Multi-Agent LLMs for science: case study of chemical applications

Nikolay Nikitin, PhD,
Associate Prof. of AI Technologies Faculty,
Head of Automated Machine Learning Lab,
ITMO University, Saint-Petersburg, Russia

MathAI, Sochi, Russia – 24.03.2025

About me

- Nikolay Nikitin, PhD, ITMO University.
- Head of Automated Machine Learning Lab, Researcher in NSS Lab, Lead of Open Source Developments.
- Research interests: AI4Science, AutoML, LLM, optimisation, scientific open source;
- Chair or AI4Science workshop at ICML conference;
- Winner of Yandex ML Prize adward (2024);
- GitHub: https://github.com/nicl-nno
- Contacts: https://t.me/nicl_nno, nicl.nno@gmail.com

ITMO AI4Science Open-Source Ecosystem

Our most popular tool is **FEDOT** - a open source data-driven tool for creating of machine learning pipelines (AutoML). 600+ "stars" in GitHub;

And much more interconnected tools:

- FEDOT.LLM, ProtoLLM – LLM-related tools
- GOLEM, GEFEST – design of physical structures;
- BAMT - search of casual relations using Bayesian networks;
- EPDE, TEDEOUS – discovery and solving of diff. equations;
- GEMCODE – AI for drug discovery of chemical co-crystals;
- SAMPO - a library for industrial process planning;
- AutoTM - NLP automation;
- **And others – more that 30 products**

URL: https://github.com/aimclub

Community (1000+ users):

ITMO OPEN SOURCE

https://t.me/scientific_opensource

Case study for automation in science: chemistry

AI + Chemistry

ИТМО

NSS Lab

Natural Systems
Simulation lab

+

Center for
AI in
Chemistry

Our papers for AI4Science in chemistry

ИТМО

- Gubina, N., Dmitrenko, A., Solovev G., Yamshchikova, L., Petrov O., Lebedev, I., Serov, N., Kirgizov, G., Nikitin N., Vinogradov, V. //. **Hybrid Generative AI for De Novo Design of Co-Crystals with Enhanced Tabletability**. 38th Conference on Neural Information Processing Systems (**NeurIPS 2024**).
- Solovev G. V. et al. Towards LLM-Driven **Multi-Agent Pipeline for Drug Discovery: Neurodegenerative Diseases Case Study** //2nd AI4Research Workshop: Towards a Knowledge-grounded Scientific Research Lifecycle, **AAAI 2025**.
- Gubina N. et al. **Generative AI for Co-Crystal Design with Property Control**. // AI2ASE workshop, AAAI 2024

Case Study for Chemistry:

GEMCODE – AI for co-crystals design

https://github.com/ai-chem/GEMCODE

[Our NeurIPS 2024 Paper]

GEMCODE – AI for co-crystals design

ITMO

Co-crystals play an important role in many industries, such as energy, electronics, optoelectronics, food, and **pharma**, especially

Tabletability of therapeutic agents can be achieved by co-crystallization

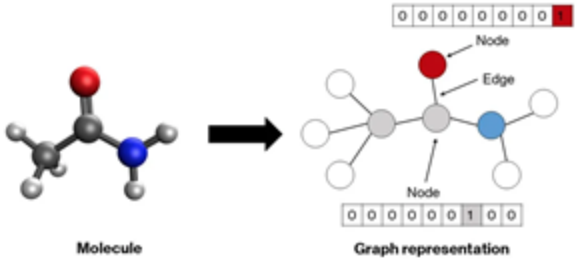
Tabletability is defined by a set of **mechanical properties**, such as plasticity

There exists no open platform for fast *in silico* screening of co-crystals with target tabletability profiles

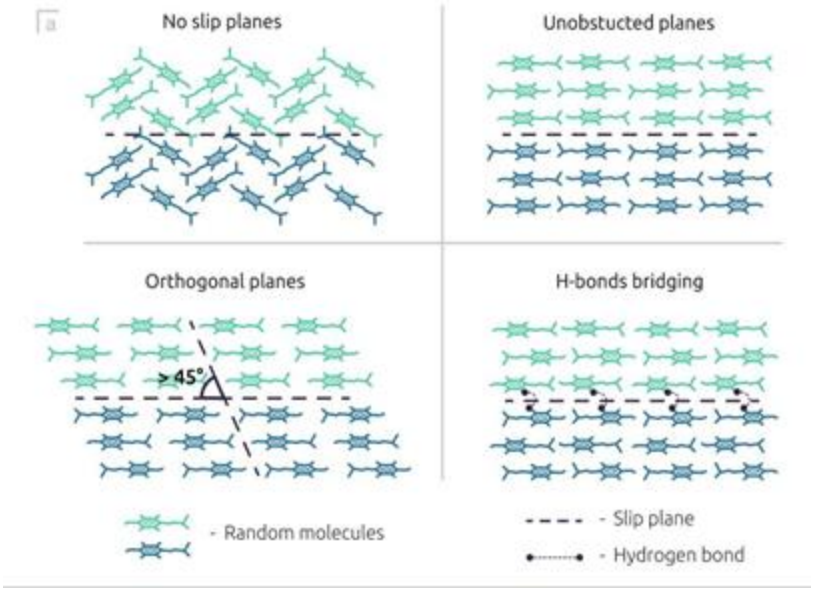
Target properties for co-crystals

Target properties:

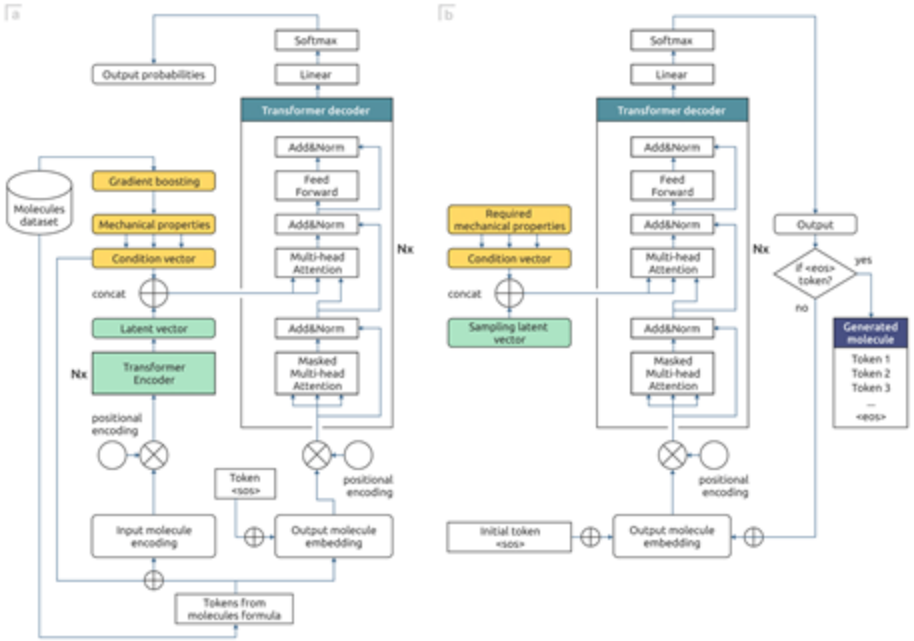
- "Orthogonal planes"
- “H-bonds bridging”
- "Unobstructed planes."



