized that we needed to develop new ways of measuring the scale of results, by being more precise about the specific influences of different dynamics.

The “emergence score” is the ratio between the number of mutations and the number of emergencies in a given generation. In other words, and in the context of emergence maximization, it's the performance of that generation. This score can be local, or averaged over the past generations. High peaks have been correlated with entrance into new semantic spaces, and low ones with the return to a previously discovered area, or a stagnation. By staying in the same semantic field for too long, the novelty possibilities are logically running out, as this field does not contain an infinity of words. This metric therefore helps to correlate the

Results are now also studied through spatial analysis. Using the UMAP graph, we want to measure trends in the distribution of points. For example we used the KDE algorithm to reveal density and visualize the exploitation of the semantic space, convex hull for the exploration, percentiles contour areas for the dispersion. In the context of our research on novelty emergence, such measures of space are highly relevant as they give us direct information on LLM tendencies to remain in a defined generation pattern or to innovate.

III. Sustainable Development and Social Responsibility

Quality of life at work

One of the things I really appreciated about this internship was to work in an environment that encouraged autonomy. There were no specific timetables and no obligation to be present in the laboratory, apart from weekly meetings. Expectations are simply focused on research results. This trust in the responsibility of the members is very pleasant, and I found that being able to fully manage my time improved my workflow. But that doesn't mean I was on my own when it comes to solving your problems. The laboratory is made up of very competent members. Even when I wasn't sure of the relevance of an idea or a lead, Pr. Suzuki encouraged me to innovate and fully extend my reasoning: in research, first you produce results and then you judge.
Finally, the atmosphere in the lab is very friendly. It's easy to talk to other people, and convivial moments are organized from time to time. Even though my Japanese was very weak, everybody tried their best to communicate and I have fully integrated myself into the laboratory’s life.

Travel impact

As for the plane, it would have taken 20,000km round-trip. At 0.15kg of CO2 per km, the total individual carbon footprint amounts to 3 tonnes of CO2.
Weekly, I came to work on site an average of 5 days, using the subway. My residence was 5km from the university, for a total of 10km per day. This type of transport needs about 0.1 kWh/km, this amounts to 1 kWh for one day. Considering that the average CO2 emission rate is 0.5kg for 1kWh, the carbon footprint for this trip is 2.5km of CO2 per week. Of course, this figure can be weighted a little, given that public transport carries many people. If we consider that at peak times there were around 500 passengers parked on the train, the individual impact can be reduced to 0.005kg of CO2 per week.
Similarly, on my business trip to the Osaka conference, I had to take the Shinkansen. For this 340km round trip, at 0.015kg of CO2 per km, my carbon footprint was around 5.1kg. I can therefore estimate the total travel impact for this internship at 3,005.145kg.

Impact of professional activities

Experiments take time to be run. Consequently, my desktop computer was on all day and night. This machine was equipped with an Nvidia RTX 4070 graphics card (GPU) and an 11th generation Intel i7 processor (CPU), the two main components providing the computing power required for the experiments.

It can be estimated that over the course of the internship, excluding days spent analyzing results and other side tasks, the computer was running experiments at a rate of 72 hours per week, with the GPU working at 100%.
On experiment hours, this model of GPU therefore consumes about 200w and the CPU about 100w. Adding a margin of 100w for the consumption of other components, we arrive at 400w/h. On regular hours, the consumption of the GPU and CPU respectively drops to 30w and 65w, for a total of 195w/h.
Considering the 72h of intensive computation and the remaining 96h, the average total electricity consumption per week is 47,52 kWh, and 427,68 kWh during the internship.

Conclusion

This internship at the ALIFE-CORE laboratory would have been a wonderful experience, professionally and personally. I had the opportunity to discover a fascinating research field, actively participate in the research of an actual promising artificial life model, and therefore experience the researcher profession. Finally, I assisted my supervisor to present it at a national scale conference.
I was able to lead a case study on the highlight of a language bias and demonstrates its usefulness for the open-endedness research through WEM.
The development of a specialized application to match the increasing needs of flexibility and new perspectives of analysis also allowed me to enhance my development skills for LLMs integration and data visualization.
To finish on a questioning note, I think there are still many case studies to be carried out on WEM, such as integrating more abstraction, and of course developing adequate visuals to measure the emergence of more complex dimensions. For instance, the development of a deeper trajectory graph using a hallucination-detection AI model over the data, was a functionality initiative that was not kept due to lack of time.
And as for the application, I personally think that the interfacing and modularization could have been better.
For sure this internship was a crucial step in my professional journey, and has opened up new career horizons.

