

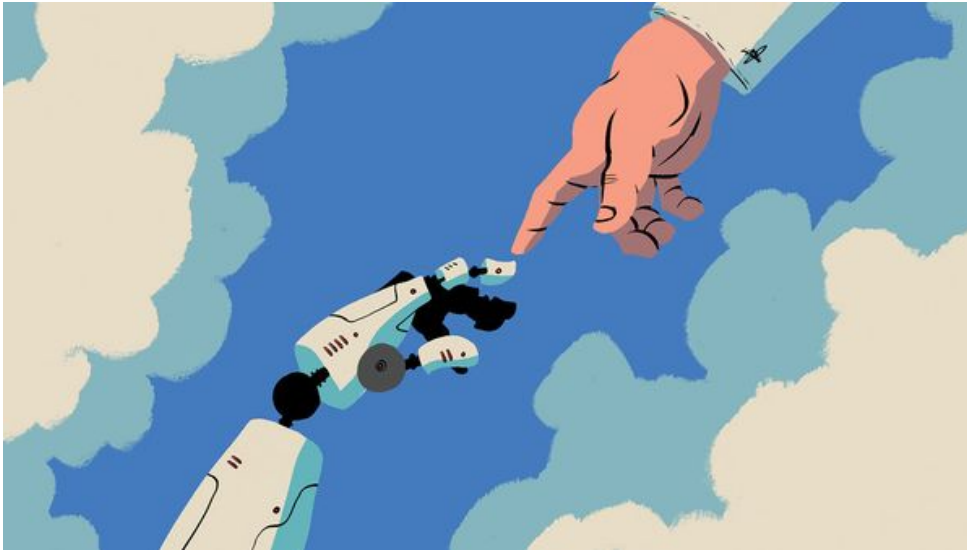
How to measure returns on AI

怎样衡量人工智能投资回报

From tokenmaxxing to something more normal

从拼命消耗词元，回到更正常的经营判断

The Economist · 2026-08-20



天空背景下，人手与机械手伸出指尖相接，构成关于人与人工智能合作的概念插画。

难度 71 / 100

建议65–85分钟；选做作文另计。

适用读者：能阅读一般英语新闻，希望读准长句、搭配与作者语气的学习者。

入门准备：能辨认基本主谓结构和常见从句，愿意结合语境查词。

About this lesson 本课导读

使用更多人工智能，是否就创造了更多价值？本文从投入指标转向结果与组织能力，借讽刺、研究实例和管理选择，说明速度、质量、调整成本和学习能力为何必须一起看。

- 区分投入、结果与组织指标
- 解释节省时间如何转化为实际收益
- 拆解插入语、长主语与指代
- 用条件和比较基准讨论回报

Before you read 读前思考

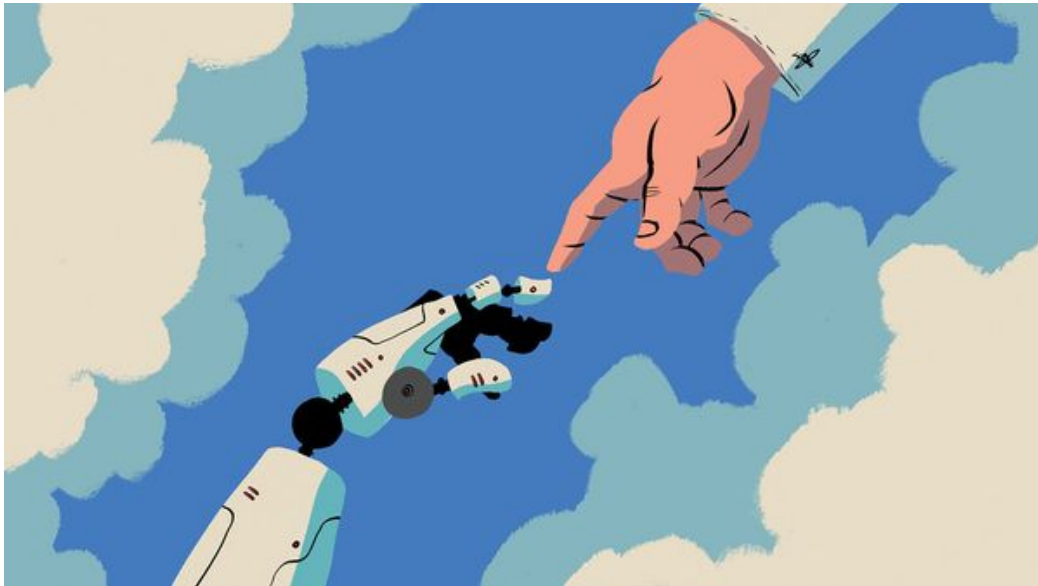
If a team doubles its use of a tool, what else would you need to know before calling that a success?