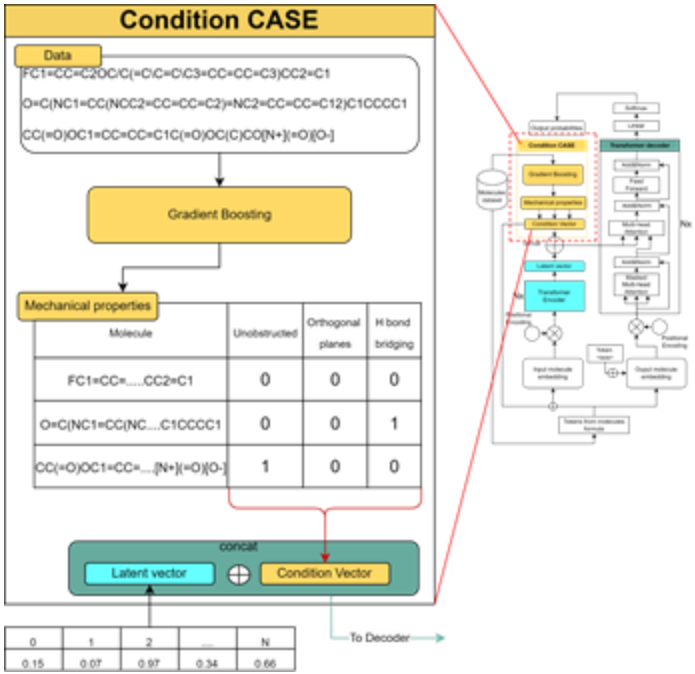
Representation of molecules



Generative models

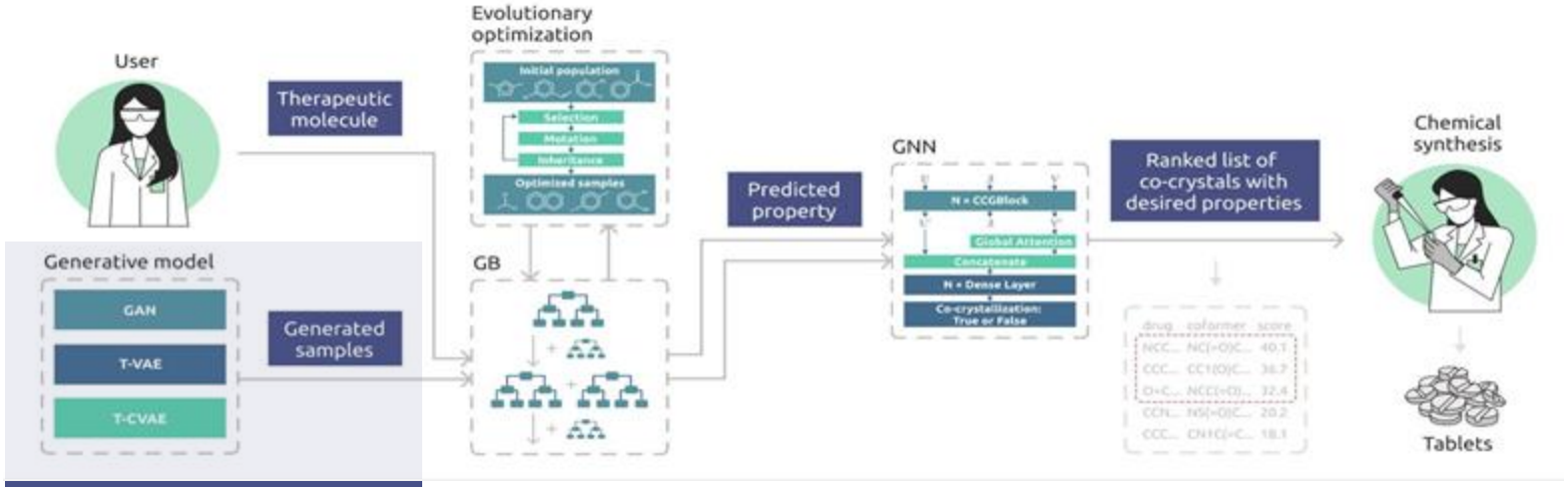


Architecture of the deep learning model during training (A) and during the generation process (B)



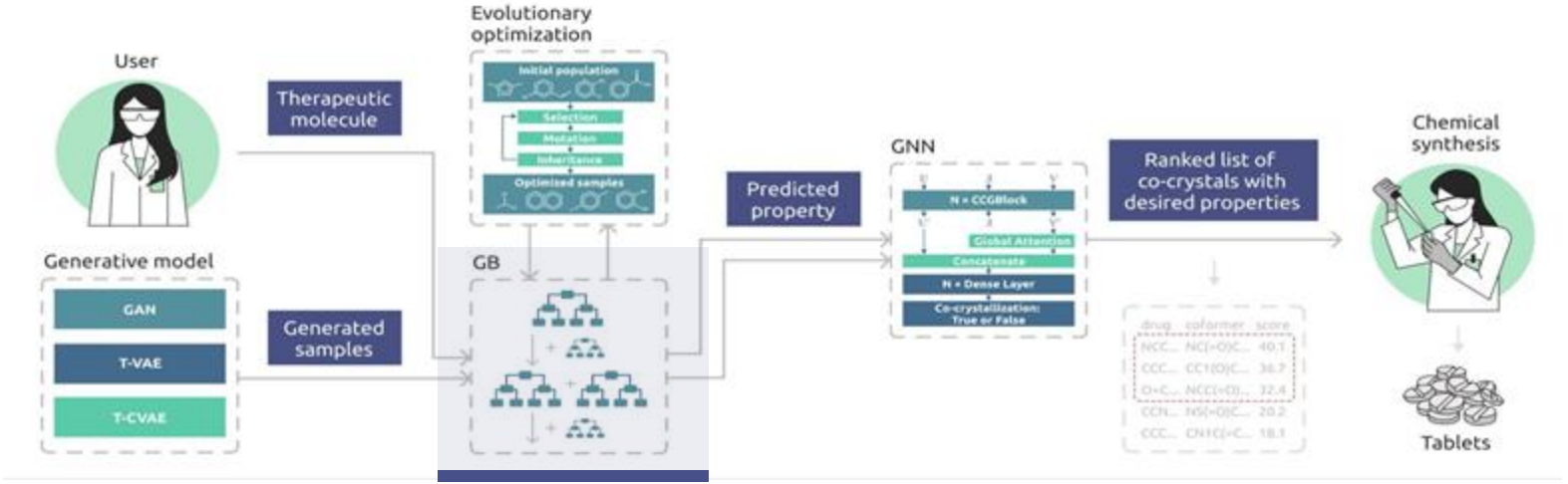
The principle of operation of the conditional block architecture that takes into account the mechanical properties of generated objects.

