THE DENARIO PROJECT: DEEP KNOWLEDGE AGENTS FOR SCIENTIFIC DISCOVERY

SCAN ME

A denario was **special** because it was ancient Rome's **standard silver coin**, essential for its economy, military, and trade, and it became a widely recognized symbol of the empire's power. It had a **consistent value, making it a reliable medium** for buying goods and paying for services

RAUL JIMENEZ, LICIA VERDE (ICREA-ICCUB) & PEDRO TARANCÓN-ÁLVAREZ (ICCUB)

ICCUB INTERNAL SEMINAR, 26/09/2025

HTTPS://GITHUB.COM/ASTROPILOT-AI/DENARIO



Villanueva-Domingo
(Computer Vision Center)



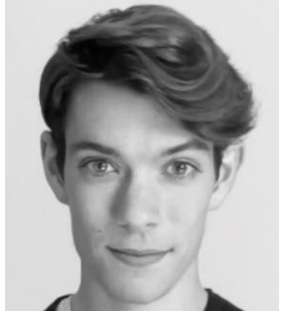
Bolliet
(Cambridge)



Villaescusa-Navarro
(Flatiron)



Bayer
(Flatiron/Princeton)



Bermejo
(DIPC/Flatiron/Google)



Bilodeau
(Virginia)



Hahn
(Arizona/Austin)



Jimenez
(Barcelona)



Zou
(SISSA)



Lee
(Flatiron/Princeton)



Meyer
(LMU)



Protopapas
(Harvard)



Roy
(Virginia)



Tarancon-Alvarez
(Barcelona)



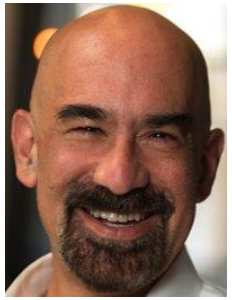
Wang
(DeepMind)



Xu
(Cambridge)



Zhu
(Flatiron)



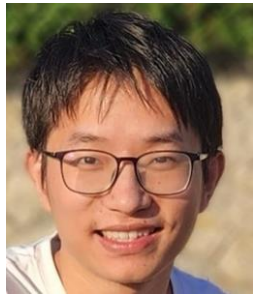
Spergel
(Simons)



Anand Ojha
(Flatiron)



França
(Harvard)



Zhou
(UTokyo)



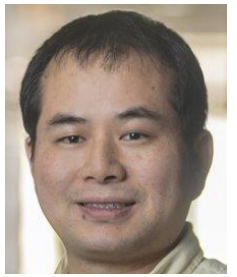
Wang
(UChicago)



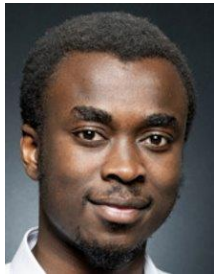
Cardenas-Ramirez
(Cornell/MIT/Harvard)



Lerario
(SISSA)



Jiang
(Flatiron)



Acquah
(Oxford)



Amancharla
(Infosys)



Wadekar
(John Hopkins)



Barzilay-Siegal
(Tel-Aviv)



Yovel
(Tel-Aviv)



Viel
(SISSA)



Cranmer
(Cambridge)



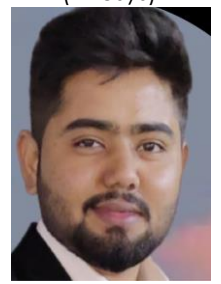
Zubeldia
(Cambridge)



Mamun
(Cornell)



Yue
(Cornell)



Tiwari
(Infosys)


FLATIRON
INSTITUTE

UNIVERSITY OF
CAMBRIDGE

CVC
Computer
Vision Center

UNIVERSITY OF
OXFORD

HARVARD
UNIVERSITY

Massachusetts
Institute of
Technology

DeepMind

Infosys

UAB
Universitat Autònoma
de Barcelona

UNIVERSITY
of VIRGINIA

Cornell University

PRINCETON
UNIVERSITY

THE UNIVERSITY
OF ARIZONA

TEXAS
The University of Texas at Austin

ma per seguir virtute e canoscenza
SISSA
Scuola Internazionale Superiore
di Studi Avanzati

Boston
Children's
Hospital
Until every child is well™

JOHNS HOPKINS
UNIVERSITY

ioa

ICSC
Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

TEL AVIV UNIVERSITY


INAF
OSSERVATORIO
ASTRONOMICO
DI TRIESTE

ifpU

Donostia
International
Physics Center

TRINITY HALL
CAMBRIDGE

Kavli Institute for Cosmology
KICC
Cambridge

東京大学
THE UNIVERSITY OF TOKYO

THE UNIVERSITY OF CHICAGO
1890

LMU
LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INFN

mcml
Munich Center for Machine Learning

APSCI
ACCELERATE
PROGRAMME
FOR SCIENTIFIC DISCOVERY

Google Research

ICCUB
Institut de Ciències del Cosmos
UNIVERSITAT DE BARCELONA
EXCELENCIA
MARÍA
DE MAEZTU
2020-2024
ICREA

Scientific dissemination and publishing

Authors do research and produce paper

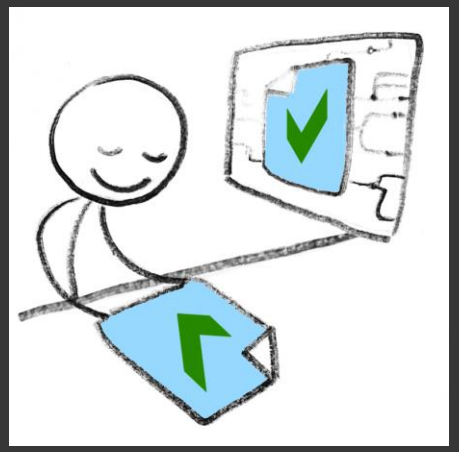


t~months/years

arXiv

t~hrs/days

Journals
Editors/referees
QA,Curation , refs etc.



t~months

Scholarly record
knowledge



JCAP

Researchers discover new ideas



Effort of scientists on this

Must be ε of the effort on this

Authors do research and produce paper

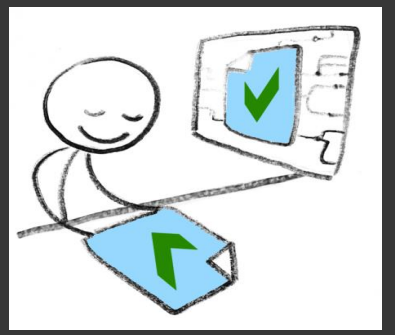


t~months/years

arXiv

t~hrs/days

Journals
Editors/referees
QA,Curation , refs etc.



t~months

Scholarly record knowledge



Researchers discover new ideas



And content must be of high value: interest and added value

WHAT IS A SCIENTIFIC PAPER?

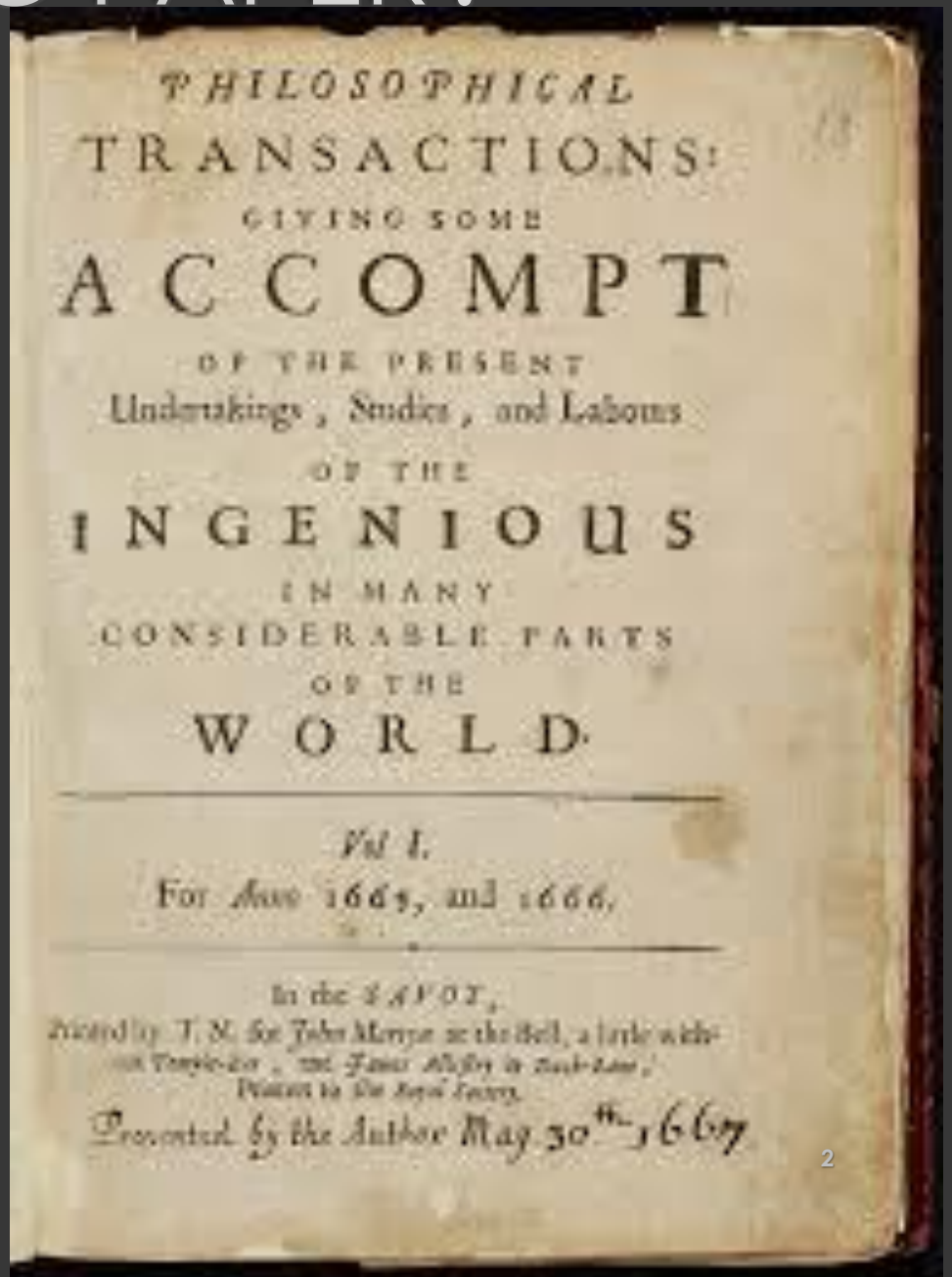
The modern concept of a scientific paper emerged with the founding of the **Royal Society of London** and its journal ***Philosophical Transactions*** in **1665**





Your editor…

Before that scientific discoveries were published in "pamphlets" and books (Galileo…)


PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol. I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane, Printers to the Royal Society.
Presented by the Author May 30th 1667

OUTLINE

1. Philosophical overview
2. Large language models and AI-agents
3. Denario's workflow
4. Writing session
5. Ethical considerations

Adverse effects of technology in “progress and thinking” are as old as humankind

Plato and Aristotle emphasized that **knowledge is acquired through active engagement and experience, not passive reception from texts**. They viewed writing as a tool for reminding rather than teaching, arguing that genuine learning involves the mind's active process of questioning and integrating information within a rich dialogue and practical life context. For Aristotle, oral transmission and interaction were superior to written works for fostering intellectual virtue and practical wisdom, making writing a potential shortcut that could hinder deeper, more engaged learning.