拓展例句与练习情境用于说明用法。

书末答案 (p.14)

Text 课文

[P01] IN THE BEGINNING there was AI, and bosses everywhere lost their minds and said you must all use this technology no matter whether it is useful and no matter how much it costs.^{K01} And they introduced AI leaderboards and came up with stupid words like “tokenmaxxing”.^{V01} And lo, usage did indeed rise.^{V02} And then the bosses remembered some very basic concepts like “budgets” and realised that this might not be such a great idea after all.^{V03} And then a different question rang through the boardrooms, and this question was about returns, and it was harder to answer.^{V04}



天空背景中，人手与机械手伸出指尖相接，构成关于人与人工智能合作的概念插画。

[P02] This potted Genesis of AI adoption is not entirely fair.^{V05V06} As long as organisations are trying to encourage usage, it can still make sense to look at things like employees’ token consumption.^{K02} “I would expect software engineers to spend more money than a finance analyst,” says the boss of one software firm. “But anyone within finance, it shouldn’t be zero.”

[P03] But usage (also known as input) measures don’t tell you whether any value is being created. And even when AI is producing benefits, points out Mika Ruokonen of LUT University in Finland, lots of them accrue quietly to individuals in the form of scattered time savings.^{V07V08K03} If an extra hour or two of an employee’s time has been freed up to shop on Vinted, someone has benefited (mainly Vinted) but it isn’t your company (unless you work at Vinted).^{V09}

[P04] Shifting the focus to returns on AI investment means concentrating more on outcomes-based measures, such as team productivity or customer satisfaction.^{V10K04} But these require careful thought. Scientists, for example, are using AI to produce more papers (yay!) of dubious merit (boo!).^{V11V12} A recent study by Tulsi Suchak of the University of Surrey and her co-authors found that papers derived from a health and nutrition data set ticked along at an average of four a year between 2014 and 2021; in the first nine months of 2024, there were 190.^{V13}

[P05] So outcomes-based measures have to be designed to capture both positive and negative

effects of AI.^{V14} At Google, for example, engineers have long been measured on three dimensions. Speed is one: how long does a code review take, say. Ease is another: how much friction is there in the system for things like onboarding new developers.^{V15V16} The last, critically, is quality: “I don’t want to just go fast and make it easy to ship terrible software,” says Richard Seroter of Google Cloud.^{V17}

[P06] Complicating matters further is the fact that it often takes a while for the benefits of AI to show up.^{K05} People have to spend time learning how to use the technology. Increased output can create bottlenecks elsewhere.^{V18} A paper published last year by Kristina McElheran of the University of Toronto and her co-authors looked at AI adoption among American manufacturing firms, and confirmed the existence of a “J-curve” in which productivity dips before it leads to an improvement.^{V19} This effect is especially pronounced at established firms and seems to be partly explained by changes to management practices that had previously helped firms to keep tabs on performance.^{V20V21V22K06} Any kind of calculation about projected returns has to take this initial period of disruption into account.^{V23}

[P07] Bosses also have to take a view on something much more fundamental: what kind of return are they interested in? One way to show financial returns on AI is to seize on time savings as an excuse to cut headcount.^{V24V25} For some jobs and in some circumstances, that might make sense. But chainsaws and employee morale go together about as well as salt and slugs.^{V26K07} And people who are no longer there cannot be redeployed in more productive ways.^{V27} A better option is to calculate the amount of money that AI saves on expanding headcount. (Working out how much extra revenue can be attributed to the technology is even more of an art; at the very least, you need good baseline data and a culture of A/B testing to isolate its effects.)^{V28V29V30K08}

[P08] On top of everything, there is wild uncertainty about where AI is heading. That is an argument for a third set of measures, to go alongside ones on inputs and outcomes. These measures, which Mr Seroter calls “organisation-based”, are more geared towards capturing how well change is being managed.^{V31K09} They could include data on levels of employee satisfaction with the way AI has been implemented, or progress towards building in-house expertise in areas that AI will not automate.^{V32} A one-eyed focus on usage ignores the importance of returns.^{V33} A rigid focus on returns ignores the need to keep learning.^{V34}

