

AI4S FUTURE SCIENCESKILLS HACKATHON · FINAL

GLOBAL GEOCHEMICAL ATLAS

从分散测定到可审计图谱

一套可追溯、可验证、如实报告局限的地球化学研究 Skill

191,715 条测定

29 个正式来源

15 项核心产物

TEAM 0066

源与置信度说明

异常区域识别结果

组采样点
2,664

候选异常
3,351

核查。

介质 (岩 / 土 / 水 / 沉积物)



全部介质

空间域



全部空间域

地图视图



二维世界地图

地图表达



组合核查 · 密度 + 样点

视图

重置全球

拖动/方向键平移 · 滚轮/



REAL RUN · 2026-08-15

数据很多，可信 workflows 很少

01

来源分散

数据散落在数据库、文献附表与 DOI 仓库中，字段规范互不一致。

02

测定值不可直接合并

介质、计量单位、报告基准、分析方法、检出限与坐标质量共同决定可比性。

03

图谱容易被过度解读

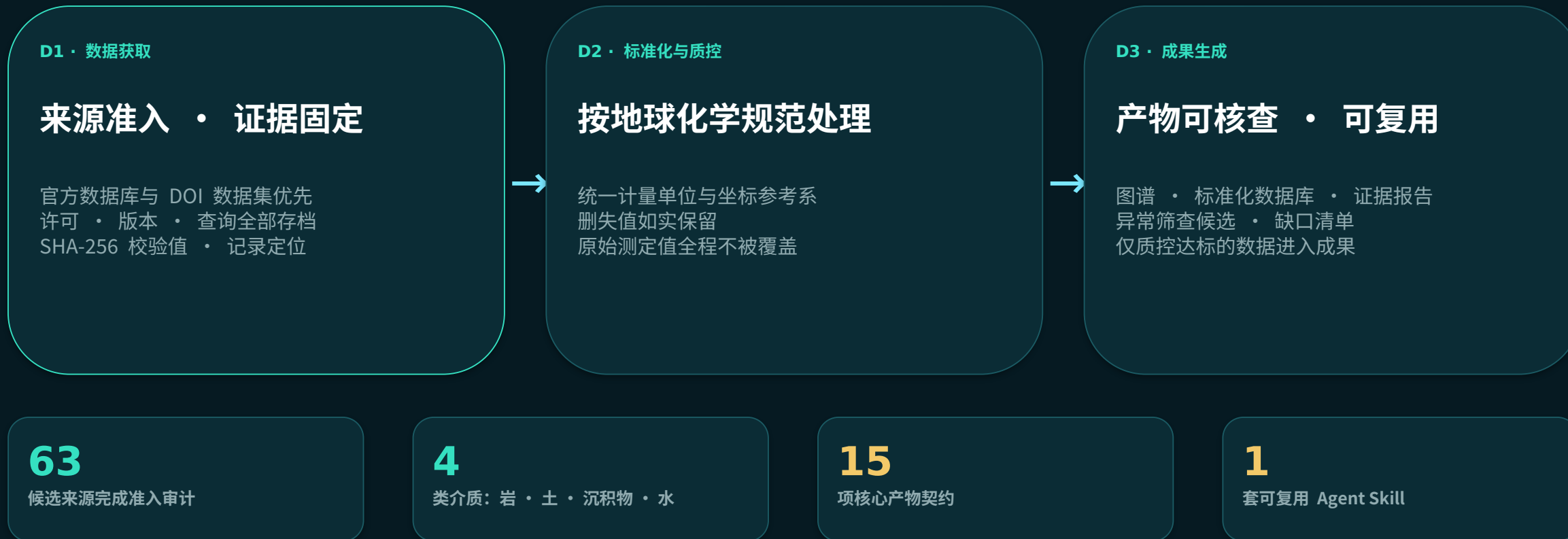
采样空洞、方法差异与背景样本不足，会制造貌似合理的结论。

研究请求

Cr（铬） × 中国东部 × 土壤 → **图谱 + 数据库 + 证据链 + 筛查候选 + 缺口清单**

目标：研究者不只拿到结果，还能核查它为何可用、哪里不可用、下一步如何补证。

计划：从请求到成果的三阶段 workflows