Circa 360 BC



The ``adverse'' effects of THE PRINT

Adverse effects attributed to the Gutenberg printing press included fears of **misinformation and propaganda**, concerns about the spread of **heretical or destabilizing ideas**, the **loss of work for scribes and monks**, a threat to the **monopoly of knowledge** held by authorities, and the potential for the press to **undermine true intellectual pursuits** by creating too many low-quality books. Scribes even attacked a printing press in 1476, fearing the new technology would threaten their livelihoods and status as knowledge custodians.



The Computer!!

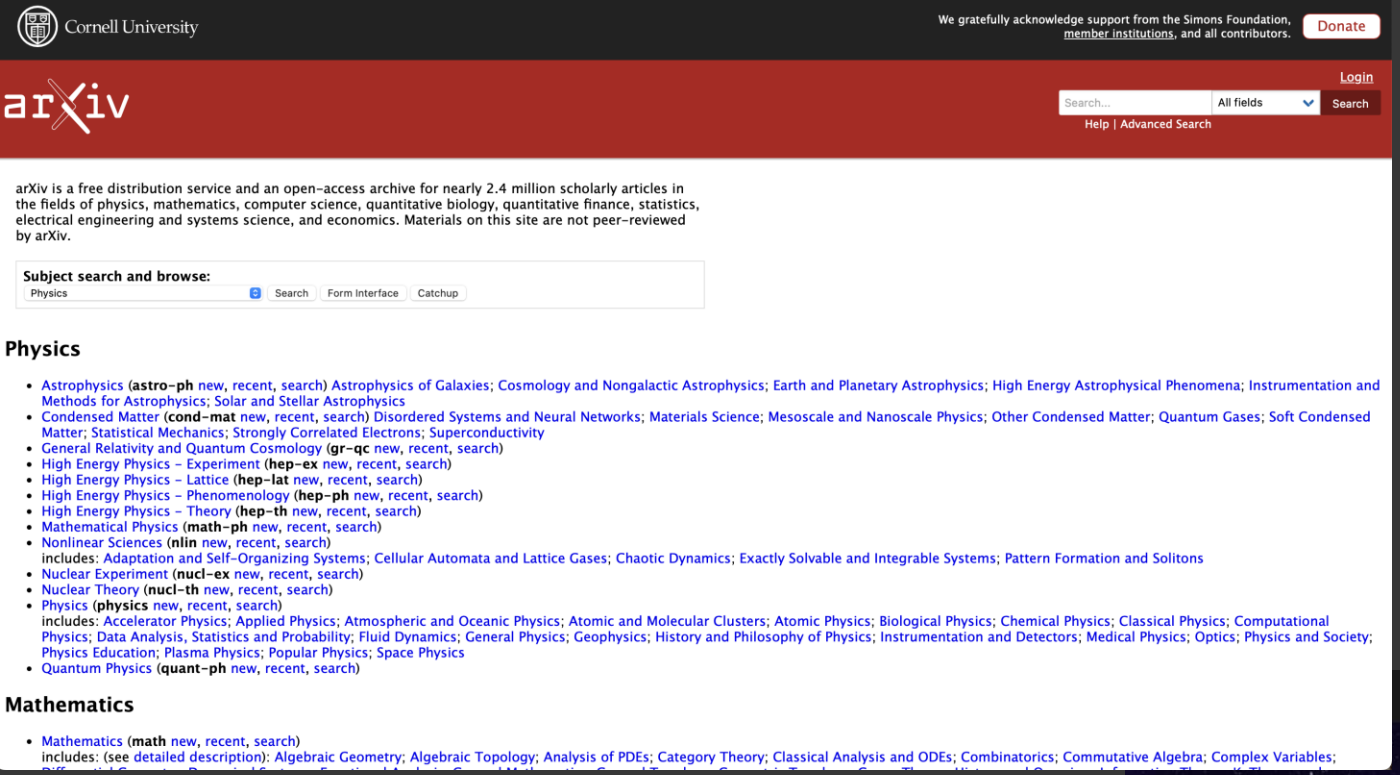
Early fears centered on potential negative impacts on conceptual understanding and reliability, as seen with widespread calculator use showing lower performance. More recently, debates focus on how computers might diminish intuition and creativity, the difficulty of verifying complex software for proofs, and the potential philosophical implications of computer-generated proofs on the role of mathematicians themselves.

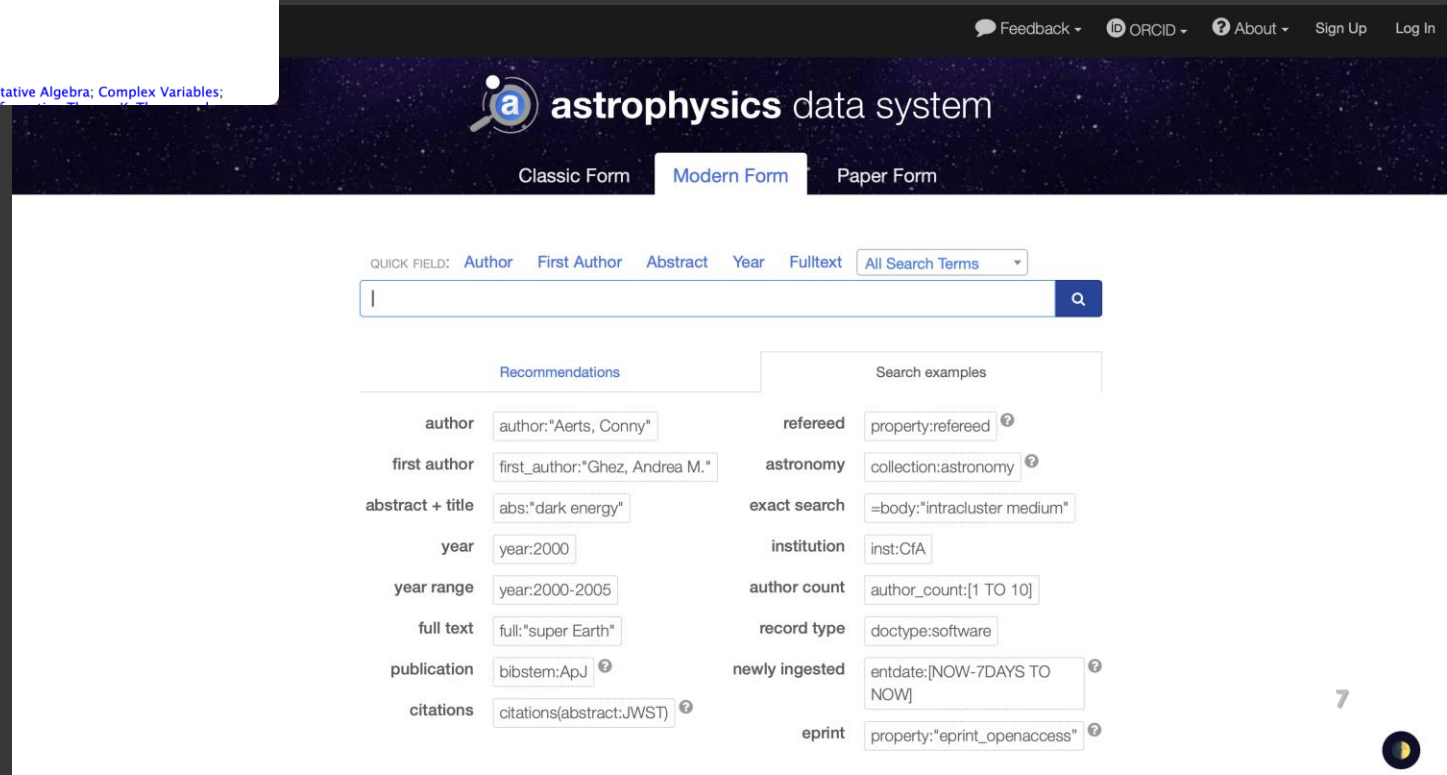
One of **Torres-Quevedo's** most surprising contributions was in the field of artificial intelligence. His most notable invention in this area was the "**Ajedrecista**," considered one of the first game-playing automata in history.
The Ajedrecista, first presented in 1912, was capable of playing a chess endgame (king and rook versus king) **completely automatically.** Although the rules were limited, the device could "think" several moves ahead and execute strategies to win the game.