Hybrid AI pipeline



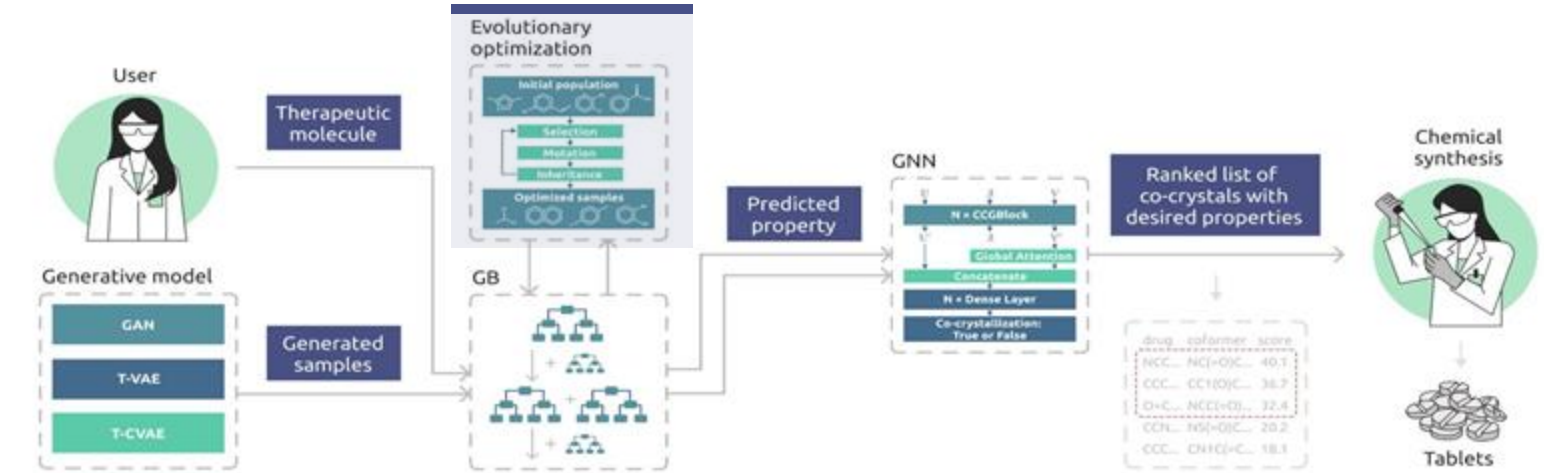
Coformer generation with generative models ▶ Prediction of mechanical properties with classical ensemble learning ▶ Coformer optimization with graph-based evolutionary algorithm ▶ Estimation of co-crystallization probability with a pretrained graph neural network

Hybrid AI pipeline



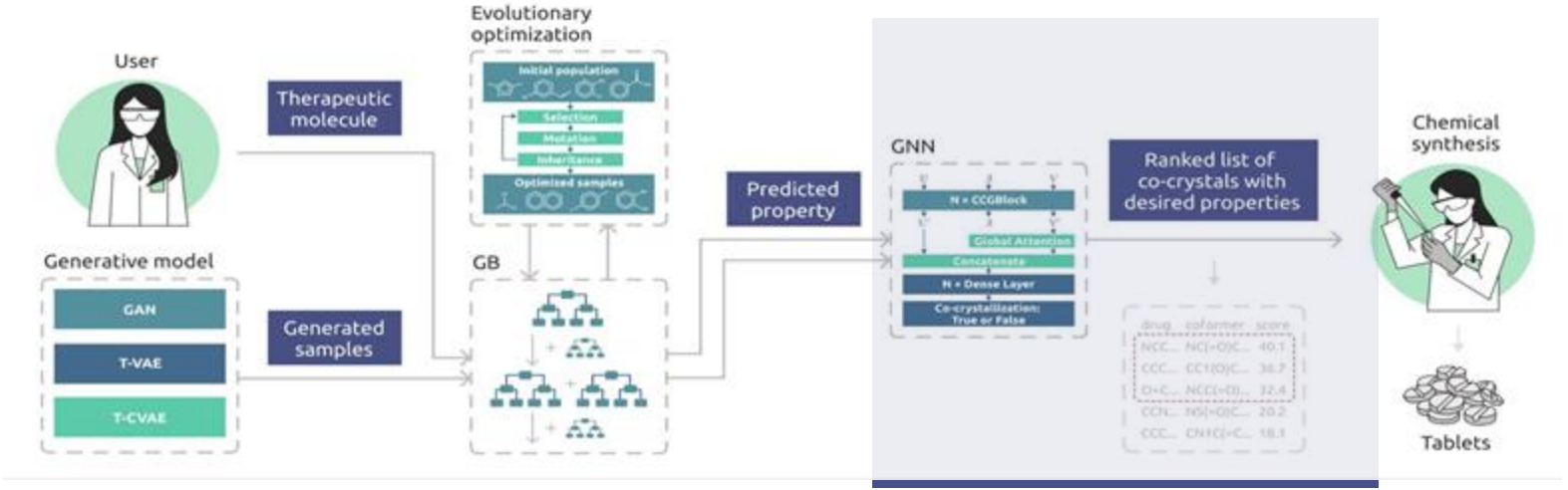
Coformer generation with LSTM-based GAN ▶ **Prediction of mechanical properties with classical ensemble learning** ▶ Coformer optimization with graph-based evolutionary algorithm ▶ Estimation of co-crystallization probability with a pretrained graph neural network

Hybrid AI pipeline



Coformer generation with LSTM-based GAN ▶ Prediction of mechanical properties with classical ensemble learning ▶ **Coformer optimization with graph-based evolutionary algorithm** ▶ Estimation of co-crystallization probability with a pretrained graph neural network

Hybrid AI pipeline



Coformer generation with LSTM-based GAN ▶ Prediction of mechanical properties with classical ensemble learning ▶ Coformer optimization with graph-based evolutionary algorithm ▶ **Estimation of co-crystallization probability with a pretrained graph neural network**

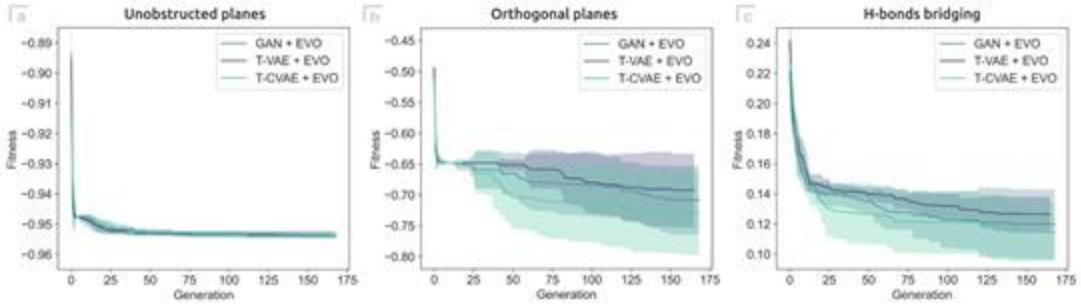
Hybrid AI pipeline

- We selected 1.75M samples from the **ChEMBL database** based on the relevant parameter distributions of the known coformers;
- We retrieved mechanical properties for 6k coformers from the **Cambridge Structural Database (CSD);**
- We used SMILES representations to extract molecular features with RDKit;
- We performed feature engineering and filtering as preprocessing steps;

Experimental results

Comparison of T-CVAEs in terms of the number of target molecules generated.