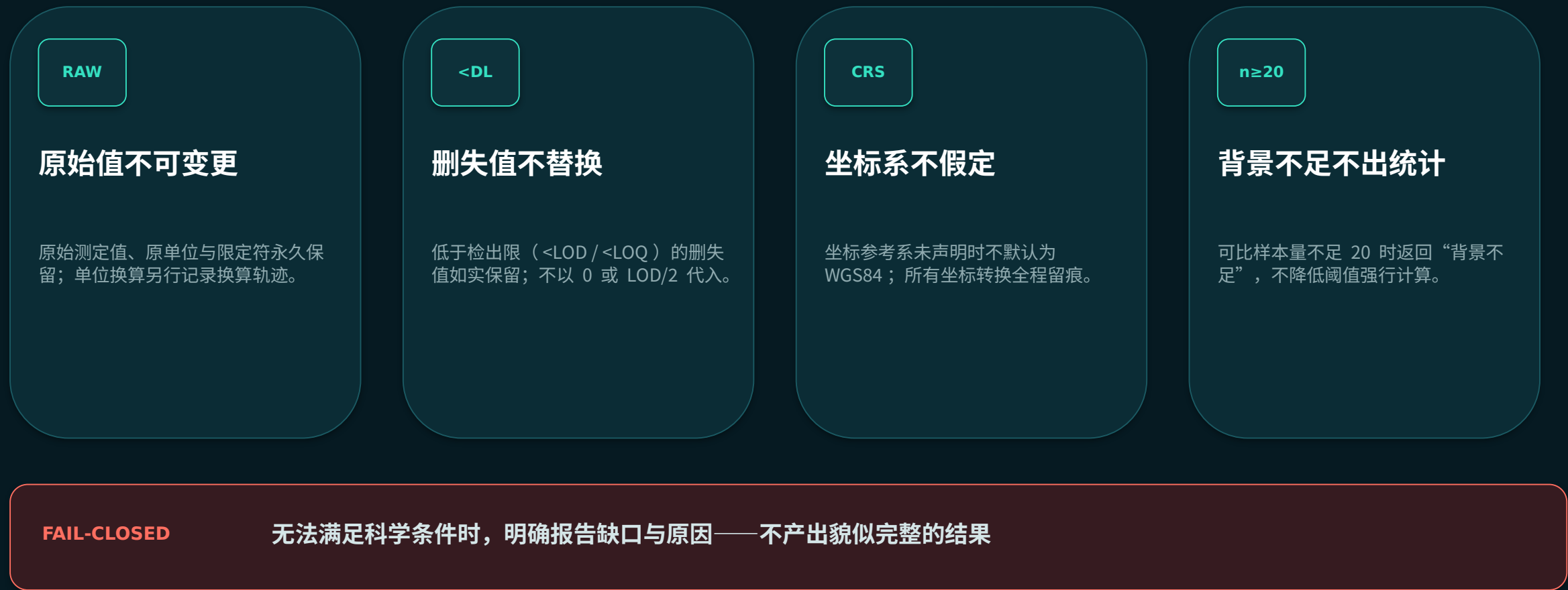
每一阶段都有输入、输出与质量门禁；模型不猜测科学元数据，缺失就如实报告。

设计：模型负责编排，计算交给确定性脚本



职责分离：模型不参与数值计算；单位换算、坐标转换、质控与异常判定全部由确定性脚本执行。

设计：四条硬约束，由执行器强制生效



失败是结构化产物的一部分，而不是被隐藏的异常。

真实运行： 8 月 15 日一次全球尺度自主采集

真实运行日志 · 2026-08-15 · 全程自主 · 在线获取公开数据

\$ python3 run_self_correction_loop.py --request ./REQUEST.json \
--online-source auto --analysis-profile production

- 【来源审计】** 63 个候选来源完成准入审计， 29 个成功入库
- 【第 1 轮】** 获取 72,346 条测定； 2 个端点持续 HTTP 500 ， 如实记为失败
- 【第 2-4 轮】** 增至 193,655 条；按空间覆盖缺口调度下一轮来源
- 【第 5 轮】** 195,505 条；主要来源已达容量上限，不重复计数扩量
- 【第 6 轮】** 新增测定为 0—— 满足零增量终止条件
- 【第 9 轮】** 终止采集，发布 191,715 条测定
- 【终态】** needs_human_review： 238 项缺口进入修复队列

