

Autonomous research agents for life sciences

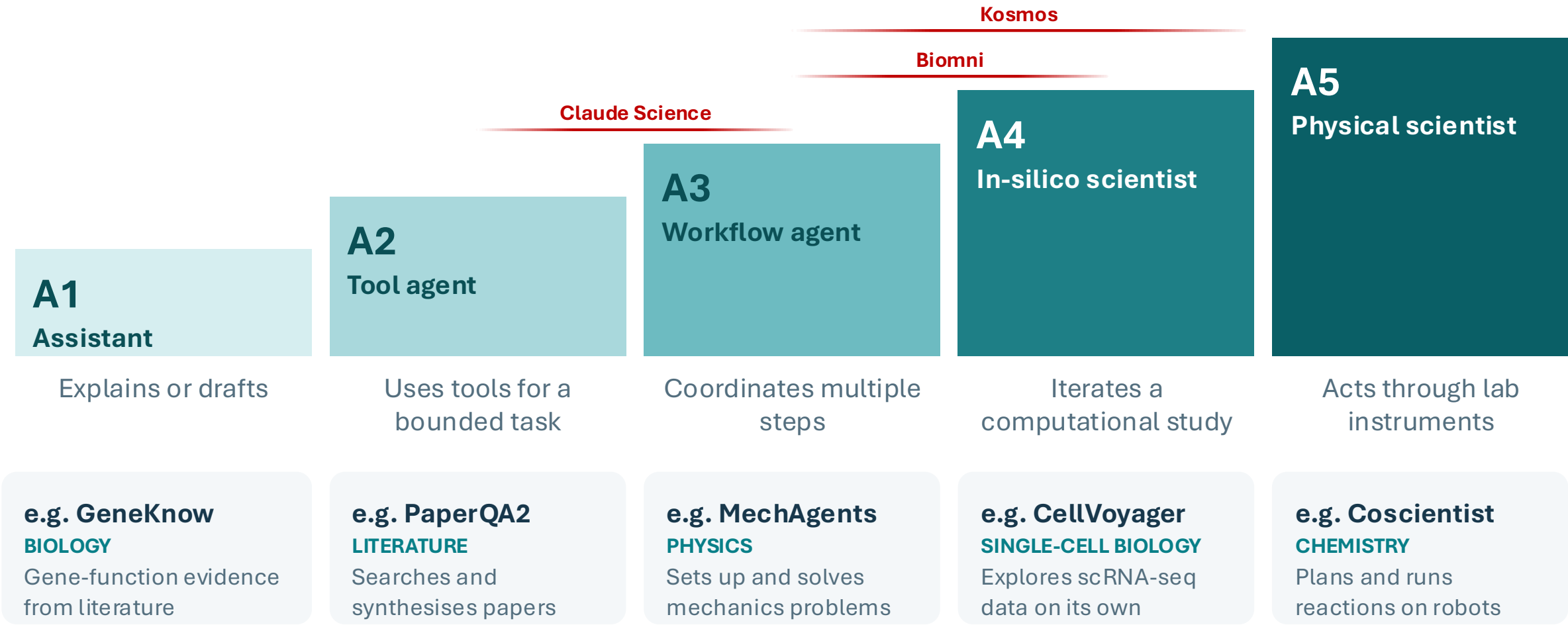
Capabilities and the limits of reliance

Michaela Liegertova

Institute of Molecular Genetics of the Czech Academy of Sciences (IMG CAS), Czech Republic

Five levels of autonomy, from drafting text to running a lab

MORE ACTIONS THE SYSTEM TAKES ON ITS OWN →



Levels and examples from my Autonomous STEM Agents index (michaliev.github.io/autonomous-stem-agents-wiki, v2.2.3). One example per level; many systems span two levels.

Two ways to work with agents

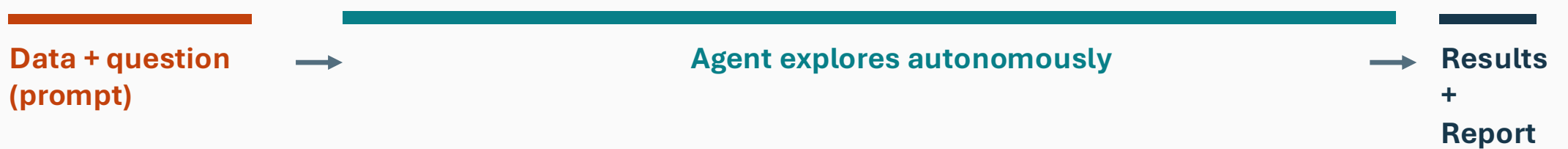
ITERATIVE RUN

you steer at
checkpoints



DELEGATED RUN

within an agreed
brief



■ Researcher

■ Agent

INTERACTIVE: steer the investigation

- + You steer key decisions
- Could miss leads outside your own framing

DELEGATED: useful for open exploration

- + Could surface leads you would not have asked for
- Could chase wrong directions into dead ends and miss important leads