New words and expressions 生词和短语

按原文出现顺序排列；*为优先练习表达。

V01 · P01 leaderboard

/ˈliːdərbɔːd/ n. 排名榜；排行榜

AI leaderboard在此用于推动员工使用，排名依据不自动代表价值。

V02 · P01 lo

/ləʊ/ interjection 看哪；瞧

带古旧或戏仿语气，配合反复And制造“创世记”式腔调。

V03 · P01 after all*

/ɑːftər ˈɔːl/ adv. phr. 终究；毕竟

这里回看先前热情并修正判断，不是列举所有事项之后。

V04 · P01 return

/rɪˈtʃːn/ n. 回报；收益

本文先谈投入回报，后又问究竟要哪种回报；不能都等同现金利润。

V05 · P02 potted

/ˈpɒtɪd/ adj. 简述的；浓缩的

potted history 一类搭配指短版本叙述，不是装进花盆。

V06 · P02 adoption

/əˈdɒpʃən/ n. 采用；应用

AI adoption 指组织采纳技术；不是收养儿童。

V07 · P03 accrue

/əˈkruː/ v. 逐渐产生；归于

benefits accrue to individuals 说明好处落到谁手中，不能只译成累计而漏掉受益者。

V08 · P03 scattered

/ˈskæɪtəd/ adj. 零散的；分布各处的

零散时间节省不必然汇成一整块可削减的岗位成本。

V09 · P03 free up*

/friː ʌp/ phr. v. 腾出；释放

free up time/resources, 强调原来被占用的时间或资源可作别用。

The new timetable frees up an hour for revision.

新时间表腾出一小时用于复习。

V10 · P04 outcome

/aʊtkʌm/ n. 结果；成效

文中举生产率、满意度为例；数量变多不保证所有结果都改善。

V11 · P04 dubious

/djuːbiəs/ adj. 可疑的；令人怀疑的

of dubious merit 质疑价值，不等于每篇都被证明造假。

V12 · P04 merit

/ˈmerɪt/ n. 价值；优点

本句是名词，与merit attention“值得关注”的动词用法区别。

V13 · P04 tick along

/tɪk əˈlɒŋ/ phr. v. 以平稳缓慢的速度持续

ticked along at ... 说明过去产出节奏，为后面的猛增作对照。

V14 · P05 capture*

/ˈkæptʃə/ v. 反映；捕捉到

指标capture effects 指有效反映影响，不是拘捕某人。

The survey failed to capture the cost of repeated repairs.

调查未能反映反复维修的成本。

V15 · P05 friction

/ˈfrɪkʃən/ n. 阻力；麻烦；不顺畅

这里是流程阻力，例如新员工上手困难，不是机器表面摩擦系数。

V16 · P05 onboarding

/ˈɒnbɔːdɪŋ/ n. 入职引导；让新人上手

包括熟悉流程和工具，不能只理解成到岗的第一天。

V17 · P05 ship

/ʃɪp/ v. 发布；交付

ship software 指把软件交付用户使用，并非用轮船运送。

V18 · P06 bottleneck

/ˈbɒtlənek/ n. 瓶颈；制约环节

一个环节加速，可能把处理不完的工作推到下游，形成新的限制。

V19 · P06 dip

/dɪp/ v. 短暂或阶段性下降

productivity dips 写曲线早段，后续能否改善要看研究范围与企业调整。

V20 · P06 pronounced

/prəˈnaʊnst/ adj. 明显的；显著的

这里不是“被发音的”，而是影响表现得更突出。

V21 · P06 established

/ɪˈstæblɪʃt/ adj. 已建立成熟业务的；老牌的

与新企业相比，现有组织惯例更多，不是刚注册的一般“已成立”。

V22 · P06 keep tabs on*

/kiːp təbz ɒn/ idiom 持续掌握；密切跟踪

keep tabs on performance 指了解表现，并非只打开浏览器标签页。

The coordinator keeps tabs on work that is still waiting for approval.

协调人持续跟踪仍待审批的工作。

V23 · P06 take into account*

/teɪk ɪntuː əˈkaʊnt/ phr. v. 纳入考虑

take this period into account, 宾语可放在take与into account之间。

V24 · P07 seize on*

/siːz ɒn/ phr. v. 抓住并利用

seize on time savings as an excuse, as 后说明把它用作什么。

They seized on the delay as a reason to cancel the order.

他们抓住这次延误，作为取消订单的理由。

V25 · P07 headcount

/hedkaʊnt/ n. 员工人数；人员规模

cut 现有人员与save on expanding 人员是两种不同调整。

V26 · P07 morale

/məˈrɑːl/ n. 士气

与moral“道德的”拼写、重音及意义均不同。

V27 · P07 redeploy*

/ri:dr'plɔɪ/ v. 重新调配

把现有员工安排到其他工作；已经离职的人无法作为内部人员调配。

V28 · P07 attribute to*

/ə'trɪbjʊ:t tu:/ phr. v. 归因于

revenue attributed to AI需要区分同时发生的其他因素。

V29 · P07 baseline

/bə'slaɪn/ n. 比较基准；基线

估计改进前须清楚与什么状态相比，不能只看投入后绝对数字。

V30 · P07 isolate

/aɪ'səleɪt/ v. 分离；识别出单独影响

这里说区分技术效果与其他变化，不能理解成孤立员工。

V31 · P08 geared towards*

/ɡɪəd tə'wɔ:dz/ adj. phr. 旨在；着眼于

be geared towards capturing ... , towards后接动词。

The training is geared towards solving unfamiliar problems.

