

超越

Beyond
Traffic

商业模式新洞察

余昌远 刘徐伟 祁琦◎著

流量

New Insights into
Business Models

 中国人民大学出版社

超越

Beyond Traffic

商业模式新洞察

余昌远 刘徐伟 祁琦◎著

流量

New Insights into Business Models

中国人民大学出版社

版权信息

COPYRIGHT

书名：超越流量：商业模式新洞察

作者：余昌远；刘徐伟；祁琦

出版社：中国人民大学出版社

出版时间：2026年03月

ISBN：9787300344874

字数：137千字

(更d书f享搜索雅 书. YabooK. org)

推荐序一：跳出流量迷思，看见商业本质

王湛

曦望联席CEO，曾任百度资深副总裁、金砖资本董事总经理、蘑菇车联首席战略官

在互联网行业深耕二十多年，看着行业从PC时代狂奔到移动互联时代，再到如今的AI浪潮翻涌，见过太多企业乘着流量红利崛起，又因摸不透商业模式的根而栽了跟头——明明手里握着用户、握着资源，却总在“怎么盈利”“怎么规模化”的迷局里打转。为什么有的互联网公司看似“不收费”却能持续运转？为什么ChatGPT引爆行业后，大模型的盈利路径仍让多数人困惑？在流量红利逐渐消退的今天，这些问题的答案决定了我们能否看懂商业的未来。这本《超越流量：商业模式新洞察》跳出常规商业分析，提供了一套深入浅出的解法。

它最核心的创新，是把复杂的商业逻辑“拆透了”——不是用“平台”“内容”“电商”这些表层标签分类，而是用“原子化思考”提炼出“一元生产、二元传递、三元激发”三大模块。比如爱奇艺早期亏损的原因，行业里多归结为“内容烧钱”，但书里用“一元生产模式”点透：是“内容投入”与“用户付费价值”的阶段性失衡，这种失衡不是运营问题导致的，而是模式成本结构的必然结果；再比如贝壳的ACN机制，很多人只看到“经纪人合作”，作者却能挖到二元商业模式的核心——这是“降低交易成本、提升供需匹配效率”的极致设计，链家从自营到贝壳平台的跃迁，都能用“模式迭代”讲得明明白白。这种“穿透表象见本质”的能力，没有多年一线的摸爬滚打，没有对行业细节的精准把控，是做不到的。

让人惊喜的是，本书没有局限于传统互联网业务。针对当下火热的大模型赛道，它专门用一章的篇幅，结合ChatGPT、DeepSeek-R1等案例，拆解AI原生应用的商业模式：“一元生产”是模型训练与数据积累，“二元传递”是API接口的价值分发，“三元激发”是企业服务的变现。本书还有一个特别实用的创新——它把“抽象分析变成了可落地的工具”。书里不光有理论、有案例，还设计了“商业模式量表”：一元模式看“价值主张、收入模型和成本”，二元模式盯“货币化率和交易成本”，三元模式算“DAU×ARPU”。这种“理论+案例+工具”的组合，让从业者能直接拿来梳理业务，投资者能对着指标判断价值，连学生都能顺着这个框架搭起系统化的商业思维，这不是虚头巴脑的“方法论”，而是真能上手用的“解码钥匙”。

而这本书最让我触动的，不是它拆解了多少案例，而是它把复杂的互联网商业拉回了“价值创造”这一本质层面。商业的根始终是“价值创造与流动”，所以新技术来了，重要的是立刻抓住它的商业逻辑。商业模式决定了一个组织的价值获取方式与发展天花板。在今天，单纯追求规模扩张已难以持续，唯有深度理解用户价值、优化成本结构、提升变现效率，才能在激烈的竞争中脱颖而出。

余昌远博士应届毕业生进入百度，就是在商业产品部，正是当时我负责的部门，也是百度商业模式的核心中枢。团队中人才辈出，很多人已经成长为各互联网大厂核心业务的领军人物。而昌远从实践中更进一步，把深厚的行业积累浓缩提炼成这样一套通透的框架。这本书里没有空泛的理论，每一个观点都带着行业里摸爬滚打的温度，用简洁的语言阐述复杂的概念，用数据支撑逻辑推演，用案例还原商业现场。现在行业早过了“抢流量”的阶段，拼的是“看懂商业”的深度，而《超越流量：商业模式新洞察》最珍贵的，就是让我们跳出“只看数据涨跌”的迷思，掌握判断商业本质的能力。希望这本书也能帮各位读者把“雾里看花”的商业逻辑变成“一用就通”的清晰路径。

推荐序二：商业模式的认知升级

叶劲

省广集团品效中心总经理

互联网浪潮席卷数十年，流量从稀缺到饱和，企业从增量争夺到存量博弈，业务模式也在不断迭代升级。很多人虽沉浸于流量数据的狂欢中，却始终困惑于“流量如何转化为实实在在的收益”和“不同互联网产品的赚钱逻辑为何天差地别”。《超越流量：商业模式新洞察》一书，恰好为破解这些困惑提供了清晰的答案与系统的框架。

作为行业深耕者，我们亲历了互联网商业模式从简单到复杂、从显性到隐性的演变。从早期的免费模式引发争议，到平台型企业崛起后的生态博弈，从长视频平台的持续亏损困境，到短视频与直播电商的爆发式增长，行业始终缺乏一套能贯穿不同业务形态的底层分析工具。本书凭借作者多年的行业实践与深度思考，跳出了“流量决定一切”的固有认知，以原子化思维拆解商业模式的核心逻辑。

书中提出的“一元生产、二元传递、三元激发”三类基础商业模式极具创新性与实用性。一元商业模式的产品与服务直接变现，二元商业模式的平台连接供需并传递价值，三元商业模式的内容场域激发潜在需求，这三类商业模式几乎可以分析当前互联网行业的所有核心业务形态。作者不仅清晰界定了各类商业模式的本质与边界，更通过对十几家知名企业的案例分析，将抽象的理论转化为可感知的商业实践，让读者能快速把抽象的理论对应到熟悉的产品与服务中，理解其盈利的底层逻辑。顺着这个思路我们就能找到市场上各种互联网公司的商业变现密码。

对于互联网行业从业者而言，本书是业务分析的“工具箱”，无论是要判断项目可行性、优化现有业务流程，还是要寻找新的增长机会，都能从中获得启发；对于投资者与创业者而言，书中的商业模式度量框架与结合数据分析的表格，能帮助他们精准把握行业趋势与项目潜力；而对于高校学生与行业新人来说，它更像是互联网商业逻辑入门的“导航图”，可以让人跳出具体产品表象，直抵商业本质。

在流量红利逐渐消退的今天，单纯追求流量规模已难以为继，理解流量背后的价值传递与商业逻辑才是破局的关键。本书摒弃了泛泛而谈的理论说教，以量化分析为支撑，以案例实践为依托，全面呈现了互联网商业化的多元路径与底层规律。相信无论是深耕行业多年的“老兵”，还是初入赛道的新人，都能从中实现

认知升级，在复杂的商业环境中精准把握机会，实现从“看懂流量”到“驾驭流量”的跨越。

在我看来，流量商业化是一门“易学难精”的学问。对大多数人而言，搭建一个基础的商业系统或许并不复杂，真正的难点在于精细化的逻辑设计与商业闭环的自洽。本书恰好为我们打开了一扇窗，让我们得以窥见各类产品与服务的商业模式大厦，其设计之精巧、逻辑之严密远超“简单投放广告”或“设置收费卡点”的浅层认知。事实上，商业模式的确立是系统性工程，若设计失当，即便手握流量“富矿”，也难免陷入“有流量却难变现”的困境。

余昌远、刘徐伟是我相识多年的好友，他们不仅逻辑缜密、擅长深度思考，更在商业化赛道深耕十多年，经验覆盖媒体侧、甲乙双方、海内外等多种场景，至今仍扎根业务一线，积累了极为丰富的实操经验。本书堪称“研学致用”的典范，是作者多年行业洞察与实战心得的结晶，对于有志于深入理解商业模式底层逻辑的读者而言，无疑是不可多得的佳作。

自序

从原始社会出现以物易物的交换行为开始，便有了商业（business）。商业与我们每个人的生活都息息相关，当我们购买商品、享受服务、玩游戏、使用搜索引擎查询信息、通过各种App和网站浏览文章或者视频时，都处在商业活动之中。只不过有一些商业活动是显性的，有一些则是隐性的。所有这些商业活动一定发生在某一个组织所构建的商业模式之中。

商业模式（business model）决定了一个组织通过什么方式来赚钱，也决定了其可规模化的程度。对于一家公司而言，一个好的商业模式组合是这家公司存在、发展和持续经营的基础。而我们也总是置身于商业模式之中，不同的商业模式通过各种方式影响着我们的生活，因此看懂公司的商业模式非常关键。

与传统的商业模式相比，互联网商业模式往往更加复杂和难以理解。很多人并不知道互联网公司是怎么赚钱的，以及能赚多少钱。哪怕是互联网行业从业者、行业研究人员和投资人等业内人士，很多时候对互联网商业模式的理解也比较模糊，更不知道该如何分析一家互联网公司的商业模式。例如，爱奇艺为什么在发展初期连续多年亏损？京东与淘宝的核心差异在哪儿？美团到家业务和到店业务在利润上有什么区别？贝壳的平台模式和链家是什么关系？什么是ACN机制？字节跳动目前一年大约4 000亿元广告收入的底层逻辑是什么？等等。对这些问题，很多人往往难以回答。

此外，随着生成式大模型的发展，大模型相关的业务该如何赚钱、我们应该如何理解大模型的商业模式等问题，也都值得思考和探讨。

本书回归本质，通过原子化的思考，对商业模式进行提炼和抽象，提出了“一元生产、二元传递和三元激发”这三类商业模式，旨在帮助读者快速理解典型的互联网业务。我们从商业模式出发，去看一个业务如何运转、如何货币化和规模化。伴随着这些深层次的思考和分析，我们会发现，在流量变现逻辑的基础上，一个业务能不能赚钱、能赚多少钱，不仅要看流量规模，更要看其对价值主张的选择和影响深度。

本书适用于互联网行业从业人员、投资行业从业人员、高校学生，以及对商业模式和人工智能感兴趣的其他读者。本书分为六章，逐步展开介绍了商业模式、商业模式画布、经营分析框架、公司案例，以及对大模型商业模式的思考。

第1章重点介绍商业模式的定义。本章从传统的商业模式出发进行原子化思考，给出了商业模式的定义：商业模式是指一个组织通过生产、传递或者激发来获得

价值的基本原理。同时，我们将互联网商业模式分为三大类：生产（一元）、传递（二元）和激发（三元）。我们日常所接触的互联网产品和服务商业化的底层逻辑是什么，可以在本章找到答案。

第2~4章针对前述的三大类商业模式分别展开，介绍不同商业模式的画布，以及属于该模式的典型公司或业务。这三章基于第1章定义的商业模式分类详细介绍了各类商业模式的分析维度，洞察它们的用户、价值主张、成本和收入等核心要素。同时针对每种商业模式，分别介绍了一些具有代表性的公司或业务。结合这三章的理论和案例，我们能够更深入地理解某些公司或业务所面临的问题的底层逻辑。

第5章重点介绍了人工智能和大模型相关的知识，并探讨了大模型的商业模式。2022年11月OpenAI发布了ChatGPT并引爆了全行业，在全球掀起一轮大模型浪潮。ChatGPT拥有强大的理解、推理、记忆和生成能力，用户可以通过自然语言灵活地与之互动，这将大模型的LUI（语言用户界面，language user interface）能力往前推进了一个时代。2025年初DeepSeek-R1横空出世，更是改变了大模型行业的竞争格局和叙事逻辑，再次引爆全球市场。本章对大模型和AI原生应用有什么特点、AI原生应用目前的发展情况以及该如何思考它们的商业模式等问题，展开做了介绍。

第6章整体介绍了商业模式度量方法，并设计了一张商业模式量表。本章首先在前述章节介绍的商业模式框架和结构分析的基础之上，明确了商业模式度量的逻辑；其次通过各案例之间的横向对比，进一步剖析了三类商业模式度量中的关键因素，包括一元商业模式中的六个因素、二元商业模式中的货币化率和交易成本，以及三元商业模式中的用户规模（DAU）和广告变现效率（ARPU），等等；最后，本章生成了一张商业模式量表，用于对商业模式关键要素的全面量化和评估，这张量表可以作为分析、诊断和设计商业模式的有力工具。

总而言之，本书对商业模式进行了原子化拆解，创新地提出了三类商业模式——生产（一元）、传递（二元）和激发（三元），并给出其逻辑图和可量化的经营分析框架，进而完成了对商业模式的度量。在理论之外，书中还通过对16家知名企业的分析和洞察，揭秘了企业盈利的秘密（此处容我特别说明一下，本书提到的所有公司和业务，仅代表作者个人观点，不构成任何投资建议）。本书并非泛泛而谈，而是融入了作者的业务经验和抽象思考，给读者带来了兼具量化逻辑和案例分析的商业模式新洞察。

由于篇幅限制，一些案例分析的细节无法尽述。同时，市场在不断变化，很多公司的业务也在持续迭代。尽管我们在写作过程中对行业信息和数据进行了深度调研和交叉验证，但由于作者水平有限，书中难免存在疏漏之处，恳请读者朋友们

多多包涵，并提出宝贵的意见和建议。如果您在阅读过程中发现任何问题或错误，欢迎随时与我们联系，我们会及时进行勘误和修正。

最后，我要感谢中国人民大学出版社的编辑老师们。在合作过程中，你们出色的专业能力、敏锐的市场洞察力、严谨的工作态度以及不辞辛劳的付出，都让我深受启发、受益良多。本书能够顺利出版，离不开你们的大力支持。由衷地感谢所有人。

希望读者朋友们在阅读本书的过程中，能够感受到商业模式的魅力。

第1章 商业模式研究

1.1 商业模式的原子化思考

商业模式是一家公司存在、发展和持续经营的基础。伴随着互联网过去几十年的快速发展，全球诞生了很多大型互联网公司，它们创造了一个又一个商业奇迹，大家耳熟能详的阿里巴巴、腾讯、字节跳动、京东、拼多多、美团、Google（谷歌）、Meta（脸书，即Facebook）、Amazon（亚马逊）、Uber（优步）、Yandex等均属此列。在观察这些互联网公司的时候，我们会发现互联网公司的商业模式与传统行业的商业模式不尽相同，前者更加复杂、更加隐性，往往很难用一个简单的模型或者框架来描述清楚，甚至有些公司的商业模式很难被非业内人士所理解。

在我们调研分析和分享交流一些互联网公司的业务模式、增长潜力、财务报告乃至发展历程时，下面四个话题经常被讨论到，恰好作者认为这些话题的本质都是商业模式。

1. 免费

互联网公司会免费给用户提供各种产品和服务，仿佛一切产品和服务只要被互联网化，就有了免费的可能性，这与传统公司的做法差异巨大。作为普通用户，你可以想一下每天使用的互联网产品和服务有多少是免费的：微信是免费的，可以免费地聊天、视频通话、发朋友圈、看文章、看视频等；Google是免费的，可以免费地搜索、看新闻、看视频、看直播等；抖音是免费的，可以免费地看视频，看直播，甚至看电影等；微博、知乎、脉脉、哔哩哔哩（B站），以及大量的工具App，如天气、日历、Wi-Fi管理、清理大师、应用商店、笔记、摄影摄像等等，全都可以免费使用。实际上，这些产品和服务都有成本，包括但不限于带宽成本、服务器成本、渠道成本、内容采购成本、员工成本、运营成本等等。为什么互联网公司能免费提供这些产品和服务？它们是如何获取收入来覆盖成本，并将利润投入再生产的？我们从后验视角来看，这些互联网公司不但能赚到钱来维持公司的运营并扩大公司规模，甚至还能赚取超额利润，造就无数的财富神话。或许大家会认为之所以能免费，是因为这些互联网产品和服务的边际成本极低，多一个用户，公司增加的成本是极少的，甚至有些产品和服务的边际成本趋近于零。边际成本低的确是某些业务能规模化的底层逻辑，但本质上免费是由商业模式决定的。在这样一套商业模式中，你的注意力、你的时间都会变成商业价值，最终转化为公司的收入和利润，而且利润率往往还非常高。

2. 瓶颈

很多互联网公司发展到一定阶段，往往在规模还没有很大的时候，业务就会出现瓶颈、增长困难，或者是迟迟无法盈利，不具备造血能力，只能靠资本续命，更有甚者主营业务出现危机，进而导致公司破产清算。我们可以来看两个典型的例子：长视频平台爱奇艺和曾经辉煌一时的ofo小黄车。（1）爱奇艺是2010年含着金汤匙出生的，2018年顺利在纳斯达克上市。爱奇艺这些年在版权、自制剧、自制综艺等方面都有着非常大的投入，也为观众们带来了不少的优质内容。但这看似风光的背后却是爱奇艺公司连续12年亏损（保守估计，爱奇艺自成立至2022年累计亏损超过400亿元，直到2023年才开始慢慢盈利），需要靠资本不断输血来维持公司的发展。（2）ofo小黄车2015年带着“以共享经济+智能硬件，解决‘最后一公里’出行问题”的理念从北京大学横空出世，受到了大量资本的追捧，短短三年内便获得了超过20亿美元的融资，一时风光无两；但2018年ofo陷入债务危机，而且更荒诞的是直到2025年还有用户的押金没有退。诚然，这些公司遇到的发展瓶颈，或者类似ofo这种非常不体面的结局受到了多方面因素综合影响，例如公司经营不善、管理混乱、财务混乱、用人不佳、合伙人内斗、恶性竞争、成本失控，甚至资本作恶等等。但本质上，这些问题只会带来一部分影响，更核心的原因是商业模式的问题。

商业模式决定了业务可规模化盈利的程度和路径，也决定了业务的天花板。

3. 基因

在谈到行业内一些公司的时候，尤其是在不同公司之间进行横向对比的时候，大家总是会提到“基因”这个词，每个公司都有自己擅长的事情，都有自己的基因。例如很多时候大家会认为腾讯的产品强、阿里的运营强、字节的推荐算法强、美团的地推能力强，所以腾讯是产品驱动、阿里是运营驱动、字节是技术驱动、美团是业务拓展（BD）驱动等等。但不同的职能驱动只是商业模式带来的结果。本质上，是商业模式决定了公司基因和职能驱动力。不同的商业模式，其关键影响因子不一样，侧重点不一样，需要解决的问题也不一样，所以各职能驱动力的强弱自然也就不一样。而且随着互联网增量红利的减少，各大公司都进入存量竞争阶段，大家的业务越来越综合，很多事情做得越来越深和越来越重，商业模式也同步延展和变化，进而影响各种职能在业务中的驱动力，各公司基因之间的差异也就越发不明显了。

4. 创始人

创始人对公司的重要性不言而喻。在讨论一些公司的发展历程时，大家很喜欢讲创始人的故事，并在创始人之间进行横向对比，他们之间的渊源和一些似有似无

的细节，让故事充满了戏剧性。但商业模式的差异对公司发展的影响，并不亚于创始人的差异。大家在对比同类或者同时代公司的时候，最先关注的应该是其商业模式的差异。最典型的例子就是雷军，雷军在金山和小米的经历很好地说明了即使同样的人带队，不同的商业模式也会给公司的发展带来截然不同的发展道路和市场结局。商业模式是公司的基石，如果没有好的商业模式支撑，不论有多么优秀的创始人带队，公司的经营都依然会举步维艰。

我们要看一个公司或者一项业务，首先要看的就是商业模式。公司的很多现象和问题，本质上都由其商业模式决定。公司的管理、增长、人才储备、组织建设、竞争力等方方面面，都有着商业模式的影子。因此在研究公司的时候，不论是为了投资、工作、求职，还是为了个人兴趣，研究和分析商业模式都是重要且必要的。

那商业模式到底是什么呢？我们先不着急给商业模式下定义，先来看看商业模式的复杂性。

一个公司或者公司里的一项业务，其商业模式往往非常复杂。我们可以来看两个例子：小米和京东。

小米，2010年由雷军创立，经过十几年的发展，成功跻身世界500强（2024年《财富》世界500强排行榜，小米集团排名第397位）。小米的业务也已经非常庞杂了，如硬件[手机、电脑、电视、智能家居、大家电（空调、冰箱、洗衣机等）]、电商（米家、有品）、金融、游戏以及2021年开始进入的智能电动车领域（小米汽车）等等，很难用一个商业模式来概括小米公司所有的业务。就算聚焦于小米的手机业务，其商业模式依然复杂：首先，手机的硬件销售会有一部分利润（值得注意的是，不同价格区间的手机，其硬件利润率也有差异）；其次，手机中会有很多小米自有的App，例如应用商店、浏览器、小米视频、小米游戏、云服务、安全中心、天气、日历等，这些App会贡献一定的广告收入、会员收入和增值收入；最后，手机在出厂时，还会自带很多第三方公司的App，这些又会贡献一定的App预装收入。手机业务的这些收入构成，其底层的商业模式也不完全一样。

京东，1998年由刘强东创立，从中关村一个线下柜台起家，发展为一个兼具实体企业基因与科创基因的集团，也成功跻身世界500强（2024年《财富》世界500强排行榜，京东集团排名第47位）。京东的核心业务包括零售（含电商和京东便利店）、物流、健康、云和金融等，同样也很难用一个商业模式来概括京东所有的业务。就算聚焦于京东的电商业务，其商业模式依然复杂，其中至少包含三种业务模式：第一，京东直接作为卖家，在京东开店，也就是消费者看到的京东超市和京东自营店；第二，京东作为电商平台，招募第三方卖家在京东开店；第三，

京东打造自有商品品牌，如京东京造，直接参与商品的设计或生产，并将产品通过京东销售给消费者。这三种业务模式背后的商业模式不完全一样，其成本结构、货币化率、利润率以及业务的侧重点和竞争力也都有所不同。

像小米和京东这样的案例比比皆是，公司和业务的规模或大或小，但商业模式都有可能很复杂。由此可见，从宏观角度去讲商业模式会非常笼统，无助于我们高效且清晰地洞察其本质，而且每个商业模式所依赖的关键点也都不一样，其中的成本、收入、价值主张、渠道和资源依赖等也不可一概而论。

既然公司的商业模式如此复杂，那么我们不妨回归本质，对商业模式进行原子化思考。

作者认为一项业务的商业模式是多种商业模式的组合，而一个公司往往又拥有多项业务。所谓原子化思考，就是指洞察到商业模式的最小单元，抽象出几类最基础的商业模式，然后就可以像搭积木一样，用几种简单的积木，组合搭建出所有公司和业务的商业模式。

基于原子化思考，这里就可以给出作者对商业模式的定义：商业模式是指一个组织通过生产、传递或者激发来获得价值的基本原理。

在这个定义中，有三个关键词：生产、传递和激发。这三个关键词就对应着互联网公司最基础的三种商业模式。本书接下来的内容，就围绕这三个关键词展开，给大家介绍互联网商业模式的原子化形态。

在此之前，我们先来看一下最传统最原始的商业模式。

1.2 传统的商业模式

最传统最原始的商业模式其实非常简单。我们可以回到商业的原点，看看以物易物的交换行为背后的逻辑。

在商业的原点，人们用以物易物的方式，交换自己所需要的物品，例如A用一只羊，和B交换两把石斧，各取所需（见图1-1）。

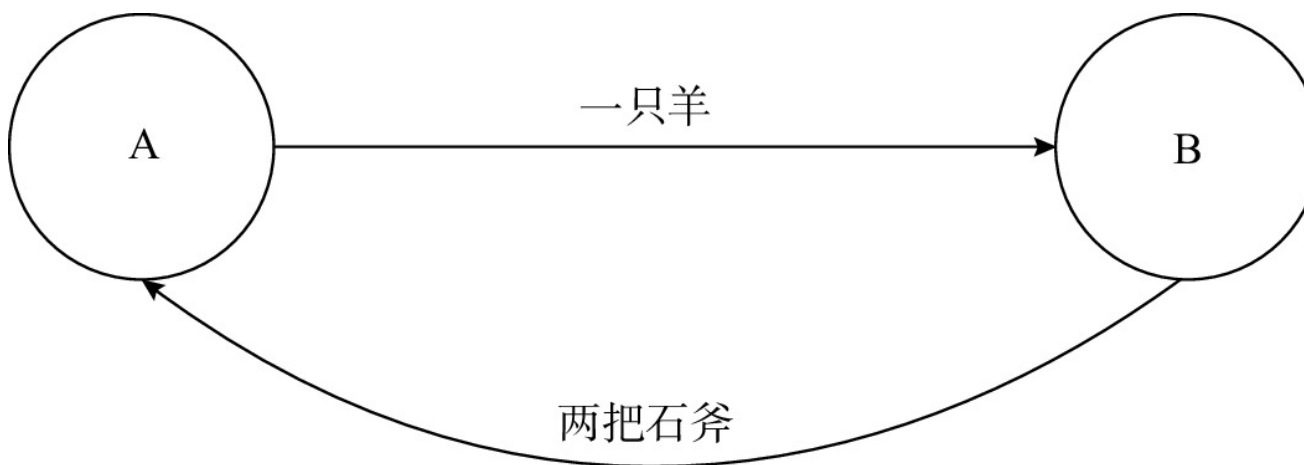


图1-1 商业的原点：以物易物

随着交换行为越来越频繁，交换物品的种类和交易对象所受的限制越发明显，于是就发展出了最早期的货币，例如A给B提供一只羊，B可以支付给A对应的货币，A拿着货币，从C那里购买两把石斧（见图1-2）。自此，商业就开始了其繁荣的发展。

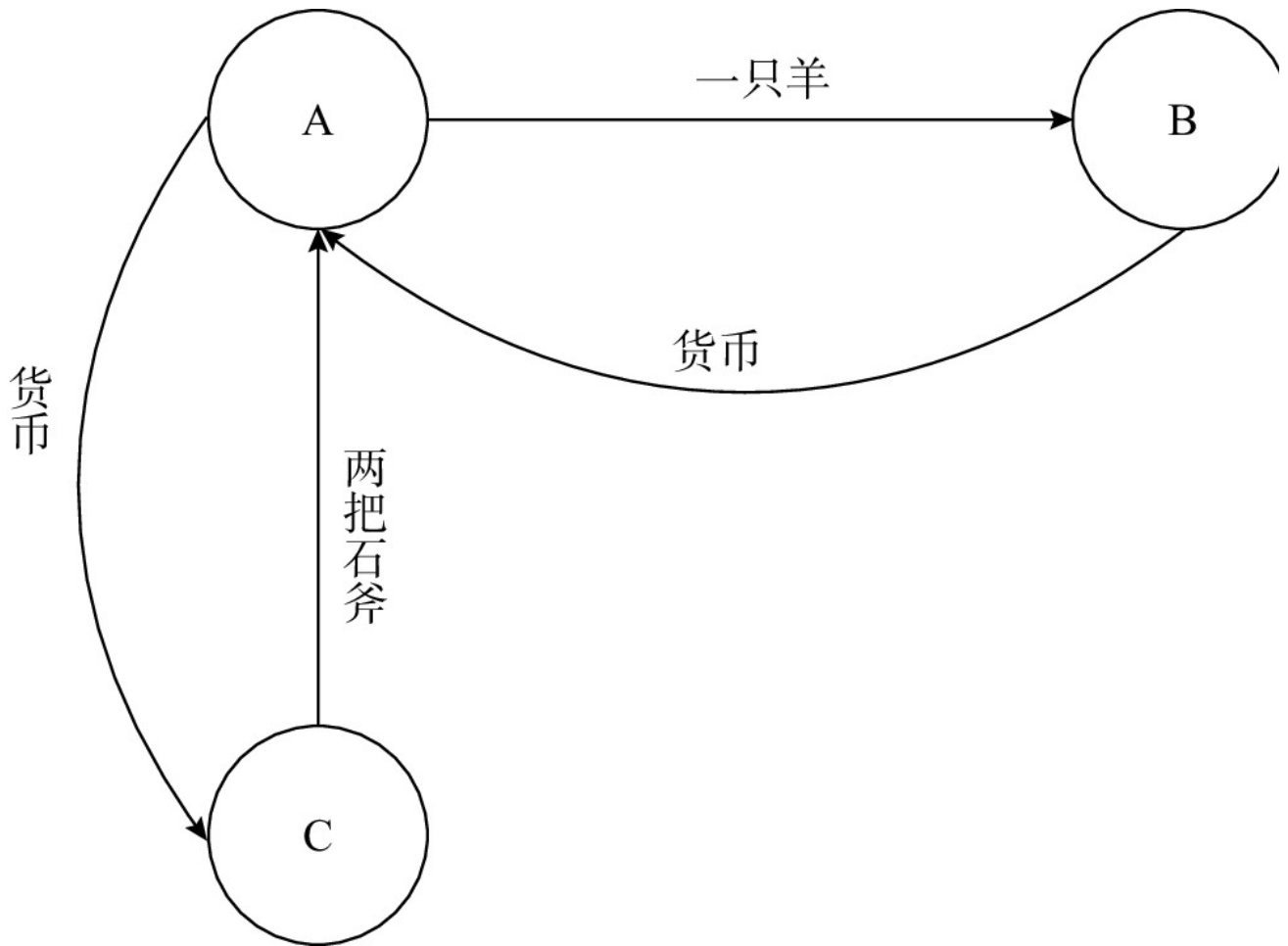


图1-2 商业的起点：以物易物→引入货币

本质上，这种以物易物，以及引入货币来协助交易的最传统商业模式可以抽象为一个组织生产商品，把商品提供给目标用户并获得价值的过程，即生产商品→用户（见图1-3）。

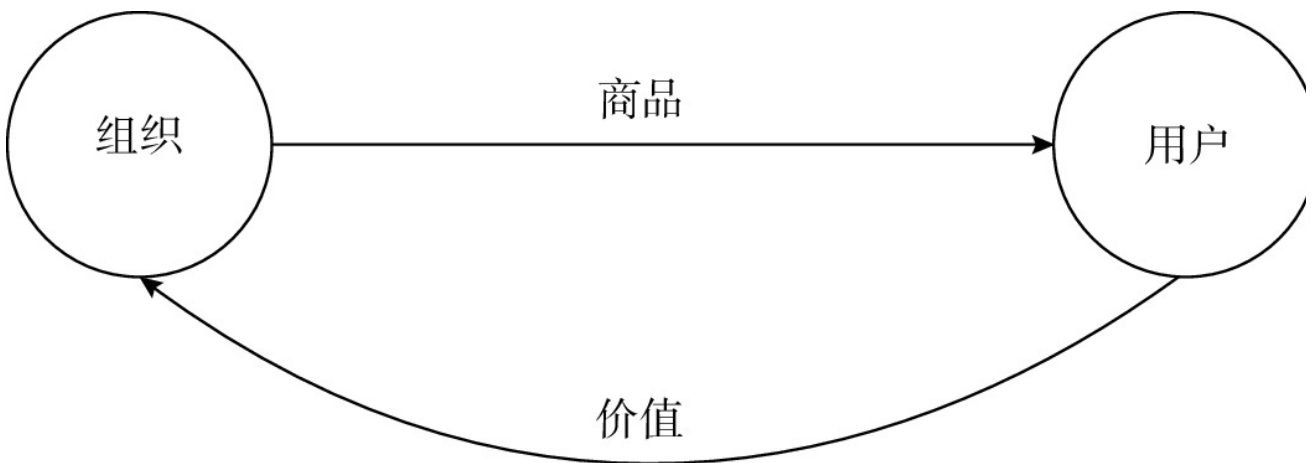


图1-3 为客户提供商品，并获得价值

这里的商品和用户都是广义的：

- **商品**：广义的商品，可以是产品也可以是服务，可以是实体的也可以是虚拟的。
- **用户**：广义的用户，可以是C端的普通消费者，也可以是B端的商家客户。