Query Based Information

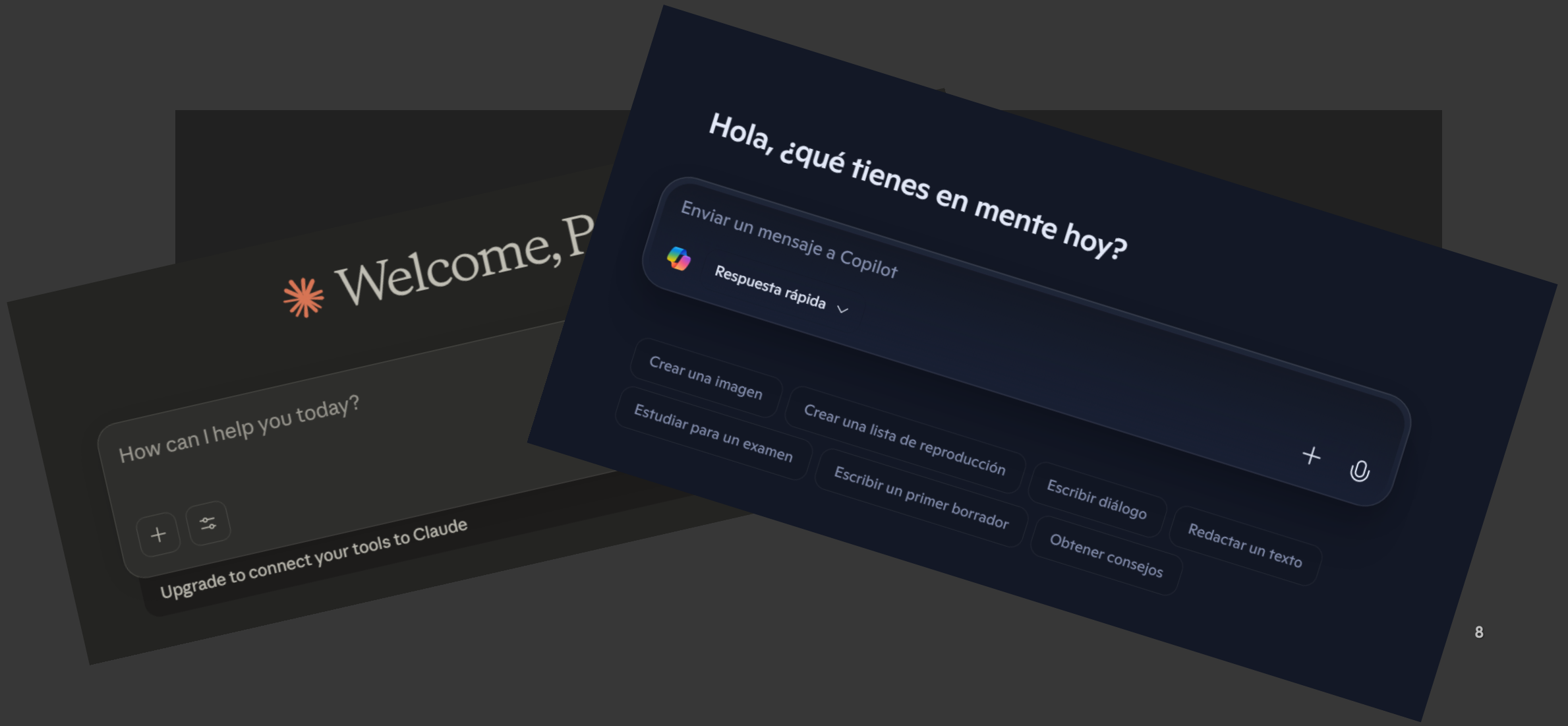
> 1992





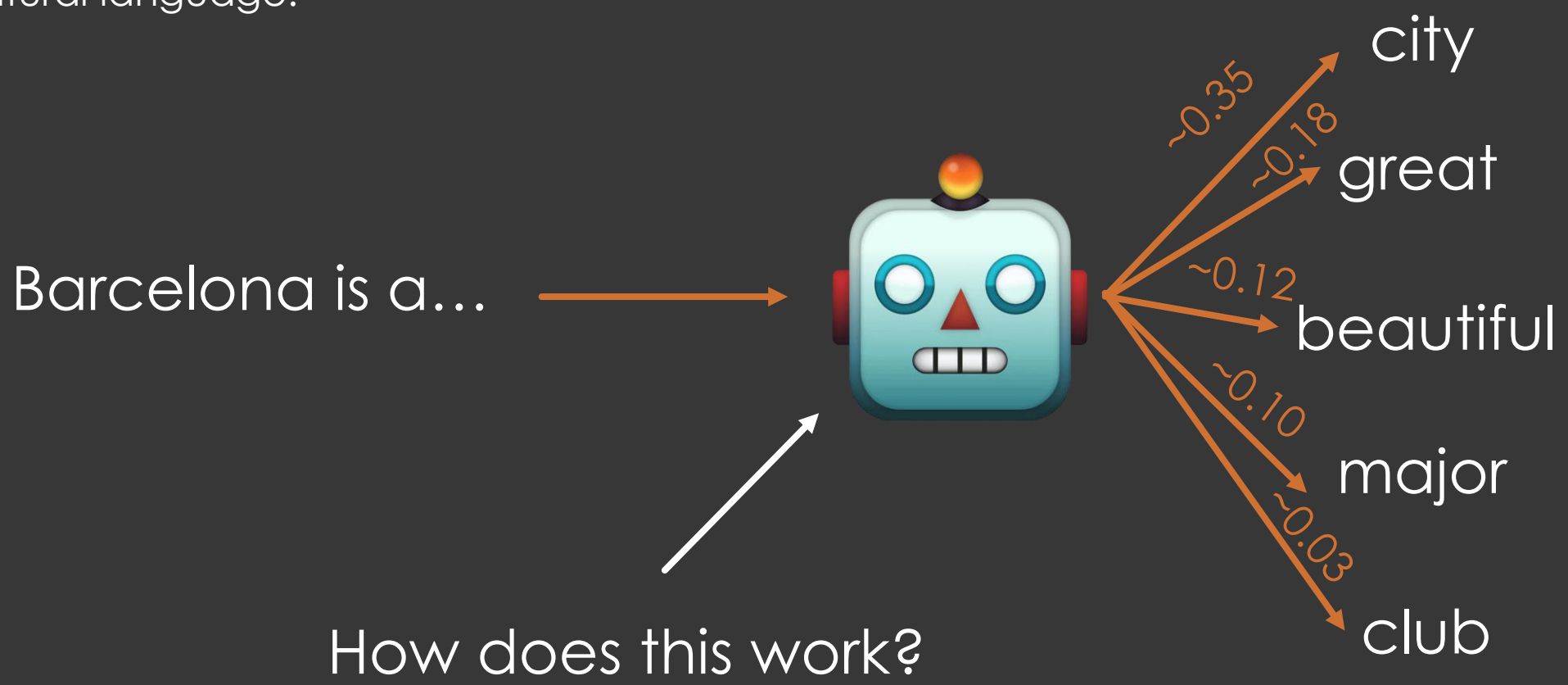
Interestingly LV is the chair of advisory panel of both of these efforts

LARGE LANGUAGE MODELS



LLM

- Type of neural network trained on massive text datasets to predict and generate natural language.



How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

- The text that we want to be continued

"The cat that chased the mouse was hungry"

How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

- Split the text into different words
- Each word is then mapped to a number depending on the algorithm used

"The cat that chased the mouse was hungry"

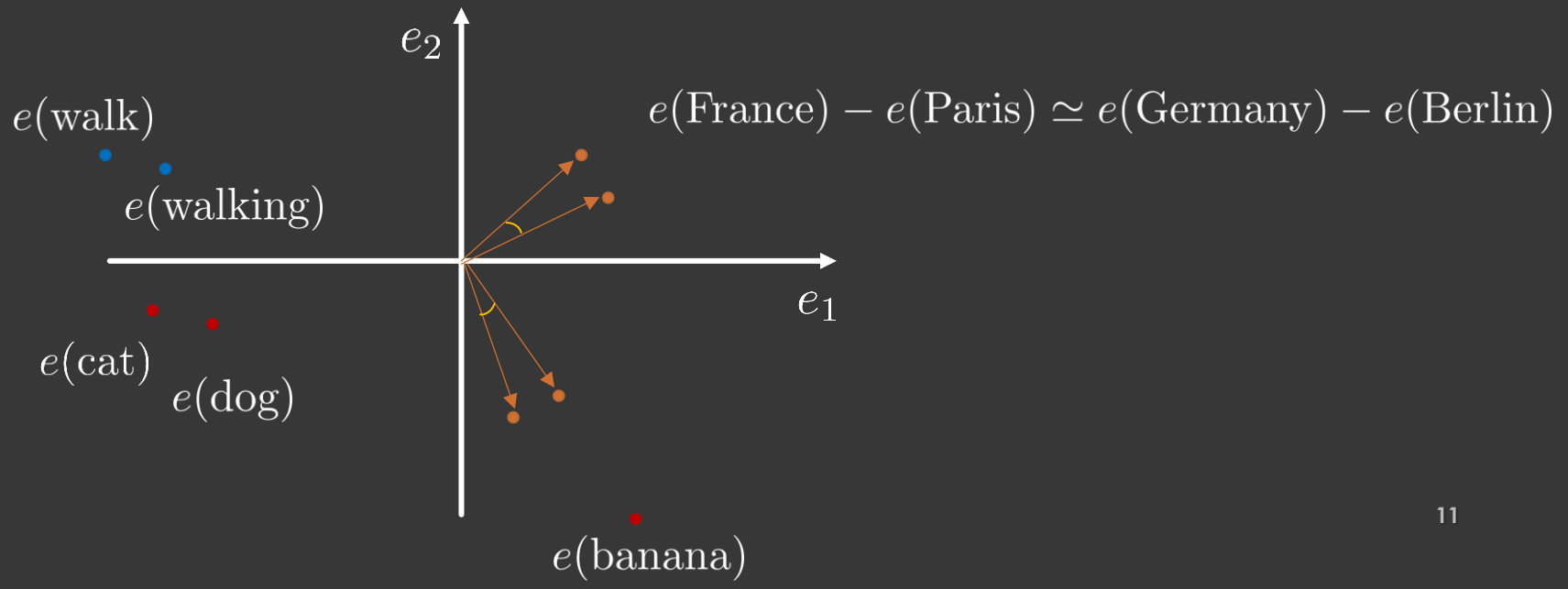
↓ BERT's WordPiece tokenizer

[101, 1996, 4937, 2008, 4025, 1996, 4301, 2001, 5823, 102]