Webography

- *R. Suzuki T. Arita, Evolutionary ecology of words (2025), arXiv:2505.05863.*
- *ALIFE-CORE, artificial life laboratory at the Graduate School of Informatics of Nagoya.*
- Tim Taylor, Evolutionary Innovations and Where to Find Them (2019), *arXiv:1806.01883v4*
- Xuhui Jiang et al., A Survey on LLM-as-a-Judge (2024), arXiv:2411.1554
- *Mistral's LLM AI was used, with proofreading, to calculate energy consumptions and carbon footprints (cf. SDSR), and search for orthographic and grammatical faults.*

Glossary

Research

- **Artificial Life** - multidisciplinary field that explores the nature of life by creating and studying systems that mimic or simulate life processes, often using computer models, robotics, and biochemistry.
- **WEM** - abbreviation used to designate the Evolutionary Ecology of Words model.
- **UMAP** - Uniform Manifold Approximation and Projection is a dimension reduction technique that can be used for visualisation; in the context of a WEM based experiment it is used to create a normalized space to plot a set of vectors, representing words.
- **Semantic vector** - the embedding process of a word into a vector uses the relationship between the meanings of words as a reference to calculate their vector. These vectors are called “semantic” because they represent the meaning of the words in the current system.
- **Sentence Transformers** - this Python library grants access to state-of-the-art embedding models, to convert words into vectors.
- **Emergence** - in the context of WEM, the term emergence designates every novelty appearing in the experiment, from a certain state at a certain step. It is mostly new words obtained from the mutation process, but it can also be more complex phenomena such as dynamics or uncommon structures of agents.
- **KDE** - Kernel Density Estimation, algorithm to calculate the density of a set of points within a defined grid.