例如某服装公司生产了服装，售卖给普通的消费者，某咨询公司能做战略咨询，为另一家公司提供咨询服务等，都体现了这种传统的商业模式。

本质上，传统的商业模式就是组织生产产品或服务，把产品或服务提供给目标用户，并获得价值的基本原理。我们将这种商业模式定义为生产。它是一元的。

生产的商业模式逻辑如图1-4所示。

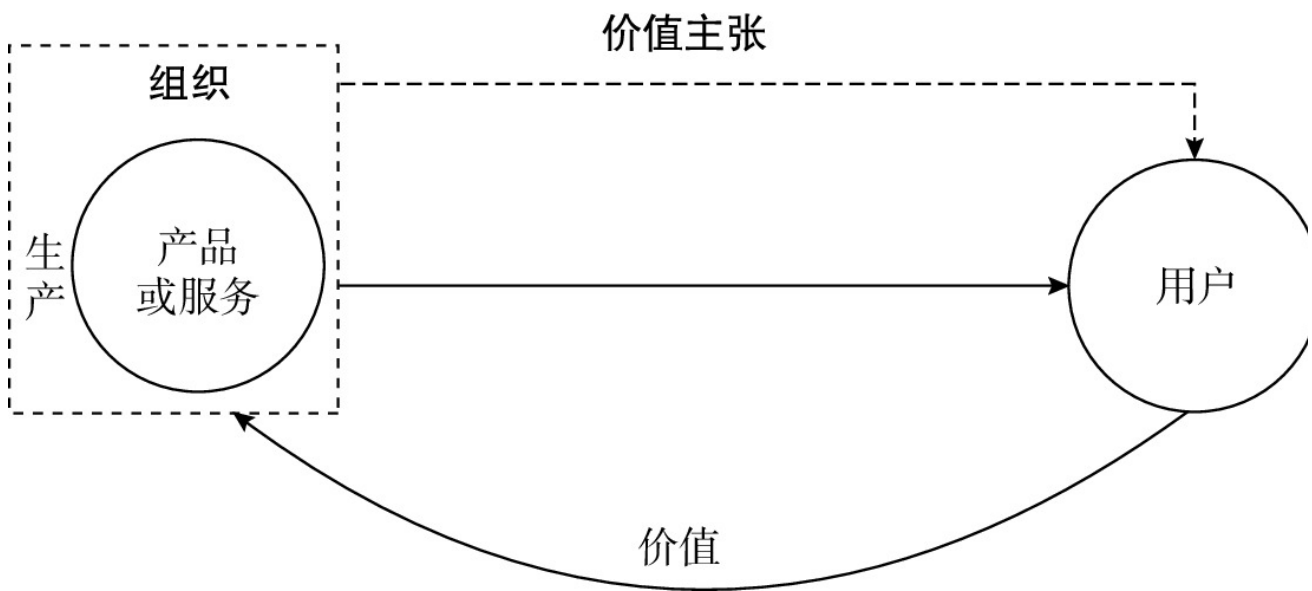


图1-4 传统的商业模式：生产（一元）

组织为了追求利润，针对其服务的目标用户，提供一定的价值主张，并通过渠道触达用户，进而获取价值。在这个过程中，当组织的收入大于成本时，组织就可以实现盈利，赚取利润。

值得注意的是，在传统业务中依然充满商业模式的创新。有些创新虽然没有改变商业模式的本质，但却大大提升了商业模式的效率；而有些创新则彻底改变了商业模式。在商业史中，有两个经典的创新，堪称商业模式发展的代表作：剃须刀模式和免费报纸模式。

1. 剃须刀模式

20世纪初，金·吉列首次向市场推出了可替换刀片的剃须刀，刀片和剃刀是组合产品，吉列的剃刀需要使用吉列的刀片。于是金·吉列就大力推广销售这类剃须刀，几乎以成本价销售，甚至在顾客购买其他产品的时候赠送剃刀，来为刀片创造长久的需求。然后，金·吉列通过售卖刀片，获取顾客长久的价值。这不仅可以赚回剃刀的成本，而且还能绑定顾客，从刀片销售中获得长期且较高的利润。金·吉列在推出自己创意后的一年内，就售出了9万个剃刀和1 200万片以上的刀片，可见商业模式创新能带来巨大的价值。剃须刀模式在商业中十分流行，如打印机厂商通过不断地销售低毛利的打印设备，提高打印机的市场保有量，带动后续配套高毛利的打印机耗材的持续销售，进而实现长期稳定的盈利。**剃须刀模式并没有改变传统商业模式的本质，但改变了其收入模型：从获取单次价值，变成获取用户全生命周期的价值（life time value, LTV，即从用户获取到用户流失所得到的全部收益的总和）。**这大大提升了传统商业模式的效率。

2. 免费报纸模式

20世纪90年代，源自瑞典的《地铁报》（*Metro*，世界上首份现代免费报纸，于1995年在瑞典斯德哥尔摩创刊）革命性地改变了传统日报的商业模式。它通过免费发放的形式，精准覆盖通勤人群，将营收核心完全转向广告收入。其成功迅速启发了全球报业，其中，英国报业巨头罗瑟米尔勋爵在考察后深受鼓舞，于1999年在伦敦创办了英国版《地铁报》，自此将免费报纸模式推广至全球，引发了报业格局的深刻变革。**免费报纸模式彻底改变了传统的商业模式：不再是组织生产报纸，用户付费购买报纸；而是组织生产报纸，免费提供给用户，再从广告商处获取收入，这些收入可以覆盖所有的成本，并且能继续投入对更多报纸的生产。**

剃须刀模式和免费报纸模式，是传统行业中最接近互联网商业模式的，甚至可以说，互联网商业模式的设计，有从传统的商业模式创新中寻找灵感。剃须刀模式提出了组织在获取价值时可以撬动LTV；免费报纸模式则直接改变了商业模式。

接下来我们展开介绍互联网的商业模式，但在此之前先看一下“商业模式的核心要素”和“商业模式画布”的相关内容。

1.3 商业模式的核心要素和商业模式画布

1.3.1 商业模式的核心要素

前面介绍传统一元商业模式的时候引入了一些概念，例如用户和价值主张等，我们称之为商业模式的核心要素。根据亚历山大·奥斯特瓦德和伊夫·皮尼厄在《商业模式新生代》中的介绍，这里给出相关核心要素的定义，这些定义在所有的商业模式描述中都适用。

- **用户**：组织服务的目标用户，是广义的，可以是一个或者多个群体。
- **价值主张**：组织能为用户创造的独特价值。
- **渠道**：组织连接用户的通路和桥梁，负责价值的传递和交付。即价值主张通过渠道触达用户，用户也需要通过渠道连接组织及其价值主张。
- **成本**：属于组织的金融能力。是指商业模式运转所需要的所有成本，包括但不限于生产成本、渠道成本、供应链成本、资源成本、组织经营成本以及获客成本。
- **收入**：属于组织的金融能力。是指组织在商业模式运转过程中所有的收入，收入来源取决于组织的商业模式。

另外，为了实现上述各要素的供给和交付，组织还需要一些**核心资源**。核心资源可能是生产所需要的原材料、供应链关系、关键技术、知识产权、人才等，也可能是渠道的合作伙伴网络等。

1.3.2 商业模式画布

在研究商业模式的时候，有一个常用的工具：商业模式画布（business model canvas, BMC）。

我们以商业模式的核心要素为基础先画一张图，再将这些核心要素放到图里面，就可以形成一张商业模式画布的底板，然后基于这张底板，把每个核心要素都拆解出来，填到对应的区域，就可以完成一张完整的商业模式画布。

商业模式画布是指基于商业模式的核心要素，在一张画布中将某项业务的商业模式拆解和描述出来，是一种用于描述、观察、评估和创新商业模式的通用工具。正如《商业模式新生代》中提到的，商业模式画布是一个非常好用的商业模式分析工具，你可以将新的或既有的商业模式在画布上描述出来，甚至可以将商业模式画布打印出来，让一大群人用便利贴和马克笔共同描画和讨论一个商业模式的各个要素，这样的使用方式可以最大限度地发挥这个工具的价值。这是一个可以帮助理解、讨论、分析和创造的操作工具。

一般可以通过三个步骤来完成一张商业模式画布：第一，在纸上画出商业模式核心要素的模块；第二，用关键词、数字或者简洁的句子来描述和拆解每个要素，并填到对应的模块；第三，做一些必要的标注，例如差异点、创新点、短板等。

商业模式画布可以帮我们统一话语体系，在梳理、分析、研究和设计商业模式的过程中，建立通用的思考框架，提升思考的质量和效率。借助商业模式画布，我们能快速拆解一个业务的商业模式，拥有全局视角，不遗漏关键点也不陷入细节。商业模式画布还可以帮我们设计出新业务的商业模式，这对于创业团队和创新方向都很有价值。另外，还有一点值得一提，商业模式画布能帮助我们评估个人作为个体的“商业模式”，是一个很好用的自我评估工具。

1.3.3 传统商业模式的商业模式画布

根据商业模式的核心要素和商业模式画布的定义，我们回到传统一元商业模式（组织将价值主张提供给用户的过程），来画一下它的商业模式画布。

第一步：在纸上画出商业模式核心要素的模块。

传统一元商业模式的核心要素有五个：价值主张、渠道、用户、成本和收入（见表1-1）。

表1-1 传统一元商业模式的核心要素

价值主张	渠道	用户
成本	收入	

第二步：用关键词、数字或者简洁的句子来描述和拆解每个要素，并填到对应的模块。

我们分别梳理并描述每个核心要素，并将其填到画布中对应的模块（以一只羊的交易为例，见表1-2）。

- **价值主张：**组织生产的产品或服务，这些产品或服务需要匹配用户的需求，一般都是实物。
- **渠道：**价值主张需要渠道来触达用户，用户也需要渠道来了解并购买组织的产品或服务。如果没有渠道，不论组织的产品或服务多么优秀，都没法产生价值，其实酒香也怕巷子深。传统业务中的渠道一般都是线下的，通常为多级经销商或经销体系。
- **用户：**传统一元商业模式中最重要的角色，组织服务的目标用户。
- **成本：**组织生产产品或服务所需的原材料成本、渠道成本、人力成本、获客成本等。
- **收入：**组织将产品或服务成功提供给用户就能产生收入。在传统一元商业模式中，收入往往是单次的，发生一次交易就获得一次收入，并赚取当次交易的利润。

表1-2 传统一元商业模式核心要素的描述和拆解

<p>价值主张 产品或服务， 一般都是实物 (一只羊)</p>	<p>渠道 线下渠道，如经销商、经销体系 (面对面交易、线下交易市场)</p>	<p>用户 C端用户、 B端用户 (需要羊肉的个人或者餐馆)</p>
<p>成本 原材料成本、渠道成本、 人力成本、获客成本 (羊苗成本、饲养成本、运送成本等)</p>		<p>收入 用户付费， 单次交易收入模型 (羊的售价)</p>

第三步：做一些必要的标注，例如差异点、创新点、短板等（这里我们主要标注其潜在的竞争力）。

在传统一元商业模式中，价值主张、渠道、核心资源、成本控制和收入结构都有可能成为业务的核心竞争力所在，例如：

● **价值主张：**产品或服务的创新性是否满足了用户新的需求或者为用户提供了更好的商品；商品所面临的市场竞争情况，商品的可替代性，以及市场是否为垄断市场；商品是否可规模化生产，生产的产能以及边际成本如何支撑商品的规模化生产。

● **渠道：**广度和深度上，渠道能覆盖多少目标用户，以及构建的壁垒高不高；效率上，渠道是否具备网络效应，能不能高效触达目标用户；与组织的关系上，渠道是自建的还是合作的，组织对渠道的把控力如何。

● **核心资源：**组织所拥有的核心资源，不论是内部的还是外部的，是否具备一定的优势；有哪些核心资源是独特的、独占的或是持续创新的。

● **成本控制：**成本结构的健康度；固定成本和变动成本的具体情况；商品生产的边际成本情况；商品的库存周转效率；等等。

● **收入结构：**组织是否拥有健康的现金流以及较高的利润率水平，是否由于具有垄断地位而具备一定的定价权，是否由于具有品牌属性而具备一定的溢价能力。

读者可以以一只羊的交易为例，思考一下如何提升竞争力。至此，我们完成了传统一元商业模式的画布。

1.4 互联网的三种商业模式

互联网的商业模式远比传统的商业模式复杂，但是通过原子化思考，我们就会发现其中蕴含着传统的商业模式。本小节会围绕三个关键词——生产、传递和激发——来展开，用这三个关键词来定义互联网最基础的三种商业模式，并给出对应的商业模式逻辑。

1.4.1 一元商业模式：生产

互联网的第一种商业模式，我们将其定义为生产，它是一元的。

我们可以想一下一些日常使用的互联网产品和服务：玩王者荣耀、原神、和平精英、梦幻西游、NBA2K、三国杀Online等游戏，玩家需要付费购买道具和角色；互联网小工具，例如印象笔记、Autosleep等，我们也会付费购买；日常使用各种正版的桌面软件，例如Xmind思维导图、Office套件、Photoshop等，需要付费；办公场景使用的钉钉、飞书、腾讯会议、Zoom等办公套件，企业也需要付费购买。像上述这些为我们提供产品或服务的企业，其商业模式与传统的商业模式本质上没有区别，即由企业生产商品，提供给目标用户，并从用户处获取对应的价值。因此这类业务本质上是在互联网的加持下，用传统的商业模式经营的。

参照前文，我们可以给出互联网一元商业模式的定义：组织生产产品或服务，把产品或服务提供给目标用户，并获得价值的基本原理，即生产商品→用户。其商业模式逻辑如图1-5所示。

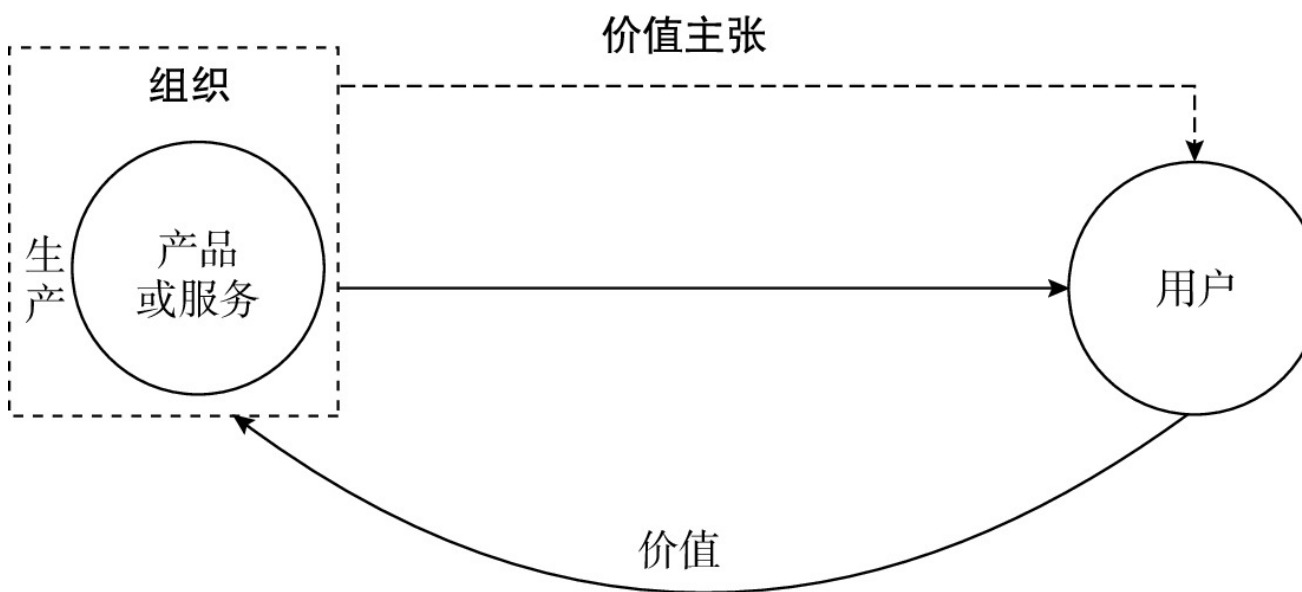


图1-5 互联网商业模式：生产（一元）

这种商业模式比较简单，与传统商业模式基本一致，有很多互联网公司或业务都是采用的一元商业模式，例如IAP和SaaS等，具体第2章会展开介绍。

透过一元商业模式，我们需要对商业模式的多元和演化进行更深入的思考。大家不妨和作者一起换一些视角来看问题：例如站在渠道的视角，看渠道的商业模式，它显然就不是一元的了。渠道虽然不生产商品，但能帮组织触达用户，并为组织提供

多元的价值，如履约、支付、售后等。又例如站在组织获取用户、打造品牌的视角，最常见也是最重要的一种方法就是**投放广告**，包括效果广告和品牌广告，那些拥有广告资源位（广告库存）、能提供营销平台、能为广告主带来广告曝光和转化的公司（在这个场景中，有着一元商业模式的组织只是这些公司的广告主），其商业模式又是怎样的呢？

带着这些从不同视角出发的思考，接下来我们继续介绍二元和三元商业模式。大家也会进一步看到广告在二元和三元商业模式中的价值和重要性。

1.4.2 二元商业模式：传递

互联网的第二种商业模式，我们将其定义为传递，它是二元的。

很多公司的一些业务模式中，组织是不生产商品的，只是在连接商家和用户。我们可以看两个典型的例子：滴滴出行和淘宝。

滴滴出行通过滴滴出行平台为用户提供打车服务，同时司机可以通过滴滴出行平台接到打车的订单并获取收入，但实际上滴滴出行并不拥有车和司机。很明显，在打车业务中，本质上是司机作为商家，为用户提供打车服务，而滴滴出行平台只是在连接商家和用户。

阿里巴巴公司通过淘宝电商平台为用户提供线上购物服务，同时淘宝卖家可以通过淘宝电商平台售卖商品。同样，阿里巴巴公司并不拥有也不生产商品。在淘宝电商业务中，本质上是用户从商家（淘宝卖家）那里购买所需的商品，而淘宝只是在连接商家和用户。

在上述这些业务中，组织不再生产商品，而是作为一个双边平台，连接商家和用户，并分别为两者提供不同的价值主张。

商家和用户，这两个重要的角色，本质上都是组织的目标用户。其中，商家能为用户提供产品或者服务，而且这里的商家是广义的，可以是B端商家，也可以是C端个人商家；用户带着商业需求来找产品或者服务，即用户是带着“一笔钱”来的。组织所做的是对用户的商业需求进行调度，在商家和用户之间进行商业价值的传递，完成“一笔钱”的传递。通俗地讲，就是用户带着“一笔钱”来寻找商品，组织负责将这笔钱传递给对应的商家，由该商家给用户想要的商品。

为了完成传递的过程，组织需要给它的目标用户——商家和用户——提供对应的价值主张。这些价值主张可能存在于多个不同的环节，需要进行大量的基础建设，例如：

- **商家：**组织为商家提供线上店铺装修、库存管理、订单管理、账单管理、用户运营平台、获客工具、选品工具、交易手续办理等服务。
- **用户：**组织为用户提供信息匹配、商品质量保障、履约保障、资金安全保障、人身安全保障、隐私保护、售后服务等服务。

在不同的业务场景（例如二手房交易业务和打车业务）中，组织所需提供价值主张的差异可能会非常大。

商业价值每发生一次传递，就会产生一次交易和对应的交易额（GMV），组织也就在传递中获得价值，完成货币化。货币化是二元商业模式中组织获取价值的逻辑，其本质上就是收取服务费。货币化（即组织获取价值）的方式有很多种，包括但不限于收取商家入驻和管理费用、交易抽佣（手续费/服务费）和收取广告费。在二元商业模式中，组织绝大多数情况下是从商家侧获得价值，但也有少数情况会从用户侧获得价值，典型的如二手房交易业务。

一般来说，这类组织会看货币化率这个指标（货币化率=服务费/交易金额）。组织在传递商家和用户之间的商业价值时，对用户的调度效率越高、对用户行为的控制力越强、提供的价值主张涉及环节越多且价值越大，货币化的潜力就越大，组织的盈利空间就越大。

综上所述，二元商业模式是指：组织建设线上场域连接商家和用户，完成商家和用户之间的商业价值传递，并获得价值的基本原理，即用户→商家。其商业模式逻辑如图1-6所示。

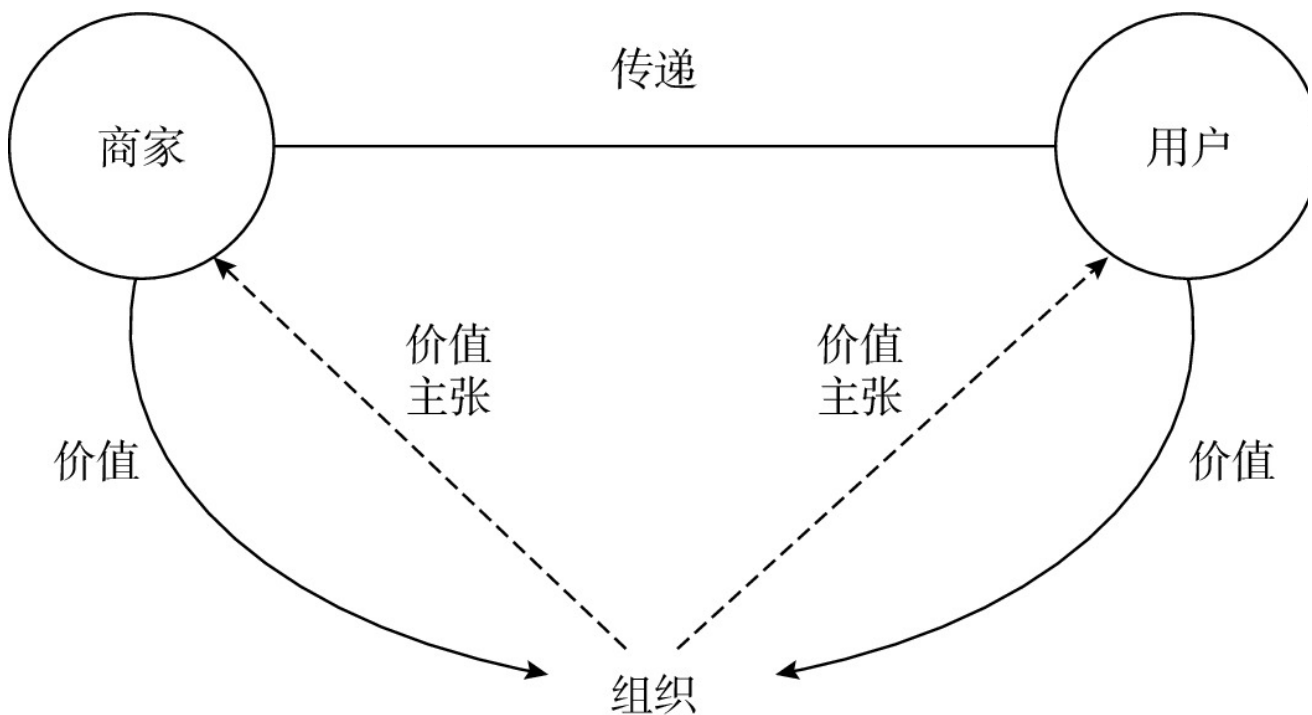


图1-6 互联网商业模式：传递（二元）

二元商业模式虽然比较复杂，但造就了很多伟大的公司。我们常说的“平台业务”很多属于二元商业模式，最典型的的就是交易平台和信息平台。二元商业模式的底层逻辑是什么，如何形成飞轮效应，其经营分析框架是怎样的，这些内容我们将在第3章展开介绍。

二元商业模式的核心是：组织面对的是带着“一笔钱”的用户。但用户并不是在使用所有互联网产品或服务时都带着“一笔钱”，例如用户只是过来获取信息和娱乐，看看新闻和视频，这类业务的商业模式逻辑是怎样的呢？接下来，我们继续介绍三元商业模式。

1.4.3 三元商业模式：激发

互联网的第三种商业模式，我们将其定义为激发，它是三元的。

对比一元生产和二元传递的商业模式，我们会看到还有很多公司的一些业务模式中，组织既不生产商品，也不仅是连接商家和用户，面对的也不是带着“一笔钱”的用户，而是连接用户和内容，免费为用户提供内容，并通过广告来获得价值。这跟前文提到的免费报纸模式如出一辙。我们可以看两个典型的例子：快手短视频和今日头条。

快手短视频最基础的形态，就是连接用户和短视频内容。用户在快手刷短视频是免费的，而在用户免费刷短视频的过程中，平台（快手）会插播广告，广告主给快手支付广告费，快手以此获取收入。同时，在这个过程中，短视频创作者还可以从快手获得分成收入，促进创作者持续生产更多更好的内容。当然，随着短视频平台不断迭代和进化，其发展出直播、电商、本地生活、游戏、短剧、长视频等不同的业务，形成了庞大的商业生态，各种商业模式混合在一起，这也体现了互联网业务商业模式的复杂性和商业模式原子化思考的必要性。

今日头条（字节跳动公司旗下产品）与快手短视频类似，其最基础的形态是连接用户和资讯内容，用户在今日头条浏览资讯内容是免费的，资讯内容的形态是多模态的，从最开始的图文，逐渐演化并涵盖了图文、视频、直播、问答等多种形式。今日头条也是通过在免费的资讯内容中插播广告，最终获取收入。

在上述这些业务中，组织免费为用户提供内容，并通过广告来获取价值。这里有三个关键的角色——内容（创作者）、用户和广告主，本质上它们都是组织的目标用户。其中，内容或者说创作者，即内容的供给方，会生产和提供满足用户需求的内容。根据内容生产模式的不同，主要分为UGC（普通用户创作内容，user generated content）、PGC（专业用户创作内容，professional generated content）、PUGC（介于普通用户创作内容和专业用户创作内容之间，professional user generated content）和AIGC（人工智能创作内容，AI-generated content）。用户，即内容的需求方，在组织提供的场域内免费消费内容。广告主，一般是B端商家，在组织提供的场域内进行企业营销活动，通俗地讲，就是投放品牌广告和效果广告（线上），触达目标用户，进行企业的品牌建设和达成效果转化。同时，随着在线营销产品不断发展，企业可以在线上场域打造经营阵地，实现企业的长效经营。

内容生产需要成本，不论是前面提到的哪种生产模式都需要成本，创作者只有能赚到钱才有可能持续地生产满足用户需求的内容。而用户却可以免费消费这些内容，例如免费看新闻、免费看视频、免费听音乐等等。从需要用户付费订阅到用户可

以免费订阅，商业模式的创新便在于此：组织引入了广告主，在用户消费内容的时候，给用户推送广告，激发用户的商业需求，为广告主带来生意和收入，并从广告主侧获取价值，即广告费。同时，组织将一部分收入通过一定的方式（例如广告分润、达人营销、带货等）分给创作者，给创作者带来收益。

三元商业模式与二元商业模式对比，最大的差异点在于：在三元商业模式中，用户并没有带着“一笔钱”来，而是带着“时间和注意力”来，组织通过商业内容，激发用户的商业需求，完成对广告主的交付。

在三元商业模式中，组织构建的是一个多边平台，连接内容、用户和广告主，并分别为三者提供不同的价值主张。为了持续激励更多的内容生产，吸引更多的用户，引入更多的广告主到平台投放广告，组织的价值主张多元且丰富，例如：

● **内容：**组织为创作者提供创作工具、创作灵感、创作方向指导、内容管理和数据分析等等，降低创作门槛，提升创作效率，管理内容生态。当然还有最重要的一点，即给创作者带来收入。组织会致力于打造繁荣的内容供给生态，激励创作者持续生产符合平台价值导向的优质内容。

● **用户：**组织为用户提供丰富的内容，而且用户消费这些内容都是免费的。随着技术的发展，平台能更好地理解用户，识别用户的兴趣偏好，在海量的内容中找到满足用户需求的内容，所谓平台比用户自己更懂用户。与此同时，随着创作门槛的降低，很多用户可以转变为创作者，正如微信公众平台的广告语所言：再小的个体，也有自己的品牌。

● **广告主：**组织为广告主提供流量、触点、营销工具、广告创意制作、数据洞察、经营阵地、私域管理、应用程序接口（API）以及各种营销服务等，帮助广告主触达、影响以及经营用户，最终实现生意和收入的增长。

组织为了提升自身的竞争力和构建“护城河”，会向各方向渗透，打造自己的生态，强化对各方的价值主张。典型的是组织有可能自己演变成内容的供给方，构建自己的内容生态和创作者生态，为创作者深度赋能。随着内容越来越多，内容的质量、丰富度和差异性成为组织的核心竞争力之一。组织同样也会构建自己的广告投放生态，从技术、服务和收益等方面全方位地赋能整个广告体系，例如开放数据能力、API生态、代理商体系建设等。

正如前文所言，与二元商业模式相比，三元商业模式的用户并没有带着商业需求来，而是带着时间和注意力来。因此本质上，组织是对用户的注意力进行调度，进而激发出商业需求。用户看新闻、刷微博、刷视频、看小说、看漫画、看短剧等等，都是免费的，而组织在这个过程中激发了用户的一些商业需求，并从对应的广告主处收取广告费，获得广告收入。通俗来讲，就是在用户消费内容的过程中，

让用户看到广告、点击广告，最终形成转化，交付给对应的广告主（让合适的用户在合适的时机看到合适的广告）。

不论是什么内容，也不论用户如何消费这些内容，只要用户的时间和注意力在平台上，平台就有了激发其商业需求的场景，此时只需要平台再引入广告主，而且广告主愿意为触达这些用户而付出广告费，则组织就能从激发中获得价值。同时，组织获取价值之后，会通过一些方式反哺创作者，例如广告分润、活动奖励、创作赋能等等，构建可持续的内容生态。广告是三元商业模式中组织获取价值的逻辑。在三元商业模式中，组织绝大多数情况下是从广告主侧获得价值，但也有少数情况会从内容侧获得价值，某些时候创作者会充当“广告主”的角色，典型的如创作者买粉丝量、买播放量、买观看人数等。

组织在激发用户商业需求的过程中，对用户注意力的调度效率越高，广告变现效率就越高，商业化的潜力就越大。

综上所述，三元商业模式是指：组织建设线上场域连接用户和内容，在用户消费内容的场景中激发商业需求，并获得价值的基本原理，即内容→用户→广告主。其商业模式逻辑如图1-7所示。