How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

- Each token is mapped to a vector that lives in a high dimensional space
- The embeddings capture semantic similarity (position encoding)



How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

- Reason why modern LLMs work [Vaswani et al. 2017]
- Each transformer has: self attention layers + feed forward layers

Self attention

- Each token looks to the rest
- How much each token should "attend" the rest is quantified by attention weights



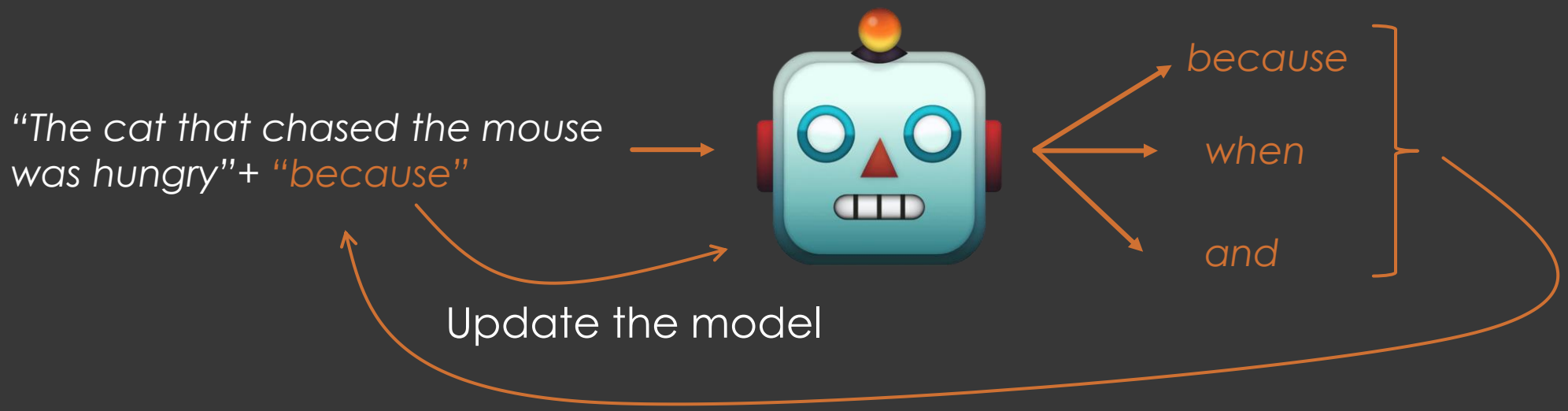
Feed forward layers

- Tokens pass through several fully-connected neural networks
- Multihead structure: each of them with their own attention

Transformers are stacked in consecutive layers

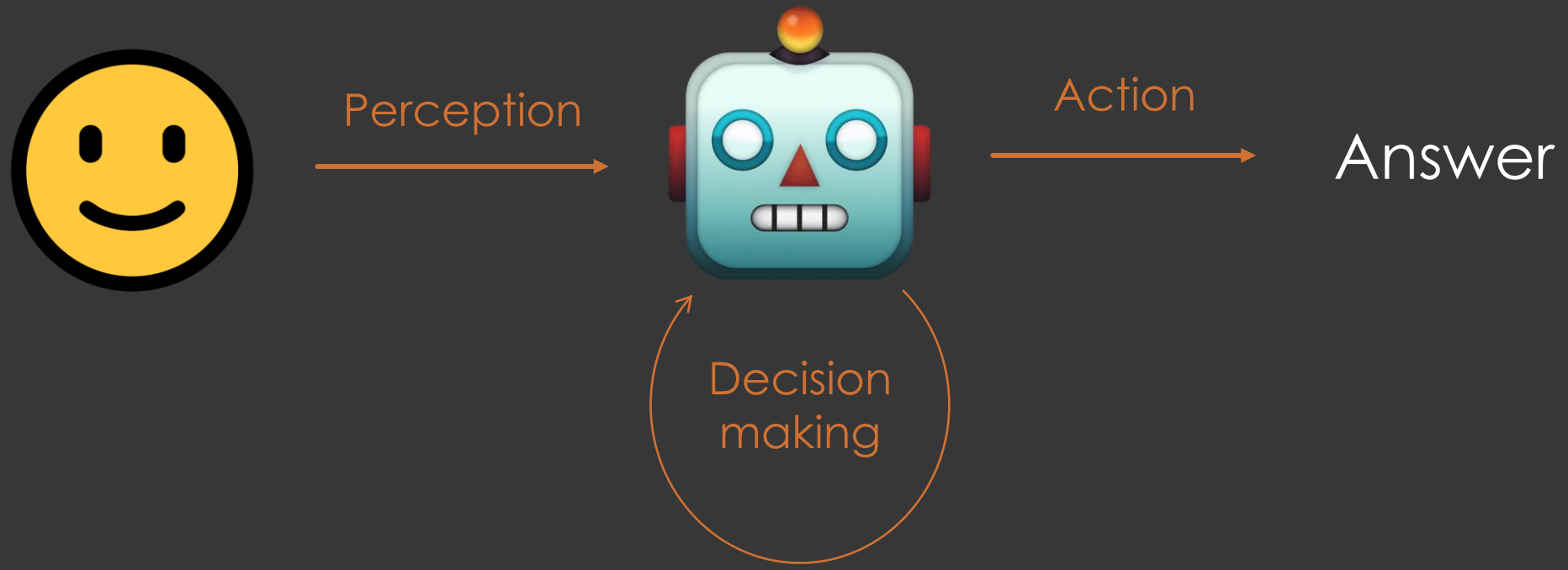
How does an LLM work?

