

Project Name	AI Agent Definition and Generation Framework
Online team meeting	<a href="https://fau.zoom-x.de/j/66967665829">https://fau.zoom-x.de/j/66967665829</a>
Production system (if any)	...
Test system (if any)	...
GitHub repository	<a href="https://github.com/amosproj/amos2024ss06-health-ai-framework">https://github.com/amosproj/amos2024ss06-health-ai-framework</a>
GitHub feature board	<a href="https://github.com/orgs/amosproj/projects/53">https://github.com/orgs/amosproj/projects/53</a>
GitHub impediments backlog	<a href="https://github.com/orgs/amosproj/projects/64">https://github.com/orgs/amosproj/projects/64</a>
Team T-shirt (white)	<a href="https://www.shirtinator.de/s/nDefN-9-QpugOXEh5oJIOA">https://www.shirtinator.de/s/nDefN-9-QpugOXEh5oJIOA</a>
Team T-shirt (black)	<a href="https://www.shirtinator.de/s/Xeu6K1B5RrSUf5lw05rzSg">https://www.shirtinator.de/s/Xeu6K1B5RrSUf5lw05rzSg</a>
Additional materials	...
Team mailing list	<a href="mailto:oss-amos-proj6@lists.fau.de">oss-amos-proj6@lists.fau.de</a>

Last Name	First Name	GitHub User Name	Email Address	Comment
Enescu	David	EnescuDavid	david.enescu@fau.de	
Karategos	Georgios	tubamos	amosproject@proton.me	used ONLY for GitHub and the Happy-AMOS tool. This is NOT the email used at the course registration form.
Negasa	Gemechis	Ghemechis	gemechismelkamu@gmail.com	
Potthoff	Jan	potthoffjan	potthoff.jan@googlemail.com	
Sandt	Eloi	eloinoel	eloi.sandt@campus.tu-berlin.de	
Varga	Lukáš	lukas-varga	lukas.varga128@gmail.com	
Zimmermann	Simon	Tims777	tim.simon.zimmermann@fau.de	
Gupta	Manik	manikg08	manik.gupta@fau.de	
Vadaliya	Preet	preetvadaliya	preet.vadaliya@fau.de	

#	Meeting Day	Product Owners	Software Developer	Release Manager	Scrum Master	Comment
1	2024-04-17	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	N/A	Simon	
2	2024-04-24	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	David	Simon	irr
3	2024-05-01	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Manik	Simon	irr
4	2024-05-08	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Preet	Simon	-
5	2024-05-15	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Jan	Simon	-
6	2024-05-22	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Eloi	Simon	irr
7	2024-05-29	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Lukáš	Simon	Mid-term due
8	2024-06-05	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	David	Simon	-
9	2024-06-12	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Manik	Simon	-
10	2023-01-11	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Preet	Simon	-
11	2023-01-18	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Jan	Simon	-
12	2023-01-25	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Eloi	Simon	irr
13	2023-02-01	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Lukáš	Simon	irr
14	2023-02-08	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Jan	Simon	Demo day!
15	2023-02-15	George, Gemechis	David, Jan, Eloi, Lukáš, Manik, Preet	Preet	Simon	Retrospective
Product owners, software developers, and Scrum Master are set and ideally don't change over time; the critical part is the Release Manager role you need to define here						

<b>Goals</b>	- learn and share learning on tech and collaboration tasks
	- keep Andi Zink happy
	- finish tasks for each sprint on time (or communicate early why it isn't possible or if you need help)
<b>Meeting norms</b>	- be on time
	- be informed on topics relevant for the meeting (eg. lecture)
	- if you are late or can't make it, give a heads up (eg. on slack)
	- active participation
	- let's not digress too much while key points haven't been gone through yet
<b>Working norms</b>	- ask for help
	- decisions made in mutual agreement
	- deliver constructive criticism with the goal of helping others and yourself grow
	- for code: settle on conventions
	- good comments/documentation of work
	- be considerate of varying sets of experiences and skills
<b>Coordination norms</b>	- the Scrum Master keeps the meeting on track and PO leads the meeting
	- the POs (at least, devs are always welcome to join) meet with the partner at regular intervals and communicate / clarify feature requests to the rest of the team.
<b>Communication norms</b>	- be welcoming and nice
	- don't interrupt others
	- communicate illnesses and expected work downtime and what it means for others
	- communication through slack student channel
<b>Consideration norms</b>	- if I disagree, I should think of an alternative/provide constructive reasoning
	- both sides should be heard and find a consensus together
	- prefer to speak in public over speaking privately
<b>Cont. improvement norms</b>	- performance evaluated based on weekly sprint target compared with git pull requests
	- automated tests?
<b>Rewards</b>	- after successful release, we can go to mensa (separately in Erlangen / Berlin)
	- consider meeting in person after the project is completed (in Berlin or Erlangen)
<b>Sanctions</b>	- Start with a warning. Persistent issues can lead to report.