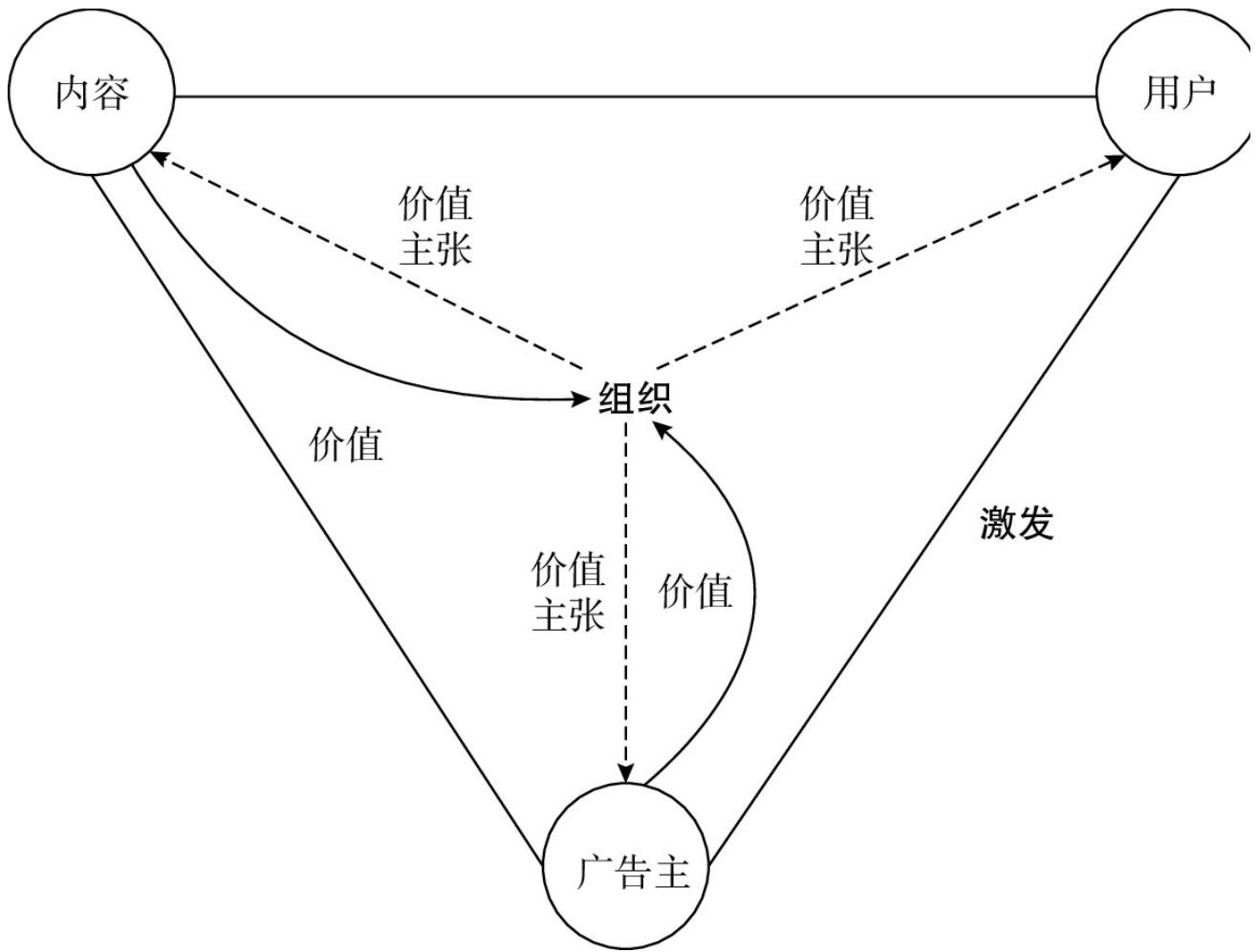


图1-7 互联网商业模式：激发（三元）

三元商业模式是一个多边平台，也具有一定的飞轮效应。具体有哪些公司和业务的商业模式是三元的，通过什么框架来分析三元商业模式，以及不同产品之间商业需求的激发效率差异等，这些内容我们会在第4章展开介绍。

至此，我们定义了互联网三种原子化的商业模式。我们再回看一下商业模式的定义：商业模式是指一个组织通过生产、传递或者激发来获得价值的基本原理。正如前文所言，公司和业务的商业模式虽然是复杂的，但都可以通过生产、传递和激发三者的组合来分析。

1.5 小结

本章强调了商业模式的重要性和复杂性，并对商业模式进行了原子化的思考，定义了互联网的三种商业模式：生产、传递和激发。

- **一元生产：**组织生产产品或服务，把产品或服务提供给目标用户，并获得价值的基本原理。
- **二元传递：**组织建设线上场域连接商家和用户，完成商家和用户之间的商业价值传递，并获得价值的基本原理。
- **三元激发：**组织建设线上场域连接用户和内容，在用户消费内容的场景中激发商业需求，并获得价值的基本原理。

接下来的第2~4章我们详细介绍这三种商业模式。

第2章 一元商业模式：生产

2.1 互联网一元商业模式研究_{2.1.1}

互联网一元商业模式的典型业务

如第1章所介绍的，我们将一元商业模式定义为生产：组织生产产品或服务，将产品或服务提供给目标用户，并获得价值的基本原理，即生产商品→用户。

很多互联网公司和业务的商业模式是生产，其中最典型的有六类业务，我们看一下这些业务的底层逻辑。

1. IAP游戏

所谓IAP，即in-App purchase，也就是App内购买，对应到游戏中，就是指付费游戏。显然，IAP游戏的商业模式是一元的：一个组织开发一款游戏，提供给游戏玩家，并从玩家处获得收入，即组织生产的商品是游戏，目标用户是游戏玩家，通过玩家在游戏中充值付费获得价值。腾讯游戏、米哈游、莉莉丝游戏、网易游戏、FunPlus等很多游戏公司或者游戏业务的核心都是IAP游戏。

2. IAP工具

顾名思义，和IAP游戏类似，IAP工具就是指付费的工具产品，一般都是面向C端用户的产品，例如摄影摄像、效率提升、健康健美、天气日历等付费的小工具产品。同样，IAP工具的商业模式也是一元的：一个组织开发一款工具产品，提供给目标用户，并从用户处获得收入，即组织生产的商品是工具App，目标用户是产品的使用者，通过用户在使用产品时付费获得价值。

3. SaaS

所谓SaaS，即software as a service，也就是软件即服务，通过网络提供软件服务。从狭义上来看，SaaS是指SaaS供应商为企业提供信息化和数据化所需要的网络基础设施及软硬件运作平台，它有别于需要本地化部署的传统ToB软件。例如OA（办公自动化）系统、CRM（客户关系管理）系统、HR系统、远程会议系统、在线协同办公软件、云服务等通用型SaaS，以及BI（商业智能）系统、工业软件、餐饮管理系统、设计协同软件、数字营销系统、金融风控系统等专业型和垂直型SaaS。大量传统企业借助SaaS完成互联网化转型，而SaaS本身，也可以理解为是

传统软件的互联网化。SaaS产品的商业模式是一元的：一个组织开发一款产品，提供给目标企业，并从企业获得收入，即组织生产的商品是SaaS产品和对应的服务，目标用户是企业，通过企业付费获得价值。Adobe、Salesforce、Oracle、Microsoft、神策数据、第四范式、Zoom、声网、AWS、阿里云、腾讯云、阿里的钉钉、字节跳动的飞书等公司或业务都是SaaS模式。

4. 流媒体付费业务

常见的流媒体内容形式有文章、音乐、有声读物（含有声书）、播客、小说、中长视频和短视频。有大量的互联网公司经营着流媒体内容相关业务，商业模式也各有不同。其中，当企业是流媒体内容的生产者或者需要通过购买版权才能提供内容服务，且通过用户付费获取收入时，其商业模式就是一元的。一个组织生产流媒体内容（直接生产或者购买版权，企业需要为内容生产支付成本），提供给目标用户（大多数情况下是C端用户），并从用户处获得收入，即组织生产的商品是流媒体内容，目标用户是内容的消费者，通过用户为内容付费获得价值。Netflix、Apple Music、QQ/网易云音乐的付费业务（会员、专辑付费等）、爱奇艺/腾讯视频的付费业务（会员、点播等）、Spotify、喜马拉雅的付费业务等，都采取了基于流媒体内容的付费模式，这些业务需要付出大量的成本来制作内容，或者购买内容的版权来获取播放权。

在这些业务中，除了个别业务是纯付费模式之外，大部分业务还同时存在着免费模式，其免费模式对应的商业模式就不是一元的了。另外，还有一些基于流媒体内容的业务，组织并不需要生产内容，而只是连接内容供给者和用户，这类业务的商业模式同样也不是一元的。

5. 租赁模式

租赁模式的代表是长租公寓和共享单车，前者是空间租赁，后者是产品租赁。同类型的还有共享办公室、共享充电宝等，本质上它们都是企业通过租赁的方式给用户提供服务，其商业模式是一元的：组织生产的是各种用来租赁的服务，目标用户是消费者（可能是B端用户，也可能是C端用户），通过用户为租赁付费获取价值。共享单车最为典型：企业定制大量单车，租赁给用户，收取租赁费。共享单车虽披着共享经济的外衣，从事的却是最简单的租赁生意。

6. 社交产品

社交产品出现在这里或许会让读者们感到惊讶，因为提到社交，大家自然就会想到腾讯的微信和QQ，它们的业务庞大且复杂，绝对不是简单的一元商业模式可以概括的。事实也是如此，不论是腾讯整个公司还是微信和QQ这两款产品，其业务

和商业模式都非常复杂。如果只看社交这一个业务点，那么主打熟人社交的微信、QQ等，主打陌生人社交的陌陌、Soul、Tinder等，以及主打职场社交的脉脉、BOSS直聘等，本质上就是企业生产了一款可以用来社交（通信、聊天、交流）的产品，提供给用户。只是有些产品并没有从用户处获取收入，也没有直接创造商业价值，例如微信的聊天功能免费提供给用户，并没有产生商业价值；而有些产品依赖从用户处获取收入，直接创造商业价值，例如陌陌、Soul等软件的陌生人聊天功能会收费，脉脉、BOSS直聘等软件的招聘功能会收费。而且职场社交产品本质上是SaaS模式，其中社交功能免费，其业务本质是面向B端用户提供付费的招聘服务，商业模式是一元的。

在社交这个功能点本身，微信和移动通信运营商没有本质上的区别。移动通信运营商给用户提供通信产品，收取电话费、短信费、流量费；而微信也给用户提供通信产品，只不过微信并没有收取通信费。即社交产品的商业模式是一元的：组织生产的商品是社交产品，目标用户是有社交需求的用户，通过用户为社交功能付费获得价值。但社交产品的强大之处在于，它可以通过社交产生的网络效应，迭代出新的业务场景，衍生出丰富的商业模式。例如在微信上生长出来的朋友圈、公众号、视频号、金融服务、小游戏等，在脉脉上生长出来的职场社区、知识付费等，在陌陌上成长起来的直播等。这些基于社交关系链生长出来的业务，会贡献新的商业价值，但它们已经不仅仅是社交属性的了。而且由于社交的网络效应极强，关系链的迁移成本极高，这些业务的壁垒都非常高，极具竞争力。

由此可见，微信如果只有聊天功能，而且聊天功能不需要付费的话，那就无法直接产生商业价值。其实这点在QQ上更为直观：只看QQ的社交功能，其聊天功能是免费的，而其获取价值依靠的是面向用户收费的QQ会员、QQ秀、QQ装扮等功能。

除了上述最典型的六类业务之外，还有互联网教育、知识付费、买手店模式（低买高卖）和外包模式（私有化部署、定制化开发、代运营、代理商）等业务，其商业模式也是一元的，这里就不赘述了。

2.1.2 互联网一元商业模式vs. 传统一元商业模式

虽然互联网一元商业模式与第1章讲的传统一元商业模式并没有本质的区别，但其依然在互联网的加持下，发生了很多变化。我们从商业模式的核心要素出发，来看一下互联网给一元商业模式带来的变化。

1. 价值主张

即一元商业模式中组织生产的商品。在互联网业务中，商品不再局限于传统的实物产品和服务，而是增加了多样的数字商品、软件商品等虚拟产品和服务，并且实现了商品高效的个性化和定制化。所谓的C2M（用户直连制造，customer to manufacturer）模式，就是构建在互联网渠道和软性供应链基础之上的、具有互联网精神的商品生产模式。

2. 渠道

互联网业务的渠道不仅包含线下渠道，还包含线上渠道。

3. 用户

互联网业务能触达的用户规模更大，有一些业务还能打破区域限制，面向全国乃至全球的用户市场。同时，借助互联网边际成本低和匹配效率高的特点，很多用户的长尾需求也能得到释放。

4. 成本

大部分互联网业务边际成本低（不绝对），具备规模经济效应。例如对于一款线上游戏而言，其边际成本会逐步降低，且具备一定的规模经济效应：一款线上游戏每增加一个玩家所带来的生产成本增加很少，随着玩家规模的增长，单个玩家所分摊的固定成本会降低，单位收入的成本会下降。因此如果一款游戏越具有竞争力，能吸引的玩家越多、生命周期越长、每个玩家的付费金额越大，那么这款游戏的成本就会越低。但这里要注意的是，游戏公司投入一款新游戏生产的成本并不低，除了有一些固定成本可以分摊之外，大部分成本都是新增的。

5. 收入

互联网一元商业模式获取价值的方式更加多元。除了用户单次付费之外，还可以通过多次付费，获取用户全生命周期的价值（LTV）；或者可以通过“免费+付费”的方式，提供部分免费服务，只获取部分用户和部分功能的价值。相较于传统一元商业模式，LTV以及“免费+付费”这两种获取价值的方式，在互联网一元商业模式中更为常见。

上述这些变化，给互联网业务带来的是效率、成本、收入模型等方面的优化，以及更大的经济效益。我们再用一张表，概述一下传统一元商业模式和互联网一元商业模式的一些差异（见表2-1）。

表2-1 一元商业模式：传统vs. 互联网

核心要素	传统	互联网
价值主张	实物产品和服务	实物产品和服务、虚拟产品和服务、个性化、定制化
渠道	线下	线上、线下
用户	规模小、区域约束强	规模大、区域约束弱（相对）、释放长尾需求
成本	边际成本高	边际成本低、规模经济效应
收入	LTV（较少）、单次付费	单次付费、LTV、“免费+付费”

2.1.3 商业模式画布

根据前文的介绍，我们可以提炼出一些关键信息，画出互联网一元商业模式的画布（见表2-2）。基于这张画布，我们在分析对应的业务时，会更加有的放矢、事半功倍。

表2-2 互联网一元商业模式的画布

价值主张 产品或者服务， 实物或者虚拟	渠道 线上渠道、线下渠道 (销售体系、直销、分销 & 代理)	用户 C端用户、B端用户、 全国甚至全球（区域约束相对弱）、 长尾需求得到释放
成本 生产成本、渠道成本、营销成本、 管理成本、经营成本、 边际成本		收入 用户付费 (单次交易收入模型、LTV收入模型、 “免费+付费”的收入模型)

综上所述，互联网给一元商业模式带来了很大变化，在互联网时代有很多业务依靠一元商业模式取得了巨大的成功，典型的比如中重度游戏（IAP模式）业务。但是，也有许多采用一元商业模式的业务发展得并不如人意，这些业务规模不大，市场占有率不高或者很难与竞争对手拉开差距，盈利困难，甚至有很多业务持续多年亏损，迟迟无法盈利。对于在分析采用一元商业模式的互联网公司时，我们要关注哪些核心要素，采用怎样的分析框架，以及业务盈利困难的核心原因是什么，我们接下来展开介绍。

2.2 商业模式经营分析框架

面对采用一元商业模式的公司，在分析其经营状况时，不论其业务包装得如何复杂、财务数据从哪些维度来披露、成本如何摊销和折旧等等，我们都只需要抓住核心点。运用下面的公式便可以洞察其经营情况，并给出经营判断：

$$\text{利润} = (\text{单用户收入} - \text{单用户成本}) \times \text{用户数}$$

公式里有三个核心指标：单用户收入、单用户成本和用户数。透过这三个核心指标，可以洞察业务的经营现状、关键问题、增长空间和核心竞争力等信息。这三个指标随着业务规模增长的变化趋势，也决定了业务的发展潜力或者瓶颈。我们分别看一下这三个指标蕴含的内容：

1. 单用户收入

单用户收入指组织从单个用户处获得的收入。单用户收入规模主要与收入模型和价值主张相关。

(1) **收入模型**。一元商业模式的收入模型主要有三大类：单次交易收入模型、LTV收入模型、“免费+付费”的收入模型。

单次交易收入模型在互联网业务中比较少见，只有比较传统的业务才会采用，例如软件激活码，一次激活终身使用。单次交易收入模型很传统，收入规模上限看用户数和定价。互联网业务比较多采用的是LTV收入模型以及“免费+付费”的收入模型。

LTV收入模型旨在获取用户全生命周期的价值，所以LT（用户生命周期）就很重要（极端情况就是单次付费，LT极短，没有复购），用户的留存和复购情况决定了业务的健康度和成长性，组织的用户运营策略和能力是决定LT的核心因素之一。

“免费+付费”的收入模型是互联网业务中最常见的，而且往往和LTV收入模型相结合，其中付费部分通常采用LTV收入模型。免费用户中也会有一部分转化为付费用户，但有一定的转化率。就算不转化为付费用户，免费用户也自有其“商业价值”，包括但不限于资本价值（估值）、数据价值和反馈价值（帮助产品迭代）。这在游戏业务中表现得更明显，游戏中有大量的免费玩家可能永远都不会付费，但是他们有很高的陪玩价值，游戏中的道具、排名、成长体系等，只有通过对比才能体现其价值，只有拥有大量的免费玩家才能支撑起付费业务。试想，

如果一个以“氪金”为核心收入来源的游戏，“氪金”和不“氪金”在游戏体验上没什么区别，那还会有玩家“氪金”吗？从某种意义上说，在游戏中，免费玩家就是付费玩家的陪玩。

(2) **价值主张**：组织提供给用户的价值主张基本决定了单用户收入的竞争力。如果组织生产的产品或者服务是差异化的、有门槛的、具有创新力的，那定价就会有竞争力，组织甚至可能拥有在市场上的定价权。

2. 单用户成本

单用户成本指组织为单个用户所付出的成本。组织生产产品或者服务的成本，以及维持组织运转和扩张的成本构成了所有成本，其中包含固定成本和可变成本。互联网业务中，占比比较高的成本为生产研发成本（人力成本、服务器成本、算力成本等）、渠道成本、营销成本和管理成本。一般而言，留存用户、复购用户的成本会低于纯新增用户。

3. 用户数

用户数指组织的用户数量。用户数主要和市场规模以及市场份额相关。

(1) **市场规模**：市场规模可大可小，与所处的赛道相关，既有面向大众市场大目标人群的生意，也有面向小众市场垂直细分人群的生意。互联网业务有一个典型的特点，即通过互联网释放长尾需求，形成一个很大的市场。

(2) **市场份额**：市场份额（market share）即市场占有率，主要看竞争，也看赛道。不同赛道的CR（行业集中度，concentration ratio，是指某行业所处的市场中，前N家最大的企业所占市场份额的总和）差异很大。我们判断行业发展的结局，需要看目标企业所处的赛道是不是一个赢者通吃的市场。在一个CR 5高达90%的市场中，排到第六名的企业注定不会活得很舒服。组织依赖产品、成本控制以及渠道的综合能力来支撑其竞争力，提升市场份额。

一元商业模式的业务，其核心竞争力一般来自以下三点：价值主张的独特性和门槛，企业的产品或者服务竞争对手模仿不了，或者进入门槛很高；生产的边际成本低，规模效应强，企业占领了一定的市场份额，构建了成本优势，反过来也拉高了竞争门槛；具有渠道优势，销售体系竞争力强，这往往是这类企业发展期最核心的竞争力。

接下来我们选择五家具有代表性的公司——爱奇艺、声网、神策数据、第四范式和Salesforce来详细介绍一下它们的业务情况。

2.3 典型公司及其业务研究_{2.3.1}

爱奇艺：长视频赛道的商业模式深度解读，努力走出亏损困境

爱奇艺，中国在线视频行业的领军企业之一，专注于提供在线娱乐服务。自2010年成立以来，爱奇艺凭借其丰富的内容库和先进的技术平台，迅速崛起成为亿万用户喜爱的视频播放和内容制作平台。爱奇艺的业务涵盖长视频内容的制作、聚合、分发，为用户提供包括电影、电视剧、综艺、动漫等在内的全方位视频服务。2018年爱奇艺在美国纳斯达克挂牌上市。

中国在线视频行业经历了从版权争夺到内容自制的转变，爱奇艺、腾讯视频、优酷三大平台的竞争日趋激烈。用户对于高质量内容的需求不断增长，推动平台不断加大内容投入。

1. 爱奇艺的财务数据（来自爱奇艺财报公开数据）

看收入结构，爱奇艺的收入主要由会员服务、在线广告服务、内容分发以及其他构成，其中会员服务和在线广告服务是其主要的营收来源，会员服务收入连续三年占比超过60%，在线广告服务收入连续三年占比在19%左右（见表2-3）。

表2-3 2022—2024年爱奇艺收入结构

营收构成	2022 年		2023 年		2024 年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
会员服务	177.11	61.08	203.14	63.73	177.63	60.78
在线广告服务	53.32	18.39	62.24	19.53	57.14	19.55
内容分发	25.62	8.84	24.59	7.71	28.47	9.74
其他	33.93	11.70	28.76	9.02	29.01	9.93
总计	289.98	100.00	318.73	100.00	292.25	100.00

看营收数据，爱奇艺2023年总营收318.73亿元，2024年下降到292.25亿元。看利润数据，在连续亏损数年之后，爱奇艺于2023年利润转正，其2023年持续经营利润为19.53亿元；2024年虽然营收年同比增长率下降，但还是保持了盈利（见表2-4）。

表2-4 2019—2024年爱奇艺营收和利润

项目	2019年	2020年	2021年	2022年	2023年	2024年
总营收（亿元）	289.94	297.07	305.54	289.98	318.73	292.25
营收年同比增长率（%）	—	2.46	2.85	-5.09	9.91	-8.31
持续经营利润（亿元）	-102.77	-70.07	-61.09	-1.18	19.53	7.91

爱奇艺作为头部在线视频平台，季度付费订阅会员人数很早就超过了1亿，那为何在早期发展阶段连续亏损、盈利困难呢？因为爱奇艺的成本挑战非常大。

爱奇艺的成本主要分为三部分：内容成本、销售和管理费用以及研发费用（见表2-5）。其中，内容成本是绝对的大头，内容成本占营收的比例非常夸张：在亏损超百亿的2019年为76.57%，在扭亏为盈的2023年和2024年分别降低到50.83%和53.72%。同时，在线视频行业竞争激烈，爱奇艺的销售和管理费用居高不下。

表2-5 2019—2024年爱奇艺成本构成情况

项目		2019年	2020年	2021年	2022年	2023年	2024年
总营收（亿元）		289.94	297.07	305.54	289.98	318.73	292.25
成本分类 （亿元）	内容成本	222	209	206	165	162	157
	销售和管理费用	52.36	51.88	47.25	34.67	40.14	36.82
	研发费用	26.67	26.75	27.95	18.99	17.67	17.78
成本占营 收比例 （%）	内容成本	76.57	70.35	67.42	56.90	50.83	53.72
	销售和管理费用	18.06	17.46	15.46	11.96	12.59	12.60
	研发费用	9.20	9.00	9.15	6.55	5.54	6.08

爱奇艺在2023年和2024年能扭亏为盈，得益于在成本上的优化，包括大幅降低内容成本，以及在销售和管理费用和研发费用上的降本增效。

2. 爱奇艺的商业模式

爱奇艺的商业模式是一元和三元商业模式的组合。一元商业模式主要体现在会员服务上，通过提供优质内容吸引用户付费订阅。三元商业模式则体现在在线广告收入上，包括横幅广告、贴片广告、暂停广告以及信息流广告等。

从爱奇艺的营收和成本构成情况可以看到，主要营收和成本都发生在会员业务中。会员模式中，用户按照不同的时长订阅会员服务，比较多的是月度会员、季度会员和年度会员，到期后需要续费才能继续享受服务。这个模式有一个很明显的问题，即对于一个付费会员，平台赚取的收入在会员周期内是固定的，用户的活跃度提升并不能带来增量收入。相反，对于广告模式，用户活跃度提升，在App内有更多的内容消费时长，会贡献更多的广告收入。会员模式导致的结果就是，如果平台想提高单用户收入，就只能提高会员价格。

会员定价上，2023年第四季度ARM（月度平均单会员收入，average revenue per member）为15.98元，而2019年ARM仅为11~12元，可见爱奇艺也在通过涨价提升会员服务的收入。

3. 长视频平台会员商业模式解读

事实上，国内的几家长视频平台，在早期发展过程中都持续面临内容版权大战和盈利困境，核心原因在于其商业模式过度依赖会员付费。这种一元商业模式不仅制约单用户收入的增长，而且随着用户规模增长，也会让单用户成本持续上升，从而导致平台随着规模扩大，逐步陷入亏损状态。接下来我们主要从单用户成本和单用户收入两个方面去分析会员模式的一些利弊。

（1）**单用户成本**：内容成本和营销成本的增加会导致单用户成本的上升。有几个因素会让长视频平台会员模式下的单用户成本越来越高，而不是越来越低。

单用户成本=内容成本/内容带来新增付费会员数+单用户营销成本

● **内容成本**：视频平台生产的视频内容，有些是平台直接投资拍摄的，例如自制剧、平台参与投资的电影电视剧；有些是平台通过购买版权或播放权等形式获得的。这里就有一个明显的问题：**这些内容的生产或采购成本并不能随着平台规模扩大而下降**。另外，平台之间的抢用户大战，会造成平台在内容采购时互相竞价，推高IP剧集的成本，而且这个问题在平台的付费用户到达一定增长瓶颈时会更严重。

● **内容带来新增付费会员数**：一个新的内容（比如一个大火的IP剧集）如果出现在平台付费用户规模较大的时期，那么能带来的付费会员数增量只会比早期更少。也就是说，单个付费会员对应的内容成本会越来越高。这里有一个关键原因：平台会员是按时间窗口售卖的，只要在会员有效期内，有效会员并不用为新引入的内容额外付费，所以实际上新内容带来的额外的付费会员会越来越少。但平台又需要不停地引入新的内容来提升会员的续费率。

● **单用户营销成本**：视频平台很大一部分新用户是依靠广告来获取的，这个方式也会逐步推高获客成本。付费会员基数越大，平台去外面获取额外的付费会员就会越困难，于是单用户获客成本会越来越高。同样，各大竞争对手之间的抢用户大战也会推高获客成本。

(2) **单用户收入**：当各大视频平台开展抢用户大战的时候，大量的会员促销会降低单用户收入。正常情况下，可以认为单用户收入这个值是比较稳定的，但一旦陷入价格战，单用户收入必然下降，全行业利润必然受损。如果各家能良性竞争，在定价博弈上形成局部的合作，则单用户收入能保持稳定，甚至能实现小幅增长来提升平台收入，例如前文提到的爱奇艺会员ARM，2023年第四季度为15.98元，相比2019年一路在上涨，但经过多轮涨价后也就趋于稳定了。

这种一元商业模式的利弊就很清晰了。这种模式的好处在于会员这个产品模式，易于用户理解，用户接受成本相对较低。弊端就在于单用户收入相对稳定，而单用户成本却不断提高，最终随着规模增长，平台很容易到达单用户亏损点（如图2-1中单用户营收线和单用户成本线的交叉点），但是为了市场竞争而不得不亏损去发展用户规模。这个亏损的根本原因在于视频平台采取了会员这个模式作为主要营收来源。因此，长视频这个赛道上的各个平台很长一段时间都处于亏损状态。

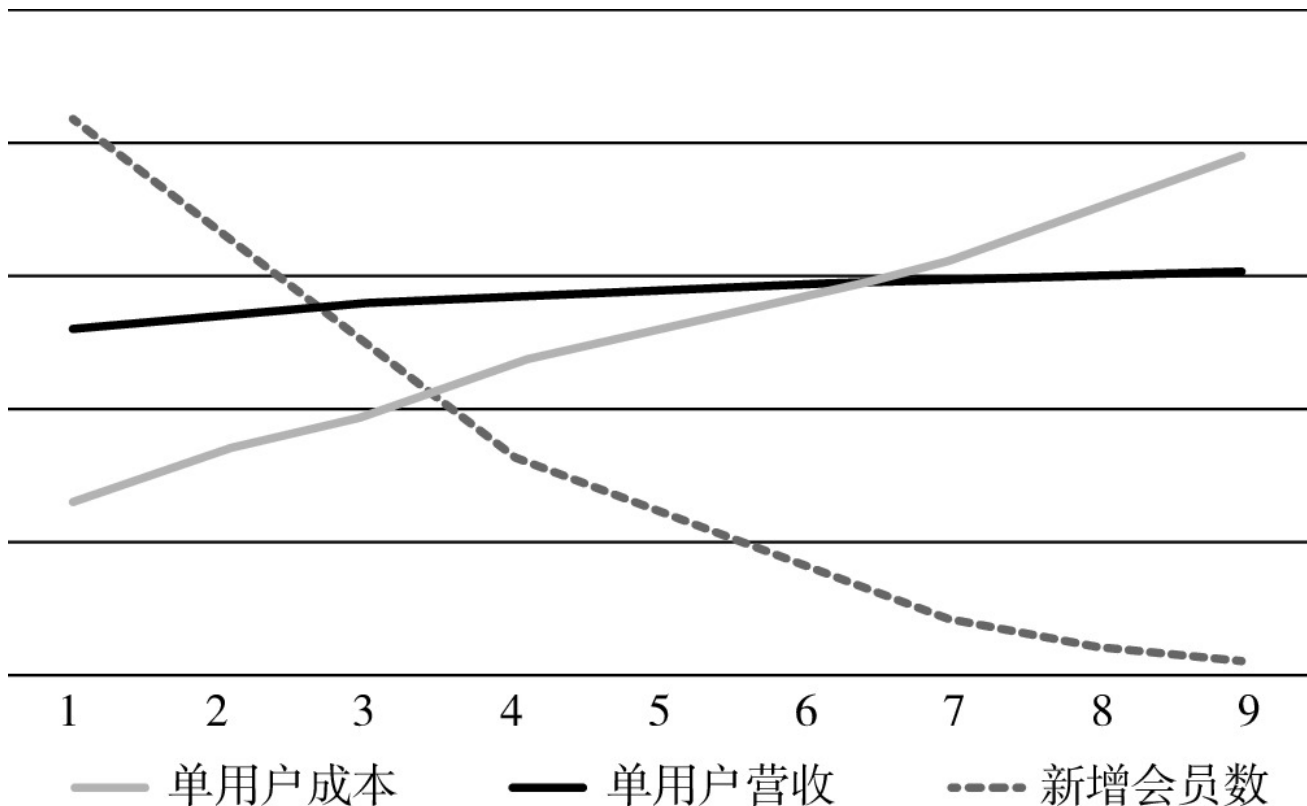


图2-1 长视频平台商业模式的关键要素

但平台需要可持续发展，长期亏损肯定不行，企业需要获取必要的利润来维持经营。我们可以看到，视频平台采取了一些改进措施来缓解亏损问题。

● **提高单用户收入：**其方式有很多，比如最直接的涨价、细分会员（如手机、平板、电视、投屏等各种权益会员，体育、儿童会员，等等）、提前点播收费等。这些方式虽然备受用户诟病，但却很有效。只要内容质量有竞争力，依然会有很多用户愿意付费。