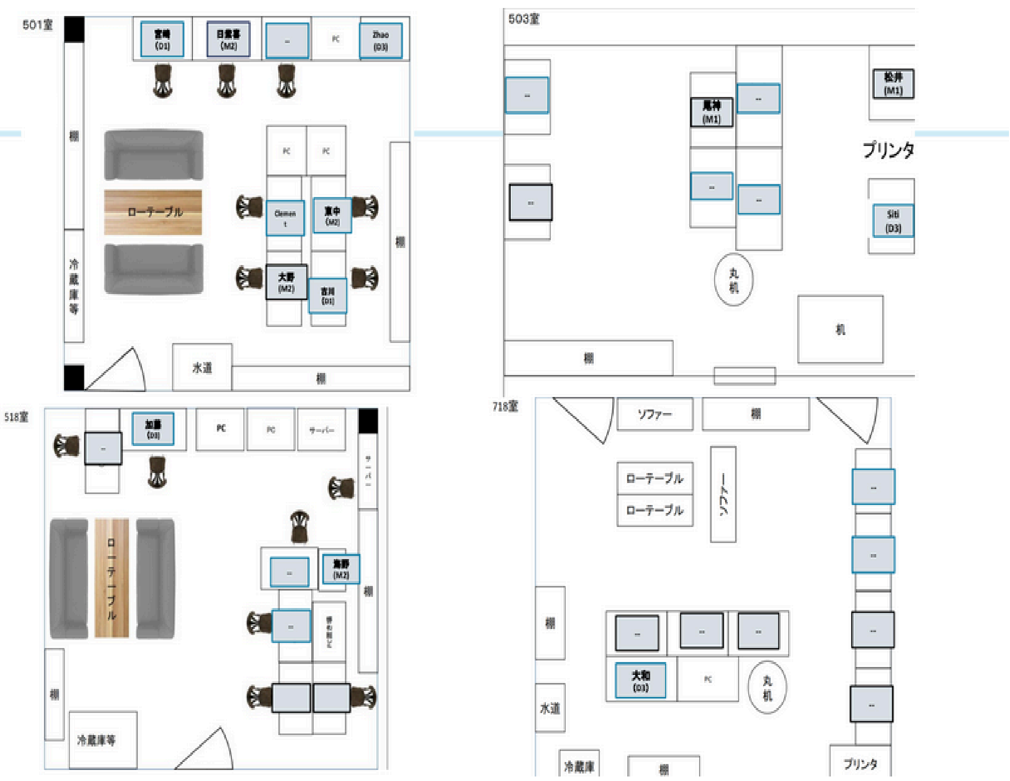
Development

- **LLM** - Large Language Model, ChatGPT like AI.
- **OOP** - Object Oriented Programming.

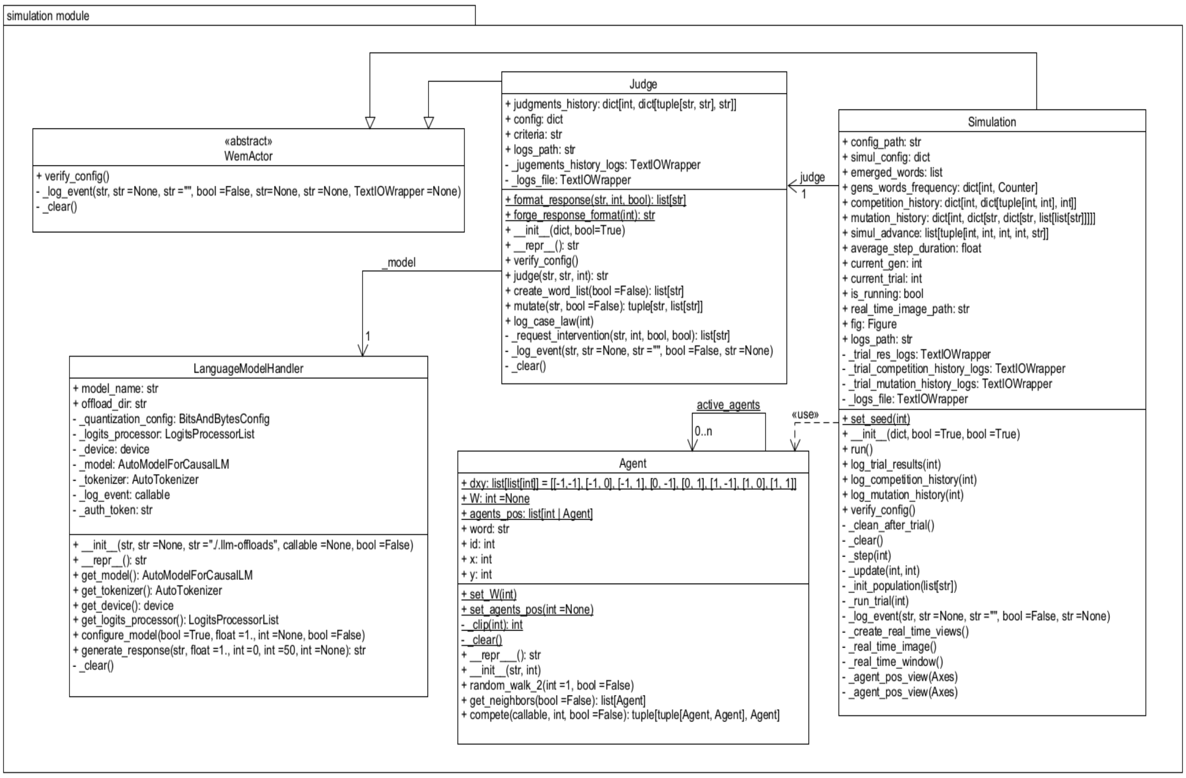
- **Local offload** - use the machine's disk to store information calculated by the LLM, and thus put less pressure on GPU memory.
- **Tokenizer** - module allowing to convert a text input into a sequence of **tokens**, in a certain format, which is prepared to be integrated to the LLM’s reflection.
- **Tensor** - multidimensional data that represent tokenized inputs within LLMs.
- **Weight quantization** - LLM weights reduction without loss of performance, allowing large models to run on smaller machines.
- **GPU** - Graphic Processor Unit, computer’s graphic card.
- **CPU** - Central Processor Unit, computer’s processor.
- **Prompt** - text input submitted to the LLM.
- **Wrapper** - layer of code that encapsulates and modifies the functionality of another piece of code, providing different interfaces for instance.
- **RAM** - Random Access Memory, computer’s memory component into which is loaded programs running on the machine, including LLMs.
- **Hugging Face** - one of the most known AI models hub.