这项培训着眼于解决不熟悉的问题。

V32 · P08 in-house

/ɪn 'haʊs/ adj. 机构内部的

in-house expertise指组织自己拥有的专长，不是员工在家办公。

V33 · P08 one-eyed

/wʌn 'aɪd/ adj. 片面的；只看一面的

比喻评价眼光狭窄，不是在讨论人的身体状况。

V34 · P08 rigid

/rɪdʒɪd/ adj. 僵化的；不灵活的

死守某类回报指标会忽视学习；结尾批评两个方向的单一关注。

Notes on the text 课文注释

1 一本“迷你创世记”怎样制造讽刺

[P01 / P02]

In the beginning、And lo和反复以And开句，把企业采用AI的过程写成庄严叙事，内容却是老板先狂热、后想起预算。庄严声调与浅显决策的反差，才是笑点；lo不是普通商务写作中必须模仿的连接词。tokenmaxxing把token与maxxing拼起来，嘲弄把词元消耗推到最高、仿佛越多越好的做法。词元是系统处理文本的单位之一，数量可以描述某类使用活动，却不直接告诉你结果是否正确或值得付费。P2立即用not entirely fair让步：在鼓励采用的阶段，使用指标仍有局部用途。

2 节省下来的时间，归了谁

[P03 / P07]

P3的benefits accrue to individuals把受益者从公司转到员工。员工多了自由时间，确实可能得到好处；作者并没有否认这一点。他提出的是另一个问题：这种好处是否已经转化为公司的价值。Vinted插句不断更换受益者，最后unless又用平台员工这个特例补一刀。

情境设定：10名员工各省10分钟，总计100分钟，但零散分布未必使一班岗位能够减少。若这些时间被用于处理更多积压订单，可能改善交付；若工资和收入均未改变，则不能仅凭100分钟乘工资就宣布现金成本已下降。时间容量、产出与现金影响要分别说明。

3 从投入转结果，还要再问结果好不好

[P04 / P05]

输入端看用了多少，结果端看工作发生了什么变化。P4马上设障碍：论文篇数更多也许是产出数量增长，但质量可疑会使这个“好消息”失色。yay与boo把赞与嘘直接塞进句子，让读者听见评价的急转弯。

P5把speed、ease与quality并列。更快通过代码审查、让新人更容易上手，仍不能抵消软件本身糟糕。评价表中既有正面结果，也应有返工、错误和下游负担等负面影响；否则团队只是学会把某个数字做得好看。

4 局部加速为何会让总流程先变慢

[P06]

学习要占用时间，旧流程要调整，上游变快又会把工作堆到下游。情境设定：每天起草量从一份增至三份，但审核每天只能完成两份，积压便增加；单看起草速度会高估整个团队的改善。bottlenecks elsewhere要求把观察范围扩展到流程。

J曲线用形状帮助记忆先低后高的路径；本文同时说seems、partly，管理惯例变化是部分解释，不是全部原因。初期低谷需要进入预计回报计算，但不能把任何失败都解释成“再等一等就一定反弹”。

P6的before it leads to an improvement，按上下文说采用AI随后带来改善；不能机械回指最近的productivity，更不能理解为生产率下滑本身保证后续改善。

5 减少现有人数，与少增加人数

[P07]

cut headcount削减现有团队；save on expanding headcount则是在业务增长或工作量增加时，可能不必像原计划那样多招人。两者都可能影响费用，但对团队能力、士气和未来配置的后果不同。

chainsaws比喻猛烈裁撤，salt and slugs用不相容的组合表现士气受伤。作者先承认某些岗位、某些条件下削减人员或许有道理，再偏向少增员的办法；不是宣称所有裁员都不可行，也不是已经证明少招人一定最优。

例如同样完成扩大的工作量，原本预计从20人增到24人，后来只需22人，差异是少增加2人，而不是现有20人里裁掉2人。这个虚构情境要求先固定工作量和比较方案，才能解释节省了什么。

6 收入增加，如何知道是AI的作用

[P07 / P08]

attribute ... to与isolate effects要求回答反事实问题：如果没有采用AI，结果可能怎样？投入前后的差额可能同时含旺季、价格、促销或人员变化；baseline是比较起点，A/B testing则是通过合适的对照更有把握地分离影响。

原文称这更像一门art，强调判断复杂，不是允许随意报数。测试分组、比较任务和质量标准也要相近。P8接着提出组织指标，避免只盯短期金额而丢掉实施体验、内部专长与继续学习的能力。

Special difficulties 难点辨析

1 三组指标，和其中一组的三个维度

[P01 / P04 / P05 / P08]

全文三组是inputs、outcomes与organisation-based measures。P5的speed、ease、quality是结果评价里举出的三个维度，不能把两套三分法拼成同一个列表。P8的ones指前面两组measures，go alongside表示并列补充。