● **增加广告模式：**一般包括前贴片广告、中插广告、后贴片广告、暂停广告、信息流广告等。为了增加用户活跃度，很多视频平台还做了短视频流，然后中间插一些广告。这些广告变现也可以增加平台盈利。广告模式与会员模式有一个很关键的差异：对于会员模式，通过用户运营提升用户的活跃度，意义没有那么大，因为付费用户来的次数多少并不会改变其对平台营收的贡献；但是用户来的次数越多，消费内容越多，看到的广告就会越多，平台通过广告模式从这个用户身上赚的钱就越多。因此，平台增加用户的活跃度，对于拉升广告营收很有帮助，从这也可以看出会员模式的一个劣势。

● **降低单用户成本：**对内容成本的控制尤其关键，例如视频平台纷纷投入制作自制剧，不再单靠大IP来获客。因为用的演员相对都不是大牌明星，自制剧的成本会明显低一些。

整体来看，爱奇艺在会员服务和在线广告服务等业务上取得了一定增长，但单用户成本上升的问题依然严峻。通过提高单用户收入、增加广告模式和降低单用户成本等策略，爱奇艺正努力改善盈利状况，并寻求更可持续的发展路径。

2.3.2 声网：实时音视频API平台行业开创者，盈利挑战巨大

声网（Agora）成立于2014年，是全球实时互动云服务开创者和引领者。公司专注于通过RTC（实时音视频，real-time communication）技术，为开发者提供即插即用的技术解决方案。开发者只需简单调用声网API，即可在应用内构建多种实时音视频互动场景。声网母公司Agora Inc.于2020年成功登陆美国纳斯达克，成为全球实时互动云第一股。截至2024年12月31日，声网全球注册应用数超过93.1万，年同比增长25%。

声网的发展可以分为两大主要阶段。

第一阶段是成长期（2014年4月—2018年9月），在这一阶段，公司初步搭建了实时音视频产品服务体系，并发布了实时音视频产品，举办了首届实时互联网大会（WebRTC大会）。随着业务的不断扩展，声网逐渐覆盖社交直播、游戏电竞、教育等多个行业。

第二阶段是发展期（2018年10月至今），在这一阶段，声网进一步深耕实时通信服务业务，同时重新定义RTE（实时互动服务，real-time engagement），致力于成为实时互动云服务商（RTE PaaS）。这一阶段，声网还推出了全球首个、迄今为止规模最大的实时音视频网络——软件定义实时网SD-RTN™；推出了全球首个、迄今为止数据处理能力最强、分析维度最广的实时音视频质量诊断分析产品——水晶球（Agora Analytics）。

声网的产品覆盖场景比较丰富，比较典型的有：实时音视频（支持一对一音视频通话、一对多及多对多连麦通话的直播间、语聊房、视频会议，以及各类直播等RTE场景）、低延时直播（适用于声网极速直播产品，可应用于各类需要低延时直播的场景）、实时录制、游戏语音、AI高级美颜等。

1. 声网的财务数据（来自声网财报公开数据）

看收入结构，声网的收入主要来自RTE，以2023年和2024年的数据为例，这两年这部分收入占比分别为94.04%和95.77%（见表2-6）。

表2-6 2022—2024年声网收入结构

营收构成	2022 年		2023 年		2024 年	
	营收 (美元)	占比 (%)	营收 (美元)	占比 (%)	营收 (美元)	占比 (%)
实时互动服务	1.53 亿	95.16	1.33 亿	94.04	1.28 亿	95.77
其他	778.40 万	4.84	844.00 万	5.96	563.00 万	4.23
总计	1.61 亿	100.00	1.42 亿	100.00	1.33 亿	100.00

注：表格中营收及占比数据来自声网财报，原数据如此。

看营收数据，声网2024年总营收1.33亿美元，但已经连续三年下滑；看利润数据，声网处于连续亏损状态，2024年持续经营利润亏损0.43亿美元（截至2024年，声网已经连续6年亏损了，见表2-7）。

表2-7 2019—2024年声网营收和利润 单位：亿美元

项目	2019 年	2020 年	2021 年	2022 年	2023 年	2024 年
总营收	0.64	1.34	1.68	1.61	1.42	1.33
持续经营利润	-0.06	-0.03	-0.72	-1.20	-0.87	-0.43

声网为什么会连续亏损呢？主因是其居高不下的研发费用和销售管理费用。作为一家技术属性的企业，为筑牢企业技术的“护城河”，确保市场领先的位置，就需要维持高研发投入；而为了拓展新客户以及维持老客户续费，销售和管理上的投入也很高。我们以2022—2024年声网的主要成本数据为例，可以看到这两项成本占营收的比例都非常高（见表2-8）。

表2-8 2022—2024年声网成本构成情况

项目	2022 年		2023 年		2024 年	
	数值 (亿美元)	占营收比例 (%)	数值 (亿美元)	占营收比例 (%)	数值 (亿美元)	占营收比例 (%)
总营收	1.61	—	1.42	—	1.33	—
研发费用	1.15	71.43	0.78	54.69	0.80	60.41
销售和管理费用	0.92	57.42	0.69	48.55	0.60	45.11

2. 声网的商业模式

从声网的业务介绍和财务数据中，我们可以很清楚地看到声网是采用一元商业模式的公司。声网生产实时互动的产品和服务，提供给有需求的开发者（主要是B端用户），并直接收取费用。以实时音视频产品为例，声网给出了免费版套餐以及不同档位的付费套餐（价格来自声网官方网站2024年第三季度信息），采用的是“免费+LTV付费”的组合收费模式（见表2-9）。

表2-9 声网的收费模式

套餐	免费版	基础版	尊享版	旗舰版	企业版
额度（万分钟/月）	1	12	40	150	150+
定价（元/月）	免费	738	2 388	8 399	定制报价

3. 声网的商业模式解读

声网是很典型的SaaS公司，采用一元商业模式，但长期处于亏损的状态，单用户收入覆盖不了单用户成本：单用户收入相对稳定，但成本居高不下，单用户成本无法下降。

单用户收入方面，如果声网要提高收入，就需要涨价，提升单次API调用的价格，即前述提到的每月的定价。但声网面临着国内外大量头部公司的竞争，包括腾讯云、阿里云、TokBox、Twilio、Zoom等，在如此激烈的市场竞争中，涨价势必会引起客户流失，不但涨价涨不动，甚至还可能面临被动的价格战。

单用户成本方面，声网需要进一步压缩成本，提升盈利能力。研发成本上，合理投入，在保持产品竞争力的同时提升效率，同时减少定制开发。营销成本上，优化渠道结构，提升渠道竞争力，降低新客获取成本，同时提升老客续费率，关注续费率指标，一般而言老客续费成本是低于新客获取成本的。非定制客户数持续增长也可以发挥规模效应，降低单用户成本。

从表2-10来看，过去几个季度声网的活跃客户数基本在3 600个左右，增长相对缓慢，结合前面提到的营收数据也基本持平甚至下跌，可见声网在客户增长和客单价提高方面都面临着很大的瓶颈，进一步限制了声网的盈利能力。受“双减”政策的影响，大量教育培训行业客户流失对声网也产生了比较大的负面影响。

表2-10 声网2023年第四季度到2024年第四季度活跃客户数

项目	2023Q4	2024Q1	2024Q2	2024Q3	2024Q4
声网活跃客户数	1 835	1 896	1 970	1 969	1 979
Agora 活跃客户数	1 683	1 720	1 672	1 762	1 723
汇总	3 518	3 616	3 642	3 731	3 702

注：业务聚焦中国市场的统计到声网，聚焦非中国市场的统计到Agora；活跃客户是指在过去12个月内，带来超过100美元收入的组织或个人开发者。

整体来看，市场竞争、成本控制、渠道竞争力与行业政策等黑天鹅因素综合约束着声网的经营，其面临的盈利挑战非常严峻。如果要快速提升盈利，一方面要提升研发效能，并快速提升客户规模，从而降低单用户成本；另一方面，也可以尝试通过技术产品创新，用新的产品价值提高产品定价，以此来拉升单用户收入。

2.3.3 神策数据：数字化浪潮中的探索者，数据服务赛道独角兽

神策数据（Sensors Data）成立于2015年4月，是国内专业的数字化客户经营软件提供商、国家高新技术企业，拥有高度开放的产品架构、灵活的被集成能力、可落地的数字化客户经营解决方案，以及完备的数据安全和合规体系，可以帮助企业实现数字化客户经营。

神策数据的发展可以分为两大主要阶段。

第一阶段是2015—2019年，在这一阶段，神策数据专注于单一分析产品神策分析（Sensors Analytics）。神策分析是一个可私有化部署的用户行为分析平台，该产品提供私有化部署和SaaS两种部署方式，可以实现基础数据采集与建模，并可作为PaaS（平台即服务）平台支持二次开发。基于用户的业务特点和多部门复合需求，神策分析助力用户建立高效的数据指标体系，用先进的事件模型抽象用户行为，提供多维度、多指标的交叉分析功能，全面支撑各个团队的日常数据分析需求，驱动业务决策。

第二阶段是2020年至今，在这一阶段，神策数据升级“产品+服务”的能力，围绕“客户旅程编排”的产品理念，构建了客户数据引擎、客户旅程分析引擎、客户旅程优化引擎三大核心引擎，赋能企业的数字化客户经营。

客户数据引擎，即客户数据平台（CDP），对应神策数界平台（Sensors Data Horizon）。除了具备全域、实时、灵活圈选三大必备能力之外，神策数界平台还拥有突出的高性能查询能力，通过整合多源数据、关联全域ID、扩展多实体的数据模型构建客户分群和标签，同时结合丰富的数据加工方式和数据输出能力，为业务分析洞察、自动化营销等场景的全域客户经营提供数据基石。

客户旅程分析引擎，即神策分析。除了具备全域、全链路、实时和灵活四大核心能力之外，客户旅程分析引擎还面向多角色提供可视化的数据分析功能，支撑企业从客户到经营，从“人”到“场”的全视角旅程分析与决策。

客户旅程优化引擎，即神策智能运营（Sensors Focus）。它作为客户旅程编排的“发动机”，囊括受众服务、用户旅程服务、内容决策服务、触达通道服务四大组件服务功能，支撑营销过程中的营销受众圈选、触发时机选择、营销内容决策、触达渠道对接，助力企业更高效地经营客户。

自成立以来，神策数据持续获得资本加持，截至2025年第一季度，共完成六轮融资，总融资金额高达2.95亿美元（见表2-11），在2022年5月的D轮融资时，其估值

达16亿美元。

表2-11 神策数据历次融资情况

融资时间	融资轮次	融资金额	投资机构
2015年	天使轮	600万元	线性资本、明势资本等
2016年	A轮	400万美元	红杉中国领投，线性资本、明势资本等跟投
2017年	B轮	1100万美元	DCM资本领投，红杉中国跟投
2018年	C轮	4400万美元	华平投资领投，红杉中国、DCM资本、襄禾资本、晨兴资本（现五源资本）、线性资本、明势资本跟投
2020年	C+轮	3000万美元	襄禾资本领投，晨兴资本（现五源资本）、DCM资本、华平投资、红杉中国、线性资本、明势资本跟投
2022年	D轮	2亿美元	Tiger Global、凯雷投资集团领投，明势资本、DCM资本、线性资本、红杉中国、华平投资、Bessemer Ventures、M31资本、襄禾资本、五源资本、纪源资本跟投

神策数据暂未上市，所以目前没有财报数据。行业内公开数据显示，神策数据的业务覆盖以泛金融、泛品牌零售、泛互联网、泛企业级客户等为代表的30多个主要行业，目前其付费客户超过2000个，年营收额超过1亿元。

1. 神策数据的商业模式

神策数据是采用一元商业模式的公司。从前述的业务介绍中可以看到，神策数据主要是面向B端客户，给客户id提供数字化相关产品和服务，是典型的SaaS模式，同时支持私有化部署，为不同规模 and 需求的客户提供定制解决方案。

在收费模式上，神策数据采用按年订阅制收费，根据客户的使用方案和接入的数据量来定价。目前神策数据产品包括标准SaaS版、私有化部署单机版以及私有化部署集群版，不同版本提供差异化的报价、产品和服务。

2. 数据类SaaS公司的商业模式解读

SaaS公司在早期定位目标客户时，要关注它们的生存周期，生存概率太小的赛道上的公司，即使成为你的客户，也只是让你付出获客成本和交付成本，很难给你带来收益。

神策数据主要是将产品定位于为企业提供数据的采集、加工和分析服务，后续发展出基于大数据的智能化决策等产品。从目标客户上来看，其最初的核心客户是一些互联网初创公司，这与这些公司对大数据的重视有关系。这些公司的管理层对高效数据分析、AB测试等基于数据的“问题发现-分析-实验-决策”工作方法论比较认可，必然对这样的数据产品服务有需求。

我们同样从单用户收入和单用户成本两个大的角度去分析一下神策数据这类数据类SaaS公司的商业模式。

（1）**单用户收入**。如前文所述，这类大数据决策分析产品的目标客户主要为互联网初创公司。从收益角度来讲，互联网初创公司有个缺点，即它们对SaaS产品的购买时间（LT）比较短，或者换个说法就是SaaS产品续费率不高。这里有两个原因：第一，这类初创公司发展经营上面临的不确定性较高，很多初创公司“死”得很快，自然续不了费。第二，这类初创公司中有极少的部分能搏杀成功、发展壮大，然而在业务规模和复杂性提高以后，很多公司会倾向于自建大数据产品。因为数据隐私和安全对于大一些的公司来说是非常关键的，可以说是生死线之一，完全交给外面的公司风险还是比较大的。而且业务复杂以后，公司对大数据产品的需求也会更多更复杂，与外部SaaS产品提供商沟通对接，效率太低了，成本也会越来越高。

但是对于一个SaaS公司来说，其产品的受众都是B端客户，很重要的特征应该是LT足够长（或者续费率足够高），这样才能持续地贡献营收，从而去摊薄获客成本和前期交付成本。如果平均LT只有1~2年，那大概率是很难获利的。作者曾经在神策数据的一个客户公司工作过一段时间，这家公司用了两年神策数据的服务，第三年就内部评估决定不用了，因为内部技术团队做了一套替代方案（这是内部技术团队必然会产生观点，团队从内部价值角度也会倾向于自己来做）。作者也曾经和神策数据的管理层交流过这个问题，他们承认其业务发展过程中确实存在这样的挑战，所以神策数据很快就开始瞄准一些传统领域的客户，比如银行。

（2）**单用户成本**。对于SaaS产品来说，理论上单用户成本应该越来越低，这也是SaaS模式成立的关键要素。然而，B端客户的个性化需求常常异常复杂，很多SaaS产品提供商并不能做到很好地管理客户，大部分SaaS公司做着做着就变成了IT外包服务商，一个客户一个版本，高度定制化，单用户成本很难降低，而且定制化的项目，费用结算也是一个难题。SaaS公司一旦走上定制化这条路，成本就越来越高，盈利就比较困难。

从业务需求的标准化程度来看，互联网公司的数据分析需求是比较复杂的，标准化程度其实没有那么高。这就意味着SaaS公司在提供产品的时候，需要投入比较多的资源去满足个性化的需求，这也会极大拉高单用户成本。而公司面对转变方向后的目标客户，如银行这类传统客户时，获客模式发生了变化。不仅销售成本比较高，还非常依赖销售资源对应的客户关系。这类传统客户的拓展具有一定的地域性，这会极大地限制客户规模。

整体来看，一方面，数据类SaaS公司在客户的获取方面需要克制对个性化定制的投入，避免付出较高的单用户成本。另一方面，对数据服务需求较高的客户（如互联网初创公司）的生命周期又较短，容易导致单用户收入不高。数据类SaaS公司需要积极拓展客户行业和类型，增加对传统公司的渗透，提升客户LTV。通过这两方面的发展策略，才能有效提升公司的营收规模和利润水平。

2.3.4 第四范式：人工智能技术与服务提供商，盈利仍需努力

第四范式（4Paradigm）成立于2014年9月，是企业级人工智能领域的行业先驱者与领导者。第四范式专注于提供以平台为中心的人工智能软件和解决方案，并运用核心技术开发了端到端的企业级人工智能产品，致力于解决企业智能化转型过程中面临的效率、成本、价值问题，提升企业的决策水平。第四范式的主要产品是先知平台（先知AIOS），这是一个端到端人工智能解决方案操作系统，可以帮助企业规模化设计、开发及操作人工智能应用，主要服务的客户涉及金融、零售、制造、能源电力、电信、运输、科技、教育、媒体及医疗等行业。

图灵奖得主、关系数据库的鼻祖James Gray在2005年以“科学发展的四个范式”为题做了一个演讲，把人类科学发展分成四个阶段，每个阶段是一个范式。第一个范式是实验科学，比如钻木取火；第二个范式是理论科学，根据观察记录的现象，总结出几个规律或者定律；第三个范式是计算科学，人类基于强大的计算力做理论的推演；而第四个范式便是数据科学，机器替代人在海量数据中发现规律，可以通俗地理解为机器学习。

第四范式的命名便来源于此，公司创立的初衷是期望通过机器学习和人工智能等技术手段，帮助各行各业发现行业发展规律，提高行业的效率。

中国人工智能行业按照应用范畴可以分为四大领域：决策类人工智能、视觉人工智能、语音及语义人工智能以及人工智能机器人。决策类人工智能可以识别数据中的隐藏规律，知道基于数据洞察的决策过程，并解决与核心业务运营密切相关的问题，而第四范式在中国以平台为中心的决策类人工智能市场中占据最大的市场份额。

第四范式创始团队具有超强的融资能力，自公司成立到上市之前，共完成11轮融资，累计融资额约70亿元，并创造了一个不小的奇迹——成为五大国有银行（中国工商银行、中国银行、中国建设银行、中国农业银行、交通银行）联合投资的唯一科技类公司。2023年9月，第四范式在港交所主板正式挂牌上市。

根据2023年9月的招股书介绍，第四范式的业务主要分为先知平台及产品、应用开发及其他服务两大板块，前者是第四范式主推的AI产品，后者是面向企业的AI解决方案定制服务。

先知平台可以被不同行业的终端客户应用，不需要大量定制，但依然支持私有化部署。先知平台及产品主要通过以下方式提供：在终端用户服务器上本地部署软件使用许可，或在终端用户服务器及其他相关硬件上预装先知软件的“一体化”解决方案SageOne。两者都支持用户利用先知平台开发他们自身的人工智能应用。软件使

用许可的定价主要基于算力消耗的情况，而SageOne的定价基于硬件单元数量、软件的使用许可费及各自型号的硬件成本。

按照客户的要求，第四范式也提供应用开发服务，帮助客户根据业务需求利用先知平台开发定制化人工智能应用。在定制开发上，第四范式按照单项目收费，项目定价主要基于提供相关服务的人力消耗成本。

1. 第四范式的财务数据（来自第四范式财报公开数据）

如前文所述，第四范式的业务分为先知平台及产品、应用开发及其他服务两部分。看这两部分的收入结构，以2023年和2024年的数据为例，这两年先知平台及产品的收入占比分别为59.61%和69.87%（见表2-12）。

表2-12 2022—2024年第四范式收入结构

营收构成	2022 年		2023 年		2024 年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
先知平台及产品	15.91	51.61	25.06	59.61	36.76	69.87
应用开发及其他服务	14.92	48.39	16.98	40.39	15.85	30.13
总计	30.83	100.00	42.04	100.00	52.61	100.00

看营收数据，第四范式2024年总营收52.61亿元，营收连续多年保持增长；看利润数据，第四范式处于连续亏损状态，2024年持续经营利润（税后）亏损2.96亿元，亏损年同比有所收窄（见表2-13）。

表2-13 2019—2024年第四范式营收和利润 单位：亿元

项目	2019 年	2020 年	2021 年	2022 年	2023 年	2024 年
总营收	4.60	9.42	20.18	30.83	42.04	52.61
持续经营利润（税后）	-7.18	-7.50	-18.02	-16.53	-9.21	-2.96

从成本构成情况来看，第四范式的销售成本和研发成本都非常高，而且销售成本最近几年甚至超过了研发成本。以2024年为例，2024年第四范式的销售成本高达32.85亿元，占营收的比例约为62%，研发成本高达21.70亿元，占总营收的比例约为41%（见表2-14）。第四范式有非常多的大客户，公司将《财富》世界500强企业

或上市公司定义为“标杆用户”（例如金融行业典型的银行客户），“标杆用户”的单用户收入一般会达到1 000万元以上，但相应地针对这类客户的销售成本也很高。研发费用高也比较好理解。第四范式为了保持技术上的领先，需要持续投入研发资源。同时公司还支持私有化部署、定制产品和服务，这些都需要大量的技术人员支撑。根据2023年9月的招股书数据，第四范式公司雇员中研究及开发人员占比超过70%。

表2-14 2019—2024年第四范式成本构成情况 单位：亿元

项目	2019年	2020年	2021年	2022年	2023年	2024年
销售成本	3.95	7.61	15.20	20.08	26.48	32.85
行政成本	2.24	2.46	5.42	5.28	3.42	1.93
研发成本	4.16	5.66	12.49	16.50	17.69	21.70

2. 第四范式的商业模式解读

第四范式本质上是一个企业服务公司，采用的是一元商业模式。第四范式为企业提供SaaS产品以及定制化产品和服务，并从企业客户那里获取价值。如前文所述，第四范式的业务主要分为两部分：先知平台及产品、应用开发及其他服务。其中先知平台及产品是标准化的产品，通过“标准化产品+私有化部署”为客户提供企业级人工智能解决方案；应用开发及其他服务则是非标准化的产品和服务，根据客户需求进行定制化开发和服务（见表2-15）。

表2-15 两种业务模式对比

业务	产品交付方式	提供的服务	收费方式
先知平台及产品	标准化产品+私有化部署 1. 在终端服务器部署软件使用许可 2. 预装“一体化”解决方案 SageOne	SaaS 产品+服务	1. 软件使用许可根据部署的AI应用消耗算力定价 2. SageOne 综合考量服务器、软件使用许可、运营和维护服务费等定价
应用开发及其他服务	定制化开发和服务 帮助客户在先知平台上开发定制化人工智能应用	定制化 产品+服务	按项目的人力消耗成本定价

3. 技术服务类SaaS公司商业模式解读

技术服务类SaaS公司在发展过程中，需要重点关注标准化和定制化的限度。很多行业标准化程度低，客户业务场景可复制性差，如果只做标准化产品会面临获客困难的问题；但定制化开发成本高，没有规模效应，公司做一单算一单，单用户研发成本居高不下。

很多技术服务类SaaS公司都面临着持续多年亏损的困境。虽然很多公司开始提供产品的定位是SaaS模式，但是其销售成本和研发成本都很高，单用户收入无法覆盖单用户成本。而定制化应用开发业务，本质上就是外包项目，按照单项目人力消耗报价的模式，就算单项目不亏损（实际很多项目是亏损的，主因是技术人员成本太高，但项目溢价能力不足），这部分收入也无法规模化，其边际成本没有下降，项目维护成本高且需求不稳定。

另外，技术服务类SaaS公司在经营上还面临一个不容忽视的挑战：回款难带来的现金流压力大。以某公司2024年上半年的财报数据为例，其应收款项为19亿元〔去掉信用减值损失（坏账准备）1.3亿元，还剩17.7亿元〕，这大概能占到公司流动资产的30%。同时，其应收账款周转天数不断拉长，已经从2022年的130天左右拉长到2024年的170天以上。从账期来看，财报中显示一般给客户90天的账期，但从实际账龄看，已经有接近64%的款项是3个月及以上的账龄了（见表2-16），回款压力可想而知。这种回款慢的问题，也给技术服务类SaaS公司带来了很大的经营现金流压力以及一定的坏账风险。

表2-16 各账龄的金额及占比

账龄	金额（亿元）	占比（%）
3个月内	6.91	36.35
3~6个月	4.83	25.41
6个月~1年	4.12	21.67
1年以上	3.15	16.57
总计	19.01	100.00

“现在几乎没有企业不用AI的技术，但是如果说哪个企业把AI拿掉后就活不下去了，那这样的企业也很少。这是我们认为AI面临的最大的挑战，就是虽然有用，但是没有特别大的用处。”这是2021年6月第四范式创始人兼CEO戴文渊在公司年度峰会上的发言。凭借着公司创始团队超强的融资能力和大客户销售管理能力，第四范

式的营收还在保持增长。从改善经营的角度看，公司需要提高研发效率，降低个性化定制研发的成本投入。同时，在这一轮大模型技术变革浪潮中，公司如果能抓住机会进行产品技术创新，提升产品力，从而吸引更多的客户，那么一定能进入更快速的发展阶段。

2.3.5 Salesforce: SaaS开创者，产品创新与服务优化的平衡艺术

Salesforce，成立于1999年3月，是一家总部位于美国旧金山的CRM（客户关系管理，customer relationship management）软件服务提供商，被誉为SaaS和PaaS模式的先驱。公司创始人马克·贝尼奥夫（Marc Benioff）当时判断随着互联网的发展和宽带的普及，会有越来越多的企业通过互联网得到各种软件服务，于是他离开Oracle（他是Oracle历史上最年轻的高级副总裁）并自主创业，开始对SaaS模式进行探索。Salesforce以“软件的终结”为口号，致力于通过云端技术为企业提供灵活、高效的客户关系管理平台，帮助企业实现数字化转型和增长。同时，Salesforce允许客户和独立软件供应商定制并整合其产品，生成他们各自所需的应用软件。对于用户而言，这可以避免购买硬件、开发软件等前期投资以及复杂的后台管理问题。Salesforce的成立标志着企业软件服务向云端迁移的重要趋势，其发展历程充分展示了技术创新与市场需求的紧密结合。

整体上Salesforce的发展可以分为四个关键阶段。

第一阶段是初创期（1999—2004年）。在这一阶段，Salesforce专注于SaaS模式的CRM订阅服务，通过免安装部署、持续更新、按需付费等创新模式，迅速吸引了大量中小企业用户。这一阶段的成功奠定了Salesforce在SaaS市场的领先地位。

第二阶段是平台化扩张期（2005—2009年）。随着市场需求的扩大，Salesforce开始布局PaaS平台，推出了世界上第一个可以在统一架构上部署应用的PaaS平台Force.com，支持应用程序在云端的开发、部署和运行，给用户包括数据对象、权限管理、用户界面等在内的软件开发基础工具。这一战略转变不仅增强了Salesforce的技术实力，还为其后续的产品矩阵扩展提供了强大的支撑。

第三阶段是产品矩阵丰富期（2010—2015年）。在此期间，Salesforce通过疯狂收购和并购，不断丰富产品矩阵，形成了包括销售云、营销云、服务云和商业云在内的全方位企业级服务解决方案。这些产品覆盖了企业运营的各个环节，为Salesforce赢得了更广泛的客户群体。同时，Salesforce还深耕特定场景化服务，有针对性地进行行业化深耕，针对各行业销售的特点和难点，提供了汽车、通信、金融、医疗保健、高科技、生命科学、制造业、传媒、公共服务、零售这些垂直行业的专业销售、客服、营销方案。

第四阶段是智能生态构建期（2016年至今）。随着大数据和人工智能技术的兴起，Salesforce开始利用这些技术优化产品性能，提升用户体验，并收购多家AI公司，积极构建产业生态。

从初创期的CRM订阅服务，到PaaS平台的SaaS生态，再到移动和社文化时代的迅速反应，以及人工智能的布局，Salesforce持续紧跟时代步伐，不断推陈出新，牢牢把握竞争优势，形成了集合多领域业务的产品矩阵（见图2-2）。

Salesforce于2004年6月在纽交所成功上市，股票代码是CRM（赛富时），这也很好地诠释了公司业务的成功。

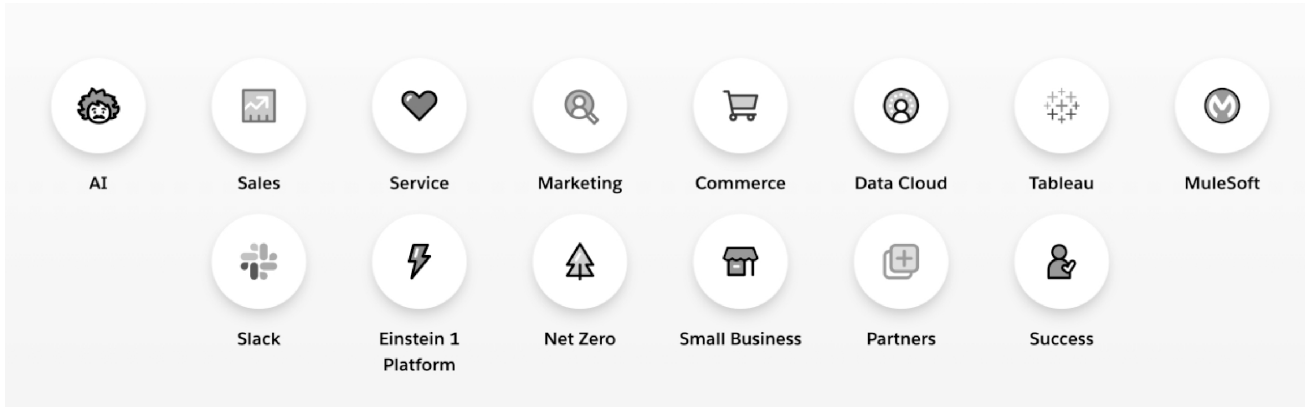


图2-2 Salesforce产品矩阵

1. Salesforce的财务数据（来自Salesforce财报公开数据）

从收入结构上来看，Salesforce的收入主要分为两部分：订阅和支持、专业服务和
其他服务（见表2-17）。其中，订阅和支持又分为服务、销售、平台及其他、市场
营销和商业、集成和分析这五大板块。我们可以看到，Salesforce的收入主要来自
订阅和支持业务。

表2-17 2022—2024年Salesforce收入结构

营收构成		2022 年		2023 年		2024 年	
		营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)
订阅和 支持	服务	73.69	23.50	82.45	23.65	90.54	23.89
	销售	68.31	21.79	75.80	21.75	83.22	21.96
	平台及 其他	59.67	19.03	66.11	18.97	72.47	19.12
	市场营销 和商业	45.16	14.40	49.12	14.09	52.81	13.94
	集成和 分析	43.38	13.84	51.89	14.89	57.75	15.24
专业服务 和其他 服务	专业服务 和其他 服务	23.31	7.43	23.20	6.66	22.16	5.85
总计	总计	313.52	100.00	348.57	100.00	378.95	100.00

看营收数据，Salesforce 2024年总营收378.95亿美元，营收连续多年保持增长；看利润数据，Salesforce成立至今大部分年份都处于盈利状态，截至目前已经连续多年盈利（如果继续回溯的话，2011—2015年Salesforce是有一些亏损的，这段时间公司处于平台化扩张期），2024年利润已经达到了76.66亿美元，盈利能力年同比大幅提升（见表2-18）。

表2-18 2019—2024年Salesforce营收和利润 单位：亿美元

项目	2019 年	2020 年	2021 年	2022 年	2023 年	2024 年
营收	170.98	212.52	264.92	313.52	348.57	378.95
利润	4.63	4.55	5.48	18.58	59.99	76.66

从成本构成情况来看，与大部分SaaS企业类似，Salesforce的主要成本为销售和管理成本以及研发成本。公司需要通过销售和渠道管理来获客，巩固市场地位；同时需要持续的研发投入来保证产品的竞争力。以2024年的财务数据为例，2024年Salesforce的销售和管理成本为160.93亿美元，占总营收的比例约为42%；研发成本为54.93亿美元，占总营收的比例约为14%。由此可见，Salesforce的研发成本投