Annexes

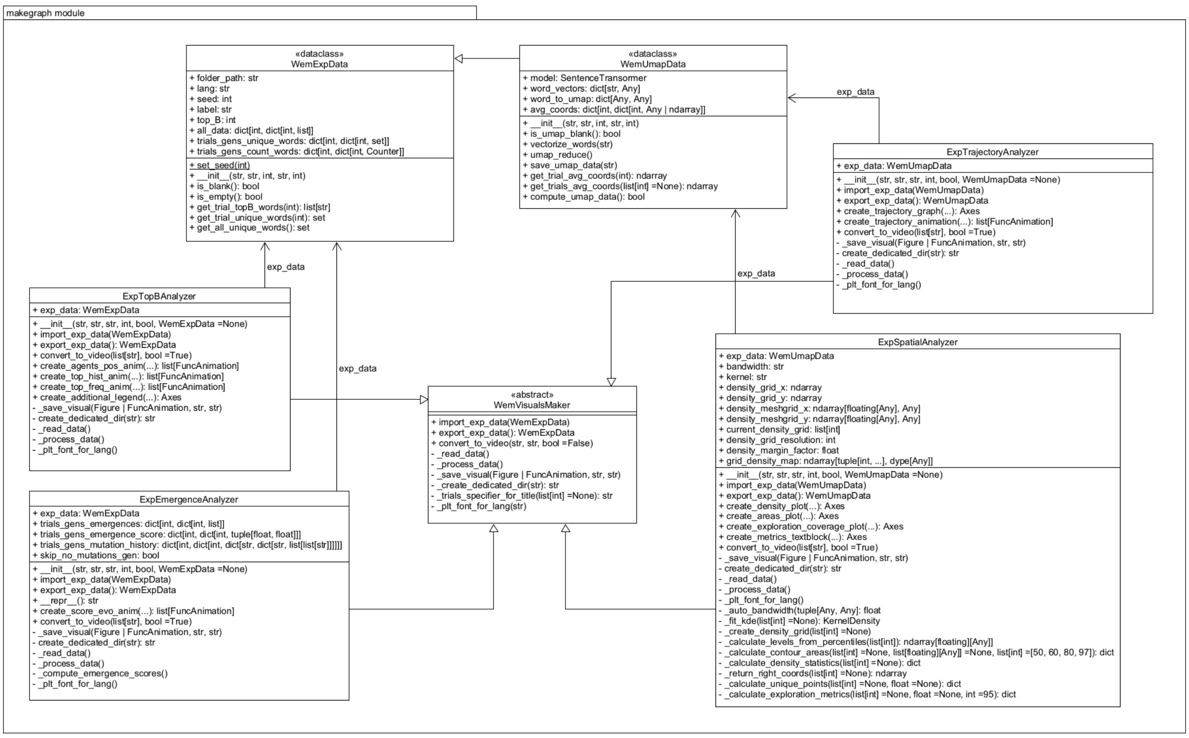
Annexe 1 - working desks disposition at the laboratory



