

GEO 实践讲义

从问题采集到有据可查的内容改写

GEO 高级研修课项目组

中文译本·版本 1.0·2026 年 9 月 8 日

摘要

本讲义介绍生成式引擎优化（GEO）的实操流程：界定用户问题、保存回答、调查信源、核验主张、改写有用页面，并透明计算观察指标。十次工作坊围绕同一个虚构校园自行车服务案例展开，学生使用本地网页、表格和简短文字模板完成任务，无需编程、付费账号或 API 密钥。每次任务均提供示例、明确步骤、预期成果、评分标准和错误恢复方法，重点训练可供他人检查的证据工作。

适用对象

面向高职、大专、成人初学者，以及首次学习数字内容证据工作流程的从业者。需要会建立文件夹、打开网页、复制文本和编辑基础表格单元格；附录提供简明表格操作指南。

组织形式	时间	最终成果
五组循环，十次工作坊	10 × 90 分钟，共 15 小时	一份报告与可复查证据包

关于案例与双语版本

松风校园自行车服务站、青禾职业学院及全部相关记录均为虚构。所有问答和测量记录专为教学编写，没有通过商业 AI 系统生成这些结果。中文版译自英文主稿，编号、价格、计数、图示和评分口径保持一致。案例原始服务语言设为中文，因此英文样例答案是翻译，不是独立的语言测试。

课程目录

1 建立项目与事实底稿	6
2 围绕用户决策整理问题库	8
3 保存问答与引用，保持原始记录	10
4 查清每个信源背后的发布者	12
5 把答案拆成主张，逐条核验证据	14
6 在事实不变的前提下改写 FAQ	16
7 检查网页，并说明检查边界	18
8 用表格计算重复观察指标	20
9 同伴复核证据，安排改进优先级	22
10 交付一个别人能重新检查的项目	24
A 使用练习工作簿	26
B 编号、标签与提交文件	28
C 独立检查参考答案	30
D 教学选择与下一步	31
E 参考文献与图像说明	32

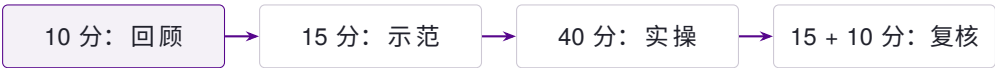
第一次课前

工作坊如何衔接

课程分为五组，每组两次课。每次形成一个可供后续使用的小成果。前期提供较充分示例，后期要求自己作出选择并解释。最后一次课将方法迁移到另一种校园服务。

循环	工作内容	任务	成果
1	界定项目，收集问题	P01-P02	项目卡与冻结问题库
2	观察答案，调查信源	P03-P04	观察记录与来源地图
3	核验主张，改写内容	P05-P06	核验表与 FAQ V1
4	检查网页，重复测量	P07-P08	页面检查与指标汇总
5	复核证据，项目交付	P09-P10	同伴复核与证据包

每次 90 分钟



最后十分钟独立完成离堂小结。

图 1 稳定的课堂节奏让操作可预期，证据判断则逐步加深。

尽量两人一组。操作员编辑文件，核对员要求查看相关来源并记录不确定项；每次交换角色。单人学习可延迟自查，但应明确标为自查。课堂交接检查不能被称为外部审计。

打开素材

完整解压学生 ZIP，用现代浏览器打开 START.html，用 Excel、WPS 或 LibreOffice 打开 GEO-practice-workbook.xlsx。输入前先另存工作副本。保持 pages 与 data 的目录结构，避免本地链接失效。templates 内有十份 Markdown 表单，可用文本编辑器填写，也可将表格复制到文字处理软件。

贯穿案例

一个细节充分、便于核查的小服务

常规换内胎合计 33 元，其中工时 15 元、材料 18 元；服务站周日休息，不提供上门取车；通常 30-60 分钟的维修时间带有限定条件。这些日常细节让学生无需专业领域知识，也能练习有实际意义的纠错。

GEO PRACTICE LAB · 离线教学材料

松风校园自行车服务站 · 虚构案例

价格与维修细则

松风校园自行车服务站 · 生效 2026-09-01

页面编号 s02 · 价格单位：人民币元。日期与适用条件必须一起记录。

服务	工时或服务费	材料与条件
基础安全检查	20 元/次	含刹车、车铃与轮胎检查
常规更换内胎	15 元/次	内胎另计 18 元/条；合计 33 元

维修时间

常规换胎通常需要 30-60 分钟，具体取决于排队、型号与零件库存。上述区间不是保证完成时限。

复检范围

同一维修项目在 7 日内可免费复检。免费复检不等于无条件换新，外力损伤不适用。

适用时间：2026 年 9 月 1 日起。价格变化以本页后续公告为准。

本页全部机构、人物、价格和事件均为原创虚构教学数据。页面日期为样例字段，非现场采集日期。
[返回练习素材目录](#)

图 2 所附本地教学页 s02-prices.html 的实际截图，所有数值均为虚构。阅读时要把工时费与材料费合在一起。

案例有六份来源文档，却不是六次独立确认。服务商拥有两页，推广资讯重复其资料；记者独立到访过地点，论坛经历较旧，公共机构指南只给一般建议，没有对服务站背书。

将研究问题落到实操

为什么要把答案与信源一起保存

GEO 奠基论文将生成式引擎抽象为问题改写、搜索、摘要和回答生成等环节 [1]。这一抽象帮助我们识别可观察材料：用户问题、显示的答案和来源链接。它并不揭示某个现有产品的实际内部实现。