Input text → Tokenization → Embeddings → Transformers → Prediction

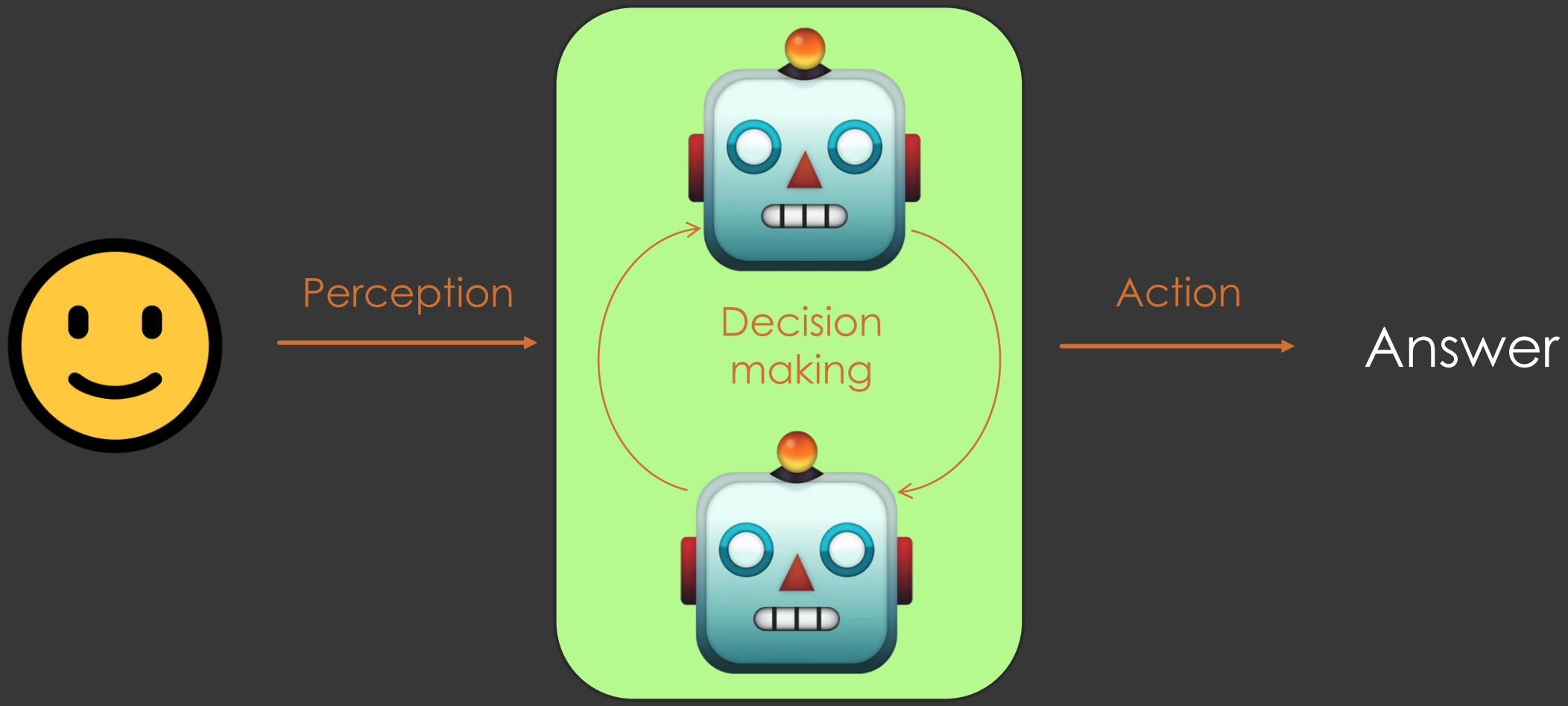


Compare the model output with the real answer

AI agent system

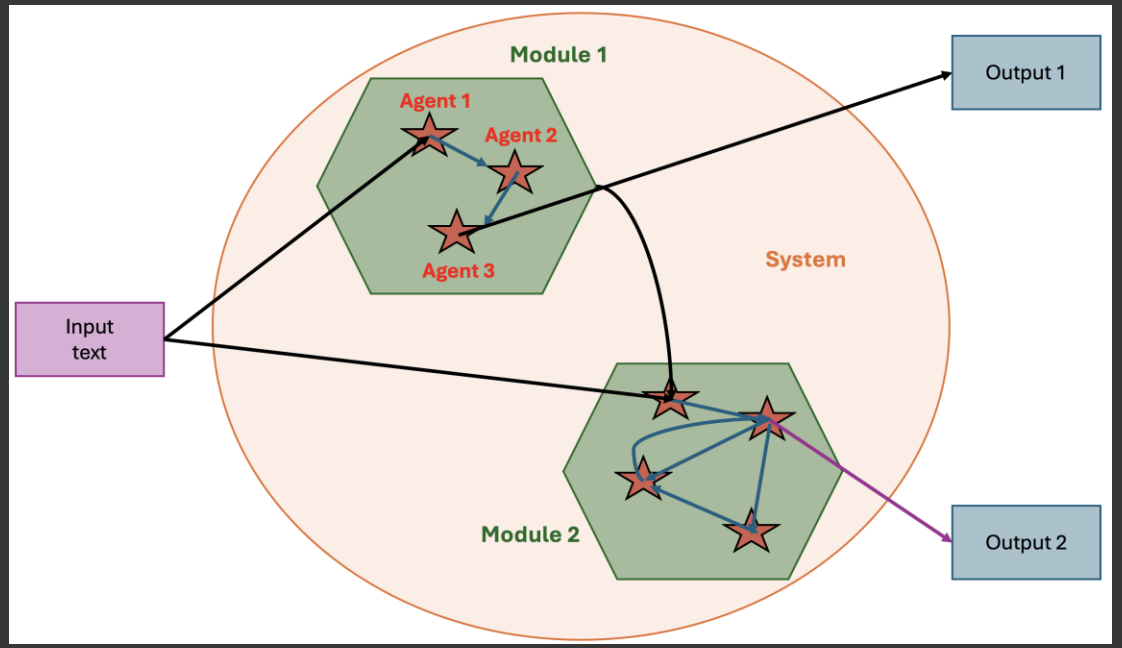


AI multi-agent system



AI multi-agent system

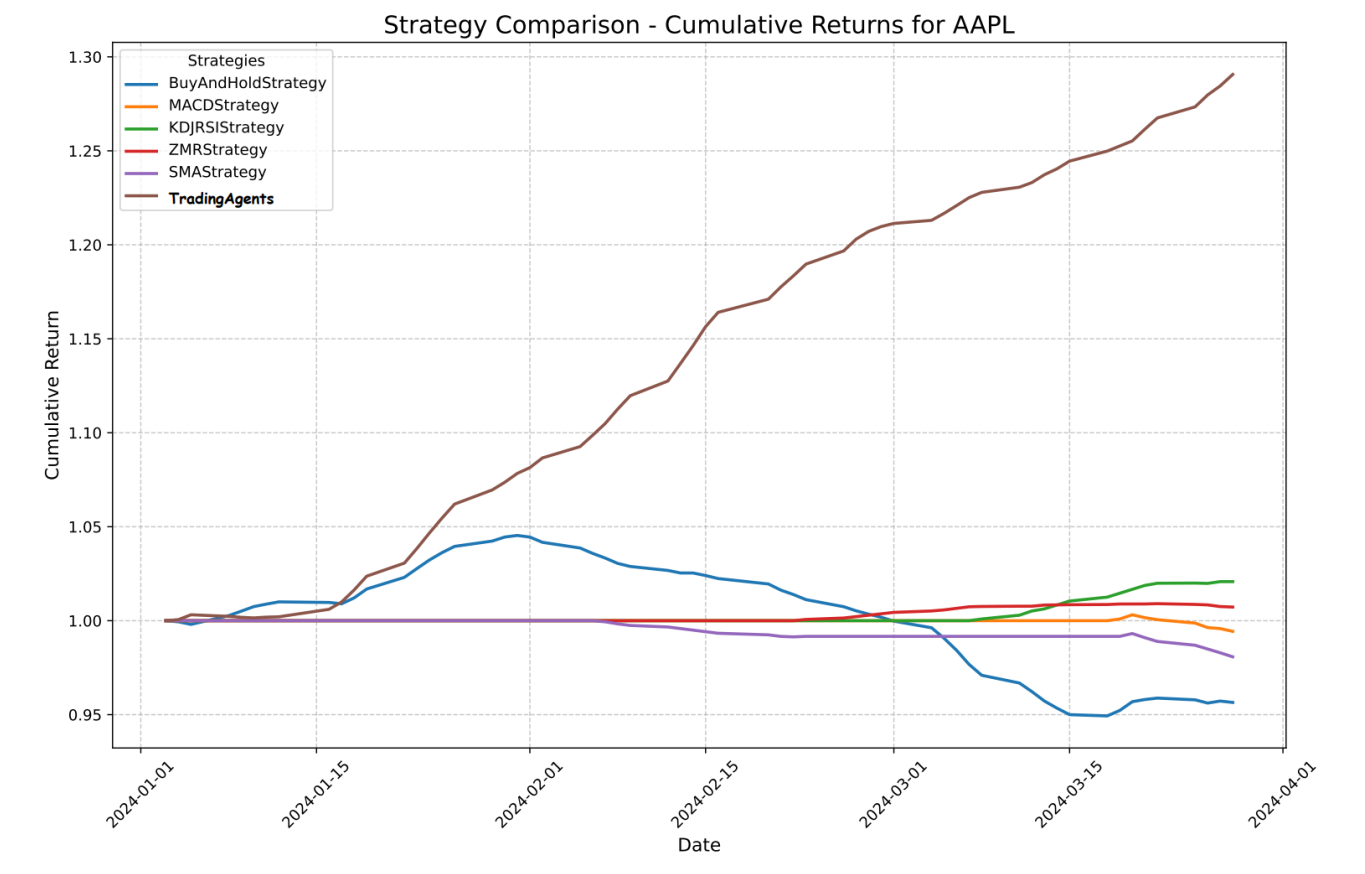
- Autonomous system composed by different number of LLMs
- Agents can
  - Perceive the environment (take text as input)
  - Make decisions
  - Gives outputs (text, code, plots, files…)
- Each agent can be specialized on different tasks
- Different relations between agents will give different responses
  - Meeting-like
  - Company-like
  - Plan and control



AI multi-agent system

[Xiao et al. 2025]

AI multi-agent system to replicate real-world trading strategies



CAN WE DO SCIENCE WITH AGENTS?

The Denario Project: Modular Automation of Scientific Research with Multi-Agent Systems

Francisco Villaescusa-Navarro[1,2,*], Boris Bolliet[3,4,*], Pablo Villanueva-Domingo[5,*], Adrian E. Bayer[1,2], Aidan Acquah[6], Chetana Amancharla[7], Almog Barzilay-Siegal[8], Pablo Bermejo[9,10,11], Camille Bilodeau[12], Pablo Cárdenas Ramírez[13,14,15], Miles Cranmer[16], Urbano L. França[17,18], ChangHoon Hahn[19,20], Yan-Fei Jiang[1], Raul Jimenez[21,22], Jun-Young Lee[1], Antonio Lerario[23], Osman Mamun[13], Thomas Meier[24], Anupam A. Ojha[25,26], Pavlos Protopapas[27], Shimanto Roy[12], Pedro Tarancón-Álvarez[21,28], Ujjwal Tiwari[7], Matteo Viel[23,29,30,31,32], Digvijay Wadekar[33], Chi Wang[34], Bonny Y. Wang[35], Licong Xu[36,4], Yossi Yovel[8,37], Shuwen Yue[13], Wen-Han Zhou[38], Qiyao Zhu[25], Jiajun Zou[39], Íñigo Zubeldia[36,4]



Francisco Villaescusa-Navarro (Flatiron Institute)



Boris Bolliet (University of Cambridge)



Pablo Villanueva-Domingo (Universitat Autònoma de Barcelona)

Denario: Modular Automation of Research

https://github.com/AstroPilot-AI/Denario

- AI multi-agent system designed to be a scientific research assistant
- Generate ideas, check the literature, develop research plans, write and execute code, make plots, and write a scientific paper
- Modular system: multi agent orchestration with AG2 and LangGraph