Model	GAN	T-VAE	T-CVAE
Validity, %	94.57 ± 0.00	$\mathbf{99.70 \pm 0.00}$	98.40 ± 0.00
Novelty, %	94.90 ± 0.08	$\mathbf{95.12 \pm 0.11}$	80.62 ± 0.25
Duplicates, %	42.29 ± 0.69	$\mathbf{24.30 \pm 0.45}$	55.70 ± 0.19
Target coformers, %	3.66 ± 0.17	2.88 ± 0.12	$\mathbf{6.52 \pm 0.22}$
Diversity of target	$\mathbf{0.93 \pm 0.00}$	0.92 ± 0.00	0.91 ± 0.00



Increasing the number of generated target molecules by a factor of x2 for the co-crystal search task

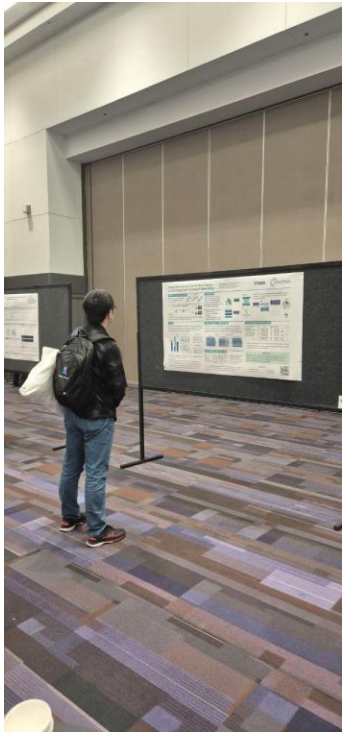
Experimental results

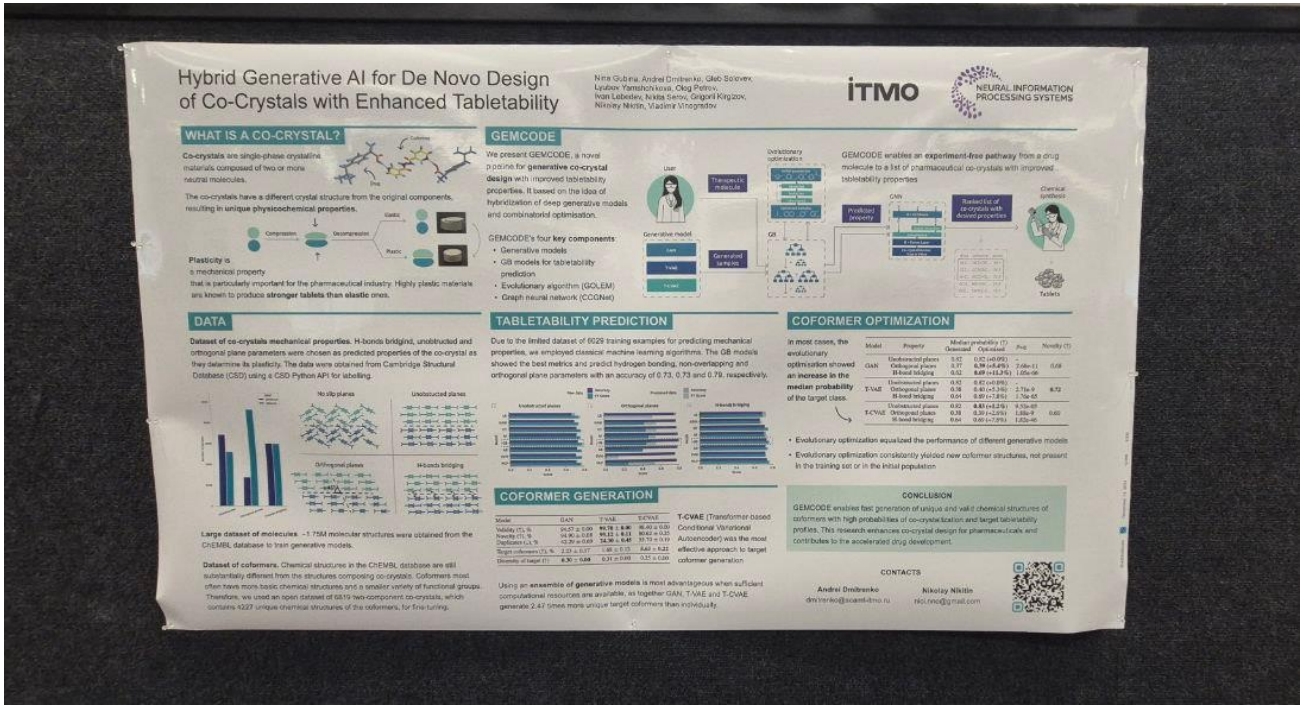
We demonstrated utility of the pipeline in the **Theophylline case** study and discussed its current limitations.

Experimentally validated coformers improving drug tabletability generated by GEMCODE:

Drug	Generated SMILES	CSD Refcode	Model
Nicorandil	O=C(O)C=CC(=O)O	WAHGEV	GAN / T-VAE / T-CVAE
Rivaroxaban	O=C([O-])CC(=O)[O-]	YORVEJ	T-VAE
Paracetamol	C1=CC=C2C=CC=CC2=C1 C[N+](C)(C)CC(=O)[O-]	LUJSIT CUQKAC	GAN / T-VAE / T-CVAE T-CVAE

Looks fine, but…





NeurIPS 2024 main-track poster

What is wrong?