入相对而言不算多，过去几年大多保持在总营收的15%左右，这也是Salesforce可以持续盈利的核心原因之一（见表2-19）。

表2-19 2019—2024年Salesforce成本构成情况 单位：亿美元

项目	2019年	2020年	2021年	2022年	2023年	2024年
销售和管理成本	96.34	153.59	144.53	160.79	154.11	160.93
研发成本	27.66	35.98	44.65	25.53	49.06	54.93

2. Salesforce的商业模式

Salesforce是典型的企业服务公司，SaaS模式的开创者，采用的是一元商业模式。Salesforce为企业提供涵盖客户管理、营销、销售、协同办公、数据等业务模块的SaaS产品和服务，并从客户侧获取价值。

在收费模式上，Salesforce也是采用订阅制收费，根据不同的企业规模以及不同的功能版本差异化定价。我们参考销售云产品的定价方式，来理解Salesforce的收费模式（见表2-20）。

表2-20 Salesforce销售云产品的定价方式

客户企业规模	客户销售团队规模	销售云产品版本	价格（美元/用户/月）
小微企业	5~20人	基础版	25
中型企业	20~100人	专业版	75
大型企业	100~300人	企业版	150
特大型企业	300人以上	无限制版	300

注：SaaS产品的定价策略变化较快，不同区域的报价不同，数据仅供参考。

3. Salesforce的公司解读

作为SaaS模式的鼻祖，Salesforce凭借着对订阅模式的坚持，以及在提供标准化产品和定制化服务之间的平衡，取得了傲人的成绩，营收规模和营业利润同步增长。

Salesforce从成立之初，就坚持SaaS模式，而摒弃传统软件的一次性买断模式，这大大提升了LTV；同时Salesforce坚持提供标准化产品，而减少定制化服务，虽然这样可能会丢掉一些大客户，且会因为部分中小客户生命周期短而LT和单用户ARPU（单用户平均收入，average revenue per user）低，但是这大大降低了Salesforce的边际成本，使其具备了规模化扩张的基础。

我们从前述提到的成本构成情况也可以看到，Salesforce的销售和管理成本大多维持在总营收的50%左右，这是个不低的数字。这与整体市场的竞争环境相关，虽然Salesforce是全球排名第一的CRM提供商（根据Statista的数据，2024年Salesforce在CRM领域的市场份额达到21.7%），但市场竞争依然激烈，Microsoft、Oracle、SAP、Adobe等都是Salesforce的竞争对手，Salesforce必须不断地投入销售和管理成本来维持其市场份额（见图2-3）。Salesforce在销售投入上采取人海战术，也印证了这一点。

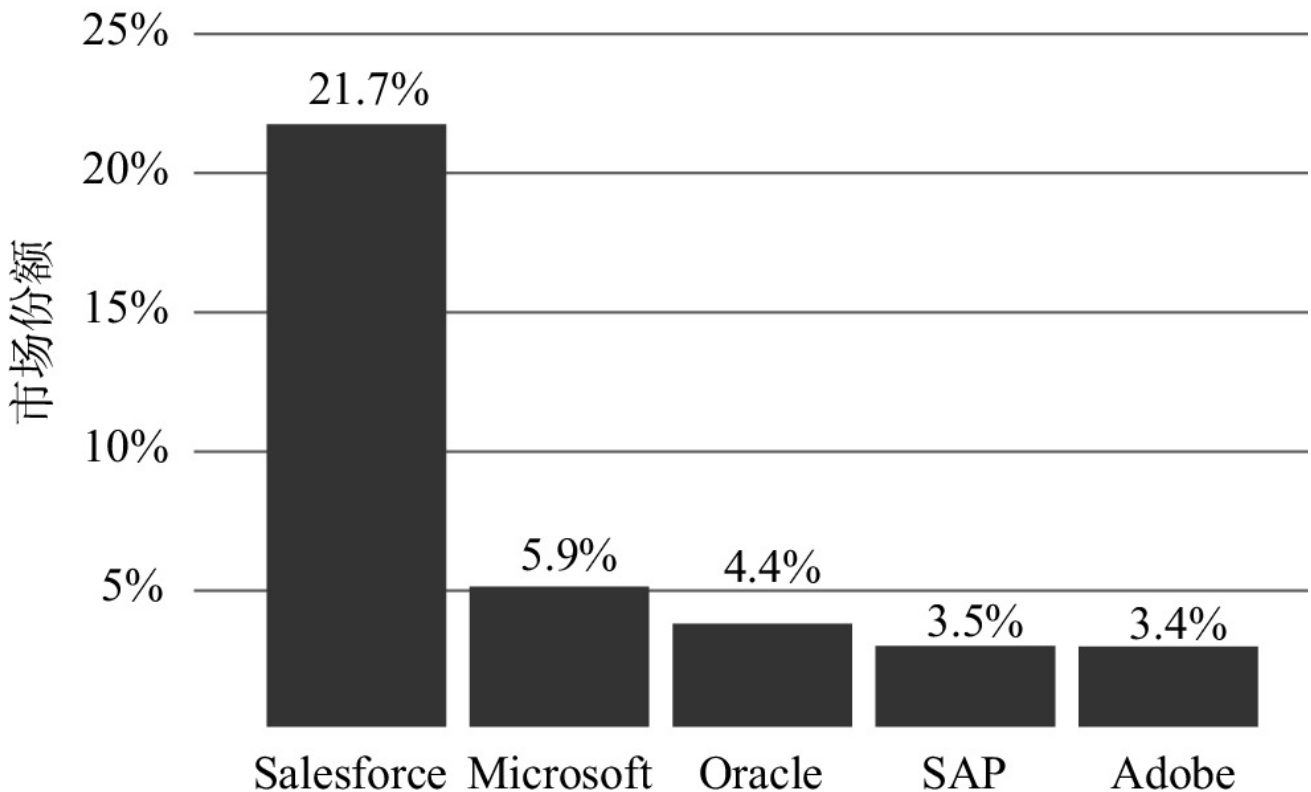


图2-3 2024年全球CRM市场份额资料来源：Statista。

面对未来的发展机遇与挑战，Salesforce需要继续深化技术创新，优化客户体验，并灵活应对市场变化，以实现可持续的增长和领导地位的巩固。

2.4 小结

在本章结束之前，我们可以再看一下ofo小黄车的案例。ofo成立于2014年，是国内首个无桩共享单车出行平台，创造了“无桩共享单车”出行模式。我们抛开ofo在资本市场上与资方的恩怨情仇不谈，从商业模式上看，本质上ofo采用的是披着共享经济外衣的单车租赁模式，甚至由于违规使用用户押金，ofo的商业模式更是在某种意义上变成了单车售卖模式。

ofo在扩张期间，占用了大量的公共空间，这部分公共空间对ofo而言没有财务成本，进而摊薄了ofo的整体经营成本。在租赁的商业模式下，随着规模的扩大，企业之间比拼的是运营效率，开城、车辆调度、车子的损耗等都在挑战企业运营效率，而且每开一个城市，边际成本都在提升，企业组织管理的成本随业务规模的扩大而膨胀。

本章我们介绍了互联网一元商业模式的典型业务逻辑、核心要素、画布和经营分析框架，并分析了几个典型的公司。互联网行业中，还有其他大量采用一元商业模式的公司和业务。读者可以参考本书对互联网一元商业模式的解读以及对部分公司的扫描，对其他公司进行分析和研究。

免责声明：本书提到的所有公司，仅代表作者个人观点，不构成任何投资建议。

第3章 二元商业模式：传递

3.1 互联网二元商业模式研究

3.1.1 互联网二元商业模式的典型业务

如第1章所介绍的，我们将二元商业模式定义为传递——组织建设线上场域连接商家和用户，完成商家和用户之间的商业价值传递，并获得价值的基本原理，即用户→商家。

很多互联网公司和业务的商业模式都是传递，其中最典型的有两大类业务：交易平台和信息平台。我们看一下这两大类业务的底层逻辑。

1. 交易平台

所谓交易平台，是指在双边平台中，用户和商家能完成交易的闭环，履约可以发生在线上或者线下。即，组织提供的价值主张中包含交易环节。典型的交易平台有电商交易平台、网约车平台、本地生活服务交易平台、OTA旅行平台、房产交易和服务平台和二手车交易平台。用户在光顾这些平台的时候，都是带着商业需求来的，即用户会带着“一笔钱”来购物、打车、点外卖、订机票、买房、买车等，组织就是完成这笔钱与对应商家之间的传递。

● **电商交易平台：**以京东、淘宝、拼多多、Amazon和闲鱼与转转等二手电商为代表的电商交易平台，其双边连接的是电商商家和买家，他们可以直接在平台上完成商品的闭环交易。一般来说，组织主要是通过商家管理费、商家广告营销费等方式，从电商商家侧获取价值。例如，商家在平台开店需要支付入驻费，商家为了获取更多的用户流量和销量、打造自己的品牌等需要支付广告费在平台进行广告投放。

在电商交易平台中，还普遍存在一种特殊的场景，即平台上的商家是组织自己的自营商家，典型例子就是京东和天猫上会有很多自营店，甚至还有京东超市和天猫超市。尤其是京东，其创业初期做的就是自营生意。在这个场景中，组织不但充当平台角色，还作为平台的商家角色出现，而这两个角色的商业模式不完全相同。自营店和自营超市可以类比为传统的商超，而商超的商业模式取决于货权是否发生转移。如果货权转移，那就是买手店的模式——通过低买高卖赚差价的模式来获取价值，这是传统的一元商业模式；如果货权没有转移，那就是货架的模式——通过提供货架来传递商家和用户之间的商业价值，这是二元商业模式。很多商超都处于一元商业模式和二元商业模式的混合态。

● **网约车平台**：以滴滴出行、Uber、Grab、美团打车和货拉拉（货运）为代表的网约车平台，其双边连接的是司机和乘客，司机为乘客提供载人或者运货的服务，他们可以直接在平台上完成打车订单的闭环交易。一般来说，组织是从司机侧获取价值，主要是通过抽佣的方式。例如，用户在滴滴出行上完成一笔快车订单，滴滴出行会从司机的车费中抽取一定比例作为平台的收入。这里有个财务现象读者可以关注一下（详情见本章3.3.4）：滴滴出行会将每一笔快车订单的费用全部计为公司的收入，然后由公司属于司机的费用统一支付给司机。这是一种给司机“发工资”的方式，但只是财务上的操作，并不影响滴滴出行快车业务的商业模式。

● **本地生活服务交易平台**：以美团外卖（即时零售）和饿了么为代表的本地生活服务交易平台，其双边连接的是提供本地生活服务的商家和消费者。本地生活服务包含生活的方方面面，像餐饮、水果、鲜花、药品、超市、休闲娱乐等，商家和消费者可以直接在平台上完成本地生活服务订单的闭环交易。一般来说，组织是从本地生活服务商家侧获得价值，主要也是通过抽佣的方式。例如，用户在美团完成一笔外卖订单，美团会从餐饮商家的外卖收入中抽取一定的比例作为平台的收入。

● **OTA旅行平台**：以携程、美团和飞猪为代表的OTA（在线旅游，online travel agency）旅行平台，给用户旅行提供相关商品的交易服务，双边连接的是旅行商品的卖家和差旅用户，旅行商品包括机票、酒店、保险和门票等，卖家和差旅用户可以直接在平台上完成订单的闭环交易。一般来说，组织是从商家侧获得价值，主要是通过抽佣和广告的方式。

● **房产交易和服务平台**：以链家为代表的房产交易和服务平台，给用户二手房买卖服务，双边连接的是房子的卖家和买家，通过平台和房产经纪人撮合双方完成交易。房产交易和服务平台的业务是比较有特点的，虽然交易闭环是发生在平台上的，但履约涉及的线下环节非常烦琐，所以房产经纪人的角色很重要。房产交易和服务平台以前一直是从买家侧获取价值，通过服务费（抽佣）的方式。例如，链家会根据房子的成交价来抽取一定的比例作为本次交易的服务费，以前一直都是由房子的买家来支付服务费，但2023年第三季度链家下调了中介服务费费率（2.7%→2%），并调整了中介服务费的收取方式，由买家支付变成买卖双方共同支付：房子的买家和卖家各支付50%。为什么服务费的支付方会发生变化呢？后续将会说明。二手房交易业务的模式比较有特点且复杂，我们从另一个思考角度切入，并留两个问题供读者思考：二手房交易业务中房产经纪人（中介）的角色体现的是商家的还是组织的价值主张？链家与贝壳的区别是什么？我们在本章3.3.6中会详细介绍。

● **二手车交易平台**：以瓜子二手车、人人车、优信为代表的二手车交易平台，给用户二手车买卖服务，双边连接的是二手车卖家和买家，其中卖家是个人车主，或者车商；而买家绝大多数情况是个人买家（同样也有车商充当买家从平台牟

利)。二手车交易平台是从买家侧获取价值，通过服务费（抽佣）的方式。例如，瓜子二手车会抽取一定比例的服务费，由车子的买家来支付。不过值得思考的是，二手车的交易佣金并不高（正常都在成交价的5%以内）。在一个完全竞争且各家公司市占率都没有明显优势的市场中，仅仅依靠这部分收入是很难支撑业务持续经营的。因此，这些二手车交易平台会衍生出大量的相关业务，例如汽车金融（金融车贷）、自己当二手车车商（一元商业模式）、新车零售（一元商业模式）等等。

2. 信息平台

所谓信息平台，是指在双边平台中，用户和商家只是完成信息的交互（用户获取商家的信息、商家触达用户），而最终的交易在平台之外完成，履约同样可以发生在线上或者线下。即，组织提供的价值主张中只包含信息聚合和匹配，不包含交易环节。典型的如分类信息平台。分类信息平台可以分为全品类信息平台和垂类信息平台，其中全品类信息平台以58同城为代表，垂类信息平台以汽车信息平台、B2B信息平台和二手房信息平台为代表。双边连接的是商家和用户，平台提供信息的分发和匹配服务。同样，用户在光顾这些平台的时候，也是带着商业需求来的：用户还是带着“一笔钱”来找本地服务、买车、买房、采购等。虽然最终交易不在平台上完成，但组织依然是在完成这笔钱与对应商家之间的传递。这一点和交易平台没有本质的区别。但是由于信息平台不具备交易能力，所以纯信息平台无法通过交易抽佣来获取价值，只能通过广告从商家侧获取价值，或者衍生出一元商业模式，通过会员模式从用户侧获取价值（例如安居客的房产中介会员）。

● **汽车信息平台和B2B信息平台：**本质上，汽车信息平台和B2B信息平台都属于垂类信息平台。以汽车信息平台为例，这类平台以汽车之家、懂车帝为代表，双边连接的是车商和用户，车商包括4S店、汽车品牌、二手车商、个人二手车主等，平台提供信息的分发和匹配服务。

● **二手房信息平台：**二手房信息平台也属于垂类信息平台，以安居客为代表，实现房源（卖家、中介）和用户之间的匹配。另外，这里值得研究的是，安居客有很大一部分收入是来自B端经纪人（中介）的会员费，安居客连接的是房子的买家和房产经纪人，在本章3.3.7介绍安居客时会展开阐述。

不论交易是否发生在组织的平台之内（更通俗地讲，就是用户的支付是否发生在平台上），交易量和GMV都真实存在，组织在二元商业模式中都是针对GMV收取服务费来获取价值。如前文所言，组织在传递商家和用户之间的商业价值时，对用户的调度效率越高、对用户行为的控制力越强、提供的价值主张涉及环节越多且价值越大时，货币化的潜力就越大，对应GMV的货币化率会越高。信息平台相对交易平

台而言，不掌握交易信息、价值主张涉及环节偏少，且在每笔交易中价值贡献度偏低（信息平台比交易平台做得更轻，信息的质量也相对差一些）。所以，对一个信息平台来说，在其他条件都不变的情况下，如果能够进一步升级为交易平台，则有希望获取更大的业务份额和更高的货币化率。就算同样通过广告从商家侧获取价值，交易平台比信息平台更有优势，因为交易平台能拿到后端转化的数据，用来优化广告投放，整体的广告变现效率会远远高于信息平台（见表3-1）。

很多信息平台都想做交易闭环，本质上也是在追求更高的壁垒和更强的竞争力。

表3-1 交易平台与信息平台对比

类型	商业模式	交易闭环	业务特点	GMV	货币化潜力
交易平台	二元	是	重	存在，好统计	大
信息平台	(用户带着“一笔钱”)	否	轻	存在，不好统计	小

3.1.2 二元商业模式的飞轮效应

二元商业模式的组织是一个双边平台，双边连接着商家和用户，本质上就是供给和需求，而平台对于供给和需求的价值源自需求和供给的存在。双边平台具有一定的飞轮效应，即双边市场效应：供给和需求之间相互形成正反馈，形成增长飞轮，促进规模扩大和价值增长。

- 供给：即商家。用户越多，需求就越多，商家能获得更多的收入，这样就会吸引市场上更多的商家加入，进而继续增加供给。商家数量增加还会加剧竞争，在良性的竞争环境下，这能促进迭代和创新，带来更优质的产品和服务。
- 需求：即用户。商家越多，供给就越多，用户就能获取更多更好的产品和服务，这样就会吸引市场上更多的用户加入，进而继续增加需求。用户增加，还可以带来更多的反馈，进一步促进供给的迭代。

最终在飞轮效应的影响下，整体GMV和组织收入持续增长。我们可以用图3-1来概述二元商业模式的飞轮效应。

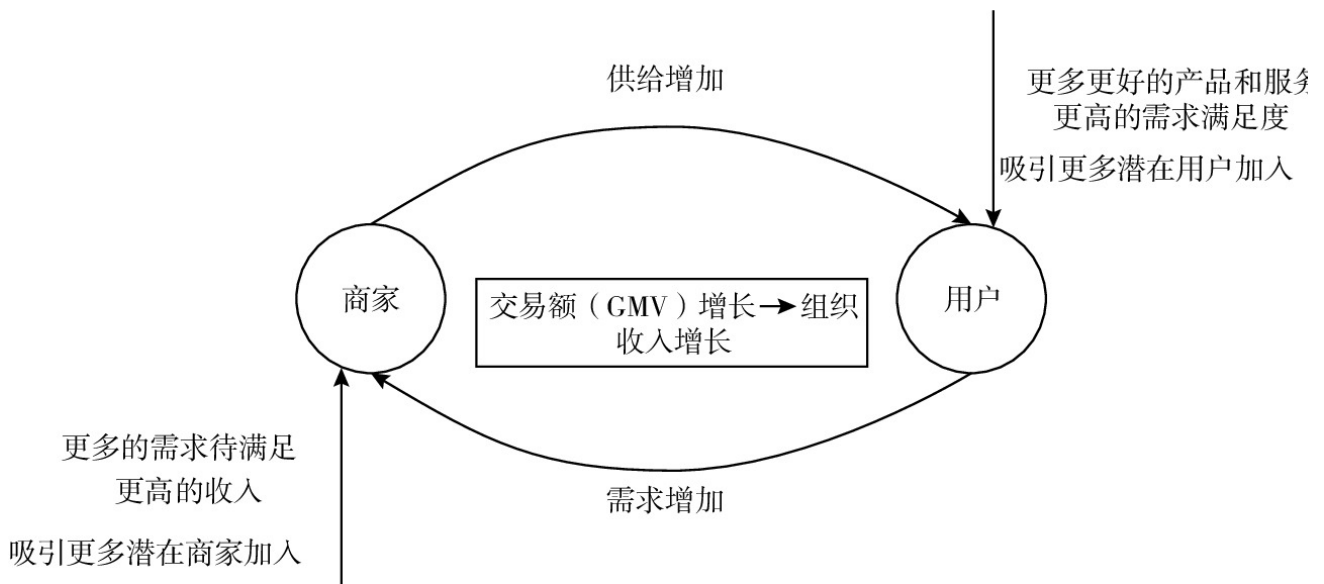


图3-1 二元商业模式的飞轮效应

在二元商业模式中，有四个关键点会影响商业模式的启动和把控，是组织在运营过程中必须考虑的。

(1) 双边冷启：从前面的飞轮效应中也可以看出来，双边平台会面临双边冷启的问题，即“先有鸡还是先有蛋”的两难困境。要解决这个问题，其中一种方式

是发放补贴。组织的运营者要清楚是一边（哪一边）需要补贴，还是双边都需要补贴，以及合适的补贴价格是多少。补贴的对象、方式、成本等，在业务发展的不同阶段也会大有不同。绝大多数双边平台应该先搞定供给。

（2）**双边平台的同边负效应**：同边负效应是指随着规模增长，同边的相同角色之间相互竞争，一笔交易发生会影响另一笔交易。例如打车，在高峰期很多用户打不到车（供给有限，运力不足，用户和用户之间竞争），在非高峰期很多司机空载（需求有限，订单不足，司机和司机之间竞争）。外卖也存在同边负效应，但低于打车，因为商家和配送人力的弹性会略大一些。而有一些业务并没有同边负效应，供给和需求接近无限，例如淘宝、拼多多等。

（3）**双边匹配的持续可调度**：所谓持续可调度，是指供给和需求之间，不会产生单个用户和单个商家在一段时间内多次达成同一个交易的现象。如果两者能形成稳定关系的交易，就不具备持续可调度的特点。例如，用户很少能在同一时间的同一地点打上同一个司机的车，但是像美容美发、运动健身等业务，用户和商家之间会在线下完成交易履约的过程中产生交流，并形成固定的搭配，用户会选择同一个理发师或健身教练持续服务。二元商业模式的基础，是双边具备持续可调度的特点。

（4）**交易平台的飞单挑战**：飞单，也可以称为水下单，即平台完成了传递的过程，但双边用户私下完成交易，最终平台无法监控到这部分交易，影响平台GMV和货币化率。例如，用户通过滴滴出行叫到车之后，司机和乘客线下交易，不走线上订单；用户通过链家找到目标房子之后，买方直接和卖方线下交易，跳过链家经纪人。这种飞单行为不利于平台管控交易，浪费平台资源，必须重点治理。组织需要通过设计合理的机制来管控飞单。

基于上述的定义和典型业务介绍，我们可以总结一下二元商业模式的核心要素。

（1）**用户**：商家和用户都是组织的目标用户。商家和用户都是广义的，存在个人商家和B端企业用户。

（2）**价值主张**：组织给商家和用户分别提供了多元的价值主张，涵盖商业价值传递的方方面面，包括但不限于信息匹配、交易撮合、履约、售后、资质审核以及交易安全保障等。提供价值主张的方式，可以由组织自建，也可以通过三方合作，例如物流、外卖配送等。组织提供的价值主张渗透在传递中的环节越多、价值越大，组织获取的价值也就越大。

（3）**渠道**：组织需要通过渠道触达目标用户。根据业务以及目标用户的不同，组织需要的渠道也大有不同：有些重线下，例如地推、开城；有些重线上，例如广告、私域。

(4) **成本**：维持商业模式运转和规模扩张的所有成本（获客成本、企业经营成本、价值主张成本、渠道成本等），都可以核算到每一笔交易中。单笔交易的成本结构也能反映出业务的健康度。对大部分采用二元商业模式的业务而言，单笔交易的边际成本会递减，业务具备规模经济效应；但也有一些业务的每笔交易成本是相对固定的，甚至随着规模扩大，管理成本会激增，这对组织的经营能力提出了更高的要求。

(5) **收入**：组织在完成传递的过程中获得价值，即从每笔交易中获得价值，完成货币化。获得价值的方式有很多种，组织会选择其中的一种或者将多种方式组合，典型的有商家入驻和管理费、交易抽佣、广告。

● **商家入驻和管理费**：这一点很好理解，组织对入驻的商家直接收取费用，例如入驻费、管理费、认证费等，虽然收费名目多种多样，但本质都是一样的。

● **交易抽佣**：组织从每笔交易订单中抽佣，抽佣可以采用固定值、固定比例、动态比例、阶梯比例等方式，具体取决于组织的经营机制设计。组织绝大多数情况下是从商家侧获得价值，但也有少数会从用户侧获得价值，典型的如二手房交易业务，是由买家或者买卖双方共同支付服务费（房屋租赁业务也是同时从双方获取价值）。具体从哪边获取价值，是由供需双方的市场关系决定的：在供给方市场中，一笔交易的成本会显性地从卖方转移到买方，即组织能从买家侧获取价值。这里就可以理解，为什么2023年第三季度链家会将二手房中介服务费从买家支付改为买卖双方共同支付了，因为房地产业的宏观环境使当时的二手房交易市场已经从供给方市场逐步向需求方市场偏移。至于这笔交易的成本是否会隐性地转移到买方（例如卖方把中介费加到房价中），则会随着房地产业整体的市场环境、政策以及供需关系的变化而动态变化。

● **广告**：组织从平台的资源（流量）中拿出一部分作为商业资源，并提供广告营销服务。商家作为广告主在平台进行广告投放，组织从商家侧赚取广告费。

3.1.3 商业模式画布

根据二元商业模式的核心要素，我们可以画出互联网二元商业模式的画布（如表3-2所示）。基于这张画布，结合业务和公司的具体情况，可以拆解所有的二元商业模式，并对其改进甚至重新设计。读者可以选择一些具体的业务来尝试拆解和剖析。

表3-2 互联网二元商业模式的画布

用户 C端用户、B端用户 全国甚至全球（区域约束相对小）、长尾需求得到释放	价值主张 To 用户：信息匹配、商品质量保障、履约、交易能力、资金安全等 To 商家：商机、需求匹配、履约、生意经营、用户运营、私域经营、交易能力、管理能力、资金安全等	商家 大商家、中小商家、个人商家 全国甚至全球（区域约束相对小）、C2M模式
	渠道 触达用户和商家 线上渠道、线下渠道	
成本 渠道成本、营销成本、用户补贴成本 管理成本、经营成本 关注单笔交易的成本结构和边际成本		收入 从每笔交易中获取收入 商家管理费、抽佣、广告 识别 GMV、关注货币化率

二元商业模式的业务，对组织的运营能力会有比较高的要求，组织需要做好供给侧和需求侧的双边运营。整体上，虽然看起来只要能完成双边的冷启动，二元模式的业务就有机会形成增长飞轮，盈利比一元商业模式要轻松很多，但实际上，也有很多二元模式的业务增长乏力、规模扩张困难、盈利困难，或者利润率极低，赚的都是辛苦钱。对于在分析采用二元商业模式的公司时，采用什么分析框架、市场规模受哪些因素的影响，以及什么会制约货币化率、为什么有些业务会以失败收场这些问题，本章接下来将重点讲解。

3.2 商业模式经营分析框架

二元商业模式远比一元商业模式复杂，其组织核心是完成商家和用户之间的商业价值传递。每发生一次传递，就会产生一笔交易和对应的交易额，而这笔交易也会有对应的交易成本。在分析二元商业模式经营情况的时候，我们可以采取下面的分析框架。

利润=市场规模_总交易量×线上化率×市场份额×（单笔交易客单价×货币化率-单笔交易成本）=市场规模_总GMV×线上化率×市场份额×货币化率-总成本

二元商业模式的盈利能力和发展潜力，主要取决于市场规模、货币化率和成本。市场规模即在企业平台上产生的交易量，或者说GMV；货币化率是企业从GMV中获取的价值。从转化为单笔交易来看，单笔交易的收入减掉成本，即单笔交易的利润，再乘以交易量即为总利润。再从市场空间、潜力以及竞争视角来看，我们可以通过企业所在行业的市场规模和市场份额，洞察企业发展状况。（说明：从财务口径来看，营收、利润和成本等都有很多统计口径，例如毛利、经营利润、净利润、Non-GAAP净利润、息税前利润等等，此处不做财务视角的拆分，重点关注商业模式的底层逻辑。）

上述公式里有六个比较核心的指标：市场规模（GMV和交易量）、线上化率、市场份额、单笔交易客单价、货币化率和单笔交易成本。这六个核心指标基本代表了业务的市场空间、潜力、竞争程度、营收能力和货币化情况。而同样都是二元商业模式，不同业务之间差异的核心也体现在这几个指标里。接下来我们具体展开讲一下这六个指标。

1. 市场规模

市场规模（market size）也可以称为市场容量（market volume），是指目标产品或行业的整体经济规模，包括产品或行业在一段时间内的产量、产值等。市场规模与产品和服务的需求和供给相关，整体会受到政治、经济、文化、人口、技术和自然条件等因素的影响。一个行业的市场规模，通常可以用产量、产值或销售额来计算。通俗地讲，一个行业的市场规模就是该行业一段时间内的经济总量，体现为GMV和交易量，例如大家熟知的商品零售、汽车、房地产、云计算、交通出行等都是万亿级的市场。市场规模未来的发展趋势，即市场增长率也很重要，受前述提到的一些宏观因素的影响。一个行业是在上升期、下降期还是横盘期，在做经营分析时也需要关注。典型的如汽车行业的细分赛道中，新能源汽车

目前处于上升期，市场渗透率快速提升，而传统燃油车处于下降期，市场被新能源汽车挤压。一个行业的市场规模和发展趋势，可以用来判断企业营收或利润的天花板、未来发展空间，以及横向对比企业在市场上的竞争力。另外，如第2章提到的，互联网还有一个典型的特点：很多行业通过互联网释放了大量长尾需求，进而提升了市场规模。

2. 线上化率

线上化率是指通过线上渠道完成的交易量（或GMV）占市场总规模的比例。研究互联网行业的商业模式时，线上化率至关重要。随着互联网行业的发展，从PC互联网时代到移动互联网时代，再到万物互联时代，几乎全行业的线上化率都在逐步提升，这也是一批批互联网企业在全中国崛起的底层支撑。不同行业的线上化率各不相同，这和行业属性、区域、文化、技术等方面都有关联。例如，根据国家统计局的数据，2024年1月到7月，全国实物商品网上零售额总计约8万亿元，占社会消费品零售总额的比重为25.6%，基本可以认为这就是商品零售行业目前的线上化率，即电商的渗透率。随着基建、技术、代际等因素的变化和发展，作者认为电商的渗透率还会进一步提升。但不同业务和品类的线上化率差异很大，比较典型的例如B2B的线上化率低于B2C，化妆品和服装的线上化率高于家居家装等。

3. 市场份额

市场份额是指一个企业的交易量（或GMV）在市场同类中所占的比重。例如，目前国内电商行业年总GMV大约为20万亿元。根据高盛发布的研究报告，在2024年“6·18”期间，从GMV来看，阿里淘天的市场份额大约为42%，京东的市场份额大约为22%，拼多多的市场份额大约为18%，抖音电商的市场份额大约为15%。国内网约车月度订单量在几亿单的规模，根据交通运输部的统计数据，2024年4月全国网约车订单量达到8.97亿单，其中滴滴出行网约车（不含出租车）的订单份额超过70%，月订单量超过6亿单，稳居行业龙头地位。市场份额直接反映了竞争格局，是企业的一个非常关键的指标。通过市场份额，我们可以判断出企业的竞争地位以及行业的竞争激烈程度。如果企业的市场份额小，规模小，那企业就发挥不出规模效应的优势，很多成本无法分摊。而且在一个充分竞争的市场中，企业的成本也会提高，补贴大战就是经典的例子。回首互联网行业过去的十几年，我们经历了团购（千团网）大战、外卖（美团、饿了么）大战、网约车（滴滴出行、快的、Uber）大战、共享单车（摩拜、ofo、青桔）大战、在线旅游（携程、去哪儿、艺龙）大战、百亿补贴（拼多多、京东、淘宝）大战等等，都是大家在砸钱