在线生产运行 · 全程自主

自主终止

零增量后停止采集，不以重复记录扩充数据量

191,715

条测定 · 来自 29 个正式入库的公开来源

238

项缺口进入修复队列，逐项标注原因

俄罗斯、南美与岩石介质覆盖不足，如实保留为缺口，未以其他区域的数据量掩盖。

运行产物： 191,715 条测定构成的全球图谱

基于公开地球化学调查汇编的标准化测定数据；覆盖岩石、土壤、水系沉积物与水体四类介质。



实际使用：每条测定都支持记录级溯源

5. 含量与证据

地球化学元素图谱

《系沉积物与水体四类介质的公开标准化地球化学测定；支持按元素、介质与区域核查分布、含量与筛选异常。

布地图

标准化地球化学数据库

元素组合对比

数据来源与置信度说明

异常区域识别结果

前下载

数据库 CSV · 置信度报告 · 异常结果 · 运行报告

地球化学数据库

是同目录下的 geochemistry.csv；页面预览仅使用当前 HTML 内嵌的可上图记录，不会丢弃完整 CSV 中的无坐标、坐标 QC 失败或区域范围外记录。

语义

样品的一项元素测定；同一样品测多种元素时占多行，统计采样点数应使用“物理采样点”而非行数。每行同时保留原始值与标准化值（含单位），并关联样品类型、测量基准、地质单元、分析方法、数据源与质量标记。

chemistry.csv

打开 QC 报告

分布

元素

自动（记录最多）

图：土壤 · n=2,108 · P10 6.33 · 中位数 15.8 · P90 34.8 mg/kg

标准化含量 (mg/kg, 对数刻度)

1.3e+3

元素 × 介质

颜色越深表示该元素

元素				
Pb				
Cu				
Zn				
Ni				
Cr	1,770	1,112	2,618	263
As	324	1,112	3,417	422
Hg	0	3,496	701	0

0.99

2.4

6.5

17

Pb 标准化含量 (mg/kg, 对数刻度)