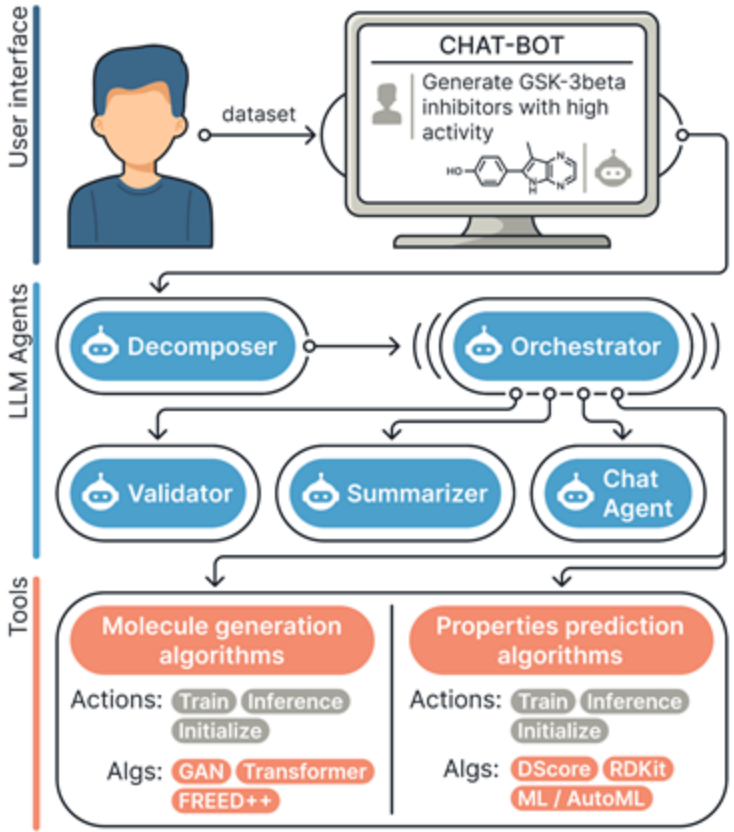
- A lot of work spent to **design the pipeline manually** for new task;

- **Chemical and programming knowledge** is required to formulate a task;

- **Hard to configure** a generative and predictive models to achieve good performance;

LLM come into play

LLM Agent for pipeline design



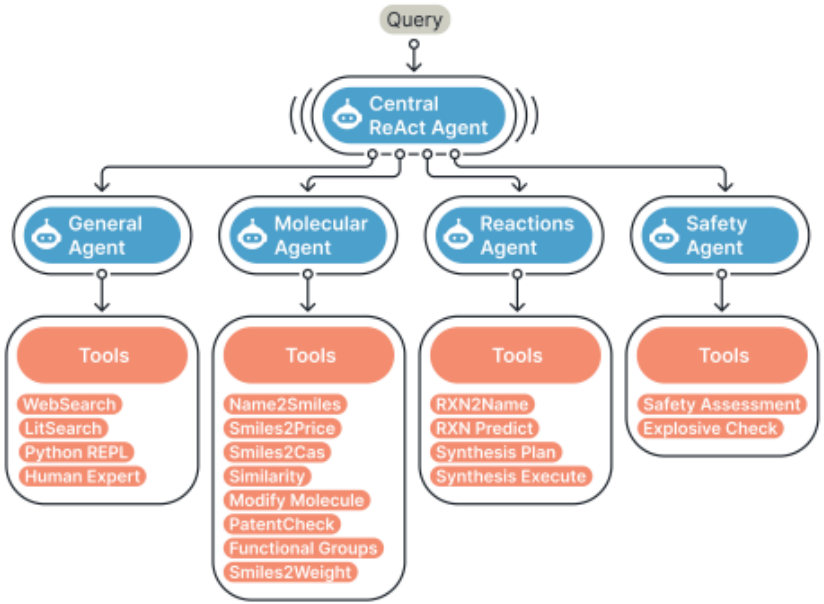
An algorithm has been developed for interaction of LLM agents with generative and predictive tools.

Functionality:

- Optimization of interaction with highly specialized generative methods via chat.
- Intelligent selection of a suitable generative method for a user-specified task, without human intervention.
- Automation of training models for calculating and predicting the properties of generated objects.
- Automatic training of generative deep learning models for new tasks.