结尾的one-eyed与rigid分别批评“只看用了多少”和“死守即时回报”，不是否定使用数据或财务回报。好的概括应写出三组互补关系，而非说指标全部无用。

2 not entirely与partly保留多少

[P02 / P06 / P07]

not entirely fair只否定“完全公允”，表示前述讽刺有不周全之处，不等于全错；seems to be partly explained表示解释有一定依据但不完整，也不是纯粹确定的因果结论。

P7的for some jobs and in some circumstances先圈定适用范围。读到作者随后提出更好方案时，仍要保留这个让步，否则就把有条件比较改成无例外禁令。

Background 背景与延伸

1 论文数量激增，究竟数的是哪一批

[P04]

Suchak等人2025年发表于PLOS Biology的研究，以特定检索条件筛出341篇基于美国NHANES健康与营养调查数据库的单因素关联研究，排除了多因素研究等类型。因此，文中的年均四篇与190篇不代表该数据库产生的全部论文。

这项研究的实际检索截止日是2024年10月9日；原文把190篇简述为前九个月，二者并非严格相同。研究讨论公式化研究、设计与发现的风险，但没有逐篇证实AI使用，也没有单独识别AI导致增长的因果效应。它支持认真检查质量，不能被改写成“这批论文全部由AI生成且全部虚假”。

2 工业AI研究与本篇管理启示的距离

[P06]

McElheran等人的2025年工作论文《The Rise of Industrial AI in America》研究美国制造业的工业AI，主要资料对应2017与2021年。它报告采用与组织调整之间的关系，以及生产率先受扰动、后续可能改善的模式；这不是直接对所有生成式AI项目做的统一回报试验。

论文的后续表现涉及存续企业，企业新旧和管理条件之间也有差异。把研究带回本文，合理启示是计入学习与组织调整成本，并继续观察实际结果；不能从一条J形轨迹推成每家公司最后都必然赚钱。

Key structures 关键句型

K01·P01 no matter whether ... and no matter how ...

IN THE BEGINNING there was AI, and bosses everywhere lost their minds and said you must all use this technology no matter whether it is useful and no matter how much it costs.

起初，有了人工智能，各地老板都失去了理智，说你们必须全都使用这项技术，不论它是否有用，也不论它花多少钱。

句子以there was AI开场，后面bosses lost ... and said ...连接老板的行为。said后的you must all use ...是被转述的命令，作者借夸张表现不顾用途与成本的热情，并非自己向读者发命令。

no matter whether it is useful引出是否有用的两种可能，no matter how much it costs引出成本高低。it都指这项技术；how much与costs配合问金额。两者共同说这些差别都不改变must use的要求。

The director insisted that every team use the new form, no matter whether it helped and no matter how long it took.

主管坚持每个团队都使用新表格，不管它是否有帮助，也不管它要花多久。

K02·P02 As long as ... , it can still make sense to ...

As long as organisations are trying to encourage usage, it can still make sense to look at things like employees' token consumption.

只要组织还在努力鼓励使用，关注员工词元消耗之类的指标，仍可能说得通。

As long as在此设置条件：只要组织仍处于鼓励使用的阶段。主句it是形式主语，真正被评价为make sense的是to look at ...这件事。can与still共同保留一种有限合理性，不能改写成永远必须用使用量评绩效。

things like ...举出指标例子，employees' token consumption是员工的词元消耗。先读出条件，再读出肯定范围，才看得见P2对P1的修正。

As long as the team is learning a new process, it can still make sense to record how often help is requested.

只要团队还在学习新流程，记录求助频率仍可能有道理。

K03 · P03 even when ... , points out ... , lots of them accrue ...

And even when AI is producing benefits, points out Mika Ruokonen of LUT University in Finland, lots of them accrue quietly to individuals in the form of scattered time savings.

芬兰拉彭兰塔 - 拉赫蒂工业大学的米卡·鲁奥科宁指出，即使人工智能正在产生好处，其中许多也会以零散节省时间的形式，悄悄落到个人手里。

主句先还原为lots of them accrue quietly to individuals. even when承认AI确有益处时也会出现这种情况，points out Mika ...为插入的观点归属，不能误当成导致收益的行为。

them指benefits，to individuals说明受益者，in the form of ...说明好处的形式。句子不是说AI没有价值，而是说有价值也未必立即归到公司账上。

Even when a project succeeds, notes the evaluator, some of its benefits accrue to people outside the organisation.

评估者指出，即使项目成功，部分收益也会归于组织之外的人。

K04 · P04 Shifting ... means concentrating ... : 两个动名词不同职务

Shifting the focus to returns on AI investment means concentrating more on outcomes-based measures, such as team productivity or customer satisfaction.

把关注点转向人工智能投资回报，就意味着更多关注以结果为基础的指标，比如团队生产率或客户满意度。

Shifting the focus ...是主语，means是主句谓语，concentrating more ...作means的宾语。这里mean doing表示意味着做某事，不是mean to do“打算做”。

returns on AI investment中的on说明什么投资的回报；outcomes-based measures后such as举例team productivity与customer satisfaction。关注结果的转向尚需下文补质量，不能把这句话当作篇章最终答案。

Improving the service means paying attention to reliability as well as speed.