https://ag2.ai/

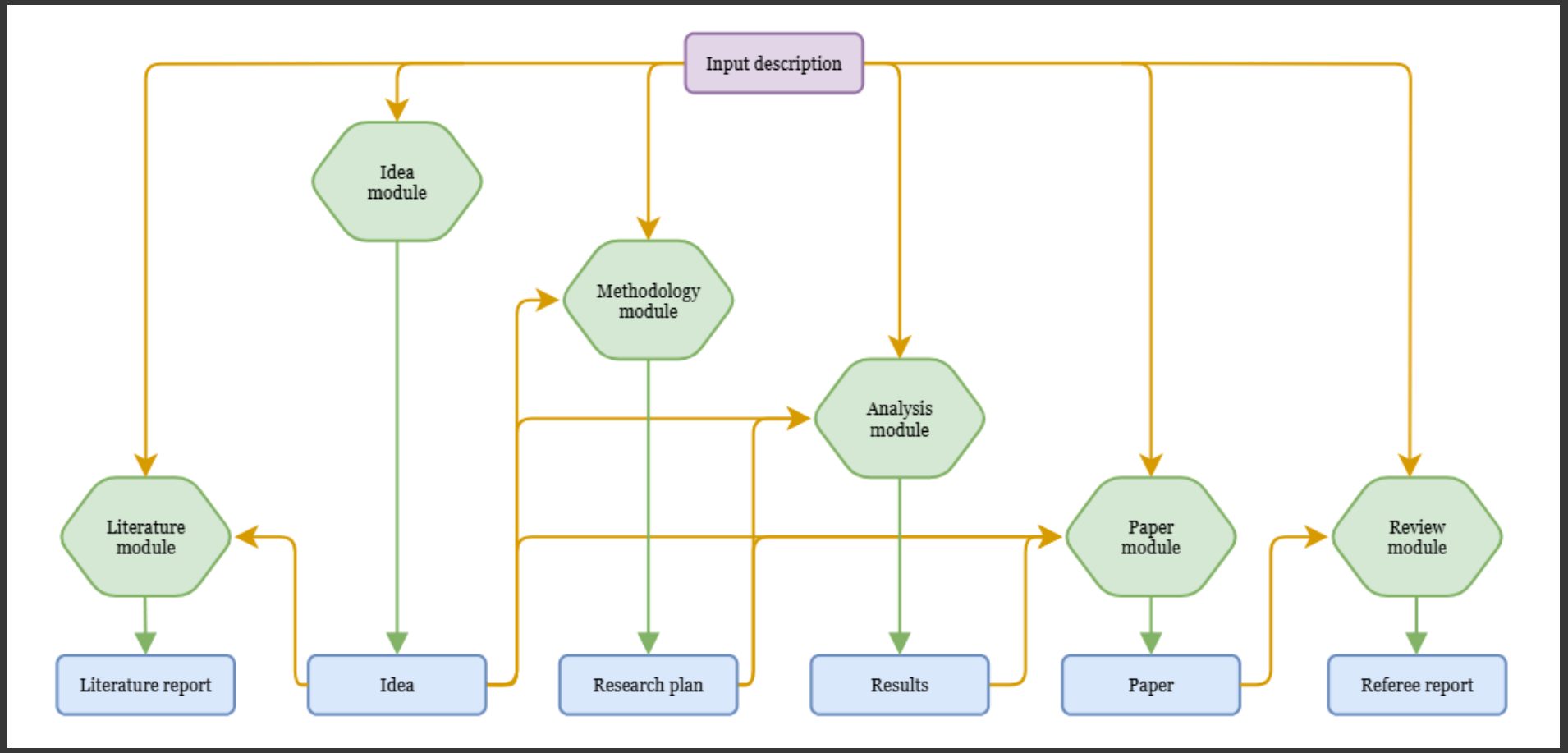


- Plan and Control is done with cmbagent

CMBAGENT

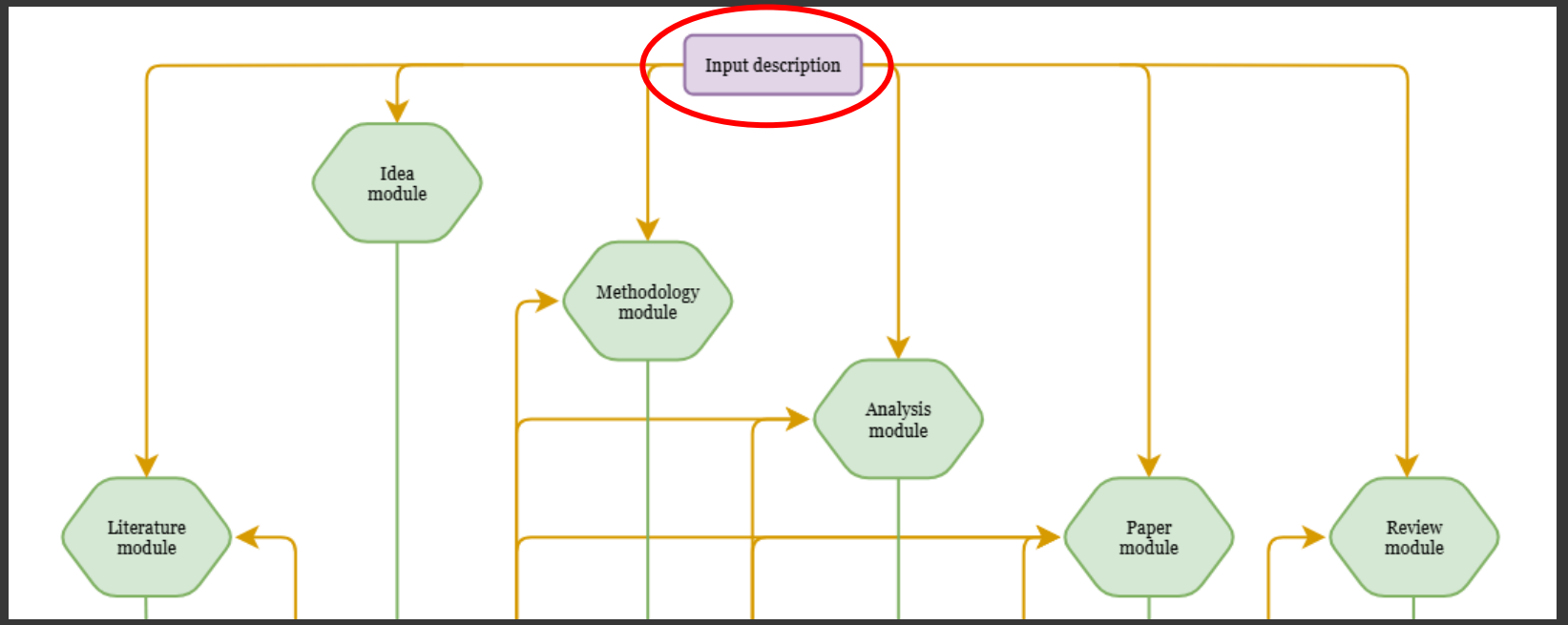
https://github.com/CMBAgents/cmbagent

Denario: architecture



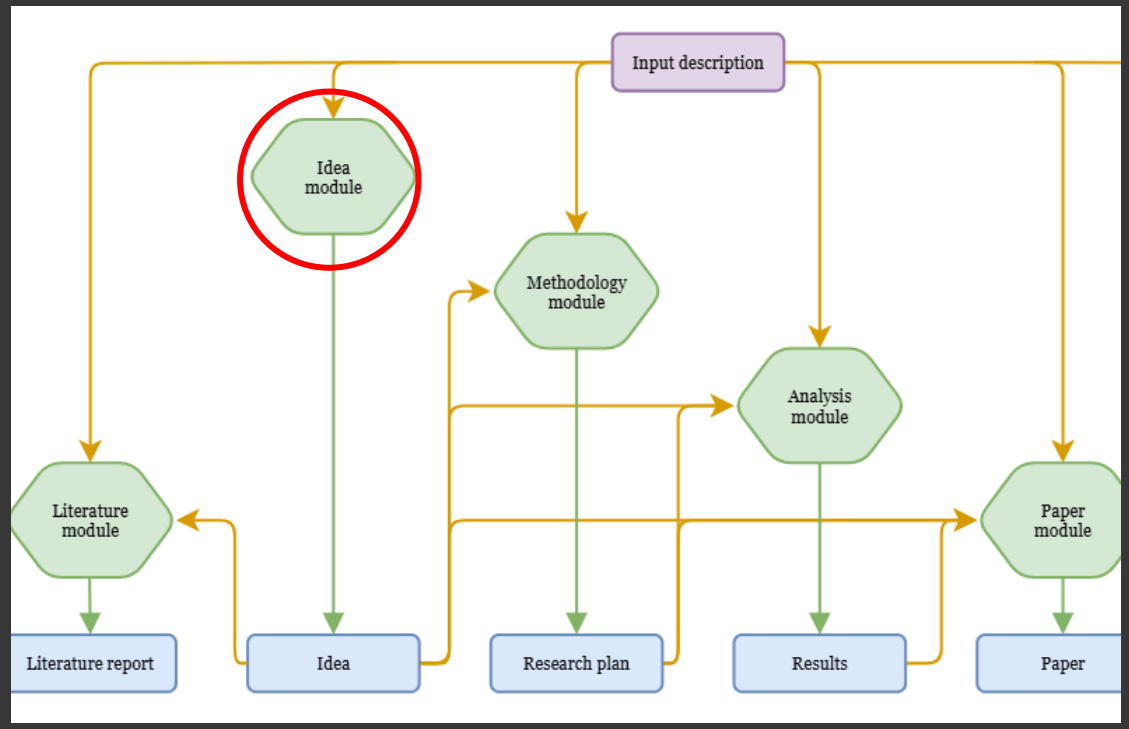
Denario: input description

- The input description will initially guide the system towards the final result
- It can contain data, topic to investigate, analysis that we want to perform, key points…
- The **input.md** file will be given to the consecutive modules

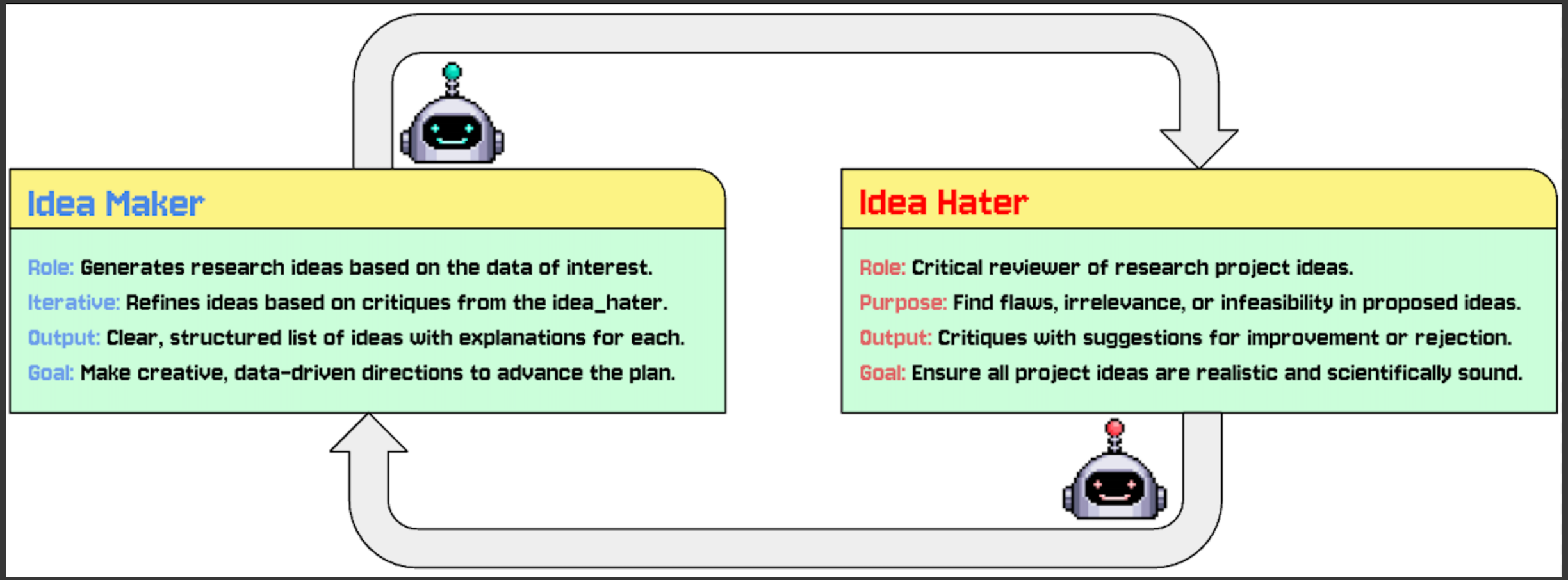


Denario: idea module

- The idea module will take **input.md** and produce a file **idea.md**
- Goal: generate a (hopefully) innovative idea given the initial data
- The pipeline tries to mirror the natural human process of idea development