原始测定值 → 标准化轨迹 → 来源定位 → 哈希校验

标准化记录

原始值 · 标准化值 · 计量单位 · 分析方法 · 质控标记

5. 含量与证据

地球化学元素图谱

《系沉积物与水体四类介质的公开标准化地球化学测定；支持按元素、介质与区域核查分布、含量与筛选异常。

布地图

标准化地球化学数据库

元素组合对比

数据来源与置信度说明

异常区域识别结果

前下载

数据库 CSV · 置信度报告 · 异常结果 · 运行报告

数据来源与置信度说明

有一个 0-1 的可用性评分，用于排序复核优先级；它衡量记录“好不好用”，不是正确概率、统计置信度或准确率。

可用性评分

个梯度加权平均得到（0-1）。低分记录优先进入人工复核，高分记录可直接用于制图与统计。它不是正确概率、统计置信度或准确率，也不能证明测量本身无误。

来源清单 JSON

全数据库评分分布

高 35,874 条 · 80.6% 中 8,573 条 · 19.2% 低 89 条 · 0.2%

全库平均分 0.87 / 1

空间精度 占15%

80

坐标是否有效、定位精度是否达标

100

质量控制 占15%

88

是否通过单位、检出限等自动检查

100

来源与许可

文献 DOI · 源文件行号 · 数据许可 · SHA-256 校验值

置信度按四个维度分别报告：来源证据 · 分析就绪度 · 空间可用性 · workflow 完备性

异常筛查：一条真实候选的完整判定过程

真实样例 · 全球运行产出的 3,351 个筛查候选均经由同一判定链路

输入 · 测定值

土壤样品的砷（As）含量

23.9 mg/kg

美国阿拉巴马州 · A 层土壤样品
样品编号 C-328917

判定 · 背景比较

仅在可比分组内比较

45

条可比记录：同介质 · 同方法 ·
同地质单元

2.30

可比组中位数（mg/kg）

10.4 倍

相对可比组中位数的富集程度

z = 3.69

稳健 Z 分数，超过 3.5 筛查阈值

判定结果：高值筛查候选

输出 · 证据链

判定依据全部可回溯

分析方法

氢化物发生 - 原子吸收光谱（HG-AAS），与比较组一致

数据来源

USGS DS 801（DOI 10.3133/ds801） · 源文件第 27 行

综合置信度 0.79

同时标注：该记录缺少坐标不确定度声明

系统仅报告：高值筛查候选

污染、矿化或成因判断需结合采样设计与独立验证，由研究者做出——筛查不构成结论。

运行验收：成果与局限，一并如实报告

191,715

条测定实际入库 · 9 轮无重复计数

0

条岩石介质测定 · 上游持续不可用，如实标注

22.3%

5° 陆地网格覆盖率（ 135/605 ）

238

项缺口进入修复队列

0 错误

15 项核心产物独立校验

delivery_ready: false

如实报告局限，是科研可信的前提

零增量后自主终止；交付回执明确标注“未达交付标准”，每项缺口附原因与修复建议——这比一张貌似完整的图谱更可信。

以上数字均可在运行产物的结构化报告中逐条核对。

15 项核心产物全部通过独立校验 · 跨环境重跑，结果逐字节一致

让每个地球化学点 可追溯 · 可验证 可重新计算

把分散的公开地球化学数据，整编为 Agent 可执行、研究者可复核的科研基础设施。

191,715

条测定

29

正式来源

15

核心产物



体验在线产品

asimfish.github.io/.../demo



检查代码与证据

github.com/asimfish/.../skill

GLOBAL GEOCHEMICAL ATLAS

TEAM 0066 · AI4S FINAL

谢谢

15 项核心产物如何对应赛题交付

交互图谱

`interactive_map.html` · `samples.geojson`

筛选、热力、样点、异常筛查候选；自包含单文件，无外部依赖

标准化数据库

`geochemistry.csv` · `batch_acceptance.csv`

原始值、换算轨迹、报告基准、分析方法、批次与质控

来源与置信度

`sources_and_confidence.json` · `source_manifest.json`
`record_evidence.jsonl` · `confidence_report.json`

URL/DOI、许可、版本、SHA-256 校验值、源记录定位；置信度分维度报告

异常筛查结果

`anomalies.geojson` · `anomaly_report.json`
`anomaly_regions.geojson` · `spatial_anomaly_report.json`

记录级稳健 MAD Z 分数 + 空间富集（精确超几何 + BH-FDR）；判定边界完整报告

质量与运行证据

`qc_report.json` · `batch_qc_report.json`
`iteration_backlog.csv` · `run_summary.json`

逐项质控、失败批次不静默删除、修复队列与产物哈希

来源目录： 63 个候选完成准入审计， 29 个正式入库

岩石

GEOROC · 太古宙克拉通 · 南极板内火山岩

土壤

USGS DS801 · GEMAS · FOREGS 表土 / 底土 / 腐殖质 · AfsIS · PANGAEA 北非 · TPDC 中国

沉积物

FOREGS 河流 / 洪泛平原 · GSJ 日本陆地 / 海洋 · NGSA · MarChem · PANGAEA · 长江 / 黄河

水体

FOREGS 河水 · GEMStat 全球内陆水 · GEOTRACES IDP2025 海水 · WQP 萨克拉门托河

EarthChem 等联邦检索站只作发现层；只有追溯到原始记录后，才可作测量证据。

评委常问：能否把结果直接当科学结论？

完全离线能用吗？

能。交互图谱与核心产物均为自包含文件，浏览无需网络、密钥或 GPU；数据采集阶段按需在线访问公开数据库。

needs_human_review 是失败吗？

不是。它表示系统自主完成采集，并如实报告 238 项缺口待人工复核——不把未达标的结果冒充交付完成。

筛查候选就是污染或矿床吗？

不是。它只是高值筛查候选；自然背景、采样偏差、方法差异与人为输入都是竞争解释，结论需独立验证。

模型会不会改科学数值？

不会。模型只负责编排与解释；单位换算、坐标转换、质控与异常判定由确定性脚本执行，可逐字节复现。

一句话边界

图谱用于筛查、比较与定位证据；成因结论需要独立的采样设计、分析质量与地质证据。