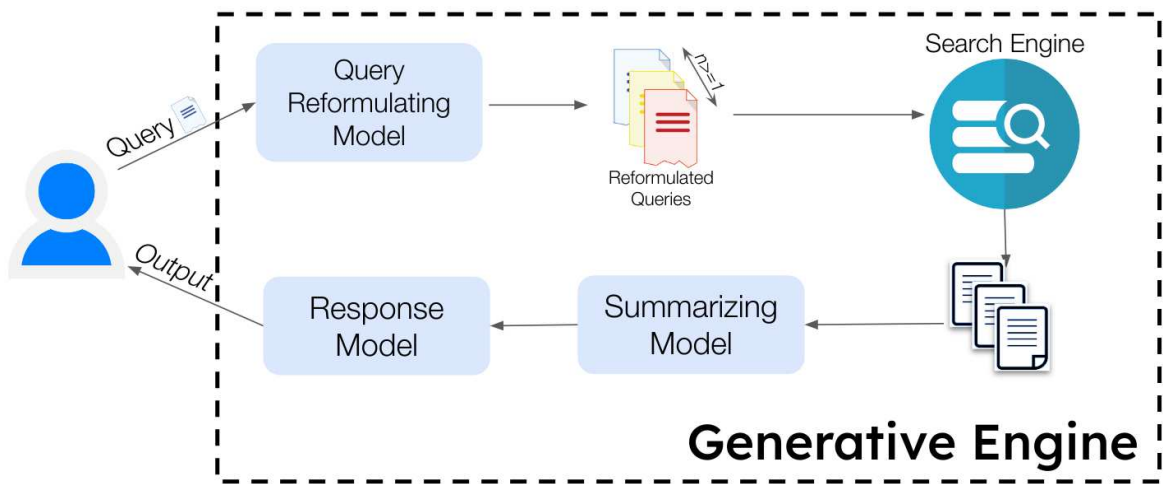


图3 Aggarwal 等, GEO: Generative Engine Optimization, 图2, arXiv:2311.09735v3, PDF 第3页 [1]。依CC BY 4.0 转载, 仅裁切原图边界, 保留原标注与结构。此为研究抽象, 不代表某个真实平台的内部机制。

四种可以检查的对象

对象	实操问题
问题	是否描述真实决策, 且不预设赞美?
答案	在什么条件下出现了哪些准确措辞?
引用	显示的链接实际打开哪份文档?
支持关系	该文档是否能证明这个具体主张?

GEO 实操可以改善清晰、有事实依据的来源内容, 也可以研究可见答案。有引用不保证获得支持, 页面整理好也不证明已被检索, 因此本讲义将采集、查源、核验与测量分别安排为任务。

1 建立项目与事实底稿

P01

90 分钟

完成后你能：限定一个服务问题，整理文件，并让每条事实都能回到来源。

本次交付：项目卡、八条核对后的事实，以及一张目录截图。

一名学生想修好自行车，同时避免超出预算或白跑一趟。你的小组要检查一段答案怎样描述松风校园自行车服务站。首先要明确：材料究竟说了什么。底层事实没有写清时，答案即使很漂亮，也难以判断是否准确。

本次使用的思路

事实底稿把一个说法与其适用范围、日期和来源一起记录。“15 元”并不完整：它只是工时费，内胎材料还需 18 元。营业时间、地点以及“保证”的含义，也应按同样的方法核对。



图 4 原件、工作稿与结论分开保存。箭头表示工作流程，不表示删除原文件。

输入材料

打开 START.html，再打开 s01-service.html 和 s02-prices.html。使用工作簿的 project、facts 页，以及 templates/01-project-card.md。保留素材包中的原始文件。

操作步骤

1. 完整解压学生素材包。在素材文件夹旁新建 submissions-group01，内部建立 originals、working、evidence、final 四个目录。把模板和工作簿复制到 working；原始素材不覆盖。
2. 从头到尾阅读两份官方说明。将价格、营业时间、服务范围的原句复制到事实底稿，保留单位与条件。不要把“通常”概括成“保证”。
3. 逐条核对 F01-F08 八个事实，记录来源编号与生效日期，并另记实际核对时间。无法确定的字段在备注中写“未说明”。
4. 填写项目卡：目标用户、具体场景、项目问题、范围、成员与分工。只围绕一个决策，例如“40 元预算不够完成一次常规换内胎”。
5. 请同伴任选两条事实，不向你询问，直接打开来源核对。补全缺失的链接和限定条件。这是在检查可追溯性，不是在考记忆。
6. 保存 project-card-v1.md 和小组工作簿，截取一张包含目录名称的截图。再写一句话：如果价格更新，怎样保留上一版记录。

P01 · 示例与验收

操作样例

项目	记录内容	判断理由
不完整的记录	15 元；修车便宜	缺单位、材料费、适用范围和日期。
核对后的记录	F02: 工时 15 元；内胎 18 元；合计 33 元	s02; 2026-09-01 生效；常规更换。
项目问题	40 元预算够不够常规更换内胎？	用户决策明确，答案可以核查。

预期结果

底稿有八个稳定编号。F02 写明常规更换合计 33 元；F03 写明周日休息；F06 写明不提供上门取送车。这些都是虚构课堂事实，不对应真实经营主体。另一位学员能找到每条记录所依据的原句。

交付与评分

项目卡、八条核对后的事实，以及一张目录截图。

维度	分值	验收证据
文件组织	20	保留原始材料；四个工作目录清晰。
事实底稿	40	八条记录齐全，来源、单位、日期和条件完整。
项目范围	25	一个用户、一种场景、一个可回答的问题。
同伴核对	15	独立打开两条来源，并记录修改。

常见错误与恢复

覆盖了原始文件 从 ZIP 恢复原件，把自己的稿件移入 working。

日期含义不明 区分生效日期、样例日期和实际观察时间。

凭空补齐未知值 写“未说明”；信息缺口本身就是发现。

独立检查

同伴写道“修车只要 15 元”。你会打开哪段来源？补上哪些内容后，这句话才可使用？

先作答，再看附录 C。写明使用的来源或规则，只有标签没有理由不算完整。

2 围绕用户决策整理问题库

完成后你能：清理小规模问题列表，同时避免只选择有利于本品的问题。

本次交付：十二条冻结问题、合并与剔除日志，以及意图覆盖检查。

原始问题中包含重复问法、没有指定服务主体的问题，以及预设好评的诱导问题。你的任务是整理出一组便于监测的问题。合适的问题库应覆盖发现服务、购买决策、实际使用和潜在误解。

本次使用的思路

保留会改变决策的差异，例如预算、星期几和服务条件；只换措辞、不换需求的问题可以合并。无品牌的发现型问题和指名品牌的问题都有价值，但它们衡量的内容不同。



保留原编号，记录每次处理。

图 5 小规模问题清理：一条探索问题和一条诱导问题移出固定集，原始行仍保留。

输入材料

打开 queries 工作表，R01-R16 是原始问题。使用 templates/02-query-log.md。这份问题列表是虚构样例，不是真实学生调查，也没有实际出现频率。

操作步骤

1. 编辑前先通读全部问题。标注发现、决策、使用、风险四类意图。先判断：用户得到答案后，会据此做出什么行动。
2. 比较 R01 与 R02、R05 与 R06。每组保留前一个编号，将后一个记为合并，并写明理由，不要直接删行。
3. 剔除 R16，因为它要求赞美且预设答案。R04 暂放探索问题区：缺少上下文，无法明确服务主体。这样得到十二条固定问题。
4. 按原始顺序给保留行分配 Q01-Q12。检查对应关系：R03 为 Q02，R05 为 Q03，R09 为 Q06。与所附回答样例对照时，既核对编号，也核对完整问题。
5. 另存一份冻结版，保留完整措辞，不删掉“含材料”或“一定”等有意义的条件。制作四行意图覆盖表，分别统计问题数。
6. 让同伴指出一个尚未覆盖的需求，将它放入探索区，不直接修改 v1。保存变更日志，并截取包含编号、问题、意图的表格截图。