改进服务意味着既关注速度，也关注可靠性。

K05 · P06 Complicating matters ... is the fact that ...

Complicating matters further is the fact that it often takes a while for the benefits of AI to show up.

更让事情复杂的是，人工智能带来的收益常常要过一段时间才会显现。

按常见顺序理解为The fact that ... is complicating matters further。作者先放Complicating matters further，再用is引出真正的主语the fact that ...；fact后的that给出这个事实的内容。

内部it often takes a while for the benefits ... to show up，它是形式主语，a while表示所需时间，for ... to show up说明等待的是收益显现。不要把时间推迟误读成不存在收益，也不要把often省成每次必然。

Making the decision harder is the fact that it takes several weeks for the test results to arrive.

更让决定难作的是，检测结果需要过几周才能出来。

K06 · P06 seems to be partly explained by changes to ... that had ...

This effect is especially pronounced at established firms and seems to be partly explained by changes to management practices that had previously helped firms to keep tabs on performance.

这种影响在已有成熟业务的企业中特别明显，部分原因似乎是管理惯例发生了变化，而这些惯例过去曾帮助企业掌握经营表现。

主语This effect指前面的J曲线式生产率变化，is especially pronounced先评价表现强弱，and seems ...再提出部分解释。seems与partly分别限定把握程度和解释范围。

by引出解释因素changes, to management practices说明哪些惯例发生改变；that had previously helped ...修饰practices。had helped把惯例的既有作用放在这些改变之前，keep tabs on说明此前它们帮助企业怎样掌握表现。

The delay seems to be partly explained by changes to procedures that had previously helped teams check urgent requests.

延误的部分原因似乎是流程发生了变化，而那些流程过去曾帮助团队检查紧急请求。

K07 · P07 go together about as well as ... : 比较句中的反讽

But chainsaws and employee morale go together about as well as salt and slugs.

但电锯与员工士气的相容程度，就如同盐与蛞蝓一样。

表面形式是A and B go together as well as C and D，比较两组东西相配的程度。about使比较近似；这里盐与蛞蝓是显然不友好的组合，于是as well as实际表达“同样糟糕地相处”。

chainsaws比喻大砍人员，morale指员工士气。不能只见well就译成相得益彰，也不要把这比喻误解成公司真的使用电锯。

Constant interruptions and careful reading go together about as well as open windows and loose papers on a windy day.

不断的打断与细致阅读，就像刮风天的敞开窗户与散放纸张一样不相容。

K08 · P07 Working out how much ... can be attributed ... is ...

(Working out how much extra revenue can be attributed to the technology is even more of an art; at the very least, you need good baseline data and a culture of A/B testing to isolate its effects.)

至于算清多少额外收入可以归因于这项技术，就更加是一门艺术了；至少，你需要良好的基线数据，以及开展A/B测试的习惯，来识别它的影响。

括号内第一句的主语是Working out ...，谓语is，表语even more of an art。how much extra revenue can be attributed ...是work out的内容，其中revenue是被归因的对象，to the technology说明归因到哪里。

分号后you need A and B并列good baseline data与a culture of A/B testing，to isolate its effects说明目的，its指AI技术。原句建议提高识别效果的把握，不能由“做了A/B测试”自动保证分组和口径完全可靠。

Working out how much improvement can be attributed to training is difficult; you need a useful comparison to separate its effects from other changes.

算清有多少改善归功于培训并不容易；你需要有用的比较，才能把培训效果与其他变化分开。

K09 · P08 which ... calls ... , are geared towards capturing how ...

These measures, which Mr Seroter calls “organisation-based”, are more geared towards capturing how well change is being managed.

塞罗特把这组指标称为“基于组织的指标”，它们更着眼于衡量变革管理得怎么样。

主句These measures are more geared towards ...说明第三组指标的重点。which Mr Seroter calls “organisation-based”是非限定修饰，which在calls后对应被命名的measures，organisation-based为名称，不是第三组指标的独立谓语。

towards后接capturing，capturing的内容是how well change is being managed。内部is being managed为现在进行时的被动形式：关心变革正在被管理得怎么样，而非只记录一次已经结束的实施。

These questions, which the team calls “process checks”, are geared towards understanding how well the transition is being managed.

这些问题被团队称为“流程检查”，旨在了解转型正在被管理得怎么样。

Exercises 练习与写作

Comprehension 理解

回答时保留作者的让步和评价范围。

02 Why does the author partly defend usage measures after mocking them?

03 What distinction does the Vinted joke illustrate?

04 What must an outcomes measure capture beyond higher speed or quantity?

05 Why should projected returns include an initial disruption period?

06 How does avoiding extra hiring differ from cutting existing headcount?

07 Why add organisation-based measures?

Words in use 词汇

用词库填空，每项一次。

词库：free up / capture / keep tabs on / seized on / geared towards

08 The revised schedule may ____ time for checking difficult cases.

- 09 A count of completed forms may fail to ____ their accuracy.
- 10 The coordinator needs to ____ the unresolved requests.
- 11 The committee ____ the delay as a reason to change the plan.
- 12 The course is ____ helping staff handle unfamiliar problems.