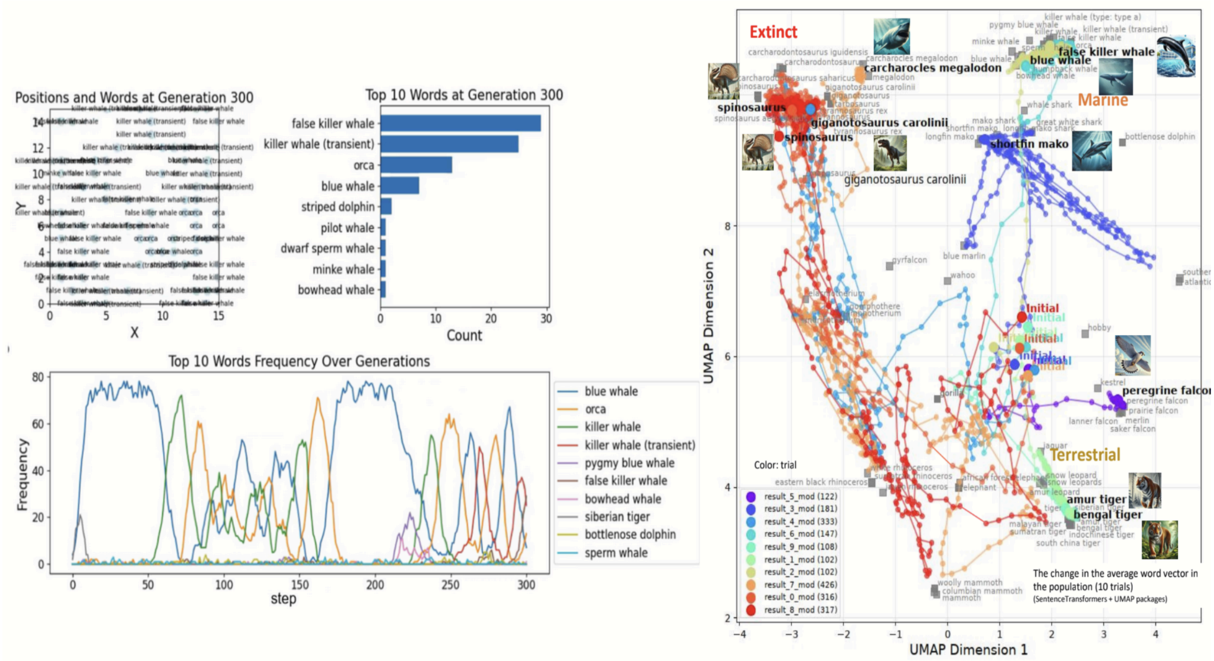
Annexe 2 - UML diagram of the simulation module



Annexe 3 - UML diagram of the makegraph module



Annexe 4 - original visuals created with experiment result data

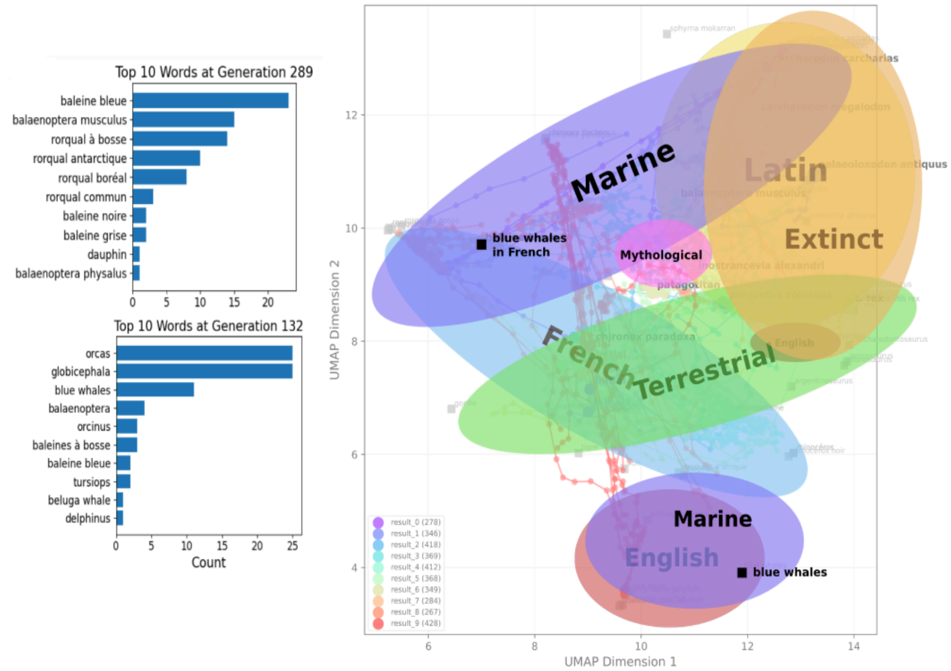


Annexe 5 - semantic gap between two translation of ‘blue whales’ in the French experiment