P02 · 示例与验收

操作样例

项目	记录内容	判断理由
合并	R01 与 R02	地点和刹车检查需求相同，只是措辞不同。
都保留	R05 与 R07	周日是否营业与周六何时关门对应不同决策。
移出固定集	R16	预设赞美，不能公平检验真实需求。

预期结果

结果为十二条保留问题、两条合并重复、一条缺上下文的探索问题、一条被剔除的诱导问题；原始十六行仍可查看。保持所附编号映射，连接问题与回答时同时核对完整文本。探索问题不混入固定测试。

交付与评分

十二条冻结问题、合并与剔除日志，以及意图覆盖检查。

维度	分值	验收证据
清理决策	30	重复组正确，剔除原因明确。
问题措辞	30	十二条问题自然，保留关键限定。
意图覆盖	20	检查四类意图，并承认未覆盖需求。
版本与映射	20	保留原始行；冻结 v1；编号对应清楚。

常见错误与恢复

把所有价格问题合并 预算和是否含材料会改变决策，应保留。

加入赞美前提 改为中性问法，保留实际需求。

测试中途改题 另存新版本，分开分析。

独立检查

“周日开门吗”和“周六几点关门”应该合并吗？用一句话解释。

3 保存问答与引用，保持原始记录

P03

90 分钟

完成后你能：完整保存答案、观察条件和显示的引用，并将自己的判断另行记录。

本次交付：六条观察记录、五份完整回答原文，以及一条保留的失败记录。

本地问答页包含正确说法、错误说法和一次失败。有些答案显示引用；有一条提到了服务站却没有引用。首先要忠实保存所见。纠正内容应放在独立标注中，不能改写原始回答。

本次使用的思路

一条观察记录对应一个问题、一组明确条件、一个时间与一次重复。提及和引用是不同标签。有链接不代表链接支持旁边的句子；没有完整回答的失败记录，结果字段应留空。



图 6 先保存，再标注。有引用与证据支持是不同变量。

输入材料

打开 <pages/answers.html>。使用 observations 页；为 A01-A06 分别复制 <templates/03-observation.md>。英文网页是固定中文样例的翻译，并非实际英语平台测试。

操作步骤

1. 记录文件名、样例语言、虚构样例日期，以及自己实际操作时间。这些字段分开填写。数据来源写“离线手工编写样例”。
2. 从 A01 开始，复制完整问题和全部回答，包含所有限定条件。按显示顺序记录引用编号，逐个打开本地来源链接，写下目标文件。
3. 将原文保存为 A01-original.md，截图同时包含问题、回答和引用。长答案使用连续编号截图，并保留完整文本。
4. 完整回答中出现松风时，提及记 1；未出现时记 0。只要显示来源链接，有引用记 1，暂不判断是否正确。正确性留到主张核实时处理。
5. 依次处理 A02-A05。A06 保留 status=failed 和故障说明，提及、有引用两列留空。不要反复重试，直到把失败从样本中“消掉”。
6. 逐条与浏览器页面核对。同伴仅凭记录，就应能还原问题、完整答案、来源目标和标注口径。

P03 · 示例与验收

操作样例

项目	记录内容	判断理由
A03	提及 1；有引用 0	提到了松风，但没有显示来源。
A05	提及 0；有引用 1	安全建议引用 s06，却未提到松风。
A06	failed；两项结果留空	未取得完整回答，不是一次品牌未出现。

预期结果

六条记录中有五条成功回答、一条失败。五条有效回答中，四条提到松风，四条显示至少一个引用。这是两个不同计数，不能解释为四条答案正确，更不能解释为有四个独立信源支持该服务站。

交付与评分

六条观察记录、五份完整回答原文，以及一条保留的失败记录。

维度	分值	验收证据
原文采集	35	保存五份完整回答、问题和来源目标。
元信息	25	区分样例日期与实际时间，记录条件。
标签	25	提及、有引用、失败口径一致。
证据文件	15	截图可读，文件名与编号匹配。

常见错误与恢复

顺手纠正原文 恢复观察到的措辞，把修订写入核验表。

只交截图 增加可复制文本，便于检索措辞和编号。

失败记为 0 保留记录、结果留空，并单独报告失败数。

独立检查

答案里有一个链接，但该页面根本没有提到服务站。能同时记“有引用 =1”和“证据不足”吗？

4 查清每个信源背后的发布者

P04

90 分钟

完成后你能：区分自有说明、独立报道、转载推广与个人经历。

本次交付：六个信源的来源地图，以及一段独立印证说明。

答案可能列出多个链接，却都在重复同一份服务商公告。如果把这些链接算作多次独立确认，就会夸大证据。把页面当作“证人”之前，应先查清发布者和信息起点。

本次使用的思路

服务商适合说明自身当前价格政策；独立记者可以核验自己到访过的地点。两者都不自动证明“技术最好”。信源是否合适取决于待核查的主张，独立性与是否支持该主张也要分开判断。



图 7 s04 重复服务商资料。s03 提供独立地点观察，证据范围限于那次探访。

输入材料

从 START.html 打开全部六个信源，使用 sources 工作表和 templates/04-source-map.md。查看发布者、日期和利益关系说明。页面故意采用一致设计，避免把视觉精美误当作可信。

操作步骤

1. 逐页找到发布者和发布或更新日期，抄录有关采编、付费、资料提供的说明。未说明的关系应记录为不确定。
2. 将 s01、s02 标为自有信息：它们是同一服务商的两份文件，不是两个独立机构。分别写明可支持哪些当前服务事实。
3. 比较 s03 和 s04。前者说明独立探访，后者说明资料由服务商提供且未独立测试。在来源图中画出 s04 指向 s01 的关系。
4. 查看 s05 时同时看日期。它只记录较早的一次个人经历。保留这段历史，但当前营业时间和价格仍应与更新的证据对照。
5. 只阅读 s06 的安全指南正文，从“出发前做三项检查”到末尾安全说明，判断正文是否推荐松风服务站。排除案例页眉和教学提示；这些位置出现名称不等于正文作出推荐。
6. 写四句话概括来源：当前政策依据、独立地点核对、转载关系、尚待处理的历史冲突。每句话都附相应来源编号。