Key structures 句型

找出真正主谓关系，保留比较和不确定性。

- 13 What two considerations are dismissed by the no matter clauses in P1?
- 14 Begin with As long as: “Recording requests may be useful if the team is still learning the process.”
- 15 What does them refer to in lots of them accrue quietly to individuals?
- 16 Which is the main verb in Shifting the focus ... means concentrating more ...?
- 17 Put the main subject first in Complicating matters further is the fact that benefits take time.
- 18 What had previously helped firms keep tabs on performance in P6?
- 19 Does go together about as well as salt and slugs imply a good match here?
- 20 In P7’s parenthetical sentence, what does its effects refer to?
- 21 Who performs the naming in which Mr Seroter calls organisation-based, and what is named?

Listening 听读

第一遍听论证转向，第二遍寻找限定与让步。

22 Listen for the three groups of measures and one limitation attached to each.

Summary writing 摘要

用90–110词概括三组指标及其互补关系。

23 Write a 90–110-word summary.

Composition 作文（选做）

选做：用100–120词为一项AI试用设计简洁的评价办法。

24 Suggest a balanced evaluation for a team testing an AI drafting tool.

Reference translation 参考译文

P01 起初，有了人工智能，各地老板都失去了理智，说你们必须全都使用这项技术，不论它是否有用，也不论它花多少钱。于是，他们推出人工智能使用排行榜，还想出了“tokenmaxxing”之类的蠢词。看哪，使用量果真上升了。然后，老板们想起了“预算”这类极其基本的概念，意识到这毕竟可能不是什么好主意。接着，董事会会议室里响起了另一个问题，一个关于回报、也更难回答的问题。

P02 这段人工智能应用的迷你《创世记》，并不完全公允。只要组织还在努力鼓励使用，关注员工词元消耗之类的指标，仍可能说得通。一家软件公司的老板说：“我会预期软件工程师花的钱比财务分析师多。”“但对财务部门的任何人来说，这个数字也不该是零。”

P03 可是，使用量指标——也称投入指标——并不能告诉你是否创造了价值。芬兰拉彭兰塔 - 拉赫蒂工业大学的米卡·鲁奥科宁指出，即使人工智能正在产生好处，其中许多也会以零散节省时间的形式，悄悄落到个人手里。如果员工因此多出一两个小时去Vinted购物，确实有人受益，主要是Vinted，但不是你的公司，除非你就在Vinted工作。

P04 把关注点转向人工智能投资回报，就意味着更多关注以结果为基础的指标，比如团队生产率或客户满意度。但这些指标需要仔细斟酌。例如，科学家正在使用人工智能产出更多论文，真棒！可论文的价值却令人怀疑，糟糕！萨里大学的图尔西·苏查克及其合著者最近的一项研究发现，基于某个健康与营养数据集的论文，在2014年至2021年间，每年平均发表四篇，数量一直不多；而2024年的前九个月，就有190篇。

P05 因此，以结果为基础的指标必须被设计成能同时捕捉人工智能的积极与消极影响。例如，谷歌长期以来从三个维度衡量工程师的工作。速度是一项：比如，一次代码审查要花多久。便利程度是另一项：在新开发者入职上手等事情上，系统存在多少阻力。最后一项尤其关键，是质量；谷歌云的理查德·塞罗特说：“我可不想只是图快，又让糟糕的软件更容易发布。”

P06 更让事情复杂的是，人工智能带来的收益常常要过一段时间才会显现。人们必须花时间学习如何使用这项技术。产出增加，还可能在其他环节造成瓶颈。多伦多大学的克里斯蒂娜·麦克埃尔赫兰及其合著者去年发表的一篇论文，研究了美国制造企业采用人工智能的情况，证实存在一条“J曲线”：生产率先下滑，随后人工智能的采用才带来改善。这种影响在已有成熟业务的企业中特别明显，部分原因似乎是管理惯例发生了变化，而这些惯例过去曾帮助企业掌握经营表现。任何关于预期回报的计算，都必须考虑最初这段受扰动的时期。

P07 老板们还必须想清楚一个更根本的问题：他们究竟关心哪一种回报？展示人工智能财务回报的一种办法，是抓住节省时间这一点，作为裁员的理由。对某些岗位、在某些情形下，这样做或许说得通。但电锯与员工士气的相容程度，就如同盐与蛞蝓一样。而已经不在公司的人，也无法被重新安排去从事更有生产力的工作。更好的办法，是计算人工智能在扩充人员方面省下多少钱。至于算清多少额外收入可以归因于这项技术，就更加是一门艺术了；至少，你需要良好的基线数据，以及开展A/B测试的习惯，来识别它的影响。