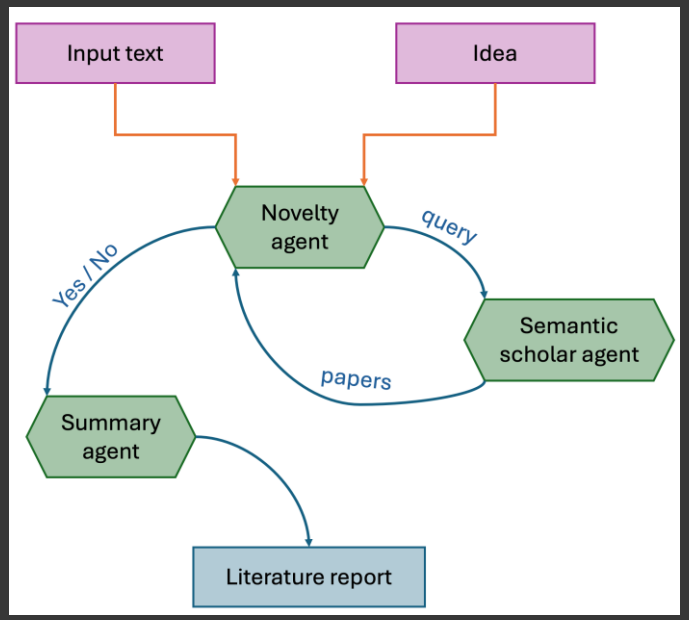


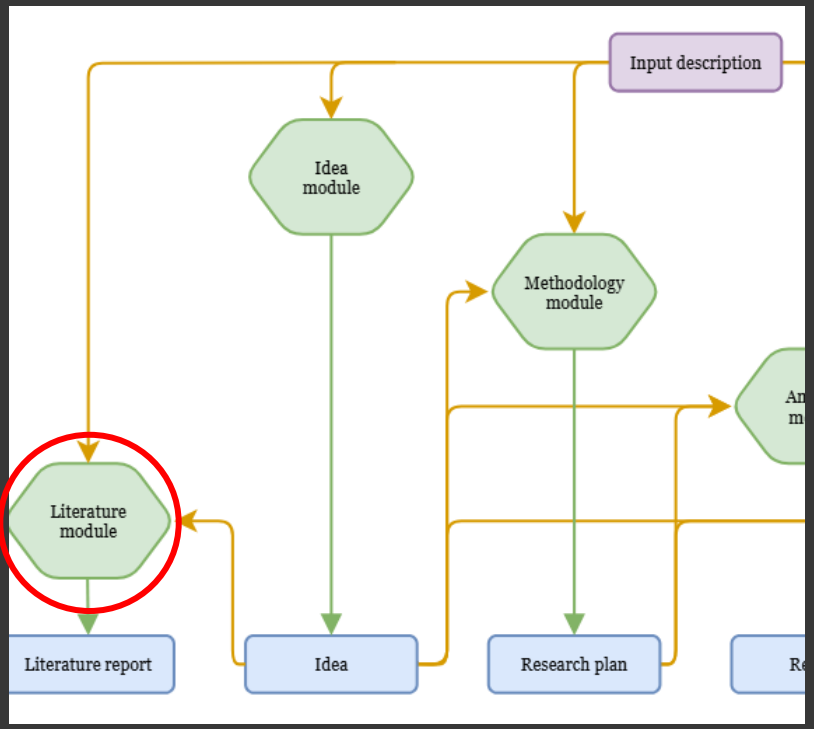
Denario: idea module



Denario: literature module

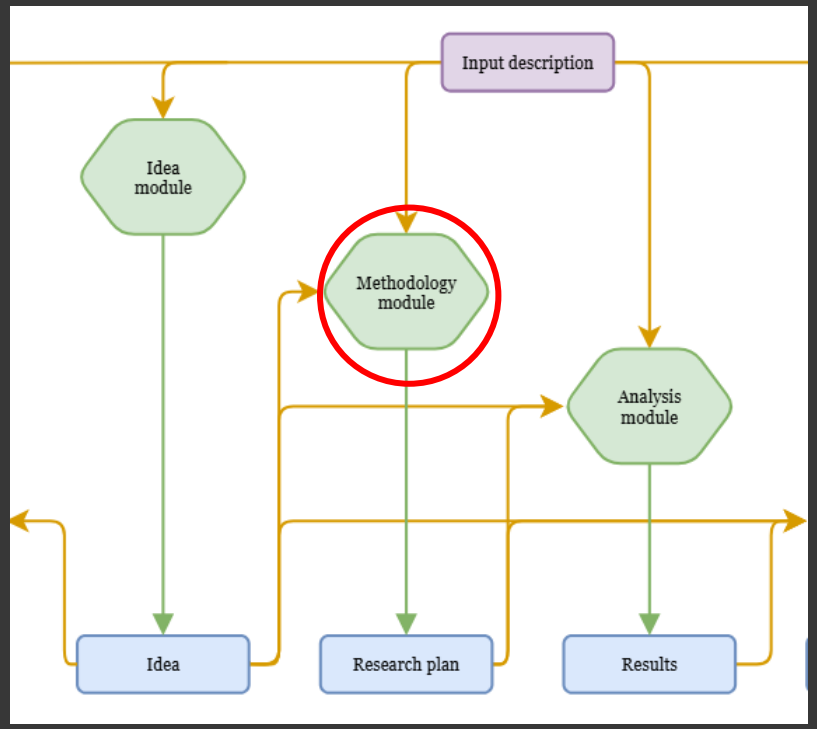
- Takes **input.md** and **idea.md** and will produce a literature review
- Goal: determine whether the idea is original or not
- Composed by three agents