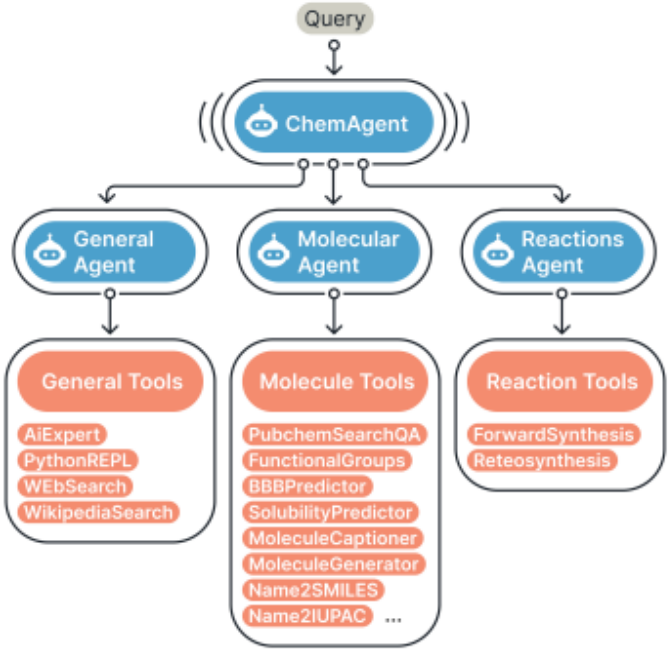
Various multi-agent setups

MolFormer

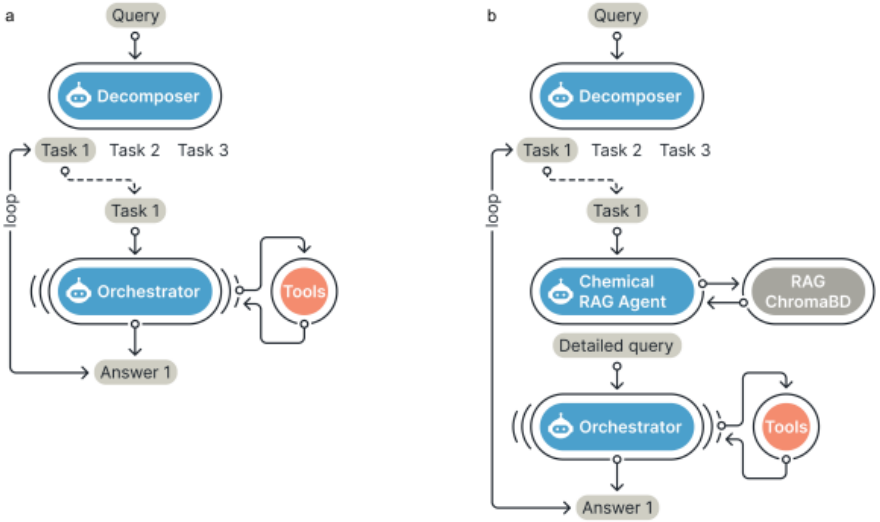


ChemCrow

Various multi-agent setups

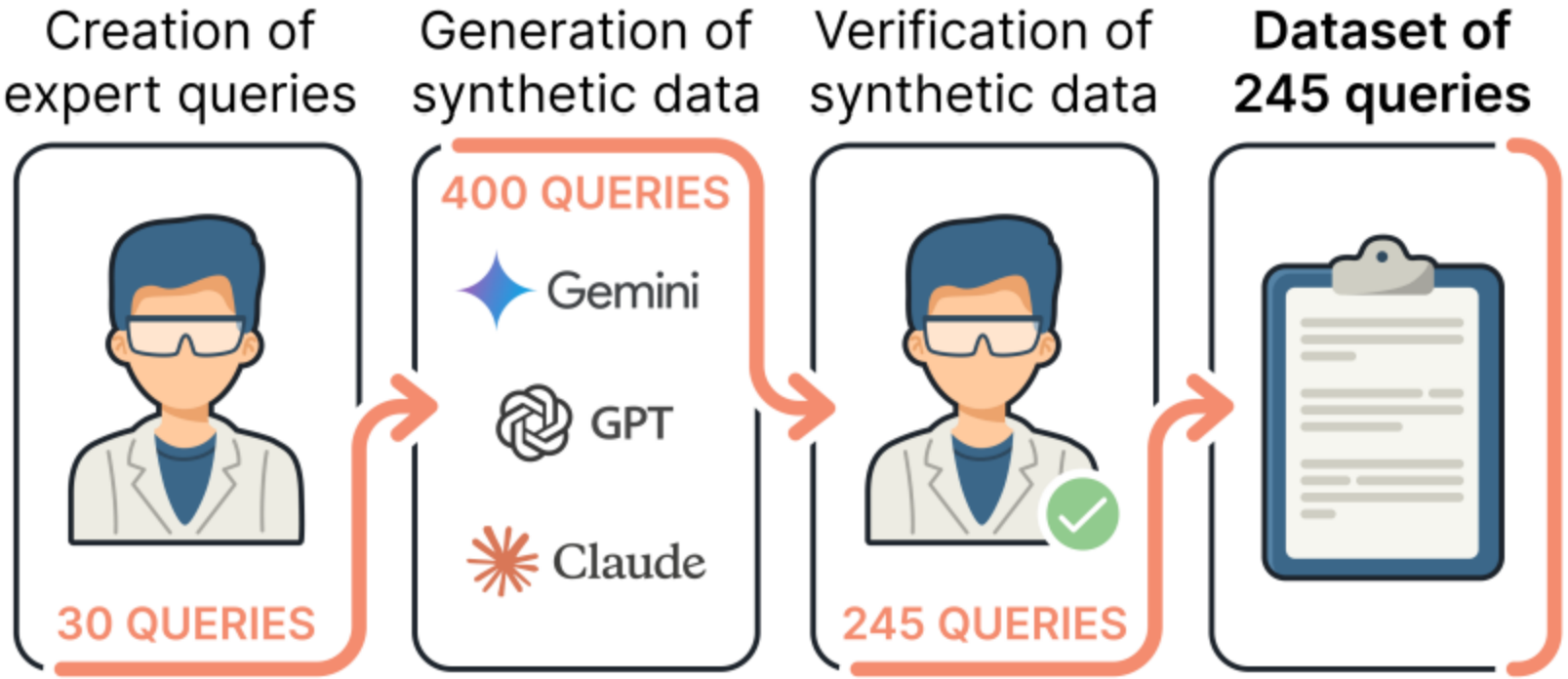


ChemAgent



Different variants of our approach

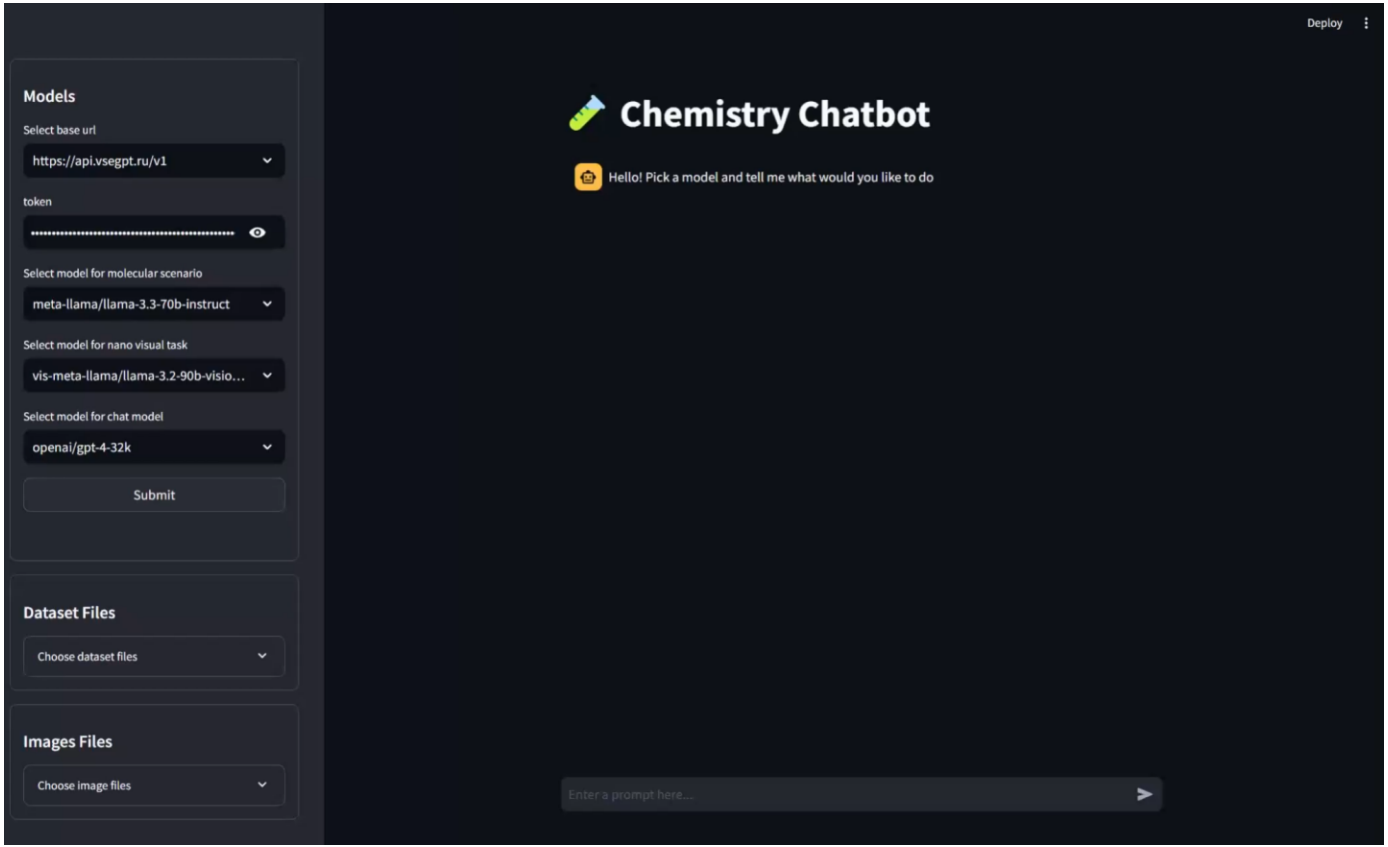
Benchmark for comparison



Benchmarking

Case	Model	Novelty %	Validity %	Mean DS	GR1, %	GR2, %	GR3, %	GR4, %	GR5, %
Alzheimer	MADD	78.21	87.47	**-7.46**	20.30	17.56	**13.72**	**13.40**	**13.40**
	Llasmol	-	64.00	-5.36	4.54	4.54	4.54	0.00	0.00
	X-LoRA Gemma	-	43.60	-4.14	0.00	0.00	0.00	0.00	0.00
	ChemAgent	-	23.50	-6.42	2.50	0.00	0.00	0.00	0.00
	ChemDFM	-	**99.84**	-6.80	**28.57**	**17.86**	10.71	10.71	10.71
Sclerosis	MADD	73.45	84.32	**-9.24**	**14.71**	**12.81**	**11.73**	**11.62**	**11.62**