P08 除此之外，人工智能将走向何方，还存在极大的不确定性。这就为第三组指标提供了理由，使其与投入和结果指标并列。塞罗特把这组指标称为“基于组织的指标”，它们更着眼于衡量变革管理得怎么样。这可以包括员工对人工智能实施方式的满意程度数据，也可以包括在人工智能不会自动化的领域建立内部专长的进展。只盯着使用量，会忽略回报的重要性。死盯着回报，又会忽略持续学习的必要性。

Reference answers 参考答案与解析

开放题不要求与参考答案逐字一致；先检查信息和意思，再修改表达。

Before you read 读前思考

01 读前思考

What work improved, whether quality changed, what the total costs were and whether the change supports the organisation's aims.

可提出不同具体指标，但至少把投入与结果分开，不只问又多用了多少。

Comprehension 理解

02

They may help an organisation understand and encourage adoption while people are learning to use the technology, though they do not measure value by themselves.

P2是有条件让步，不推翻P3的价值问题。

03

An employee or another business may benefit from freed-up time without the employer receiving a corresponding benefit.

不能把有人受益改写成无人受益。

04

Quality and negative effects, such as poor work or friction elsewhere, as well as positive effects.

P5同时有ease，不必把流程便利一概排除。

05

Learning, organisational change and new bottlenecks may absorb time or reduce productivity before any later benefits appear.

说明机制；不要承诺每个项目随后必然回升。

06

The first compares staffing growth with what would otherwise have been needed; the second removes people already in the organisation, affecting opportunities to redeploy them.

两个比较基准不同；都要先说明所需工作量。

07

Because AI's future is uncertain and the organisation needs to know how well it manages change, how employees experience implementation and whether it builds useful expertise.

补充投入与结果指标，而非宣布财务回报不重要。

Words in use 词汇

08

free up

may后用原形；腾出时间的用途仍需明确。

09

capture

数量指标未必反映正确性。

10

keep tabs on

后接需持续跟踪的对象。

11

seized on

词库给过去式；as后说明如何利用延误。

12

geared towards

towards后为helping，不接help原形。

Key structures 句型

13

Whether the technology is useful and how much it costs.

这是被讽刺的命令条件，不是作者自己的建议。

14

As long as the team is still learning the process, recording requests may be useful.

保留may；条件成立不等于永远必须记录。

15

The benefits of AI.

individuals是受益者，不能把them误读成员工。

16

Means.

Shifting为主语核心，concentrating为means的宾语核心。

17

The fact that benefits take time is complicating matters further.

that从句属于the fact的内容，不能删掉后说facts自行改变。

18

The management practices.

变化作用于这些惯例；had helped说明其在变化之前的作用。

19

No. It uses an unfriendly combination to suggest that aggressive cuts and employee morale fit badly together.

well在表面比较框架中，实际评价需看比较对象。

20

The effects of AI or the technology.

不是基线数据的效果，也不是收入本身的效果。

21

Mr Seroter names the measures.

which承接measures，为calls的宾语；主句仍是measures are geared ...。

Listening 听读

22

Inputs can indicate adoption but not value; outcomes need quality and adjustment costs as well as quantity; organisation-based measures track change and learning alongside, rather than replacing, the other groups.

先归组再说明限制；不要把speed、ease、quality误当全文三组。

Summary writing 摘要

23

AI usage measures may encourage adoption, but they cannot establish that an organisation receives value. Time savings can benefit individuals without improving company results. Outcomes measures should therefore include quality alongside speed and ease, while allowing for learning costs, bottlenecks and organisational disruption. Financial assessments also need clear choices about staffing and credible comparisons for attributing additional revenue to AI. Because the technology's direction remains uncertain, organisations should track how well change is managed and useful expertise is developed. The author recommends combining input, outcome and organisational measures, avoiding both obsession with usage and a rigid focus on immediate returns.

应覆盖时间归属、质量、调整与归因，并收束到互补指标。勿把J曲线写成保证回本，把少增员写成裁员。

Composition 作文

24

A support team could test an AI tool on comparable groups of routine requests. It should record usage to see whether staff actually try the tool, but assess results through resolution time, accuracy and customer satisfaction. Review time and corrected mistakes should count as costs, rather than disappearing behind a larger number of drafts. The team should also ask employees whether they understand the workflow and can handle cases without automatic assistance. Early results may reflect learning, so the trial needs a stated period and a follow-up review. Its purpose is to discover whether service improves at an acceptable total cost, not merely whether more text is produced.

原创试用方案；交代比较、正负结果、学习与复查。参考段落提供思路，具体任务可以不同，但不能只统计输出篇数。