Agents widen the search, but their shortcuts can stay hidden

WHAT AGENTS ADD

Broad hypothesis mapping

Many comparisons and explanations tested in one run

Speed and flexibility

Weeks of analysis in days; methods and questions can change mid-run

Leads you would have missed

Connections across data and literature you did not think to check

WHERE THEY CUT CORNERS

Could work from abstracts only

Full-text access varies by platform; paywalled papers are mostly out of reach

Could pick the outdated or wrong method

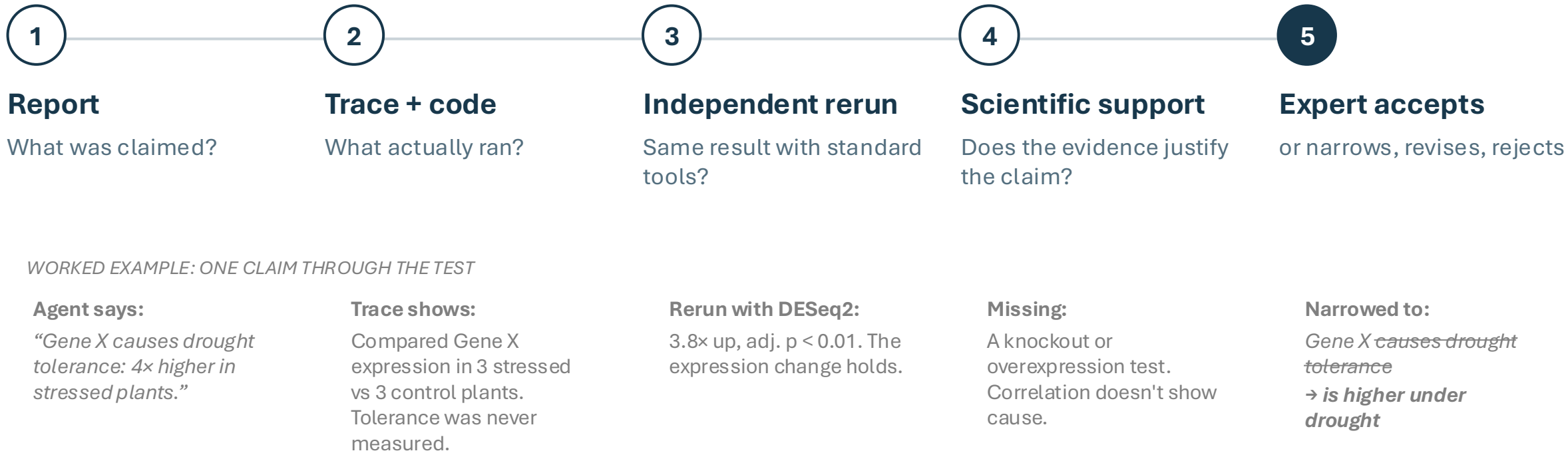
Default tools and code that runs, not necessarily the best fit

Writes fluent, plausible reports

Final reports read like an almost finished paper

Treat every agent output as a claim to test

FROM AGENT OUTPUT TO ACCEPTED RESULT



For important claims: trace the evidence, check key results and judge what they support

Where your data go ...

SETUP	CODE + FILES RUN	MODEL REASONS	WHAT LEAVES	FITS DATA THAT IS
Managed platform Biomni Lab, Kosmos	Vendor's cloud	Vendor's model provider	Your files and everything derived from them	Public or already open
Your HPC + commercial model Claude Science; open-source Biomni with an API model	Your computer / HPC	Model provider (API)	Prompts: table excerpts, errors, code, figures	Unpublished, not personal check the product's data terms
Your HPC + self-hosted model Open-source Biomni + open-weight model (e.g. e-INFRA CZ)	Your HPC	Your HPC	Can keep data within approved infrastructure; more setup	Personal, clinical or under contract
Outside Inside your institution				

NOT ALL DATA ARE SENSITIVE
more control costs speed, flexibility and capability, so for open data a managed platform saves time, people and resources