	Llasmol	-	58.00	-6.08	0.00	0.00	0.00	0.00	0.00
	X-LoRA Gemma	-	43.60	-5.81	5.12	2.56	2.56	2.56	2.56
	ChemAgent	-	29.20	-6.50	1.04	0.52	0.52	0.52	0.52
	ChemDFM	-	**85.70**	-8.13	11.11	5.56	5.56	5.56	0.00
Parkinson	MADD	61.21	78.21	**-6.04**	**5.93**	**4.86**	**3.87**	**3.74**	**3.74**
	Llasmol	-	68.00	-6.18	0.00	0.00	0.00	0.00	0.00
	X-LoRA Gemma	-	45.50	-5.01	0.00	0.00	0.00	0.00	0.00
	ChemAgent	-	27.70	-3.41	0.00	0.00	0.00	0.00	0.00
	ChemDFM	-	**94.60**	-6.03	2.50	2.50	0.00	0.00	0.00

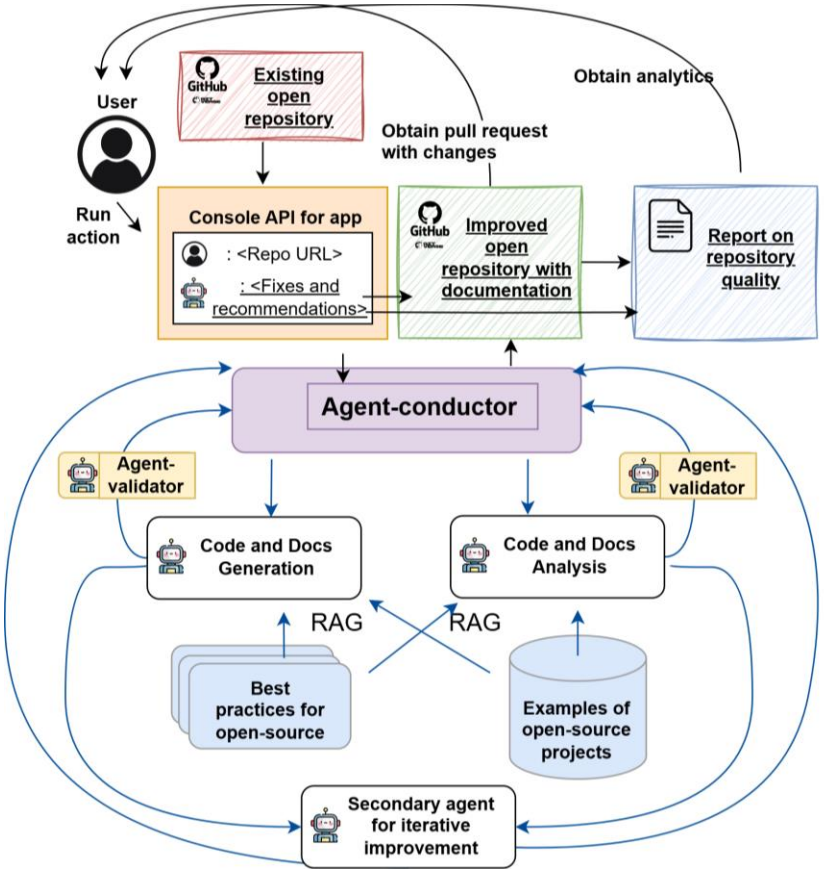
Interactive demo



ИТМО

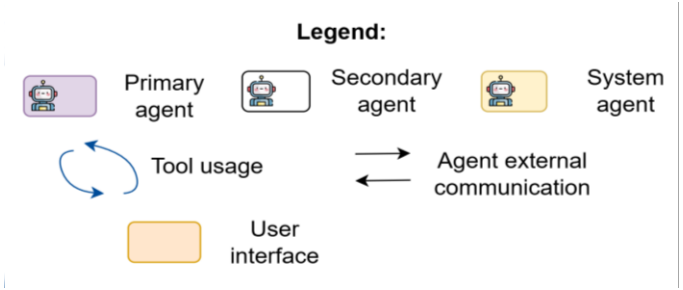
But what about open source?

AI Tool for Open Source Scientic Software Improvement



OSA (Open Source Advisor) - LLM-agent based tool for automated improvement of open-source repositories.