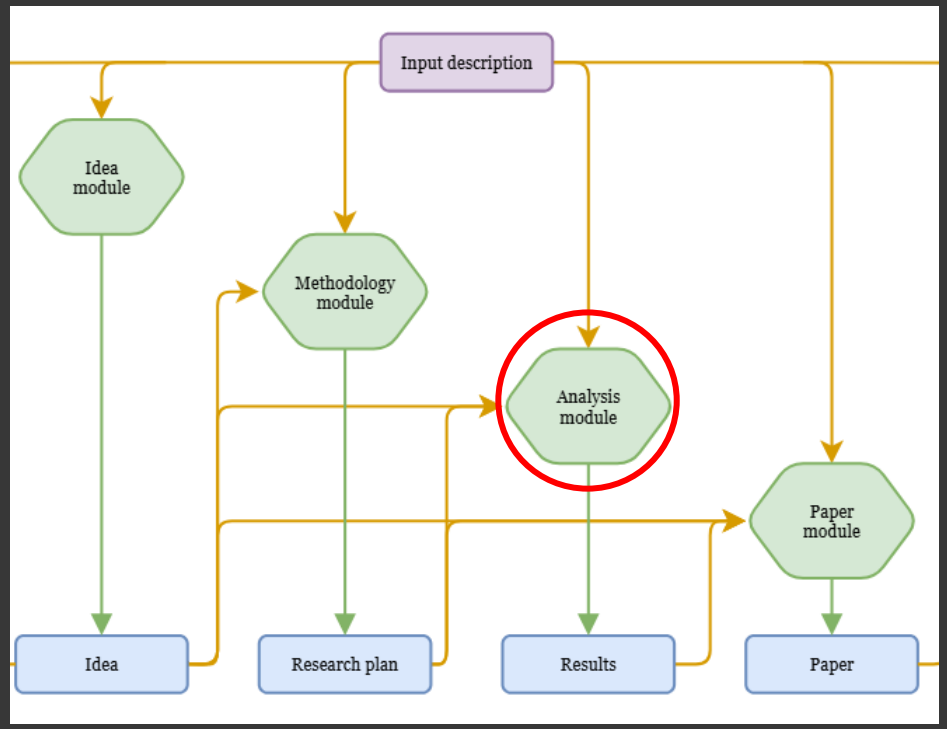
Denario: methodology module

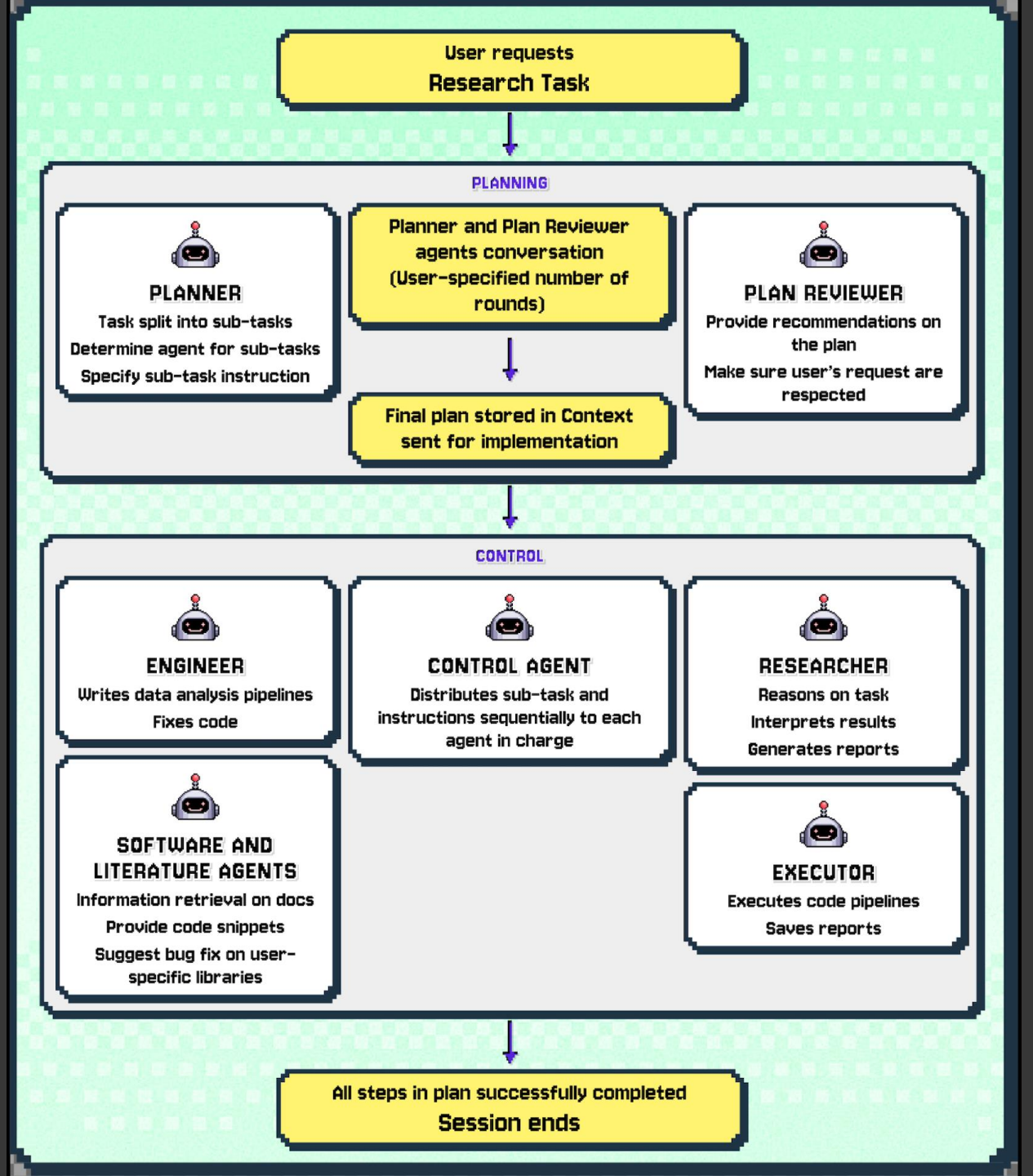
- Takes **input.md** and **idea.md** and will produce **methods.md**
- Goal: produce plan and methodology to tackle the problem
- The output will be a detailed research plan
- It will focus on the methods and workflow, rather than generating results



Denario: analysis module

- Takes **input.md**, **idea.md** and **methods.md**.
- Goal: generate the results as stated in the methodology
- It can write and run code, generate plots and interpret the results
- The output will be a **report** and **plots**
- Framed out in research-specialized **cmbagent** [Xu et al. 2025]

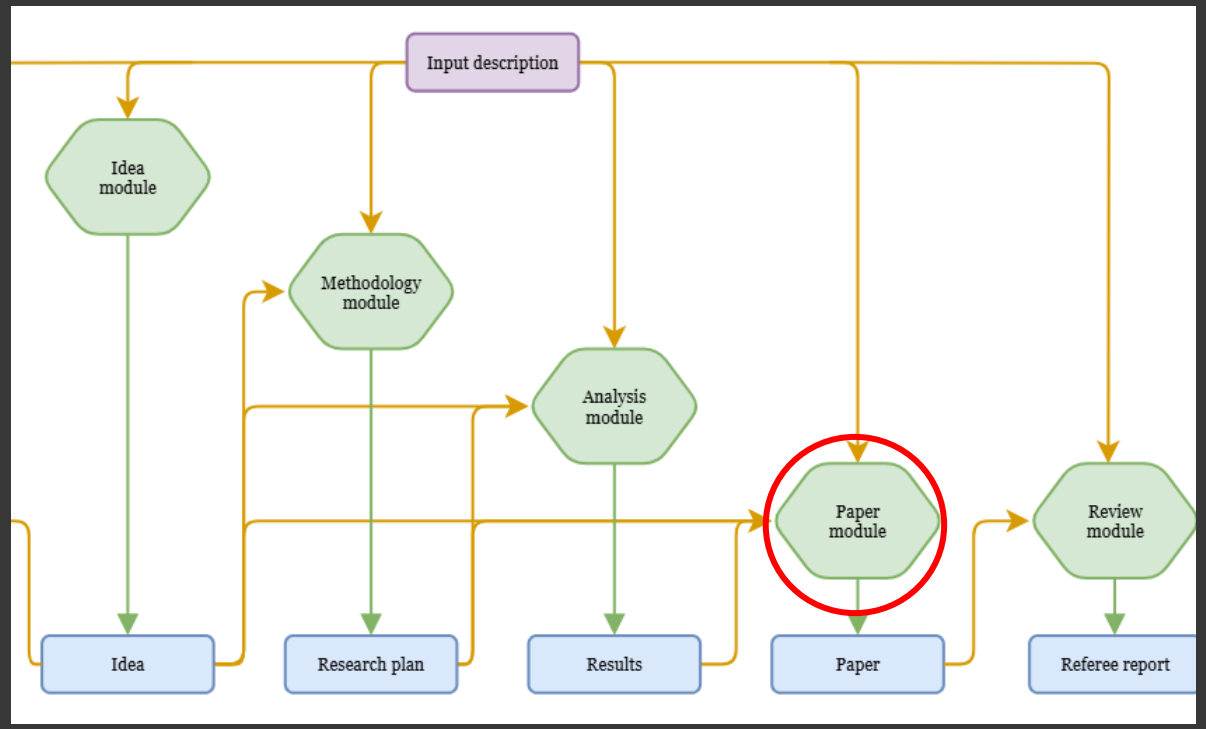



User requests
Research Task
PLANNING
PLANNER
Task split into sub-tasks
Determine agent for sub-tasks
Specify sub-task instruction
Planner and Plan Reviewer agents conversation (User-specified number of rounds)
Final plan stored in Context sent for implementation
PLAN REVIEWER
Provide recommendations on the plan
Make sure user's request are respected
CONTROL
ENGINEER
Writes data analysis pipelines
Fixes code
CONTROL AGENT
Distributes sub-task and instructions sequentially to each agent in charge
RESEARCHER
Reasons on task
Interprets results
Generates reports
SOFTWARE AND LITERATURE AGENTS
Information retrieval on docs
Provide code snippets
Suggest bug fix on user-specific libraries
EXECUTOR
Executes code pipelines
Saves reports
All steps in plan successfully completed
Session ends

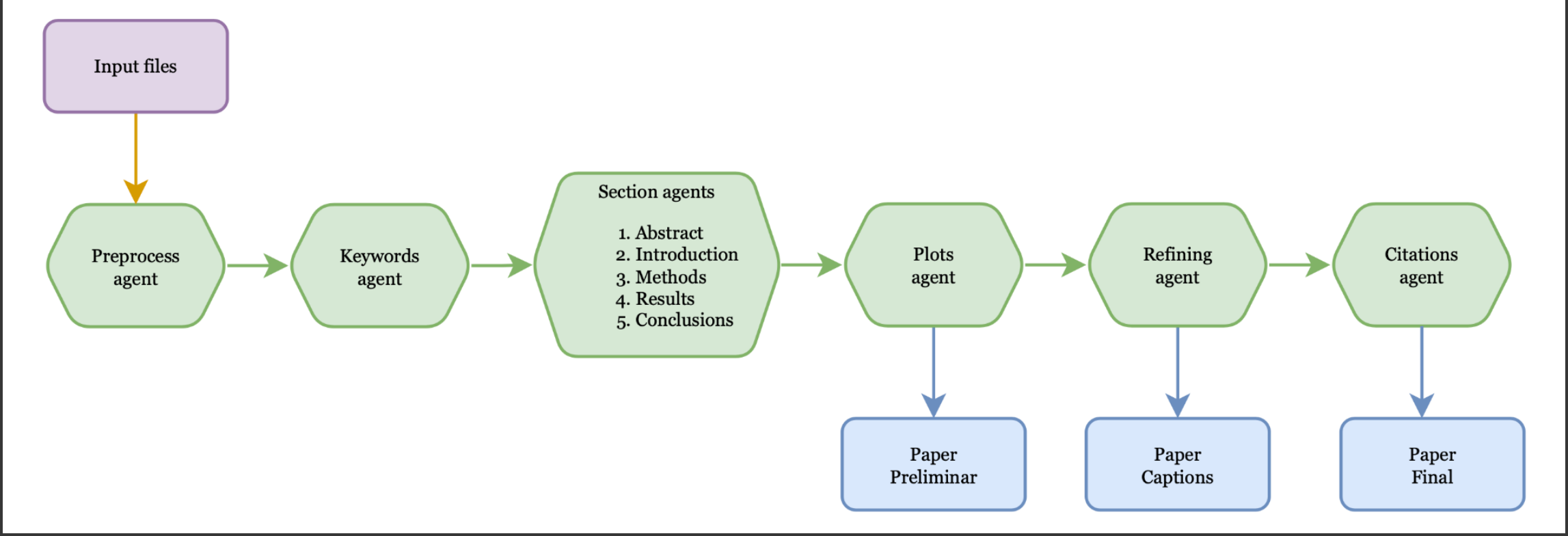
[Xu et al, 2025]

Denario: paper module

- Takes **input.md**, **idea.md**, **methods.md** and **results.md**
- Goal: put everything together and write the paper
- Different agents will take care of different abstract/sections/plots.
- The output will be a .tex file with the final paper.



Denario: paper module



Denario: conclusions

- Given an input text, DENARIO can generate ideas, develop research plans, Write and execute code, make plots, and summarize results and write a scientific paper
- Validation is crucial. Mimicking a scientific paper structure doesn't mean it is valuable
- Advantages: replicability and modular structure. It can be used as a tool to speed up tedious processes
- Denario is publicly available, although OpenAI API key is mandatory to run the whole pipeline

Ethical Considerations

1. Quality of Research
2. Plagiarism
3. Implications for scientific training and education
4. Assessment of quality and legitimacy
5. Homogenization towards what?
6. Access
7. Authorship
8. Misalignment of incentives in scientific publishing
9. Environmental concerns
10. Malicious use





Countermeasures being considered

1. Must declare how AI was used
2. Only specific use being allowed
3. Penalty may include being banned from publishing and contacting host institution/funder

WHAT IS A SCIENTIFIC PAPER?

HOW SHOULD WE DISSEMINATE RESEARCH?