抢占市场份额。一旦市场份额稳定了，很多补贴退场之后，实际交易量并不会下降，但企业的成本会下降，而利润会提升。

4. 单笔交易客单价

单笔交易客单价是指单笔交易对应的GMV。通俗点儿说，就是一笔交易的订单金额。这笔交易可以是发生在线上的，即用户在线上完成支付，也可以是发生在线下的。但不论是线上还是线下，交易量和GMV都真实存在。回到我们前面提到的二元商业模式的定义，组织就是在商家和用户之间进行商业价值的传递，完成“一笔钱”的传递，这笔钱就是单笔交易的客单价。客单价有高有低，与产品或服务的定价相关，不同行业和品类差异很大。例如，房地产、汽车、服装、外卖等的客单价范围非常宽。

5. 货币化率

货币化率（take rate）代表了组织在交易中获取的价值。例如，一笔交易的客单价为X，企业从这笔交易中获取的收入为Y，则货币化率=Y/X。二元商业模式的业务营收就是针对GMV的货币化。货币化率非常重要，直接决定了企业的盈利能力和潜力。

6. 单笔交易成本

单笔交易成本是指在一笔交易中组织需要支付的成本。一般而言，单笔交易的成本包括B端（商家端）成本和C端（用户端）成本。例如，在外卖业务中，B端会有履约成本，包括骑手成本、售后成本等；C端会有获客成本，包括补贴、超时补偿等。

上述这六个核心指标可以分为两组：**战略层面指标**和**日常经营层面指标**。

战略层面指标包括市场规模和线上化率，指导企业长期投入的方向。这两个指标主要由赛道选择决定。企业要扩大市场规模和提高线上化率，就要推动整个赛道的发展。例如，房地产行业交易的线上化率目前还不够高，从战略层面看，链家需要继续提升线上交易的渗透率。

日常经营层面指标包括市场份额、单笔交易客单价、货币化率和单笔交易成本，决定企业当前的财务状况，支撑企业可持续发展。对于这几个指标，企业的控制

力更强，它们是企业产品运营和技术优化的重点，且相对容易被企业影响和决定。

综上所述，采用二元模式的企业的发展，综合受到市场规模、线上化率、市场份额、单笔交易客单价、货币化率和单笔交易成本的影响。接下来我们选择了七家具有代表性的公司，详细介绍一下它们的业务情况。它们是京东、阿里巴巴、拼多多、滴滴出行、美团、贝壳和安居客，它们都是二元商业模式。

3.3 典型公司及其业务研究

3.3.1 京东：从线下柜台起家，以自营为核心，以供应链为基础的技术与服务企业

京东是从北京中关村一个线下柜台起家的。1998年6月18日，刘强东在中关村创业成立京东公司。最开始京东只是一个光碟、磁盘等产品的代理商。2003年，京东在北京已经开了12家门店，是光磁产品领域最具影响力的代理商之一，销售量和影响力在行业内首屈一指。但2003年“非典”这个“黑天鹅”对线下生意产生了剧烈的冲击，刘强东开始研究线上零售和电子商务，并在CDBest论坛上迈出了线上成交的第一步。由此，刘强东决定关闭线下所有门店，全力做线上。2004年1月1日，京东多媒体网站上线，标志着京东正式转型为电商。与同时代的淘宝、eBay等电商不同，为了适应电子类产品客单价高、质量不易得到保障等特点，京东从上线之日起做的就是自营模式。

经过持续的经营和发展，如今的京东已经是集零售、物流、健康、云、金融等业务于一体的大型技术与服务企业。2014年5月，京东集团在美国纳斯达克证券交易所正式挂牌上市，是中国第一个成功赴美上市的综合型电商平台。2020年6月，京东集团在香港联交所二次上市，募集资金用于投资以供应链为基础的关键技术创新，以进一步提升用户体验及提高运营效率。2021年5月，京东物流于香港联交所主板上市。

京东目前的业务可以分为四大板块——零售、物流、健康和新业务（见图3-2），收入主要还是来自零售业务。京东物流也从一开始的成本中心变成了利润中心，从2023年开始盈利。京东物流财报显示，2024年京东物流营收为1 828.38亿元，其中一体化供应链客户占比47.78%，外部其他客户占比52.22%，持续经营业务税后利润达70.88亿元。

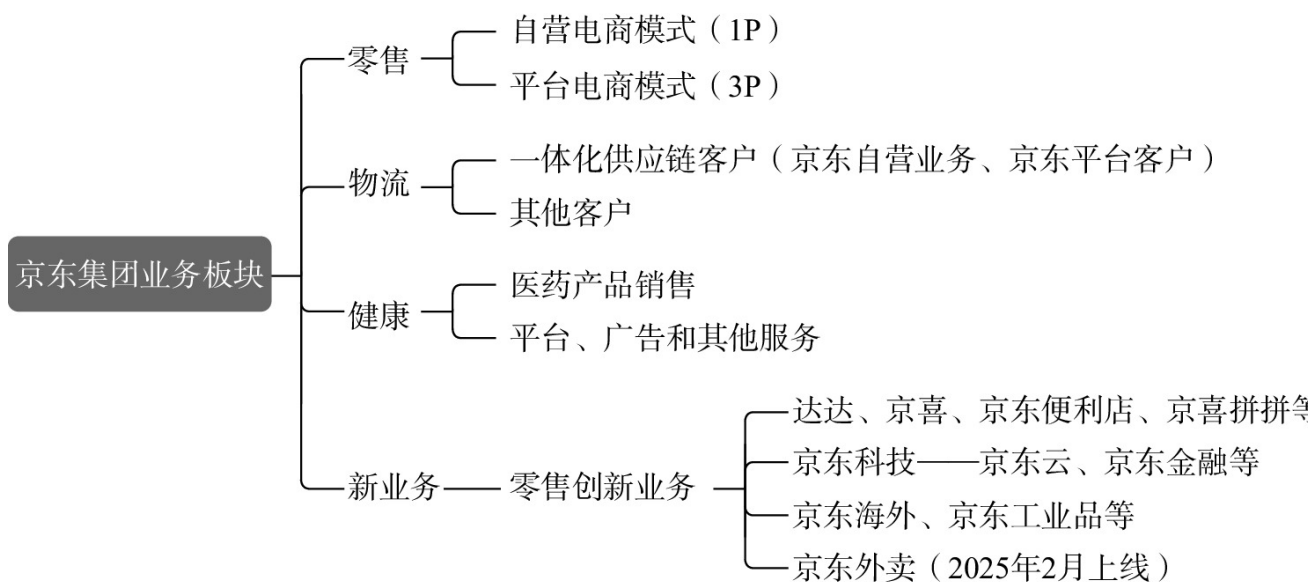


图3-2 截至2025年第一季度京东集团四大业务板块

结合京东的业务板块，再回顾京东的发展，有三个非常关键的战略决策：**2004年转型电商、2007年确定自建物流和全品类扩张**。围绕这三大战略决策，在落地执行上有几个标志性的事件。

- 2007年设立物流服务部门，第一个仓库在北京凤凰岭落地，第一个配送站在北京潘家园站建成。
- 2008年上线大家电产品和日用百货类产品，并在2012年与苏宁展开家电大战。
- 2010年上线图书类产品，并在2011年与当当网展开图书大战。
- 2010年上线平台电商模式，从自营电商走向综合电商，继续扩充品类和SKU（库存量单位，stock keeping unit）。
- 2014年在上海启动首座“亚洲一号”大型智能仓库。
- 2017年正式成立京东物流集团。

在核心战略决策的牵引下，京东20年如一日地坚持落地执行，逐步成为行业内最具代表性的企业之一，并推动了电商和物流全行业的共同发展，创造了巨大的社会价值，助力实体经济高质量发展。

京东坚持“成本、效率、产品、价格、服务”的核心经营理念。京东之前的品牌口号“多、快、好、省”堪称是电商的**第一性原理**，即商品种类多、物流速度快、

商品质量好、产品和服务性价比高。只是以“自营模式+自建物流”为基本盘的京东的核心竞争力为“快和好”，而“多和省”方面比竞争对手略有不足。

1. 京东的财务数据（来自京东集团财报公开数据）

从收入结构上看，京东的收入主要来自京东零售。以2023年和2024年的数据为例，这两年京东零售营收占比分别为87.16%和87.67%（见表3-3）。

表3-3 2022—2024年京东营收构成

营收构成	2022年		2023年		2024年	
	营收（亿元）	占比（%）	营收（亿元）	占比（%）	营收（亿元）	占比（%）
京东零售	9 299.29	88.88	9 453.43	87.16	10 159.48	87.67
京东物流	1 374.02	13.13	1 666.25	15.36	1 828.37	15.78
新业务	217.79	2.08	161.11	1.49	94.93	0.82
达达	80.30	0.77	105.06	0.97	96.64	0.83
分部间抵销	- 509.04	- 4.87	- 539.23	- 4.97	- 591.23	- 5.10
总计	10 462.36	100.00	10 846.62	100.00	11 588.19	100.00

看营收和成本数据，京东2024年总GMV约3.66万亿元，总营收约1.16万亿元，均连续多年保持增长，但总GMV和总营收的增速已经快速下滑；看利润数据，京东已经连续盈利多年，2024年归母公司净利润达413.59亿元（其中2021年由于股权投资亏损、新业务亏损以及物流亏损，当年京东归母公司净利润亏损35.60亿元，见表3-4）。从成本结构上看，京东的主营业务成本非常高，主要是因为京东自营电商模式本质上是买手店模式，自营模式中GMV基本等于营收，而与之对应就会有很高的商品采购成本。

表3-4 2019—2024年京东营收与成本等数据

项 目	2019 年	2020 年	2021 年	2022 年	2023 年	2024 年
总 GMV (万亿元)	2.09	2.61	3.30	3.48	3.54	3.66
GMV 年同比 (%)	—	24.88	26.44	5.45	1.72	3.39
总营收 (亿元)	5 769	7 458	9 516	10 462	10 847	11 588
总营收年同比 (%)	—	29.28	27.59	9.94	3.67	6.83
主营业务成本 (亿元)	4 924.67	6 366.94	8 225.26	8 991.63	9 249.58	9 749.51
毛利 (亿元)	844.21	1 091.08	1 290.66	1 470.73	1 597.04	1 838.68
销售和管理费用 (亿元)	277.24	335.65	503.05	488.25	498.43	568.41
研发费用 (亿元)	146.19	161.49	163.32	168.93	163.93	170.31
其他营业费用 (亿元)	347.46	461.55	565.73	630.11	645.58	704.26
营业利润 (亿元)	73.32	132.39	58.56	183.44	289.10	395.70
归母公司净利润 (亿元)	121.84	494.05	-35.60	103.80	241.67	413.59

注：2019—2022年的GMV数据来自京东财报；2023年和2024年的GMV数据来自行业研究报告。

国内电商行业的渗透率已经接近30%，并且竞争非常激烈，除了京东、阿里巴巴和拼多多三家持续竞争多年，以抖音和快手为代表的兴趣电商和直播电商快速崛起，微信视频号电商蓄势待发，都在抢占市场份额。近两年京东的GMV和营收增速下降，与竞争环境息息相关。从GMV份额来看，2024年京东的市场份额大约为16.5%，已经被拼多多反超，并且与抖音电商差距非常小，排名随时有可能跌到第四位（见表3-5、图3-3）。

表3-5 2024年中国主要电商GMV及其市场份额

中国电商业务	GMV (万亿元)	市场份额 (%)
阿里淘天	8.20	37.02
拼多多	5.20	23.48
京东	3.66	16.52
抖音电商	3.50	15.80

续表

中国电商业	GMV (万亿元)	市场份额 (%)
快手电商	1.39	6.28
微信视频号电商	0.20	0.90
汇总	22.15	100.00

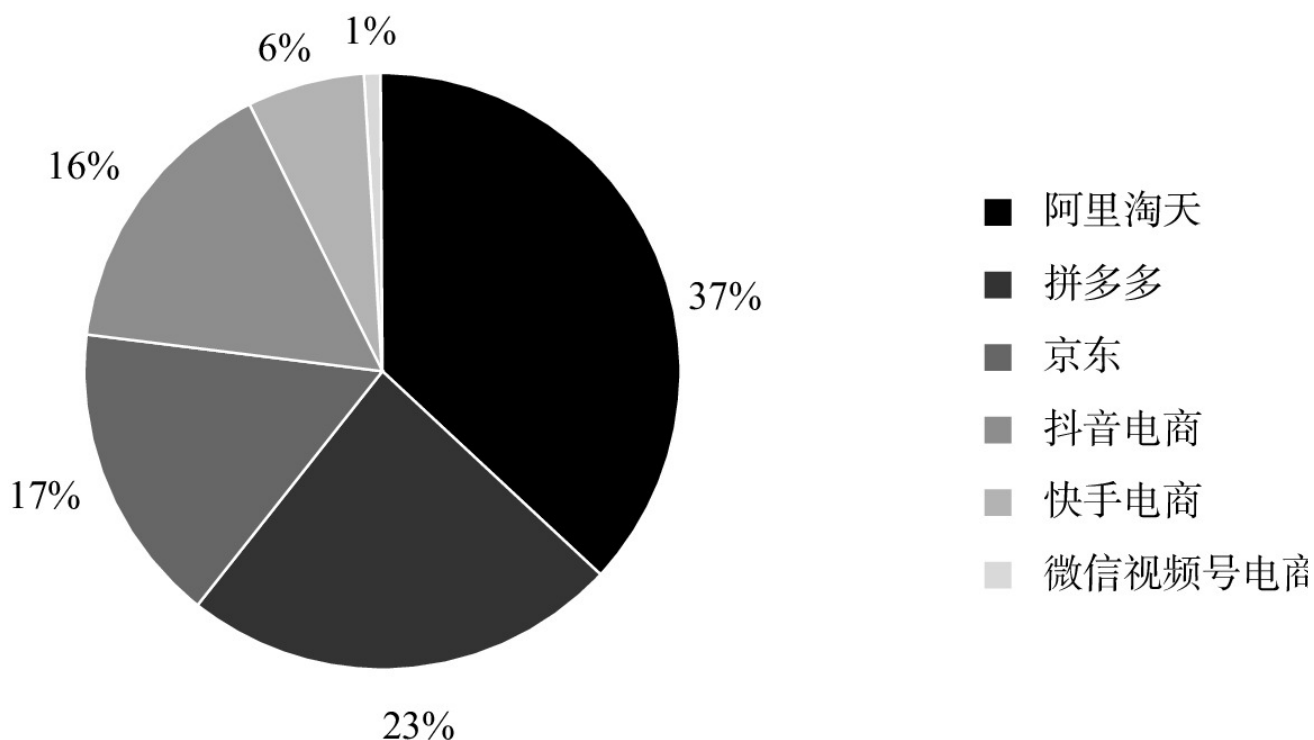


图3-3 2024年中国电商业GMV市场份额注：此处份额与前文提到的2024年“6·18”期间的份额略有差异，主要因时间窗口不同；此处未统计美团、唯品会等公司的数据。

2. 京东的商业模式解读

京东集团四大板块中各业务的商业模式各不相同，这里我们主要讲一下零售业务的商业模式。京东的零售业务可以分为自营电商模式和平台电商模式：自营电商模式，即1P模式，商家是京东的供应商，或者京东自己制造产品（京造模式）；平台电商模式，即3P模式，商家是京东平台上的零售商。这两种模式的商业模式也不一样：自营电商是一元商业模式，京东自采（或自产）自销，赚差价；平台电商是二元商业模式，京东通过平台GMV完成货币化（见表3-6）。

表3-6 自营电商和平台电商对比

	自营电商 (1P)	平台电商 (3P)
商业模式	一元	二元
核心指标	GMV、收入、成本	GMV、货币化率、成本
经营特点	需要稳定的供应、高效的周转、低毛利换销量	为供需两端提供价值主张、打造可持续发展的生态
品类	适合标品，如 3C 数码产品，但 SKU 相对较少	标品和非标品都适合，商家规模大，能满足多样化需求

京东从创业的第一天就开始做自营模式，京东自营模式做得大、做得好，并占据了用户“快、好”的心智，其中有两个不能忽视的重要因素：**周转效率和履约费用率**。

京东在需求预测、供应链管理和物流等方面的持续优化，使其存货周转天数持续下降，在业内处于领先地位。根据京东财报数据，其存货周转天数从2018年的40天下降到2024年的31.5天（在自营商品SKU超过1 000万的水平上，京东的存货周转天数处于全球领先水平，见图3-4）。

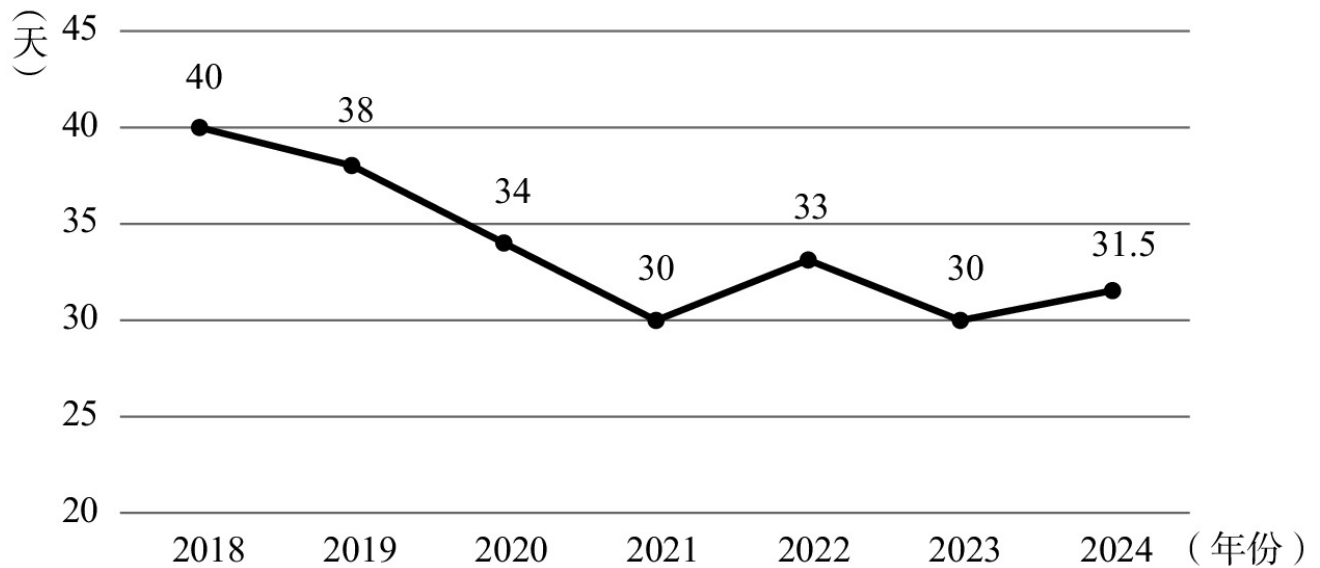


图3-4 2018—2024年京东存货周转天数

作为对比，同样是自营零售的优质代表企业，Walmart和Costco在管理的SKU数量远低于京东的情况下，周转效率都比京东要差，可见京东在存货周转上的竞争力和优势非常明显（见表3-7）。

表3-7 京东、沃尔玛、Costco存货周转天数和管理SKU数量对比

经营能力对比	京东	Walmart	Costco
存货周转天数	约 30 天	约 45 天	约 30 天
管理的 SKU 数量	1 000 万	5 万	5 000

京东的履约费用率也从2018年的7%以上下降到2024年的大约6%的水平。履约费用是自营模式中成本优化的重要指标，费用率的下降一定程度上得益于京东物流的发展。京东物流在仓网资源、仓储架构、仓内技术和管理等方面都处于业内领先水平，并且京东物流也面向外部其他客户开放服务，进一步释放规模效应的优势。

履约费用率=履约成本÷营业收入=单笔交易履约成本÷单笔交易客单价

对于平台电商模式，我们重点关注货币化率。此处需要注意，在看京东零售业务的GMV和货币化率的时候，只能看3P部分，因为只有3P的GMV才有货币化率的逻辑，1P部分本质上是一元商业模式，没有货币化率的逻辑。财报会议信息显示，京东3P业务的GMV目前占比大约为50%，预计长期会对标Amazon，提升至60%左右。

目前，京东3P业务的货币化率大约为8%，我们以2024年的数据为例，可以推算出京东3P业务的GMV和营收情况（如表3-8所示）。

表3-8 2024年京东3P业务的GMV和营收情况

总 GMV (万亿元)	3P 业务 GMV 占比 (%)	3P 业务 GMV (万亿元)	货币化率 (%)	营收 (亿元)
3.66	50	1.8	8	1 440

影响京东的货币化率除了抽佣和广告之外，还有物流履约费用，这些都是平台价值主张的体现。

从基本盘来看，京东的发展是比较稳健的，但是京东面临的挑战是巨大的。如前述数据所示，京东零售的GMV和营收增速都在下滑，GMV份额也被拼多多反超，而且还持续面临各大电商的挑战。前面我们提到，京东的核心竞争力是“快和好”，而“多和省”方面对比竞争对手略有不足。但在当前的市场发展和竞争环境中，京东“快”的优势在缩小，而“多和省”的短板却在放大。

(1) “快”的优势在缩小：以美团闪购为代表的即时零售业务的发展，将电商履约的时效性要求进一步提高，从天级别升级为小时甚至分钟级别（30分钟万物到

家），这导致京东会流失一部分时效性要求高的用户。京东快递的优势也在被蚕食。

（2）“省”的短板在放大：新冠疫情之后，用户收入和消费能力恢复缓慢，消费理念也发生了较大的转变，价格因素在消费决策中的影响力更大了。加之以拼多多为代表的平台对“价格力”的追求，用户跨平台比价的行为越来越普遍，京东在获取中低线市场的增量用户时劣势明显（京东在高线市场的用户增长已经接近天花板），而且存量用户也会向更低价的平台转移。

（3）“多”的短板在放大：以自营电商为基本盘的模式决定了京东的SKU数量与其他3P模式平台有较大差距，虽然京东也在扩大3P业务的规模，3P业务的GMV增长也很快，但目前京东平台上的商家仅100万量级，而淘宝上有超过1 500万商家。这方面的差距不仅体现在品类丰富性上，还体现在价格竞争力上。

目前，京东在积极地采取行动应对这些挑战。例如，加大即时零售业务的投入，通过达达和京东到家业务来对抗美团闪购的冲击；在2025年2月上线京东外卖业务，通过对美团核心业务的进攻来带动防守；通过百亿补贴和京喜等业务来提升“价格力”；引入更多的第三方商家，并在流量分发机制上促进自营商品与3P商品的平权，从考核收入改为考核GMV，减少商品在资源分配上的倾向性，给予3P商家和商品更多的曝光机会，进一步提升平台SKU量级。目前这些行动的成效还参差不齐，但作者认为电商的竞争是个无限游戏，各大平台会在竞争中不断迭代和成长，为用户和商家带来更大的价值，进一步推动全行业的发展。

京东集团定位于“以供应链为基础的技术与服务企业”，“快和好”为其核心竞争力，京东应该持续巩固“快和好”的核心优势，提升壁垒，拉开与竞争对手的差距；同时，在补齐“多和省”短板的过程中，不能陷入竞争对手的游戏规则中，避免深陷泥潭。

选择了什么样的商业模式，本质上也就选择了对应的优势和不足，商业就是这样。

3.3.2 阿里巴巴：从B2B发展成庞大的阿里生态系统，追求成为一家活102年的好公司

1999年，曾担任英语教师的马云与其他来自不同背景的伙伴，一共18人，在浙江省杭州市西湖区湖畔花园的一间公寓里创立了阿里巴巴集团。集团的首个网站是英文全球批发贸易市场阿里巴巴国际站。同年，阿里巴巴集团推出专注于国内批发贸易的阿里巴巴中国交易市场（现称“1688.com”）。这18个创始人后来被称为“阿里巴巴十八罗汉”。

阿里巴巴从B2B业务起家，不断地探索和创新，在互联网的浪潮中逐浪而上，陆续推出了C2C电商平台淘宝、第三方网上支付平台支付宝、B2C电商平台淘宝商城（后改名为天猫）、云计算业务阿里云、物流业务菜鸟等，并通过不断地投资并购扩大业务范畴，使业务遍布全球，渗透到用户生活的方方面面。2007年11月阿里B2B业务在香港联交所挂牌上市，后于2012年6月退市。2014年9月，阿里巴巴集团于美国纽约证券交易所正式挂牌上市，融资金额超过250亿美元，成为当时全球史上规模最大的IPO。2019年11月，阿里巴巴集团在香港联交所主板二次上市。

2025年5月，阿里进行品牌焕新，将大文娱集团升级为“虎鲸文娱集团”。

如今阿里巴巴作为控股公司拥有六大业务集团——淘天集团、云智能集团、国际数字商务集团、菜鸟智能物流网络有限公司、本地服务集团、数字媒体和娱乐集团（大文娱）¹²，以及各种其他业务。我们可以看一下阿里庞大的商业版图（见图3-5）。

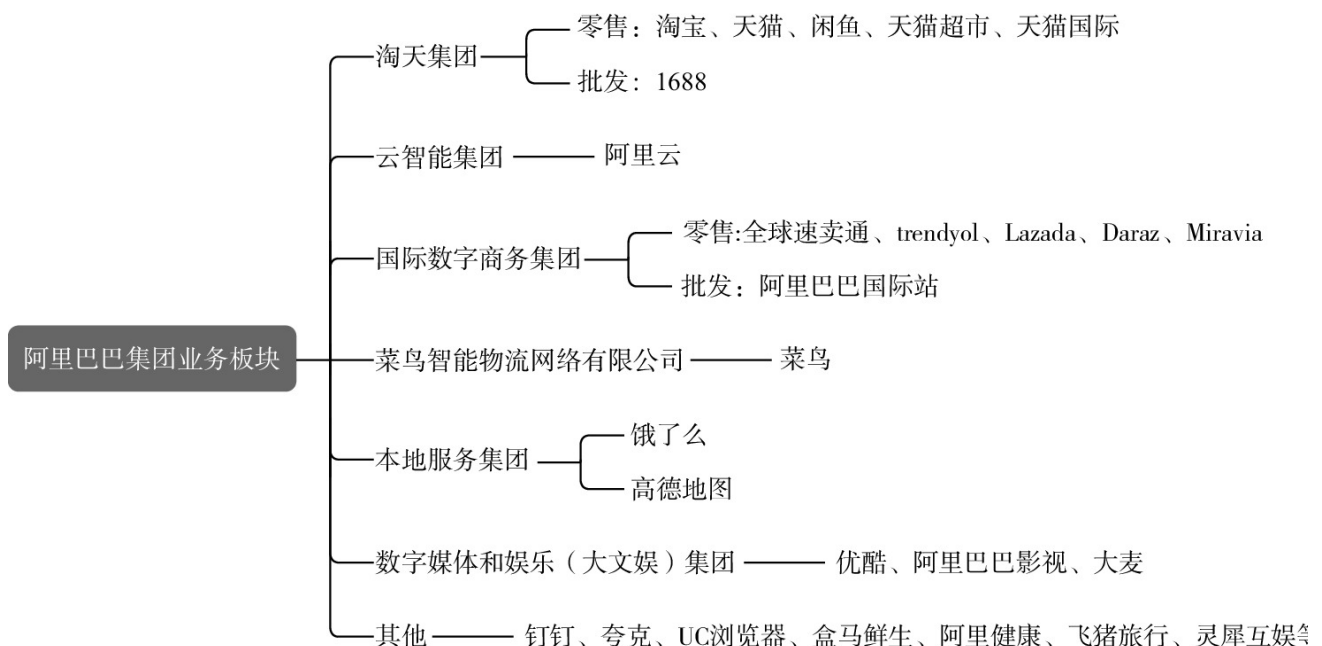


图3-5 截至2025年第一季度阿里巴巴集团六大业务板块

阿里巴巴的发展史堪比中国互联网行业的发展史，其中有一些关键事件值得我们回顾。

- 1999年阿里巴巴创建，同年推出阿里巴巴国际站和1688.com。
- 2003年上线淘宝网，与eBay展开C2C电商平台之争（淘宝网2005年在C2C市场的份额达到57%，超过eBay；eBay于2019年退出中国市场），首次推出担保交易服务。
- 2004年推出支付宝，作为淘宝网的信用中介，负责担保交易业务，进行网络购物资金第三方托管，解决了电子商务中交易信任的问题，对推动中国电子商务和移动支付的发展具有划时代的意义。
- 2008年推出面向品牌与零售商的第三方商业平台**淘宝商城**（2012年更名**天猫**），发力B2C平台电商业务。
- 2009年成立阿里云，并举办**第一届“双11”全球狂欢节**。同年支付宝推出手机客户端，开启**移动支付时代**。
- 2010年推出**全球零售交易市场之一全球速卖通**，同年上线手机**淘宝App**。
- 2013年与商业伙伴共同创立**菜鸟网络**。
- 2014年完成对数字地图公司**高德**的投资；**淘宝旅行**成为独立业务——**阿里旅行**（现称**飞猪**）；在**纽约证券交易所**正式挂牌上市。
- 2015年推出**钉钉**，为现代企业和组织提供了**工作、分享和协同新方式**；成立**本地生活服务平台——口碑**。
- 2016年**盒马鲜生**第一家门店开业，阿里巴巴新零售业务中的**生鲜食品及日用品零售连锁品牌**首次亮相；取得**优酷**的控股权；取得**东南亚领先电商平台Lazada**的控股权。
- 2017年对**银泰百货**进行私有化，并增持**菜鸟网络**股权至51%，成为控股股东。
- 2018年收购**即时配送及本地生活服务平台饿了么**，并将**饿了么和口碑**整合成**阿里巴巴集团旗下的本地生活服务业务**。

● 2019年获得阿里影业控股权，在香港联交所主板二次上市（当年阿里巴巴集团整体上市的首选地是香港，但由于其采用的“同股不同权”机制与港交所当时的上市规则不符，阿里最终只能舍港赴美上市）；同年，集团CEO张勇接替马云担任阿里巴巴集团董事会主席一职。

● 2020年推出**淘宝特价版**，主打工厂直供和定制化的商品服务；控股中国领先的商超运营商高鑫零售，以推进其新零售战略。

● 2023年阿里云推出自主研发的生成式AI大语言模型**通义千问**，陆续接入阿里巴巴生态体系的商业应用中，并于年内发布通义千问2.0和多款全新行业模型；集团组织及管理团队调整，进行“1+6+N”组织改革。

阿里巴巴从创业之初推出首个让中国中小企业接触全球买家的网站以来，逐步发展成以“电商+云计算”为基石的、全球领先的电商互联网生态巨头之一。根据易观国际的数据，在截至2024年3月31日的过去12个月里（阿里巴巴的财年为前一年的4月1日到次年的3月31日），按GMV计算，阿里巴巴是世界上最大的零售商业企业。当然，阿里巴巴的发展之路也并非一帆风顺，外界的挑战持续不断，尤其在电商业务上，京东、拼多多、美团和抖音电商的崛起和持续发展，在用户侧、商家侧以及人才竞争上都给阿里巴巴带来了极大的挑战。

1. 阿里巴巴的财务数据（来自阿里巴巴集团财报公开数据）

虽然业务众多，但阿里巴巴目前的核心业务依然是**中国商业零售**，即淘天集团，包括淘宝、天猫、1688.com、闲鱼、直营业务（主要包含天猫超市、天猫国际）等。从收入结构中可以看到，淘天集团2023财年和2024财年的收入占比分别为47.57%和46.21%，在集团收入中占比最大。从利润看更是如此，2024财年淘天集团经调整的EBITA（息税前利润，earnings before interest, taxes and amortization）约1 948亿元，占比**118.06%**（国际数字商务集团、本地服务集团、数字媒体和娱乐集团都是亏损状态），EBITA margin（摊销前利润率）为44.79%，是集团当之无愧的现金奶牛。