https://github.com/ITMO-NSS-team/Open-Source-Advisor



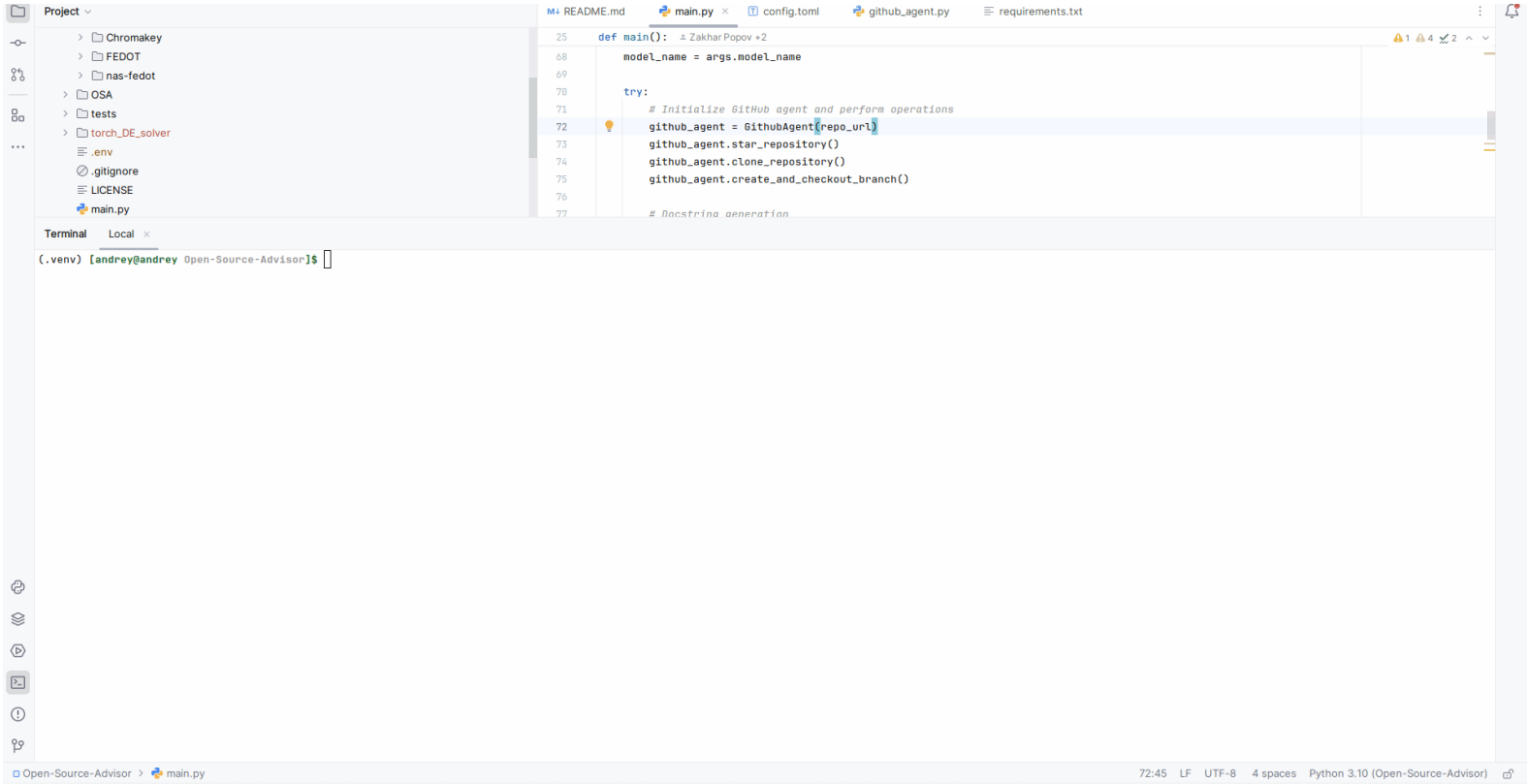
Technical detais

Implementation of AI part:

- For code analysis and generation - open-source LLM (LLama 3.2/CodeLlama/CodeQwen - 7 to 90B) in in-context mode + external level models
- GPT4o-mini (for complex tasks).
- To improve quality - multiagent collaboration between multiple LLMs.
- To account for existing knowledge - RAG on open-source best practice descriptions.

Computational resources: LLM on ITMO servers (or local) + external APIs (for quality improvement)

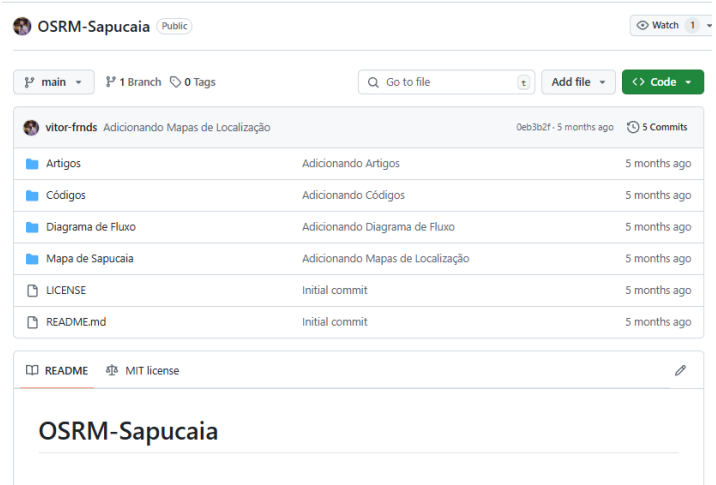
How it works?


Project
Chromakey
FEDOT
nas-fedot
OSA
tests
torch_DE_solver
.env
.gitignore
LICENSE
main.py
README.md
main.py
config.toml
github_agent.py
requirements.txt
def main(): Zakhar Popov +2
model_name = args.model_name
try:
# Initialize GitHub agent and perform operations
github_agent = GithubAgent(repo_url)
github_agent.star_repository()
github_agent.clone_repository()
github_agent.create_and_checkout_branch()
# Docstring generation
Terminal
Local
(.venv) [andrey@andrey Open-Source-Advisor]$
Open-Source-Advisor
main.py
72:45
LF
UTF-8
4 spaces
Python 3.10 (Open-Source-Advisor)

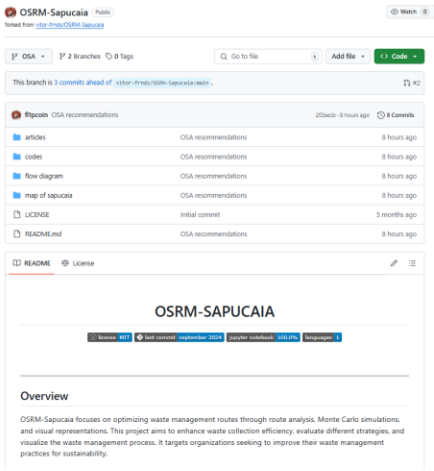
Examples

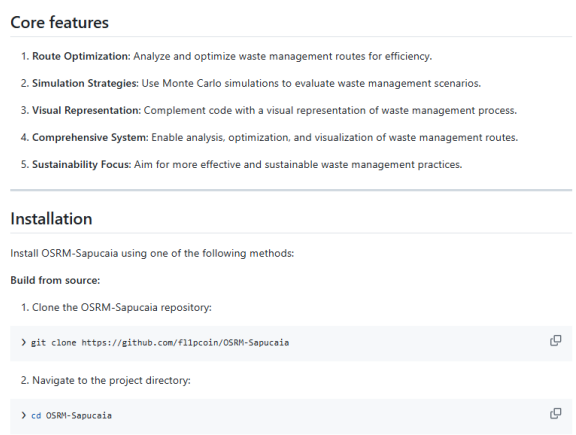
Repository of Dr. Leonardo Goliatt (Brazil):
https://github.com/fl1pcoin/OSRM-Sapucaia/tree/OSA

Before:

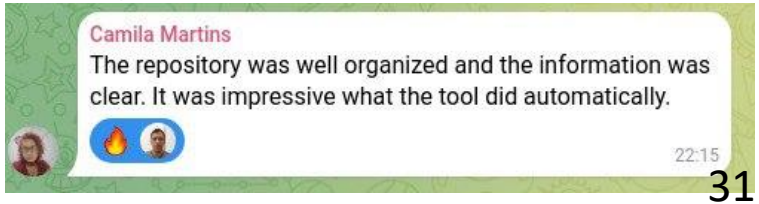


After:





Auto changes: README and docstring generated, translated to Portuguese to English.



Use OSA to improve your repos!

ИТМО

- **Repository** - create issues with suggestions;
- **Telegram chat** - ask questions about usage;
- **Github bot** - start and improve your repositories;
- **Stars** - rate and support the project.

https://github.com/ITMO-NSS-team/Open-Source-Advisor

Thanks!

ITMO more than a UNIVERSITY

nnikitin@itmo.ru