P04 · 示例与验收

操作样例

项目	记录内容	判断理由
当前价格政策	s02	适合说明本服务主体当前如何收费。
地点独立印证	s01 + s03	服务商说明与独立现场探访一致。
不能新增的证人	s01 + s04	转载不等于增加了一次独立核验。

预期结果

完成六份文档的来源地图。s01、s02 同属一个主体；s04 回溯到服务商资料；s03 独立报道一次探访；s05 是较旧个人经历；s06 仅讨论通用安全。结论应对应具体主张，不是给六个网站排一个通用可信度名次。

交付与评分

六个信源的来源地图，以及一段独立印证说明。

维度	分值	验收证据
发布者与日期	25	六个来源均依据可见信息记录。
独立性	35	识别 s04 的原始资料关系及 s01/s02 同主体。
适用范围	25	不虚构学校背书或质量排名。
来源小结	15	四句说明有来源编号，并指出不确定项。

常见错误与恢复

第三方就等于真实 继续核对证据原句、日期和采编方法。

网址多就等于证人多 追溯转载内容的原始来源。

公共机构就意味着背书 核对正文究竟有没有评价这个服务主体。

独立检查

同一新闻稿出现在五个网站上。还需什么证据，才能称它们为五次独立确认？

5 把答案拆成主张，逐条核验证据

P05

90 分钟

完成后你能：将答案拆成可核查的主张，依据原文判断支持关系。

本次交付：八条主张核验、三条事实纠正，以及一条证据不足示例。

A03 虽短，却包含两个不同承诺：完成时限和上门取车。它们需要分别核验。本次检查五条成功回答中的八个主张，并解释为什么标为支持、矛盾或不足。

本次使用的思路

证据能够在原条件下证明完整主张，记“支持”；相关证据与说法冲突，记“矛盾”；现有材料无法确定，记“不足”。同一个句子里的两个分句，可能分别得到支持与矛盾的判断。

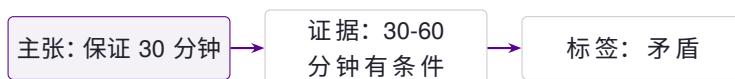


图 8 核验对象是包含限定条件的完整说法。

输入材料

打开 claims 工作表、answers.html 及 s01-s06，保持 P03 的回答原文不变。较长证据摘录和纠错内容填写在 templates/05-claim-check.md。

操作步骤

1. 在 A01-A05 中标出事实说法。如果连接词两侧可能有不同真值，就分开核验。至少完成所附 C01-C08 八个主张。
2. 逐条打开相关来源，不依赖答案里显示的引用标签。摘录能够同时包含证据及必要条件的最小原文片段。
3. 核对数字和单位。C01 计算 $15 + 18 = 33$ ；C02 保留“没有额外维修项目”的条件。数字正确但条件缺失，仍会误导。
4. 核对 C03 的日期：当前官方页说明周日休息，较旧论坛经历不能确定当前营业时间。当前主张记为矛盾，同时保留历史记录。
5. 将 A03 的两个承诺分开检查。C04 与“通常 30-60 分钟，非保证时限”冲突；C05 与“不提供上门取送车”冲突。分别写出准确修订句。
6. 再增加一个测试主张：“松风全校最好”。检索全部六份来源。因为没有比较评价材料，记为不足；不要把缺少证据误当作反面结论的证明。

P05 · 示例与验收

操作样例

项目	记录内容	判断理由
C03：周日营业	矛盾	s01 写明周日休息；s05 时间较旧。
C04：保证 30 分钟完成	矛盾	s02 给出有条件的 30-60 分钟区间。
全校最好	不足	没有比较材料，也没有排名方法。

预期结果

依照案例中的现行政策日期，八个给定主张为五条支持、三条矛盾。额外的“全校最好”证据不足。这是可复查的课堂答案口径，不是商业 AI 平台准确率。好的纠错应把措辞收窄到证据确实能证明的范围。

交付与评分

八条主张核验、三条事实纠正，以及一条证据不足示例。

维度	分值	验收证据
最小主张	20	八条独立主张均关联原回答。
证据摘录	35	原句相关，限定条件保留。
标签与纠错	30	三处矛盾修正，不添加新承诺。
不确定性	15	“最好”保持不足，并记录依据。

常见错误与恢复

整段只标一个标签 拆开主张，以表达一段中的不同支持关系。

信源权威就支持一切 仍要核对具体句子和适用范围。

没证据就判假 没有相关反证时，应标为不足。

独立检查

来源说“7 日内免费复检”，能支持“7 日内任何损坏零件免费更换”吗？

6 在事实不变的前提下改写 FAQ

P06

90 分钟

完成后你能：写出直接有用的回答，并让价格、条件和来源保持可核查。

本次交付：四问 FAQ、事实锁定核对表，以及修改前后对照记录。

旧 FAQ 充满笼统赞美，价格也不完整。你要把它改成能帮助用户安排到店的页面。本任务改善的是可靠信息的可获取性，并不能证明搜索引擎一定会检索或引用改稿。

本次使用的思路

有用的服务内容先回答真实问题，再说明条件和来源。可以改变措辞与组织方式，但要锁定底层事实。独立报道只能印证它实际核查的事实，不能加工成并不存在的背书。



图 9 改写改变信息呈现方式，本身不证明 AI 可见性提升。

输入材料

打开 faq-before.html 和八条已核对事实，使用 templates/06-faq-draft.md。完成自己的初稿后，再打开 faq-after.html 作为自查参考。

操作步骤

1. 标出 V0 中每条事实和宣传说法，对应到事实编号，或标为缺少支持。“随时来”“总共 15 元”“30 分钟”“学校推荐”都需要处理。
2. 从冻结问题库选出四个实际问题：预算、周日到店、地点与取送、维修时长。用问题作标题，首句直接回答。
3. 紧接回答写出必要条件。15 元工时费与 18 元内胎费同时出现；时间说明保留排队、型号和库存限制。
4. 每个回答旁标注来源编号，让来源链接的含义清晰。写明维护角色、核对日期和更新规则，不虚构个人资质或专家姓名。
5. 在变更日志中把删除句、新句和事实编号配对。如果借助 AI 起草，只提供虚构的已核事实，并逐项检查它提出的说法。
6. 通读自己的稿件，再与参考修订稿比较。请同伴只凭你的页面回答四个问题。保存 V1 与事实核对记录，同时保留 V0。