阿里巴巴营收构成和利润构成分别如表3-9、表3-10和图3-6所示。

表3-9 2023—2024财年阿里巴巴营收构成

营收构成	2023 财年		2024 财年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
淘天集团	4 132	47.57	4 349	46.21
云智能集团	1 035	11.92	1 064	11.31
国际数字商务集团	705	8.12	1 026	10.90
菜鸟智能物流网络有限公司	775	8.92	990	10.52
本地服务集团	502	5.78	598	6.35
其他	1 971	22.69	1 923	20.43
数字媒体和娱乐集团	184	2.12	211	2.24
未分配	9	0.10	13	0.14
分部间抵销	-627	-7.22	-763	-8.11
汇总	8 686	100.00	9 411	100.00

表3-10 2023—2024财年阿里巴巴利润构成

经调整的 EBITA	2023 财年		2024 财年	
	利润 (百万元)	占比 (%)	利润 (百万元)	占比 (%)
淘天集团	189 140	127.87	194 827	118.06
云智能集团	4 101	2.77	6 121	3.71
国际数字商务集团	-4 944	-3.34	-8 035	-4.87
菜鸟智能物流网络有限公司	-391	-0.26	1 402	0.85
本地服务集团	-13 148	-8.89	-9 812	-5.95
其他	-9 388	-6.35	-9 160	-5.55
数字媒体和娱乐集团	-2 789	-1.89	-1 539	-0.93
未分配	-12 143	-8.21	-6 190	-3.75
分部间抵销	-2 527	-1.71	-2 586	-1.57
汇总	147 911	100.00	165 028	100.00

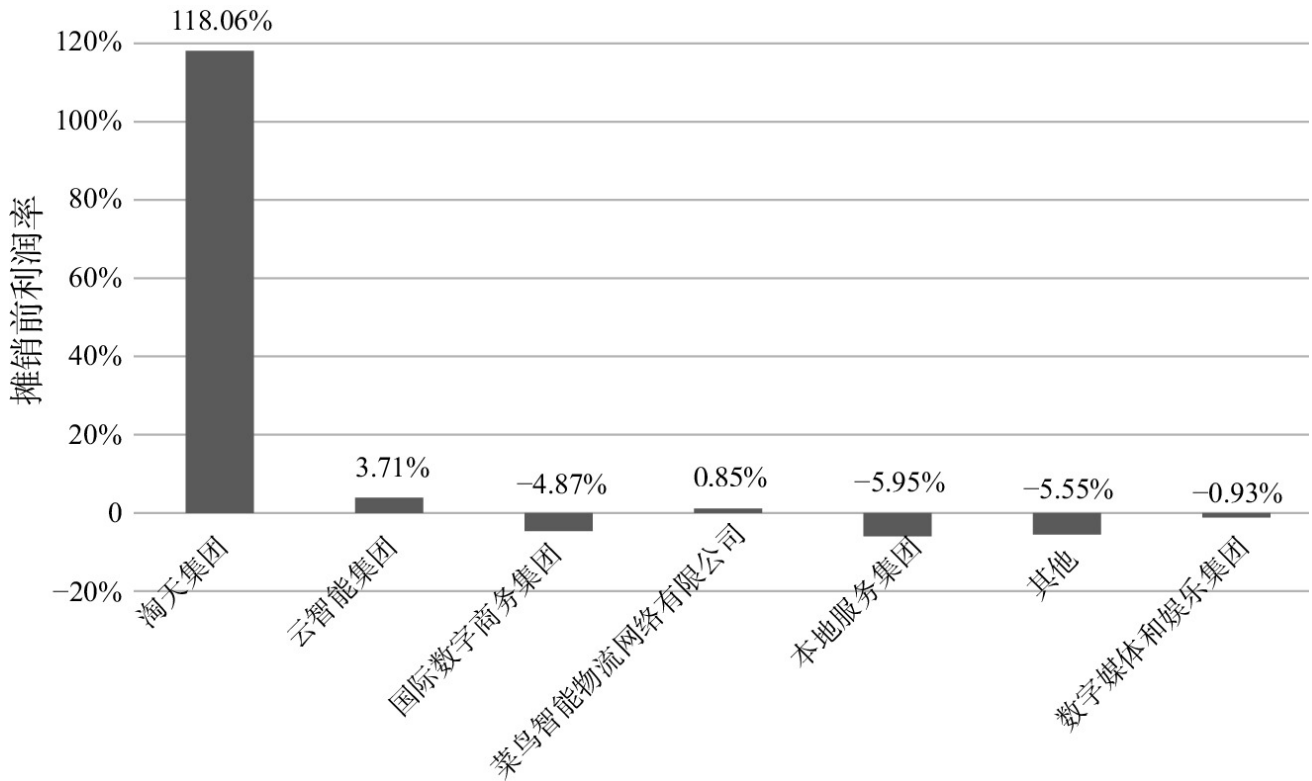


图3-6 2024财年阿里巴巴经调整的EBITA结构

如前文所述，按GMV来计算，阿里巴巴现在是全球最大的零售企业。看中国的电商数据，2024年淘天中国商业零售业务的年GMV接近8.2万亿元，市场份额约为37%，大幅领先竞争对手。但从2017年开始，拼多多和抖音电商先后崛起，老牌对手京东依然强势，还有微信视频号蓄势待发，淘天面临的挑战越来越大，最近几年增速明显放缓。

2. 阿里巴巴中国商业零售业务的商业模式解读

我们可以从淘天集团的营收结构中洞察其商业模式。

淘天集团的收入分为两部分——中国商业零售、中国商业批发，其中中国商业零售占比超过95%，为绝对的大头。我们聚焦于中国商业零售部分，其收入也分为两部分：客户管理、直销和其他。直销和其他收入占比约27%，主要是天猫超市、天猫全球和其他直销业务；而客户管理收入占比约73%，为平台电商模式（见表3-11）。

表3-11 2023—2024财年淘天集团营收构成

营收构成	2023 财年		2024 财年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
淘天集团	4 132	100.00	4 349	100.00
中国商业零售	3 953	95.68	4 144	95.29
客户管理	2 915	73.74	3 040	73.36
直销及其他	1 038	26.26	1 104	26.64
中国商业批发	179	4.32	205	4.71

这样我们就可以看出来，中国商业零售的商业模式是一元和二元的组合，以二元为主（看利润更是如此），这点与京东恰好相反，京东目前还是以一元商业模式为主。

阿里巴巴电商业中的非直销部分是典型的二元商业模式——不卖货、不存货、不送货，只是做电商交易平台，将用户的购物需求传递给商家，完成交易，并通过广告和佣金来完成货币化。淘天集团2024财年的货币化率接近4%，其中广告的货币化率约为2.8%，佣金的货币化率约为1.1%。过去几年这部分的货币化率趋势如图3-7所示。

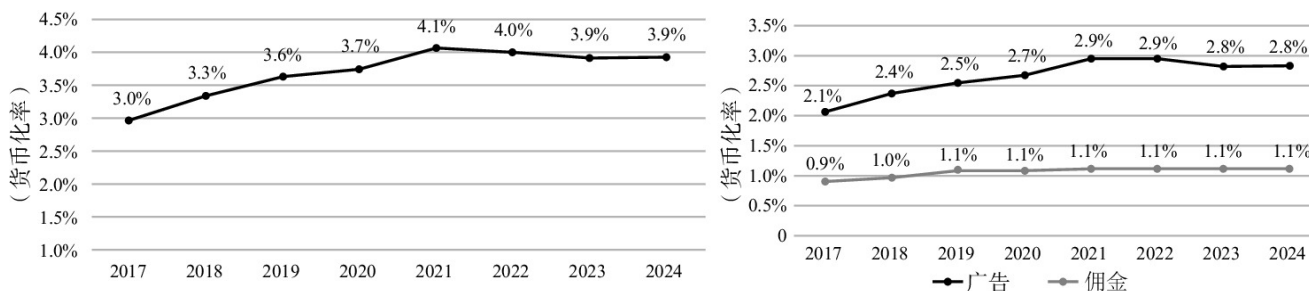


图3-7 2017—2024财年淘天集团的货币化率趋势图

这样在中国商业零售的GMV（2023年约7.7万亿元）中，大约有1 100亿元是直销GMV；接近7.6万亿元是平台电商GMV，即3P业务GMV，这部分靠着4%的货币化率贡献了约3 000亿元的客户管理收入。

在以3P业务模式为核心的支撑下，淘天集团的商家和商品数远超京东。在用户心智中，“万能的淘宝”便是对其SKU丰富度最好的诠释。毕竟淘宝的商品数过亿，遥遥领先于京东和拼多多。

回看淘天集团业务的发展，我们会洞察到三个核心问题：第一，淘宝的战略在追求商品品质和服务与追求商品便宜和种类繁多两条路线之间有些摇摆。从最开始的“万能的淘宝”，到2015年前后为了应对京东的竞争，开始主打消费升级，全力扶持天猫，清退小商家。包括“双11”战略，一定程度上也是在挤压小商家的生存空间和流量份额。过度追求GMV目标也让商品价格一路走高。这些打法都给拼多多留下了填补市场空白的战略机会。所以，拼多多这样的主打价格便宜的用户价值的产品出现后就迅速赢得了市场。第二，天猫和淘宝的关系不清晰。理论上，如果要与京东竞争，也许更好的思路是天猫独立出来，然后在商家、商品、供应链、服务等各个环节都做品质化升级，然后和京东去竞争相对高端的消费群体和品类；淘宝仍然保持对中低消费人群的吸引力。但实际上阿里巴巴并没有这么做。第三，物流等服务环节、平台模式能否真正做好值得思考和讨论。和京东自主把控物流不同，淘宝的物流走的是第三方平台模式。极致的竞争导致物流公司的利润不断被压缩，这样背后的服务始终得不到保障。在电商市场膨胀阶段，这样的思路没什么问题，因为有太多的增量用户带动规模增长。然而，到了存量用户竞争阶段，服务体验有竞争力，能带来更多的复购才是真正的增长动力。这个时候，淘宝在服务体验维度的竞争力不足就体现出一定的劣势。

虽然在竞争对手的全面冲击和监管挑战下，淘天集团的业务面临内忧外患——GMV和营收增速放缓，市场份额被拼多多和抖音电商抢夺，但在竞争策略上，淘天集团还是保持着一定的“多和好”优势：得益于淘宝C2C模式吸引了大量个人及中小商户入驻，天猫B2C模式又吸引了大量品牌商家入驻，淘宝的商品和品牌覆盖面之广形成了很高的壁垒，用户长尾小众的商品需求都可以在淘宝满足；而天猫丰富的品牌资源和正品保障，也满足了用户对好商品的诉求。

阿里巴巴集团以“让天下没有难做的生意”为使命，自创立之日起，不断自我进化，给社会的方方面面都带来了积极正向的影响，推动了很多行业的发展。

正如阿里巴巴官方提到的，其二十多年的发展史也是不断突破的创新史。阿里巴巴集团过去二十多年取得的成就已经足够伟大，但是想成为横跨三个世纪的企业，必然会经历高光和低谷，面临挑战、周期和传承问题。唯有不忘初心，坚持做正确的事，方能基业长青。

3.3.3 拼多多：C2M模式创新，全网最低价心智占领者，从创立到上市仅用三年

“省”字当头，2015年成立的拼多多席卷了整个电商行业。便宜、下沉市场、性价比高、低价、“砍一刀”、百亿补贴、“山寨货”、仅退款，拼多多身上有太多的标签，这些标签能让我们快速了解并记住拼多多，而拼多多的确也牢牢占据了低价的用户心智，并以此为基础一路狂奔，在国内电商市场上，拿到了超过20%的市场份额。

黄峥2015年4月在上海创立拼多多，以农产品零售平台起家，深耕农业，开创了以“拼”为特色的电商新模式，逐步发展成全品类综合电商平台。拼多多将“多实惠”和“多乐趣”融合起来，正如其官网介绍的，拼多多是“Costco+迪士尼”模式：拼多多新电商是一个将网络虚拟空间和现实世界紧密融合在一起的多维空间，一个由分布式智能代理网络驱动的“Costco”和“迪士尼”的结合体。互联网解决的已经不只是效率问题，拼多多在提供超高性价比的同时，将乐趣融入每个购物环节。

拼多多最早被用户熟知的是微信生态中“砍一刀”的社交裂变活动，即“社交拼团”模式。凭借着微信社交流量的红利，拼多多走出了一条差异化发展的道路——从2015年9月拼多多公众号上线运营，到2018年7月在美国纳斯达克正式挂牌上市，仅仅用了三年时间。但拼多多能吃到微信的红利，在微信生态内“病毒式”裂变扩张，目前来看是无法复制的。微信对分享裂变的管控非常严，尤其是对偏诱导性质的分享裂变的管控更严，当年拼多多能如此肆无忌惮，微信是开了绿灯的，因为腾讯投资了拼多多。截至2024年第二季度，腾讯仍是拼多多的第二大股东。

拼团是一种创新的C2M模式，将消费者的需求集合起来，直接传递到制造商，实现按需生产。C2M模式的核心是定制化生产，工厂能够按照用户的订单要求进行生产，这和传统模式相比，具有更低的成本、更低的库存风险和更高的效率。而C2M模式能够实现离不开两个关键因素：一是需求集合，即C端的需求可以通过互联网平台高效地收集并汇总；二是柔性供应链，即B端的供应链可以实现小规模定制化生产，即小单快反模式，这得益于国内强大的柔性供应链能力。

拼多多是这样介绍拼团模式的：拼多多首创了一种创新的“团购”模式。鼓励买家在社交网络上分享产品信息，并邀请他们的朋友、家人等社交联系人组成购物团队，享受“团队购买”选项下更具吸引力的价格。然后，完成分享行为的买家将得到团队购买选项提供的更具吸引力的购买价格。嵌入的社交元素有助于培养具有高度参与度的用户基础。

C2M模式目前在国内已经比较普遍了，尤其是很多抖品牌、快品牌、淘品牌都采用C2M模式来加快周转、降低库存成本和风险。作者几年前去广州的十三行和万佳服装批发市场实地考察过，大部分批发对单量的要求都很低，十件、五件都可以走批发，下单后工厂很快就能生产好并发货。很多商家卖的货，尤其是直播间卖的一些爆款单品，实际在卖的时候是没有库存的，只有样品，等订单确认后再按需生产。除此之外，美团在2022年上线的“拼好饭”，也是C2M模式。

拼多多不是黄峥的第一个创业项目，但拼多多创始团队的核心成员从黄峥第一次创业起就一直跟着他。

黄峥1980年出生于杭州，2002年从浙江大学竺可桢学院毕业后到美国威斯康星大学麦迪逊分校攻读计算机专业，2004年获得硕士学位后进入Google工作，2006年与李开复一起回中国拓展业务，建立Google中国办公室。黄峥还有一段传奇的经历——2006年段永平在和巴菲特共进午餐时带了一个人，这个人就是黄峥。

2007年黄峥离开Google开始创业，先后创办了欧酷网、乐其电商和寻梦，而拼多多是黄峥团队在做寻梦过程中的第四次创业。

● 2007年，欧酷网成立，是对标京东的消费电子B2C平台，主要通过销售步步高教育产品和OPPO蓝光播放器起家，三年内GMV达到6 000万元，最终黄峥将欧酷网卖给了兰亭集势。

● 2010年，乐其电商成立，是一家电商代运营公司，主营业务为天猫和京东店铺代运营，其中包括店铺运营、选品、仓储物流及售后服务，并涉足跨境电商业务。在三年内团队规模扩张至100多人，年销售额过亿元。这段经历锻炼了黄峥团队的供应链能力和电商运营能力。

● 2013年，寻梦成立，聚焦东南亚下沉市场的游戏业务，做游戏研发和发行。

● 2015年，在做游戏的过程中，黄峥团队将游戏和电商结合起来，孵化了拼好货——以销定采的自营农货电商。拼好货通过拼购模式打破农产品销售在时间与空间上极度分散的情况，通过拼单购买，整合供应链资源，实现了C2M模式。用户可以在拼好货内发起拼团或参团，再通过微信社群把拼团、参团信息扩散出去。当一个订单凑够成团人数时，拼好货团队再去采购农产品发货。这种“社交+拼团”的病毒式打法迅速打开了市场，在不到一个月的时间内平台就达到了日均1万单的水平。同年9月，拼好货内部孵化电商平台拼多多，继而拼好货和拼多多合并，形成了今天的拼多多。

拼多多的发展大概可以分为三大阶段。

● **第一阶段：创立和狂奔（2015—2018年）。**2015年，拼多多起步并设计好底层逻辑，打造低价心智，在微信生态内病毒式传播，迅速积累大量用户并爆炸式增长。2017年9月拼多多用户超过1亿，月GMV超10亿元；2018年7月拼多多在纳斯达克上市。

● **第二阶段：品牌上行（2019—2020年）。**拼多多采用“农村包围城市”的打法，从深耕下沉市场转向拓展一二线城市。2019年6月，拼多多启动百亿补贴，开始补贴iPhone，并与中高端品牌合作。自此拼多多开始渗透更广泛的消费者，逐步摘掉山寨、假货、商品质量差的标签，实现了品牌升级。2020年7月，黄峥辞任拼多多CEO，公司原CTO陈磊出任CEO。

● **第三阶段：多元发展（2021年至今）。**拼多多为了寻求新的增长点，在国内电商平台的业务持续增长时，推出了多多买菜和TEMU，进入社区团购和海外电商市场，在更大的舞台上继续奔跑。2021年3月，黄峥将拼多多的董事长职位也交给了CEO陈磊。

抛开全球化和新业务不谈，拼多多在国内发展电商平台业务的过程中，有两个关键词很重要：低价策略和百亿补贴。

低价策略是拼多多的基石。拼多多秉着“好货不贵”的理念为消费者提供各类产品，自创立以来便主打低价心智，将“省”做到了极致。低价策略能践行和落地，有历史的机遇，也有平台的设计：

历史的机遇：当年阿里巴巴在做消费升级的时候，拼多多抓住了下沉市场崛起和淘天系供应链转移的机遇，完成了低价策略在供需两端的双边“冷启”，尤其是吸引了大量白牌商家入驻拼多多。

平台的设计：拼团模式可以形成“以订单量撬动商家”的低价飞轮效应，同时平台围绕低价形成了一系列策略和机制，给低价商品流量，激励商家在价格上互相竞争，保障足够多的低价品供给（见图3-8）。

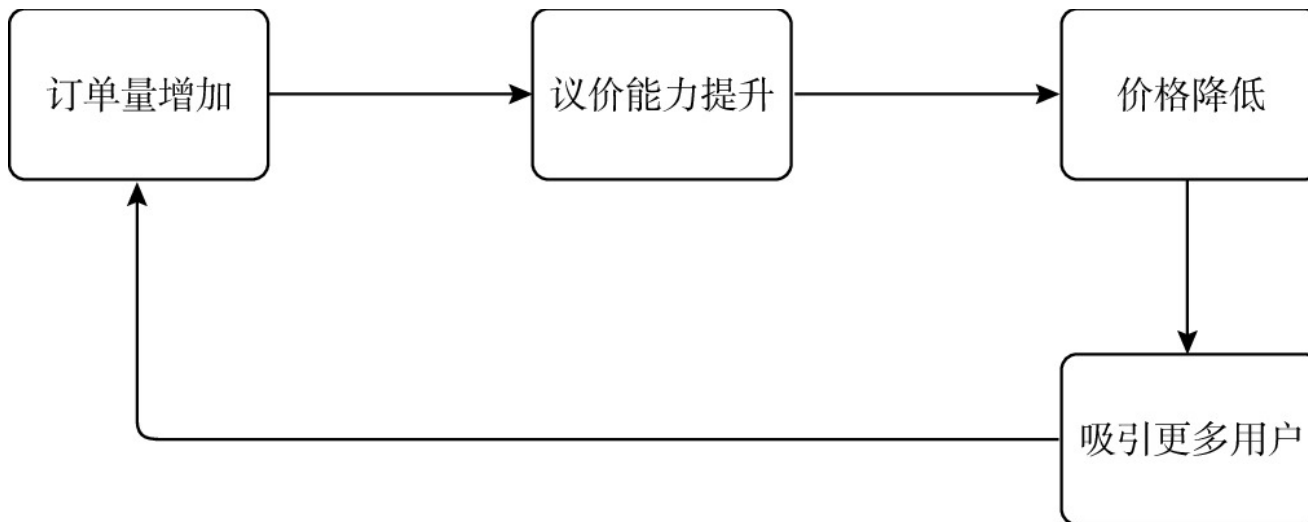


图3-8 基于低价策略的平台设计

百亿补贴是拼多多品牌上行的发动机。拼多多创造了百亿补贴，使之成为国内电商行业最成功的营销模式之一，到目前为止基本上国内每家做交易相关业务的互联网公司都做过百亿补贴，包括但不限于京东、阿里巴巴淘天系、抖音电商、美团、饿了么等，而做得最好、最成功的无疑就是拼多多。百亿补贴是拼多多低价策略的延续，是其在品牌上行过程中低价策略的极致打法。2019年6月，拼多多首次补贴iPhone，上线百亿补贴专区，正式拉开百亿补贴的帷幕。拼多多真金白银投入的百亿补贴在用户端实现了人群破圈，从下沉市场逐步渗透到一二线城市，同时改变了用户认为拼多多上都是山寨货的刻板印象；在商家端吸引大量中高端供应商入驻，大家发现在拼多多上品牌商品、贵的商品也有销量，这大大改善了平台的商品结构，使之能服务于不同消费能力的用户。随着参与百亿补贴的订单越来越多，就又回到了前述的拼团逻辑：拼多多拿着百亿补贴大量的订单，反向要求供应商给更低的价格，以及从平台出钱补贴变成平台和商家共同补贴，甚至平台仅仅提供流量补贴。根据行业研究调研数据，2023年百亿补贴的GMV在拼多多总GMV中的占比超过20%，有超过6亿的用户购买过百亿补贴的商品。百亿补贴的成功，驱动了品牌上行，拼多多实现了从白牌到品牌的转型，形成了新的增长曲线。

在业务经营上，拼多多还是一个把直营模式的经营思路和平台模式的运营思路结合得非常好的典范。对比来说，都是平台模式，但是淘宝的运营更多关注的是通过选品和流量运营提高成交额、转化率，很少涉及对供应商进行压价这样的操作，包括后来的仅退款，也是借鉴了Amazon的服务思路。但是拼多多在这一点上结合得特别好，凭借平台巨大的流量优势，进行直营公司的经营操作，创造了巨大的用户价值和商业成功。

我们再来看拼多多的“Costco+迪士尼”模式：Costco聚集了对于“低价实惠”有相同需求的不同消费能力的用户，通过更高效的方式连接供需两侧，通过C2M模

式、低价策略、品牌上行等方法源源不断地带来各种品类的高性价比商品；迪士尼则打造了有趣的购物平台和环境，通过社交和游戏化的方式，让购物体验充满娱乐性，给消费者带来更多乐趣。

1. 拼多多的财务数据（来自拼多多财报公开数据）

拼多多的财务报表披露得非常简单，虽然公司除了国内电商平台业务之外，还有多多买菜和海外电商平台的TEMU业务，但财报中没有分业务的经营数据。但是从公司整体的营收和利润就足以看到拼多多陡峭的增长曲线：营收从2016年的5.05亿元左右一路高歌猛进到2024年的3 938.36亿元左右；利润方面，从2021年扭亏为盈之后持续盈利，2024年营业利润达到1 084.23亿元，远超京东（见图3-9）。

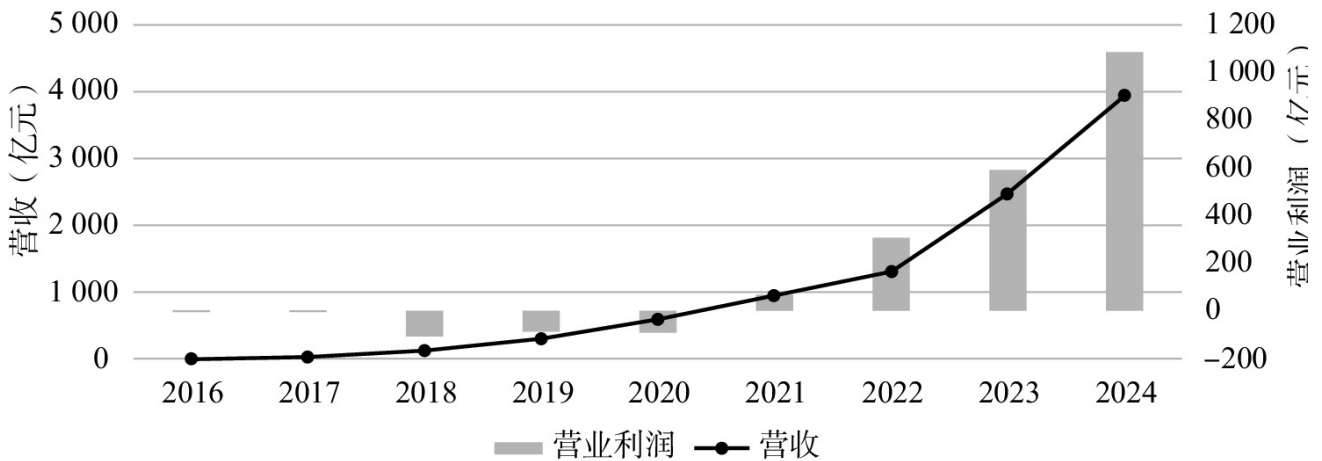


图3-9 2016—2024年拼多多的营收和营业利润趋势

2016—2024年拼多多的营收和利润详情如表3-12所示。

表3-12 2016—2024年拼多多的营收和利润

项目	2016年	2017年	2018年	2019年	2020年	2021年	2022年	2023年	2024年
营收 (亿元)	5.05	17.44	131.20	301.42	594.92	939.50	1 305.58	2 476.39	3 938.36
毛利 (亿元)	-0.73	10.21	102.15	238.03	402.13	622.32	990.95	1 559.16	2 399.36
销售费用 (亿元)	1.69	13.45	134.42	271.74	411.95	448.02	543.44	821.89	1 113.01
管理费用 (亿元)	0.15	1.33	64.57	12.97	15.07	15.41	39.65	40.76	75.53
销售和管理 费率 (%)	36.39	84.75	151.67	94.46	71.78	49.33	44.66	34.83	30.18
研发费用 (亿元)	0.29	1.29	11.16	38.70	68.92	89.93	103.85	109.52	126.59
研发费率 (%)	5.83	7.40	8.51	12.84	11.58	9.57	7.95	4.42	3.21
营业利润 (亿元)	-2.86	-5.86	-108.00	-85.38	-93.81	68.96	304.01	586.99	1 084.23

营收和利润增长的背后是拼多多用户数量的快速增长。2024年，拼多多年活跃买家数量超过了9.6亿，2024年第四季度日均订单量达到1.2亿单。因为拼多多主打低价，所以拼多多的客单价远低于阿里淘天和京东，虽然百亿补贴驱动品牌和人群消费能力上行，也带动了拼多多客单价的提升，但与阿里淘天和京东还有一定的差距。我们根据GMV和买家数量预估了各家的客单价，可以看出拼多多2024年活跃买家人均年GMV大约在5 400元，对比2018年的1 100元有大幅提升，但对比阿里淘天和京东依然有一定的差距（见表3-13）。

表3-13 阿里淘天、京东和拼多多的活跃卖家人均年GMV 单位：元

活跃买家人均年 GMV	2018年	2022年	2023年	2024年
阿里淘天	8 700	8 200	8 200	9 000
京东	5 500	5 800	6 000	6 300
拼多多	1 100	3 300	4 300	5 400

在如此高速的增长下，拼多多的市场份额也越来越大，如前文所提到的，2024年国内电商市场的份额按照GMV口径来计算，拼多多已经超过了23%，超过京东。市场格局的变化就更明显了：时间回到2017年，阿里淘天和京东在国内电商市场的份额加起来超过90%，拼多多处于其他行列；而到了2024年，拼多多的份额已经反超京东，而阿里淘天的份额从70%被蚕食至37%。在阿里淘天和京东激烈竞争时，拼多多通过另外一种方式完成了弯道超车。

2017年与2024年国内电商市场GMV份额对比如图3-10所示。

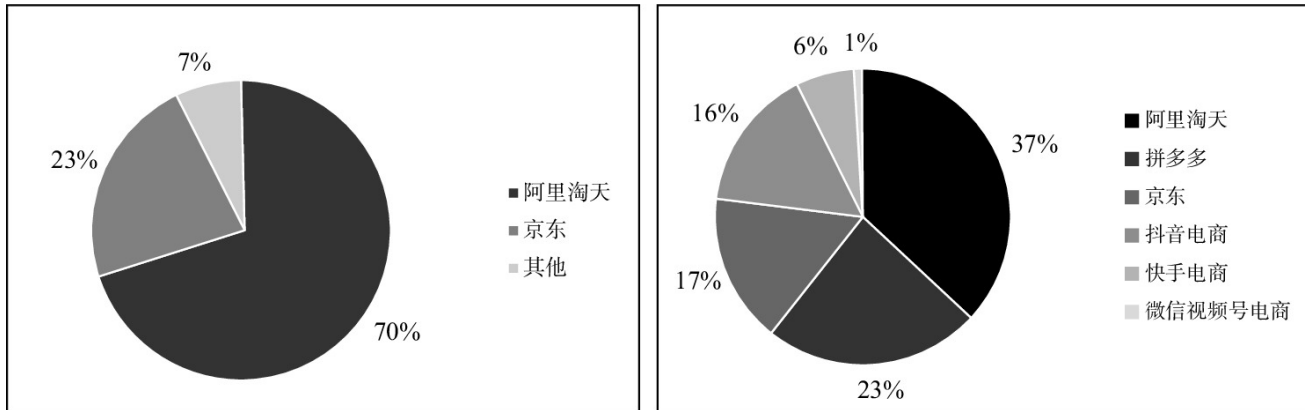


图3-10 2017年（左）和2024年（右）阿里淘天、京东和拼多多的GMV市场份额

从财报中我们还可以看出拼多多超高的经营效率。最近几年拼多多的经营费用率持续下降，其中销售和管理费用率、研发费用率可以参考图3-11。随着业务规模的扩大，经营成本会释放出规模效应。虽然拼多多这些年在不断地扩张以及探索新的业务，但成本却依然能控制得很好，经营费用率持续下降，这与它整体的高效率是分不开的。从拼多多超高的人力资源效能中便能管中窥豹。截至2024年末，拼多多只有约2.3万名员工，远低于竞争对手。阿里巴巴的员工数接近20万；而京东因为有物流业务，员工数超过60万（仓储配送相关人员超过40万）。