Same question, different investigations

SAME INPUT

~4,500 draft genomes
+ community profiles
processed data only

BRIEF, PARAPHRASED

*“What is distinctive about
this microbiome, and what
could explain it?”*

BROAD DELEGATION

Agent chooses directions

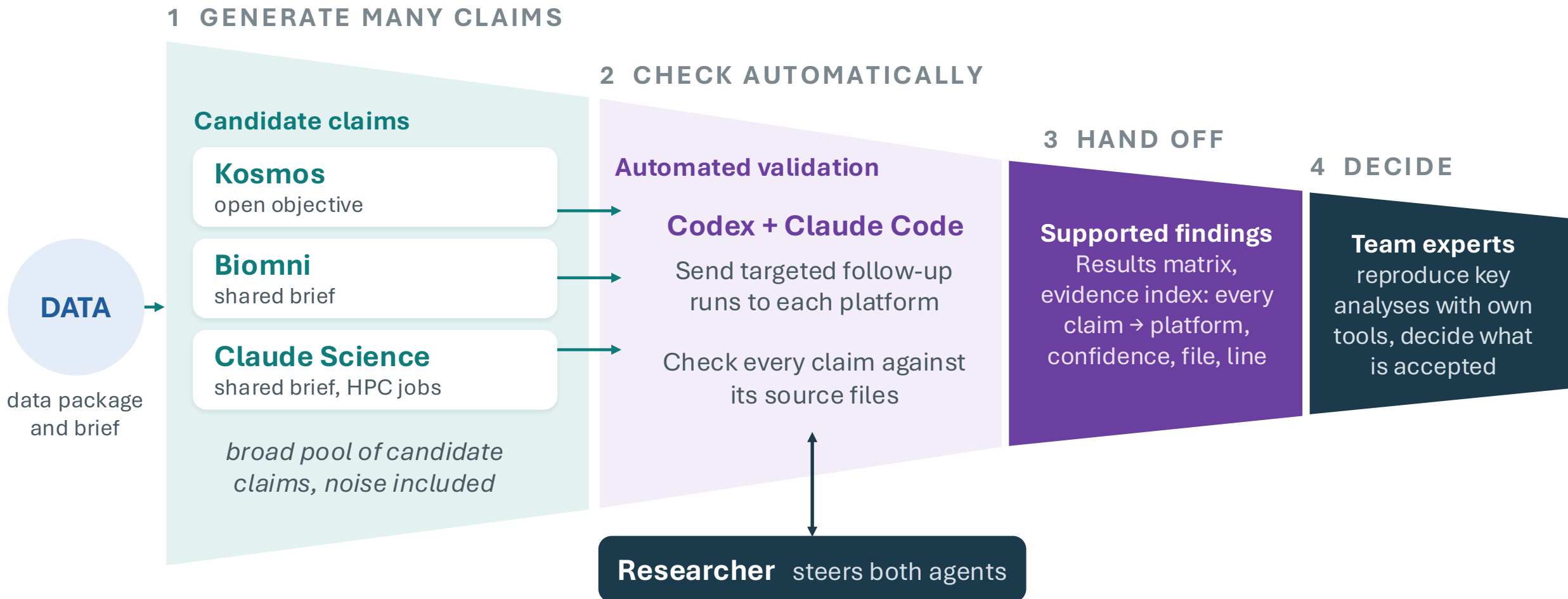
	PLATFORM · FOCUS	METHOD	EXAMPLE ANALYSIS SIZE	SELECTED CLAIMS REVIEWED
	Biomni Lab Broad screening	Screened genomes for gene clusters	780 genomes	 2 kept with limits 4 revise
	Claude Science External comparison	Compared genomes with public relatives	501 clusters + 1,021 public	 5 kept with limits 1 revise
	Edison Kosmos Ecological strategies	Grouped genomes by metabolic traits	219 genomes	 3 kept with limits 2 revise · 1 skipped

Each tool contributed useful findings

Each also produced conclusions that needed revision

Experimental setup:

automated validation reduces a broad claim pool to what experts need to judge



3 platforms **4** follow-up runs **188** claims · 1,875 source files traced

One workflow, swappable blocks: the ELIXIR ecosystem runs most of them

THE SLIDE 8 WORKFLOW AS BLOCK TYPES

BLOCK	Data + access rights	Compute + models	Tools	Agent platforms	Orchestration + validation	Evidence + provenance
IN MY PROJECT	team data package, data terms, accounts	MetaCentrum, vendor clouds and LLMs	GECCO, DIAMOND, skani, antiSMASH	Kosmos, Biomni Lab, Claude Science	Astra + Claude: follow-up runs, cross-audit	evidence index: 188 claims, 1,875 source files
ELIXIR / E-INFRA	Life Science Login (LS AAI), ELIXIR data resources	e-INFRA CZ / MetaCentrum, open-weight LLM API	bio.tools, BioContainers / Bioconda	FEASIBLE TODAY open-source Biomni + e-INFRA model	FEASIBLE TODAY coding agent on the e-INFRA model API	RO-Crate (run provenance), WorkflowHub

Agents are one block; trust comes from the blocks around them, and ELIXIR already runs those

Open data are mined first by whoever has the best workflow; the community that opens data should be able to mine it too



Generation is cheap.

Validation is the bottleneck.

Explore the Autonomous STEM Agents index



THE BIOLOGICAL BUILDING BLOCKS

Biological Foundation Models 486

Pretrained models, platforms, and benchmarks across sequences, structures, molecules, cells, tissue, and multimodal biology.

v2.3.0 · updated 10 Oct 2026

[Open the index →](#)



THE RESEARCH SYSTEMS

Autonomous Science Agents 542

Scientific agents, self-driving laboratories, simulation systems, and evaluation harnesses with explicit autonomy and access labels.

v2.3.0 · updated 10 Oct 2026

[Open the index →](#)



THE PRACTITIONER TOOLS

Coding & Data Agents 71

Privacy-first AI coding and data-analysis tools matched to data sensitivity, deployment control, and practical research work.

v2.3.0 · updated 10 Oct 2026

[Open the index →](#)



Michaela Liegertova
IMG CAS, Czech Republic
michaela.liegertova@img.cas.cz