P06 · 示例与验收

操作样例

项目	记录内容	判断理由
旧句	“换内胎只要 15 元。”	漏掉材料费。
改写	“工时 15 元，内胎 18 元，常规更换合计 33 元。”	对应 s02，不添加虚构优惠。
旧句	“学校权威推荐。”	删除：来源没有证明这项背书。

预期结果

V1 有四个问题、四段直接回答、相邻来源、核对日期和维护角色；事实与 F01-F08 一致。不含虚构评分、排名、评价或保证。参考文件只是一种合格写法，只要意思准确，可以采用不同的自然表达。

交付与评分

四问 FAQ、事实锁定核对表，以及修改前后对照记录。

维度	分值	验收证据
事实忠实	40	数字、营业时间、服务范围 and 条件与底稿一致。
读者可用	25	四个问题用平实语言直接回答。
来源与维护	20	来源邻近，日期、角色和更新规则完整。
变更记录	15	保留 V0，说明每项重要修改。

常见错误与恢复

添加漂亮的统计数 只使用实际测得且有来源的数字，否则删掉。

把条件藏到远处脚注 将限定条件放到受它影响的回答旁边。

把 AI 写得流畅当作核对完成 重开来源，逐条检查事实、单位和例外。

独立检查

本任务能够确认 V1 的什么性质？哪些结果还需要另行观察？

7 检查网页，并说明检查边界

P07

90 分钟

完成后你能：发现可见内容与链接缺陷，区分本地检查和真实抓取测试。

本次交付：八项页面检查、按优先级排列的修复建议，以及一次复查。

页面在浏览器中能打开，正文却可能错误，链接可能失效，元信息也可能指向别处。故障页让你不用改动真实网站就能观察这些问题。每一项都写下所见，也包括本地无法确定的内容。

本次使用的思路

分清三件事：人能不能读到页面、文档中实际有什么、远端系统究竟处理过什么。本地文件能回答部分内容问题，不能说明服务器 HTTP 响应、爬虫访问或引擎索引状态。



本地可检查前两项；实际抓取和索引需要其他证据。

图 10 三个不同检查层次；本离线练习未测试最右侧阶段。

输入材料

打开 `page-broken.html` 和 `templates/07-page-audit.md`，以 `faq-after.html` 作为内容修复参考。损坏图片和失效链接都是故意设置的练习缺陷。

操作步骤

1. 查看浏览器标签标题与可见标题。“测试页面”没有说明服务问题。将“15 元”和“每天开放”分别与 F02、F03 对照。
2. 尝试把关键段落复制到文本编辑器，记录正文是否可读。点击维修细则，查看缺失图片。先描述发生了什么，不要悄悄修掉原始证据。
3. 使用浏览器“查看网页源代码”，搜索 `title`、`robots`、`canonical`。不要求写代码，只需保存相关行的文本或清晰截图。
4. 观察 `noindex` 和指向 `example.invalid` 的 `canonical`。这些是文档设置，不能证明某个真实引擎已排除该本地文件。保留域名 `example.invalid` 在此故意不可用。
5. 按用户受影响程度排序：先处理价格与营业时间错误，再修详情链接，再完善标题和元信息。每项给出具体动作与验证方法，图片美观排在事实准确之后。
6. 复查 `faq-after.html` 的可见改进：价格、条件、来源链接和日期。明确填写“服务器状态、真实抓取与索引尚未测试”，不要把整个网站判为通过。

P07 · 示例与验收

操作样例

项目	记录内容	判断理由
价格不完整	“换内胎：15 元”	补充工时、材料及常规合计 33 元。
细则链接失效	not-here.html	改为 s02-prices.html，再次打开确认。
本地元信息	noindex；canonical 指向 example.invalid	记录设置，不推断线上索引结果。

预期结果

检查表区分已观察缺陷和未测试的服务器行为。能够找出价格不完整、每天营业的错误、图片与链接失效、标题笼统、缺更新时间和元信息问题。可在本地复查 FAQ 的内容修复，但不能据此确认真实平台反应。

交付与评分

八项页面检查、按优先级排列的修复建议，以及一次复查。

维度	分值	验收证据
可见检查	30	核对价格、时间、正文、图片、链接和日期。
源码检查	25	找到标题和元信息，证据可读。
修复建议	25	动作具体，按用户影响排序。
边界与复查	20	复查本地内容，服务器/抓取/索引保留未测试。

常见错误与恢复

能打开就说明已索引 只能报告本地文件可以渲染。

noindex 证明过去发生过排除 指令是设置证据，不是历史抓取事件证据。

源码截图字太小 只截相关行，或粘贴可复制文本并注明文件。

独立检查

双击本地 HTML 后，能报告“HTTP 200，已被 AI 搜索收录”吗？

8 用表格计算重复观察指标

P08

90 分钟

完成后你能：使用正确分母计算指标，报告波动，并避免把差异解释为因果。

本次交付：核对后的汇总表、两阶段比较图，以及结果边界说明。

测量表有 36 条合成记录：六个问题、各三次重复、两个阶段；每阶段都有一次失败。你要用筛选和公式复算汇总结果，并准确说明这份数据里发生了什么变化。

本次使用的思路

提及率等于提到服务站的有效回答数，除以有效回答总数；失败另行报告。百分点差异描述当前样本，不证明 FAQ 改稿导致真实平台发生变化。

输入材料

打开工作簿 measurements 与 summary 页，使用 templates/08-measurement-report.md。这里的“前/后”是合成数据中的教学分组，不是实际干预实验的采集时间。

操作步骤

1. 核对行数：前后各 18 条。检查每阶段有六个问题编号、轮次 1-3。筛选 status=failed，确认每阶段恰有一条失败。
2. 在“前”阶段筛选 status=ok，再数 mention=1，得到 17 条完整回答、8 条提及。“后”阶段为 17 条完整回答、11 条提及。
3. 查看 summary 公式。COUNTIFS 先按阶段与状态计数，再增加 mention=1 条件求分子，最后除以有效回答数。原始结果保持数值 1/0，失败留空。
4. 计算 8/17 与 11/17，设为一位小数百分比，再用未四舍五入的数值求差。报告提及率由 47.1% 到 64.7%，上升 17.6 个百分点。
5. 复算有引用率：两阶段都是 7/17，即 41.2%。提及率变化时，这个指标没有变化；有引用仍不等于证据支持。
6. 查看每题三次重复，而不挑最好的一次。保存汇总结果，写明三个限制：样本小且为合成数据、失败未进入结果分母、不能推出真实平台或因果结论。