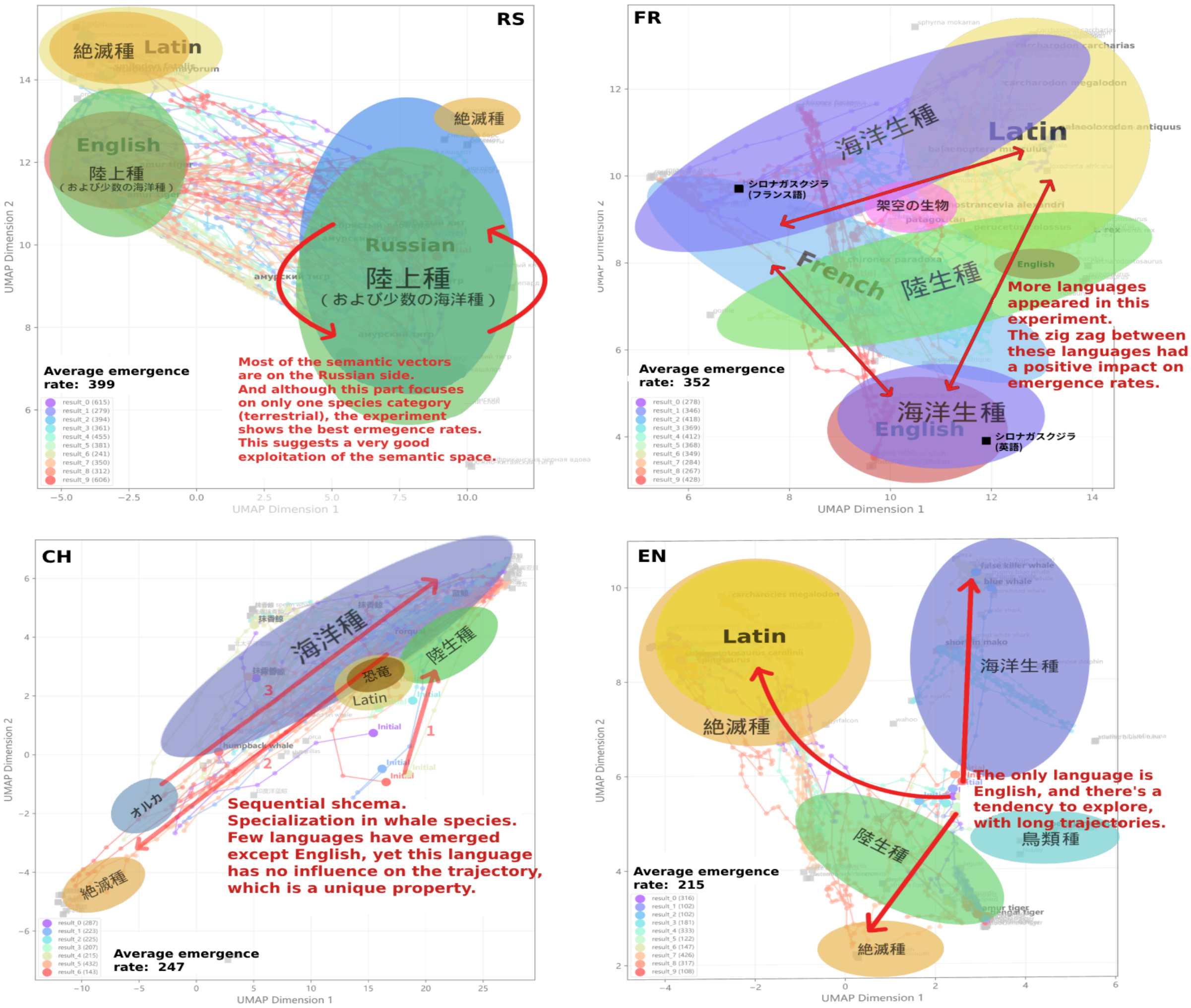
The two top 10s are not that different.

The names are simply translated into French or English.

However, the graph seems to show a semantic divide large enough to separate French and English marine species



Annexe 6 - visualization of specific characteristics of languages, presented at the JSAI2025 conference, partially translated in english



Annexe 7 - food chain dynamics between top B words