<b>Signatures</b>	
Scrum Master	Simon Zimmermann
Product owner	Georgios Karategos
Product owner	Gemechis Negasa
Software developer	Manik Gupta
Software developer	David Enescu
Software developer	Eloi Sandt
Software developer	Lukáš Varga
Software developer	Preet Vadaliya
Software developer	Jan Potthoff

Product Vision	Project Mission
<p>Ailixir is a prototyping tool that produces custom AI agents. The tool enables an entrepreneur-developer user to evaluate the quality of the results that are produced by different models when they have been provided with selected input data from a handpicked list of inputs that belong to a specific knowledge domain / vertical. The custom agent that is produced as the end product for each knowledge domain helps the entrepreneur-developer user to decide whether the specific knowledge domain is worth investing more effort into building a better agent.</p>	<p>The mission of this particular project (in the context of the product vision).</p>

Sprint #	Sprint goal
1	None
2	None
3	None
4	Optional
5	Write your sprint goal here
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

Term	Definition
json-config	a json with a list of handpicked sources to be used as targets for data scrapping
orchestrator	A piece of logic that connects the json-based input list of sources with the scrapping targets.



Sprint	Goal	Feature Name	Est. Size	Est. Remaining	Real Size	Real Remaining
Release						
Total			0	0		
Sprints						
1			0	0	0	0
2			0	0	0	0
3			0	0	0	0
...				0		0
Features						
1						
2						
3						

Sprint	Goal	Feature Name	Est. Size	Est. Remaining	Real Size	Real Remaining
Release						
Total			0	0		
Sprints						
1			0	0	0	0
2			0	0	0	0
3			0	0	0	0
...				0		0
Features						
1						
2						
3						

Type	Link / reference

[illegible]

#	Context	Name	Version	License	Comment
1	development tooling	ruff	0.4.2	MIT	linter and code formatter
2	development tooling	pre-commit	3.7.0	MIT	pre-commit hooks
3	development tooling	python-dotenv	1.0.1	BSD License (BSD-3-Clause)	reads key-value pairs from a .env file
4	data pipeline development	python-youtube	0.9.4	MIT	youtube data scrapping, wrapper around for YouTube Data API V3.
5	data pipeline development	bs4	0.0.2	MIT License	Website scraping. Parses html and provides functionality to process the html
6	data pipeline development	beautifulsoup4	4.12.3	MIT License	Website scraping. Parses html and provides functionality to process the html
7	data pipeline development	google-cloud-speech	2.26.0	Apache 2.0	Transcription. Provides functionality for transcription - not used yet, due to access issues
8	data pipeline development	SpeechRecognition	3.10.4	BSD License (BSD)	Transcription. Provides functionality for transcription (replaced by vosk in recent commit)
9	data pipeline development	pydub	0.25.1	MIT License (MIT)	To work on audiofiles
10	data pipeline development	ffmpeg	1.4	LGPL 2.1+/ GPL 2+	Needed to convert mp3 to wav
11	data pipeline development	sox	1.5.0	BSD-3-Clause	work with audio files and concatenate them
12	data pipeline development	wave	0.0.2	GNU Library or Lesser General Public License (LGPL)	work with wav files
13	data pipeline development	vosk	0.3.45	Apache Software License	Provides models for speech recognition and transcription
14	data pipeline development	pocketsphinx	5.0.3	BSD License	continuous speech recognition engines
15	data pipeline development	pymongo	4.7.2	Apache 2.0	For pipelining between database and orchestrator
16	data pipeline development	paperscraper	0.2.11	MIT	Scraping text data from pdfs
17	data pipeline development	biopython	1.83	Freely Distributable	Pubmed API access, search queries, metadata scraping
18	data pipeline development	pypdf	4.2.0	BSD	Downloading pdfs from doi links (pubmed)
19	deployment tool	apptainer	1.3.1	BSD	container tool for HPC (potentially) or Google Cloud
20	data pipeline development	arxiv	2.1.0	MIT	Web scraping tool API for Arxiv
21	retrieval augmented generation	openai	1.30.1	Apache 2.0	Access to the OpenAI REST API
22	retrieval augmented generation	langchain	0.2.0	MIT	Framework for developing LLM
23	retrieval augmented generation	langchain-community	0.2.0	MIT	Framework for developing LLM
24	retrieval augmented generation	langchain-google-genai	1.0.4	MIT	Google integrations with LangChain
25	retrieval augmented generation	chromadb	0.5.0	Apache 2.0	AI-native open-source embedding database
26	retrieval augmented generation	vertexai	1.49.0	Apache 2.0	Platform for data science and machine learning
27	retrieval augmented generation	tiktoken	0.7.0	MIT	Tokeniser for OpenAI's models
28	retrieval augmented generation	sentence-transformers	2.7.0	Apache 2.0	Multilingual Sentence & Image Embeddings with BERT

Last Name	First Name	Value					
Enescu	David			#DIV/	#DIV/		
Karategos	Georgios			0!	0!		
Negasa	Gemechis						
Potthoff	Jan						
Sandt	Eloi			0	No size		
Varga	Lukáš			1	Trivial size		
Zimmermann	Simon			2	Small size		
Gupta	Manik			3	Medium size		
Vadaliya	Preet			5	Large size		
				8	Very large size		
				13	Too large (size)		