P08 · 示例与验收

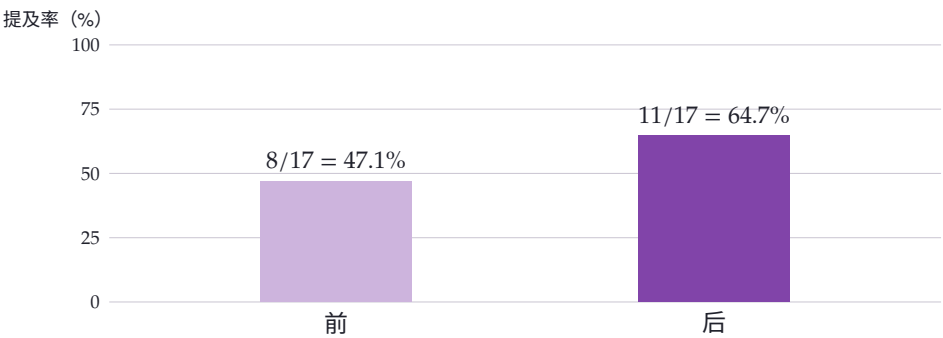


图 11 合成课堂数据。每阶段计划 18 次、失败 1 次，结果分母为 17。该图不是 FAQ 改稿效果的实际测量。

操作样例

项目	记录内容	判断理由
前阶段	8 / 17 = 47.1% 提及率	计划 18 次，失败 1 次。
后阶段	11 / 17 = 64.7% 提及率	计划 18 次，失败 1 次。
差异	17.6 个百分点	有引用率保持 7 / 17 = 41.2%。

预期结果

每阶段计划记录 18 条、有效回答 17 条、失败 1 条；提及数分别为 8、11，有引用数都是 7，缺失结果共两条。每个分子都能通过筛选原始行复算。两种语言版本使用相同编号和数值。

交付与评分

核对后的汇总表、两阶段比较图，以及结果边界说明。

维度	分值	验收证据
计数与缺失	30	36 条记录、两次失败，分阶段分母正确。
公式	30	按阶段和状态计数，输入变化时结果更新。
结果表达	25	同时报告计数、比例和百分点差异。
解释	15	明确合成样例范围以及缺少因果证据。

常见错误与恢复

直接除以 18 次尝试 按本任务口径除以 17 条有效回答，失败另报。

把 17.6 个百分点叫作增长 17.6% 说明单位；百分点是两个比例相减。

只报告最好的一次 保留所有计划重复，并查看波动。

独立检查

一次失败后来重试成功，能覆盖原始行并悄悄改变分母吗？

9 同伴复核证据，安排改进优先级

P09

90 分钟

完成后你能：用来源原句解决标注分歧，并将发现转化为具体行动。

本次交付：复核记录、简单一致率计数，以及一周改进清单。

两个人读同一个答案，可能产生不同判断。双方先独立记录标签，再讨论，才便于发现分歧。你要比较四个容易产生争议的主张，回到原文处理差异，并决定哪些问题先修。

本次使用的思路

一致率只说明标注规则的应用是否相近，不自动说明标注正确。优先级看能避免什么用户问题、需要多少工作，而不是承诺增加多少提及。营业时间错误的重要性，在于它可能让用户白跑。

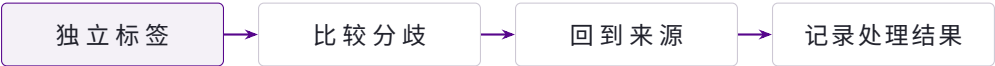


图 12 讨论前保存初始判断，不能用讨论后的一致替换初始一致率。

输入材料

使用 P05 标签、`templates/09-peer-review.md` 和来源页面。单人学习时，先隐藏原标签，完成另一任务后重新阅读，并明确记录为自查，而不是独立复核。

操作步骤

1. 选择 C02、C03、C04、C07。只把主张和来源文件给同伴，不给自己的标签。双方分别记录支持、矛盾、不足和一段证据原句。
2. 两套标签保存后再比较，统计四条中有几条完全一致。即使讨论后达成一致，也保留最初标签。
3. 每处分歧写明涉及什么规则：时间范围、限定措辞、信源独立性，还是主张拆分。重开来源，记录用来解决分歧的原句。
4. 材料仍无法判断时，保留未解决状态，说明还需要什么证据。不要为了提高一致率而强行给正向标签。
5. 列出三个一周内可做的动作，每项填写证据编号、受影响用户、责任人、期限和复查方法。错误价格、营业时间和虚构承诺优先于装饰性修改。
6. 请复核人检查一项已完成修订，填写复查记录。保留边界说明：本地复核通过，不代表平台已经处理了新内容。

P09 · 示例与验收

操作样例

项目	记录内容	判断理由
初始标签	四条中三条一致	讨论前的一致率为 75%。
分歧	“40 元预算够”	检查是否保留“无额外维修”条件。
具体行动	纠正 FAQ 的周日营业说明	负责人：编辑；对照 F03 与 s01 验证。

预期结果

复核表保留双方初始判断和讨论后标签。3/4 只是示例，要报告你们自己的实际计数。改进清单包含三个小而可查的动作；只要说明证据缺口，可以保留未解决事项。

交付与评分

复核记录、简单一致率计数，以及一周改进清单。

维度	分值	验收证据
独立初评	25	比较前保存标签，或说明单人自查路线。
分歧处理	35	对应判断规则与证据原句。
行动清单	25	三项动作关联证据，责任人和验收清楚。
如实报告	15	保留初始一致率和未解决项。

常见错误与恢复

先给对方答案 隐藏标签重新做，并记录新的过程。

一致就是正确 继续对照原文和规则核验两方判断。

只写“优化 GEO” 说明哪句话、依据何在、谁负责、怎样算完成。

独立检查

两位复核人都把无依据排名标为支持。100% 一致率能证明什么？

10 交付一个别人能重新检查的项目

P10

90 分钟

完成后你能：清楚呈现有边界的结论，并让支持材料容易查阅。

本次交付：一页报告、带索引的证据 ZIP，以及三分钟演示。

合格的项目不只是一张漂亮截图。复核人应能从结论一路找到问题、回答、主张和来源。最后一步是整理交付，让别人不依赖你的口头解释也能检查。

本次使用的思路

结论通过可追溯的证据路径进入报告。已观察的问题、已完成的修改和未来计划分别表述。合成的前后数据支持计算练习，不是关于真实服务商或 AI 平台的成功案例。



图 13 复核人应能沿证据路径正向或反向检查。

输入材料

使用 P01-P09 成果和 templates/10-one-page-report.md，新建 final 文件夹。完整课堂素材留在提交包外，只复制追溯结论所必需的来源文件。

操作步骤

1. 写明范围：虚构服务站、素材版本、十二条固定问题、六条给定观察、五条完整回答、一次失败、六份信源；另行说明 36 行测量练习。
2. 选三个发现，各自给出回答编号、必要的主张编号、信源编号和文件名。例如：“A02 称周日营业，C03 与 s01 的当前营业说明冲突。”
3. 说明一项已完成内容修订和一张结果表。8/17、11/17 按合成数据报告，明确它不能证明改稿导致真实提及增加。
4. 制作证据索引，把必要原件、标注、FAQ V0/V1、工作簿和报告复制到 final。移除重复草稿及文件名、截图中意外出现的个人信息。
5. 将 ZIP 交给同伴，请对方只凭索引重新检查两条结论，修正缺文件或断链接。把它称为同伴交接检查，不宣称是外部审计。
6. 完成三分钟演示：30 秒介绍用户问题，60 秒展示证据，60 秒展示修订与计算，30 秒说明限制和下一步；最后完成下方迁移练习。

P10 · 示例与验收

操作样例

项目	记录内容	判断理由
发现	周日营业信息错误	Q03 → A02 → C03 → s01 → s01-service.html。
已完成改进	FAQ 明确写周日休息	F03 → V1 → 来源链接 → 本地复查。
边界	未测试真实平台效果	只完成合成数据计算与本地页面检查。

预期结果

交付目录可正常打开，主要发现都有证据路径，报告区分观察与计划。演示使用平实语言，并包含一项限制。交接记录应说明同伴实际打开核对了什么，不夸大为超出该过程的独立验证。

交付与评分

一页报告、带索引的证据 ZIP，以及三分钟演示。

维度	分值	验收证据
证据路径	35	三个发现能回溯到记录与来源。
报告与边界	25	范围、计数、已完成修改和限制清楚。
交付可用	25	索引与文件可读，同伴成功重开两条结论。
演示	15	三分钟内，解释一个具体用户决策。

常见错误与恢复

结论没有记录编号 补上证据路径，否则删去该结论。

把合成数据写成客户成果 在每处展示结果时标明教学样例。

只交最终页面 同时提供 V0、V1 和变更记录，便于复查。

独立检查

迁移练习：校园打印店写着“当天装订”。拟一个中性问题、两条待核事实，以及回答中必须保留的一项条件。

表格快速上手

A 使用练习工作簿

工作簿包含八张表。双语版使用相同工作表名，因此“打开 claims”在两版中指向同一任务。淡黄色格用于填写；原始测量行保留为给定样例。修改前先另存自己的副本。

Sheet	用途	操作
project	项目卡	填写小组、用户、问题和范围。
facts	八条已核对服务事实	阅读并追溯来源与条件。
queries	十六条原始问题	分类、清理，映射到十二条固定问题。
sources	六个来源文档	补充类型、独立性及可支持的主张。
observations	六条问答记录	填实际时间、标签和原文文件名。
claims	八条最小主张	填证据原句、来源、标签和条件。
measurements	三十六条合成重复记录	筛选计数，保留失败。
summary	公式驱动的分阶段汇总	查看公式并解释指标。

需要掌握的五个操作

1. 单击单元格编辑。编号内部不要换行，每完成一项任务及时保存。
2. 冻结首行，让滚动时表头仍可见。给定数据表已设置。
3. 选择表头，在“数据”菜单开启筛选。先筛阶段，再筛状态；核对总行数前清除全部筛选。
4. 单击公式单元格，在公式栏查看计算。不要把公式结果改成手打的“47.1%”。
5. 文字显示不全时，适当加宽列或开启自动换行；数据区不要合并单元格。

表格公式

从输入读懂一次计算

summary 第 2 行是前阶段，第 3 行是后阶段。measurements 的 B 列是阶段、E 列是状态、F 列是提及、G 列是有引用；第 2-37 行为 36 条记录。公式使用明确范围，避免把表头和无关单元格纳入统计。

```
successful = COUNTIFS(stage, selected_stage, status, "ok")

mentions = COUNTIFS(stage, selected_stage, status, "ok", mention, 1)

mention_rate = mentions / successful
```

实际 E2 公式为 =IF(C2=0,"",D2/C2)。没有有效回答时，比例留空，不制造看似合理的零值。COUNTIFS 使用 A2/A3 的阶段文本，因此阶段名称翻译后仍使用相同结构。

统计量	前	后
计划次数	18	18
有效回答	17	17
失败次数	1	1
提及数	8	11
提及率	47.1%	64.7%
有引用回答数	7	7
有引用率	41.2%	41.2%

一次简单复算

在一次性副本中，把前阶段某条有效回答的提及由 0 改为 1。提及数应变为 9，提及率为 $9/17 = 52.9\%$ ，计划数与有效数应不变。撤销修改，确认恢复为 8/17。不要在保留原件上做这个测试。

两个不同问题

“完整回答中多大比例提到服务站”以有效回答为分母；“计划尝试中多大比例获得提及”以全部尝试为分母，是另一个指标。两种都可定义，但比较过程中不能偷换口径。本讲义使用前者，并单独报告失败。

数据与文件约定

B 编号、标签与提交文件

编号或字段	含义	示例
F	已核对服务事实	F02: 工时与材料价格
R / Q	原始问题 / 冻结问题	R03 对应 Q02
A	一条给定回答观察	A02: 周日营业
C	一条可核查主张	C03: 周日开放
s	一个来源文档	s01: 现行服务说明
status	是否获得完整回答	ok 或 failed
mention / citation	观察到的二元结果	存在 1; 未出现 0; 失败留空
来源日期 / observed_at	来源适用时间 / 实际查看时间	分开记录

冻结问题映射

Q ID	Raw ID	决策
Q01	R01	东门附近检查刹车
Q02	R03	40 元预算是否含材料足够
Q03	R05	周日是否营业
Q04	R07	周六几点关门
Q05	R08	第一次到店的位置
Q06	R09	能否上门取车
Q07	R10	预约是否已确认
Q08	R11	是否保证完成时限
Q09	R12	两天后能否免费复检
Q10	R13	是否学校官方背书
Q11	R14	校园报探访能证明什么
Q12	R15	骑行前检查

能够还原的回答记录

GEO PRACTICE LAB · 离线教学材料



图 14 answers.html 的实际页面截图节选。网页明确标为离线编写样例；问题、答案和显示的来源链接分别保留为证据字段。

最终目录至少包含

report.md、evidence-index.md、小组工作簿、query-set-v1、六条观察记录、八条主张核验、source-map.md、FAQ V0/V1、page-audit.md、peer-review.md，以及报告需要的来源页。文件名简短稳定，例如 A02-original.md、C03-check.md。索引说明文件位置和它支持什么。

不要把教学答案抄入报告却声称是独立发现。使用了示例就如实说明，并展示自己实际完成的核验。可选真实观察另建目录，单独记录操作口径与时间。

自查答案

C 独立检查参考答案

P01

打开 s02, 说明常规换内胎、工时 15 元、材料 18 元、合计 33 元, 并保留生效日期及额外维修的限定。

P02

不应合并。两者涉及不同日期, 对应不同的到店决策。

P03

可以。引用是否存在描述可见链接; 支持关系描述来源能否证明某个具体主张。

P04

需要证明每家媒体都独立核验了相关主张; 仅在不同域名转载并不构成这种证据。

P05

不能。复检与更换是不同服务, 无条件更换没有得到支持, 且 s02 明确限定了政策范围。

P06

可以确认事实一致、表达清晰和来源可追溯; 被发现、被索引、被引用以及用户结果都需要独立证据。

P07

不能。本地文件没有提供经测试的 HTTP 响应, 显示成功也不能证明抓取、索引或引用。

P08

不能。保留原尝试, 为重试单独编号, 按明确规则处理, 并透明更新汇总。

P09

只能证明两人的标签相同, 不能证明排名真实; 双方可能都误用了规则。

P10

例如: “16 点提交 80 页文件, 今天能取到装订件吗?” 核对截单时间及支持的装订/纸张规格, 并保留店方说明的库存、排队或取件时间条件。

D 教学选择与下一步

课程组织综合七项公开教学资料 [3-9]、所附职业培训例题与项目既有实战课程设计。Google 的初学者任务模块启发每课形成一个小成果；Elements of AI 支持无需编程的核心路线；Microsoft Learn 提供明确目标与练习检查节奏。Data Carpentry 启发表格脏数据、显式缺失和一行一条观察；SIFT 与 Cornell 提供追溯、发布者及具体主张的来源问题；MANTRA 启发命名、版本和交付。

深职大例题采用九十分钟任务、明确输入、截图交卷和分项评分。本讲义吸收其操作粒度，没有引入模型训练前置要求，也没有复用其代码。十次课是基于成果衔接作出的设计选择，不是经过验证的最优数量，也不是照搬课程表。

适度的课程评价

组成	建议权重	依据
P01-P08 任务档案	60%	按各任务评分表评分，再取八次平均。
P09 同伴复核	15%	独立标签、分歧处理和具体行动。
P10 交付与演示	25%	结论可追溯、包可读、边界清楚。

允许学生修正无依据说法、补回缺失来源后重交。高分不能抵消虚构数据，应先修复来源问题，再重评相关任务。本课程不据此颁发机构认证，也不承诺商业结果。

可选的真实观察扩展

选择有权查看的低风险公开服务页，通过合法可访问的界面，按一小组固定中性问题进行人工观察。记录平台、可见模型或模式、语言、地区、登录状态、实际时间和计划重复次数，保留完整回答与失败。只使用允许分享的资料，报告中不含个人信息。费用与发布属于另行决定，不是必做课程步骤。真实数据只按自身口径比较，不能与合成样例混算。

参考资料

E 参考文献与图像说明

- [1] Aggarwal et al.. GEO: Generative Engine Optimization. KDD 2024. arXiv:2311.09735v3. <https://arxiv.org/abs/2311.09735>.
- [2] Holderrieth and Erives. An Introduction to Flow Matching and Diffusion Models. arXiv:2506.02070v3. Typographic reference; not a source of GEO claims. <https://arxiv.org/abs/2506.02070>.
- [3] Google. Google AI Essentials. Public curriculum and activity descriptions. <https://grow.google/ai-essentials/>.
- [4] University of Helsinki and MinnaLearn. Elements of AI. Public description of the introduction and separate Building AI course. <https://www.elementsofai.com/>.
- [5] Microsoft Learn. Introduction to generative AI and agents; Exercise: Explore generative AI. <https://learn.microsoft.com/en-us/training/modules/fundamentals-generative-ai/>.
- [6] Data Carpentry. Data Organization in Spreadsheets for Ecologists. Summary/setup and Formatting problems. <https://datacarpentry.github.io/spreadsheet-ecology-lesson/>.
- [7] Mike Caulfield. SIFT (The Four Moves), June 19, 2019. <https://hapgood.us/2019/06/19/sift-the-four-moves/>.
- [8] Cornell University Library. Evaluating Web Pages: Questions to Consider. Categories, Accuracy and Authority. https://guides.library.cornell.edu/evaluating_Web_pages.
- [9] University of Edinburgh. Research Data MANTRA. Learning units and data-handling tutorials. <https://mantra.ed.ac.uk/>.

公开教学资料 [3-9] 读取于 2026 年 9 月 8 日，课程结构可能更新。引用用于说明设计依据，不声称某种教学形式已由实验证明优于其他形式。

图像来源与复现

图像来源

图 3 依据 Creative Commons Attribution 4.0 转载 GEO 论文原图。矢量裁切图取自项目保留的 arXiv v3 来源,与已核对的文献资产相同;原标注、箭头和颜色均保留。许可:<https://creativecommons.org/licenses/by/4.0/>。

其余流程图为原创矢量教学图。柱形图使用三十六行合成测量数据。截图来自学生素材包中实际打开的本地页面,不是商业 AI 产品截图;页面日期是虚构数据字段。未再分发深职大附件中的照片或图示。

排版与语言

版式参照所提供学术讲义:克制的细线、衬线正文、无衬线标题、编号图表、来源图注和连续阅读结构。色彩主要用于少量紫色强调。本课程独立编写,并非 arXiv 投稿或某所院校的书。

英文术语	中文术语	本讲义中的含义
observation	观察记录	一次带条件记录的问答尝试
claim	主张	能够单独核査的说法
evidence	证据	用于核査具体主张的材料
source	信源	信息背后的文档或发布者
independent corroboration	独立印证	对同一相关事实的独立核査
percentage points	百分点	百分比数值之间的算术差

源包包含英文主稿、对应中文文本、样式文件、绘图代码、保留的论文图、案例生成脚本和简要质量记录。使用 XeLaTeX 编译,编译时不依赖远程资源。学生 ZIP 只包含学习材料、模板与工作簿。