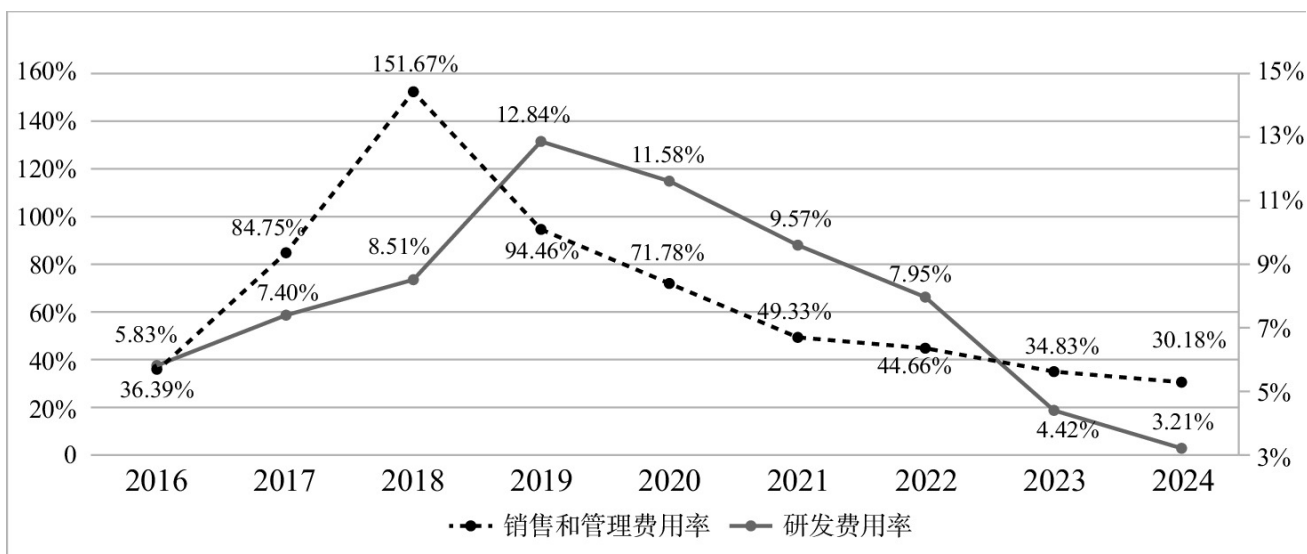


图3-11 2016—2024年拼多多经营费用率趋势

2. 拼多多国内电商业务的商业模式解读

拼多多是典型的二元商业模式，这与阿里淘天和京东的3P业务一致。拼多多也是通过广告和佣金来完成货币化，在财报中也是按照这两部分披露营收构成：2024年在线营销服务和其他的收入达到1 979亿元，占比50.25%；交易服务的收入达到1 959亿元，占比49.75%（见表3-14）。

表3-14 2021—2024年拼多多营业收入

拼多多营收构成	2021年		2022年		2023年		2024年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
在线营销服务和其他	798	84.98	1 029	78.85	1 535	62.00	1 979	50.25
交易服务	141	15.02	276	21.15	941	38.00	1 959	49.75

拼多多也在逐步提升广告和佣金费用。随着品牌上行的成功，商家结构中品牌商家占比越来越大，平台广告产品和机制不断优化，货币化率持续提升，这些都驱动着平台收入和利润增长。拼多多主站的货币化率2023年已经提升到了约5.2%，其中广告的货币化率约为4.2%，佣金的货币化率约为1.0%。过去几年拼多多主站及其中广告和佣金的货币化率趋势如图3-12所示。

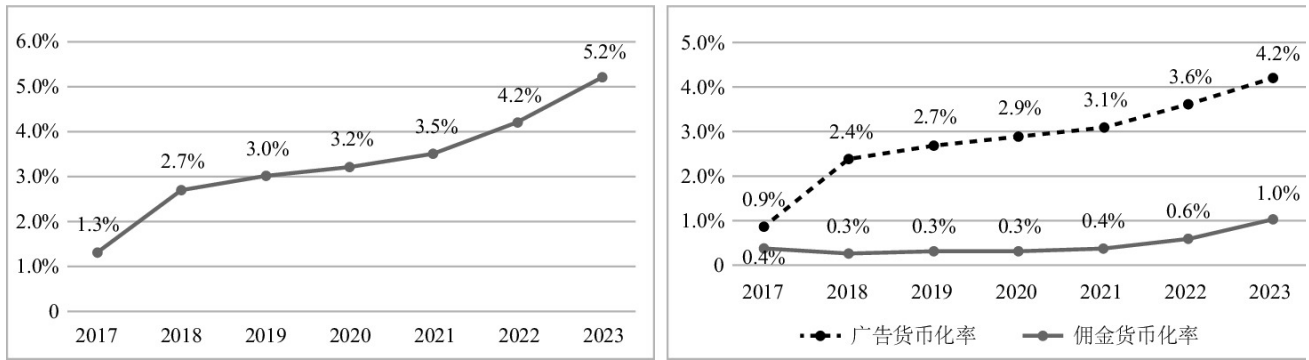


图3-12 2017—2023年拼多多主站的货币化率（左）与主站中广告和佣金的货币化率（右）注：根据数据推算，2024年拼多多主站的货币化率也在5%以上。

同时，拼多多在国内电商市场已经形成了一定的竞争壁垒，这也是其加速提升货币化率的基础。如今拼多多依然保持着远高于行业平均水平的营收增速，而且因为拼多多牢牢占据了“省”的用户心智，所以当前全球经济环境也有利于拼多多的发展。虽然拼多多推出的拼团模式、百亿补贴和低价策略等都已经成为竞争对手争相学习的对象，百亿补贴更是成为很多行业 and 企业的“标配”，但拼多多在供需两端建立起来的优势，以及围绕价格着力打造的整个系统，让拼多多仍然在竞争中处于领先的位置。

在阿里巴巴和京东已经在国内电商市场形成竞争“定局”的时候，拼多多通过一种新的方式席卷了整个行业，让市场再一次相信奇迹。这家坚守“本分”价值观的公司，像一台不知疲倦的机器一样，依然在高效运转，继续创造着奇迹。

3.3.4 滴滴出行：聚焦主业，千帆历尽，勿忘初心，不负“美好出行”的使命

“嘀嘀打车”2014年5月更名为“滴滴打车”，2015年9月升级为“滴滴出行”。

2012年，程维从阿里巴巴离职，开始创业。2012年7月，北京小桔科技有限公司成立；同年9月，滴滴打车 App 上线，提供在线出租车服务，借助移动互联网的LBS（基于位置的服务，location-based service）能力连接司机和乘客，在线上完成司乘双方的匹配。这是一种现代化的出行方式，乘客可以通过手机App叫车，而不需要在路边拦车或者打电话叫车，滴滴出行迈出了颠覆传统打车方式的第一步。滴滴出行的模式极大地提升了出行效率，改善了出行，也打开了一个巨大的市场——网约车市场。仅看国内，网约车的市场规模就有几千亿元。

滴滴出行从出租车业务出发，在网约车大战中战胜快的和优步中国，成为国内网约车的龙头企业，当年一度占据国内网约车（含出租车）市场超过90%的份额。

● **滴滴出行vs. 快的：**2014年网约车大战在滴滴出行和快的之间展开，它们背后站着的分别是腾讯和阿里巴巴。双方疯狂对司乘进行补贴，抢占市场，4个月烧掉了20亿元，直到2015年初双方握手言和。2015年2月，滴滴出行和快的签订合并协议，快的与滴滴出行合并，并且并入滴滴出行，即滴滴出行收购了快的。

● **滴滴vs. Uber：**2014年3月Uber正式进入中国，投入到国内网约车市场的竞争中。在滴滴出行和快的合并之后，新一轮网约车大战的对垒双方就变成了滴滴出行和Uber，双方烧掉了上百亿元，直到2016年8月双方达成战略协议，滴滴出行收购了优步中国。根据协议，滴滴出行和Uber全球相互持股，成为对方的少数股权股东。

随着业务不断扩张，如今滴滴出行已经发展成全球卓越的移动出行科技平台，业务遍布全球，在亚太、拉美、非洲提供出租车、网约车、顺风车、公交、共享单车、共享电单车、代驾、汽车后市场服务、配送、货运及物流、金融等多元化的业务。这些业务大概分为三部分：中国出行业务、国际业务以及其他业务。

● **中国出行业务：**主要包括国内共享出行业务和部分相关的汽车服务业务，即网约车、出租车、代驾、顺风车等。滴滴出行还探索过外卖和社区团购（橙心优选）业务，不过目前均已关停。在网约车服务中，根据用户差异化的需求，滴滴出行提供不同的车型，例如针对高端用户的专车和豪华车服务，针对主流快车用户提供的优享、特快、特惠快车等服务，针对低价用户提供的拼车、花小猪打车等服务。

- **国际业务：**主要包括国际网约车业务、外卖送餐业务和金融服务业务。
- **其他业务：**主要包括共享两轮车（单车）、汽车服务、同城货运、自动驾驶以及金融服务等业务。

滴滴出行自创立以来，一直很受资本市场欢迎，获得过二十多轮融资，但是数据安全事件也给滴滴出行带来了巨大的影响。

- 2021年6月11日，滴滴出行正式向美国证券交易委员会递交了IPO招股书；2021年6月30日，滴滴出行正式在纽交所低调挂牌上市。

- 2021年7月2日，为防范国家数据安全风险、维护国家安全、保障公共利益，依据《国家安全法》《网络安全法》，网络安全审查办公室按照《网络安全审查办法》，对滴滴出行实施网络安全审查，为配合网络安全审查工作，防范风险扩大，审查期间滴滴出行停止新用户注册；2021年7月4日，因存在严重违法违规收集使用个人信息问题，滴滴出行App在应用商店被下架；2021年7月16日，国家互联网信息办公室等七部门进驻滴滴出行科技有限公司开展网络安全审查；2022年7月21日，滴滴出行因违反《网络安全法》《数据安全法》《个人信息保护法》被处以80.26亿元的罚款。

- 2022年6月13日，滴滴出行从纽交所退市。

- 2023年1月16日，经网络安全审查办公室同意，恢复滴滴出行新用户注册，滴滴出行结束了长达五百多天的整改。

经过这一轮风波，滴滴出行减少了对新业务的投入，逐步下线了外卖、社区团购等业务，并且在2023年第三季度将智能汽车业务出售给了小鹏汽车。虽然以曹操出行、T3出行为代表的竞争对手们在滴滴出行被整改期间加速抢占市场，但滴滴出行靠着供需两端的巨大优势，在下架整改期间依然保持着强大的竞争力：在供给侧，滴滴出行存量司机多、运力足，用户在滴滴出行更容易叫到车；在需求侧，用户心智被滴滴出行占领，已经习惯了用滴滴出行，同时滴滴出行还有微信和支付宝两个巨大的流量入口，保障了用户可以继续使用滴滴出行。

随着滴滴出行整改结束以及App恢复上架，公司主营业务逐步回到正轨，目前国内网约车市场形成了一超多强的竞争格局。截至2024年第二季度，滴滴出行在国内网约车（含出租车）市场的份额超过70%，日均订单量超过3 000万单。

1. 滴滴出行的财务数据（来自滴滴出行财报公开数据）

滴滴出行2024年的总营收为2 067.99亿元，同比连续两年增长，逐步消除了2021年和2022年监管和疫情的影响。但整体的盈利情况依然不乐观，2023年还是处于亏损

状态，2024年终于盈利了，2024年全年的营业利润为19.42亿元，阶段性扭亏为盈（见表3-15）。

表3-15 2019—2024年滴滴出行营收和利润

项目	2019年	2020年	2021年	2022年	2023年	2024年
总营收（亿元）	1 547.86	1 417.36	1 738.27	1 407.92	1 923.80	2 067.99
年同比增长（%）	—	-8.43	22.64	-19.00	36.64	7.50
营收（亿元）	1 499.14	1 359.70	1 641.43	1 317.94	1 828.76	1 967.84
毛利（亿元）	151.21	159.12	169.64	249.93	294.46	378.06
毛利率（%）	9.77	11.23	9.76	17.75	15.31	18.28

续表

项目	2019年	2020年	2021年	2022年	2023年	2024年
销售费用（亿元）	74.95	111.36	169.61	97.56	104.33	115.14
管理费用（亿元）	62.15	75.51	287.15	170.05	84.12	88.63
研发费用（亿元）	53.47	63.17	94.15	95.36	89.24	77.54
其他费用（亿元）	40.78	46.96	75.25	64.44	45.68	77.32
营业利润（亿元）	-80.13	-137.88	-456.52	-177.02	-28.38	19.42

虽然滴滴出行整体的盈利情况并不乐观，但中国出行业务在2023年和2024年都实现了盈利，是国际业务和其他业务亏损导致集团整体亏损。从营收构成上可以看到，中国出行业务2024年和2023年的营收分别为1 857.41亿元和1 750.34亿元，占滴滴出行总营收的89.82%和90.98%（见表3-16）；从利润构成上看，中国出行业务2024年经调整EBITA为91.84亿元，再扣除相关股份薪酬和无形资产摊销，经营利润也是正的。而且虽然2022年中国出行业务经调整EBITA为-13.98亿元，但这里包含了前述提到的80.26亿元罚款（见表3-17）。所以从业务经营情况来看，中国出行业务过去三年（2022—2024年）都是盈利的。国内网约车市场如今竞争格局基本稳定，滴滴出行也有绝对的份额优势，业务规模效应明显，边际成本低，所以盈利也在情理之中。

表3-16 2022—2024年滴滴出行营收构成

营收构成	2022 年		2023 年		2024 年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
中国出行业务	1 259.31	89.44	1 750.34	90.98	1 857.41	89.82
国际业务	58.63	4.16	78.42	4.08	110.43	5.34

续表

营收构成	2022 年		2023 年		2024 年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
其他业务	89.98	6.39	95.04	4.94	100.15	4.84
总计	1 407.92	100.00	1 923.80	100.00	2 067.99	100.00

表3-17 2022—2024年滴滴出行利润构成 单位：亿元

经调整 EBITA 情况	2022 年	2023 年	2024 年
中国出行业务	-13.98	53.23	91.84
国际业务	-40.14	-23.02	-18.46
其他业务	-72.12	-51.06	-30.11
经调整 EBITA 总计	-126.24	-20.85	43.27
基于股份的薪酬费用	-34.47	-25.90	-22.53
无形资产摊销	-16.31	-10.03	-1.39
出售或视为出售子公司 所产生的收益或损失	0.00	28.40	0.07
营业利润	-177.02	-28.38	19.42

2. 滴滴出行中国出行业务商业模式解读

滴滴出行中国出行业务主要就是网约车、出租车和代驾业务，是典型的二元商业模式。滴滴出行构建了一个在线平台，连接乘客和司机，完成司乘两端的匹配和交

易。每完成一笔订单交易，就会有对应的GMV [在滴滴出行的业务中被称为GTV（总交易额，cross transaction value），本部分后续统一用GMV]，滴滴出行通过佣金完成货币化。

从经营逻辑来看，滴滴出行中国出行业务的营收应该是GMV货币化之后的数字，即收入=GMV×货币化率，而GMV=订单量×客单价，这和前述电商平台3P模式的逻辑一致。但在滴滴出行的财务口径中，有个情况值得读者注意：在网约车业务中，滴滴出行会将GMV扣除一部分费用、税费和C端补贴之后（约等于有效GMV，后续为了方便说明业务逻辑直接用GMV来表述）直接计为业务收入，而将司机的收入计为成本，即滴滴出行先收了所有的车费，再给司机发工资；在出租车业务和代驾业务中，滴滴出行是按照经营逻辑来处理的，将GMV货币化之后的部分计为收入。即在滴滴出行的财务口径中：

滴滴出行收入=网约车业务GMV+（出租车业务GMV+代驾业务GMV）×货币化率

所以，中国出行业务一年超千亿元的营收实际上有很大一部分是GMV，读者们不要被这个数据误导了。我们以2024年的数据为例，中国出行业务GMV为3 014.36亿元，收入为1 857.41亿元（见表3-18），如果直接计算的话，会得到货币化率约等于62%，这是个错误的数字。滴滴出行实际上综合的货币化率大约为18%，不同的车型佣金不一样：快车约为20%、顺风车约为10%、专车和豪华车约为25%~30%、优选出租车约为15%~25%。

表3-18 2022—2024年滴滴出行中国出行业务运营指标

运营指标	2022年	2023年	2024年
交易量（亿元）	77.33	108.09	123.92
日均交易量（亿次）	0.21	0.30	0.34
GMV（亿元）	1 861.74	2 707.21	3 014.36
平均客单价（元）	24	25	24
收入（亿元）	1 259.31	1 750.34	1 857.41

滴滴出行为什么要这么计算呢？按照财报的说明，滴滴出行会根据自己是提供服务的委托人还是代理人，以不同的方式确认收入。

根据相关规定，网约车服务平台必须获得牌照，并对相关服务承担全部责任，持牌平台应确保从事网约车服务的驾驶员和车辆符合规定。滴滴出行作

为一家在线叫车服务平台，由于控制向乘客提供的服务，因此认为自己是网约车服务的委托人。

● 对于网约车，滴滴出行是为消费者提供移动服务的主体，是提供服务的委托人。^④滴滴出行从消费者为网约车服务支付的金额中获得毛收入，其收入等于GMV减去通行费、税费以及消费者激励，而将司机的收入和奖励计入经营成本。所以，网约车业务的收入基本等于GMV。

● 对于出租车业务和代驾业务，滴滴出行是提供服务的代理人，向需要在线出租车和司机服务的用户，以及提供在线出租车和司机服务的司机或合作伙伴提供便利。滴滴出行的收入主要来自司机或合作伙伴支付的佣金。所以，出租车业务和代驾业务的收入基本等于GMV×货币化率。

我们以一笔订单为例，直观地对比一下不同财务口径的差异，假设这笔订单的车费为10元（见表3-19）。

表3-19 不同财务口径下滴滴出行订单的差异 单位：元

款项	按 GMV 逻辑计收	按货币化率逻辑计收
车费	10	10
通行费和税费等	1	1
乘客补贴	0.9	0.9
消费者支付	10.1	10.1
GMV	11	11
司机激励	1	1
司机收入	7.5	7.5

续表

款项	按 GMV 逻辑计收	按货币化率逻辑计收
收入录入	9.1 (GMV-通行费和税费等 -乘客补贴)	1.5 (GMV-通行费和税费等- 司机激励-司机收入)
成本-司机收入	7.5	0
成本-司机激励	1	0
成本-乘客补贴	0	0.9
利润	0.6	0.6

注：本表中的数字只是为了举例说明财务口径和实际的业务情况，不代表真实的收入和成本。

从业务的经营逻辑来看，假设这笔订单的抽佣率（ tr ）为25%，则实际的情况可以拆解如下：

平台收入=有效GMV×货币化率=(GMV-通行费和税费等)×货币化率=(11-1)×25%=2.5(元)

平台成本=司机激励+乘客补贴=1+0.9=1.9(元)

平台利润=收入-成本=2.5-1.9=0.6元

司机收入=有效GMV×(1- tr)+司机激励=10×(1-25%)+1=8.5(元)

支付GMV=车费+通行费和税费等-乘客补贴=GMV-乘客补贴=11-0.9=10.1(元)

从滴滴出行的案例中，我们要学会识别什么是真正的GMV和收入。

抛开滴滴出行的财务口径，从业务模式看，中国出行业务是典型的二元商业模式，滴滴出行通过佣金的方式从平台GMV中完成货币化。如前文所言，滴滴出行真实的货币化率大约为18%，实际上并不低，这与市场竞争以及调度效率都有关。目前，滴滴出行在国内市场的份额优势明显，而且在在线叫车场景中拥有派单权和部分调价权，所以滴滴出行对平台上的订单调度效率很高。

为什么滴滴出行的货币化率这么高，但中国出行业务的利润并不高呢？核心有两点：一是客单价低，2024年，每单平均客单价大概是24元，货币化之后的收入约为4.4元；二是补贴大，滴滴出行还在发放大量的司机激励和乘客补贴，2024年每单平均经调整EBITA仅为约0.7元。因此，中国出行业务虽然已经扭亏为盈，但利润

并不高。作者认为鉴于目前国内在线出行市场的竞争格局基本稳定，滴滴出行的市占率优势明显，而且供给侧偏饱和，滴滴出行可以适当降低补贴，以提升经营利润，尤其是用户侧补贴可以适当降低，因为滴滴出行的运力优势明显。司机端补贴短时间可能没办法降低，核心原因有两点：一是竞争对手也在争夺运力；二是运力过剩，司机收入不高。但补贴力度也取决于竞争对手的策略，例如如果高德持续补贴，滴滴出行还是需要跟进，以稳住市场份额。

“滴滴一下，美好出行。”滴滴出行作为传统打车方式的颠覆者，为用户提供了更好的出行体验，也为司机和车主提供了灵活就业和增加收入的机会。同时，滴滴出行作为一家科技企业，也在积极推动智慧交通创新，立志于应对全球交通、环保和就业方面的挑战。如果时间回到2012年，我们可以在北京看到程维带着团队，在火车站一个一个地找出租车司机安装车主端App的场景。如今千帆历尽，滴滴出行仍持续聚焦移动出行业务，不忘企业初心。

3.3.5 美团：科技零售巨头，从千团大战开始，在千锤百炼中锻造一身本领，尽显英雄本色

美团成立于2010年3月，最开始做的是团购业务，创始人是王兴。美团不是王兴的第一个创业项目，在此之前王兴最知名的两个项目是2005年创立的“校内网”和2007年创立的“饭否”。校内网（后来的人人网）是一个实名制的社交网络平台，创立之初主要面向在校学生，当年风靡一时，堪称一代人的回忆。由于没有足够的资金增加服务器和带宽，2006年校内网被千橡互动收购，王兴也加入了千橡互动，后于2007年4月卸任校内网CEO，并在同年5月创立了饭否。饭否是国内第一个提供微博服务的网站，可以让用户在自己的饭否页面发布消息和图片，并通过关注、私信、转评赞等方式进行互动，被称为“中国版Twitter”。但由于政策监管等原因，饭否在2009年直接关闭了服务器，微博鼻祖自此退出了历史舞台。

2010年，王兴创办美团，从团购入局，深耕互联网和本地生活领域，不断拓宽业务边界，逐步发展出外卖、到店、酒旅、交通票务、闪购、社区团购、餐饮供应链、网约车、共享单车、共享充电宝等业务。从千团大战到后来的外卖大战、酒旅大战、社区团购之争和共享单车大战等，美团一直都在“打仗”，并且一直都在“打胜仗”。美团每次都能在竞争厮杀中笑到最后（而且面对的是阿里巴巴这种巨头级别对手），凸显了其团队超强的战斗力，其“地推铁军”更是行业内教科书级别的典范。

美团的发展史，就是过去十几年中国互联网本地生活赛道的“战争史”，大致可以分为四大阶段。

● **第一阶段：“千团大战”时期（2010—2012年）。**2010年3月，美团上线第一个团购项目，正式进入团购赛道。据统计，2011年第三季度，市场上的团购网站已经超过5 000家，团购几乎成为互联网公司的标配业务。以拉手网、窝窝团、满座网、高朋网和美团为代表的平台为了抢夺市场份额相互竞争，掀起了中国互联网历史上最著名的烧钱大战之一：千团大战。2011年7月，阿里巴巴领投美团5 000万美元，这笔融资是美团在千团大战之中最终存活下来的关键因素之一。2011年第四季度，团购商家陆续退出市场。2012年第三季度，美团成为团购赛道中市场份额排名第一的公司，赢得了千团大战的最终胜利，至此美团的基石业务在“战争”中完成了冷启和规模化。

● **第二阶段：团购延伸，美团外卖登场，美团和大众点评合并（2013—2015年）。**2013年起，美团在团购的基础上将业务扩展到电影和酒旅，并在2013年11月正式上线美团外卖。其实在美团外卖上线之前，有一家从高校做起来的外卖平台已经落地十几个城市，它就是饿了么，其老板是张旭豪。2013年7月，美团找到张旭

豪想收购饿了么，未果，于是4个月后，美团便上线了外卖业务，拉开了外卖大战的帷幕。结局很明显，美团在认知、管理和战略上全面领先，疯狂开城，对饿了么在全国城市中发动了战争（美团比竞争对手做了更多的“地图作业”，更早理解了中国的城市结构，并认识到外卖市场的纵深），最终在外卖大战中大获全胜。2015年，美团战略合并大众点评，确立了行业龙头地位。虽然饿了么在阿里巴巴的支持下，持续在外卖市场与美团抗衡，但美团2016年便拿下了外卖市场最大的份额，到2024年第二季度市场份额更是达到约70%，优势显著。美团外卖作为后来者，在王兴团队的带领下，完成了一次经典的后发先至式超越。

● **第三阶段：夯实到家业务，加码到店和酒旅业务，完成上市（2016—2019年）。**2016年1月，美团获得由腾讯、DST资本和挚信资本领投的33亿美元的融资，同期阿里巴巴出售其在美团的股权，外卖赛道竞争格局已定。美团继续强化在外卖领域的领先优势，加强网络技术、人力与餐饮供应链等方面的基础能力建设，进一步加固“护城河”。同时，美团加码酒旅业务，与携程展开酒旅大战，并布局生鲜、跑腿、打车、买菜等业务，参与共享打车大战，收购摩拜单车，在本地生活赛道全面布局。2018年9月20日，美团在港交所正式挂牌上市。各品类生活服务在美团平台铺开，借助高频餐饮业务对细分低频业务的流量输送，即用高频打低频，美团形成了低获客成本、长用户生命周期的显著比较优势。

● **第四阶段：以“零售+科技”为核心战略，实现万物皆可到家（2020年至今）。**从2020年开始，美团优选、美团买菜、美团闪购陆续上线，开展即时配送、次日达等实物零售服务，成功将LBS和餐饮配送的能力升级为万物到家。2021年，在美团闪购数字零售大会上，时任美团高级副总裁王莆中提出在线零售市场正在从“Everything Store”（万货商店）向“Everything Now”（万物到家）转变。截至2025年4月，美团闪购日均订单量已经突破1 800万单（含非餐饮品类）。2021年9月，美团将战略升级为“零售+科技”。2022年初，抖音推出生活服务板块，正式进军本地生活赛道，当年便实现约770亿元的GMV，2023年这一数字就迅猛增长到约2 000亿元，接近美团的三分之一。抖音挑起了本地生活赛道的又一场战争，美团再一次迎来了新晋对手的挑战，目前，这场流量与铁军之争还未见终局。

从美团的发展历程中可以看到，互联网本地生活赛道总是硝烟弥漫，美团也从不缺少对手。在千锤百炼中一路走来美团，如今已成为科技零售巨头。随着业务和组织的不断迭代，截至2025年第一季度，美团的业务大致可划分为“核心本地商业”和“新业务”两部分（见图3-13）。

● **核心本地商业：**包括餐饮外卖、到店、酒店和旅游、闪购、民宿和交通票务业务，是美团营收和利润的主要来源。餐饮外卖及闪购业务主要帮助用户通过美团购买商家准备的餐饮和食杂，并提供即时配送服务；到店、酒旅相关业务主要帮助用户购买商家提供的众多到店类的本地消费服务或预订酒店、景点门票及交通票。

● **新业务**：其中既包括美团优选、美团买菜、网约车、共享单车、共享充电宝等业务，满足用户在不同生活场景和消费场景下的不同需求；也包括B端服务，如餐饮供应链快驴、餐饮管理系统以及金融服务等，为美团平台的商家提供对应的服务。

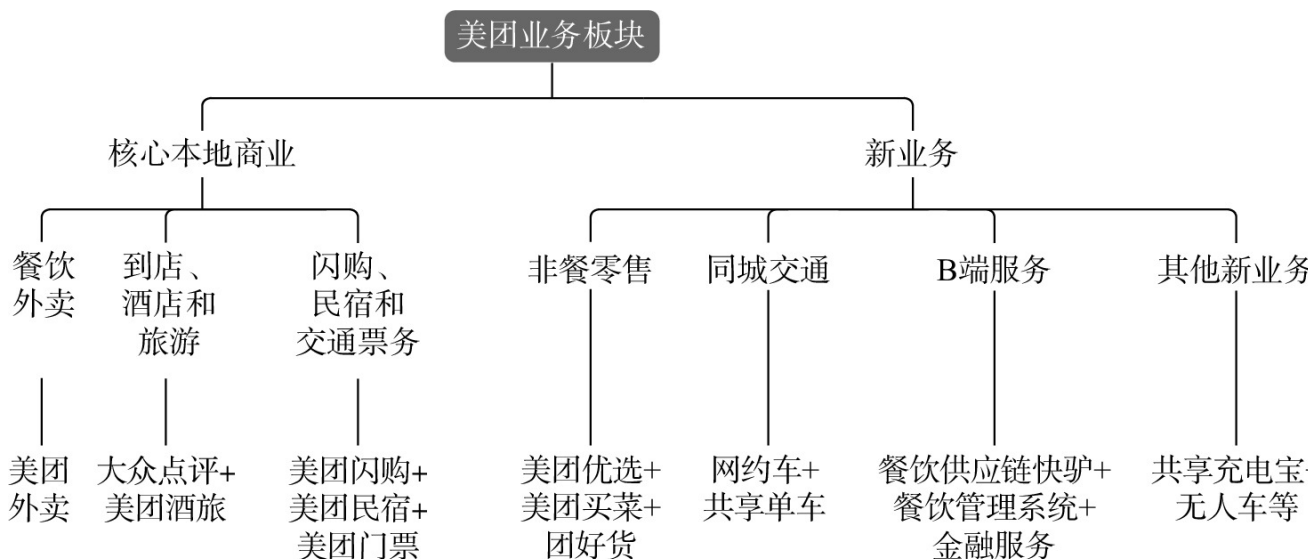


图3-13 美团主营业务构成（截至2025年Q1）

互联网本地生活赛道也是个万亿级赛道，经过十几年的发展，已经覆盖社会生活的方方面面。每天在大街小巷中穿梭的外卖小哥，便是这万亿级赛道最好的具象化呈现。美团作为行业的龙头企业，持续驱动行业发展，完善基础设施建设，在供给、需求和履约三方面共同发力，践行“帮大家吃得更好，生活更好”的使命，不断地创造更大的社会价值。

1. 美团的财务数据（来自美团财报公开数据）

美团在过去几年营收稳定增长，2023年和2024年都保持了20%以上的年增长率：2023年营收2 767.45亿元，年同比增长25.82%；2024年营收3 375.92亿元，年同比增长21.99%。利润方面，美团2019年就已经盈利了，但2021年却由盈转亏，全年经营溢利亏损高达231.27亿元，主要是因为这一年美团加大了对新业务的投入（新业务年亏损383.94亿元），同时新冠疫情反复，美团外卖对商家的佣金优惠和骑手成本双双上涨，影响了美团的利润。2023年美团再度扭亏为盈，并在2024年继续扩大盈利规模，全年经营溢利达368.45亿元。

美团营收和利润详情如表3-20所示。

表3-20 2019—2024年美团营收和利润

项目	2019年	2020年	2021年	2022年	2023年	2024年
营收（亿元）	975.29	1 147.95	1 791.28	2 199.55	2 767.45	3 375.92
毛利（亿元）	323.20	340.50	424.74	617.53	971.91	1 297.85
毛利率（%）	33.14	29.66	23.71	28.08	35.12	38.44
销售费用（亿元）	188.19	208.83	406.83	397.45	586.17	639.75
销售费率（%）	19.30	18.19	22.71	18.07	21.18	18.95
行政费用（亿元）	43.39	55.94	86.13	97.72	93.72	107.29
研发费用（亿元）	84.46	108.93	166.76	207.40	212.01	210.54
研发费率（%）	8.66	9.49	9.31	9.43	7.66	6.24
经营溢利（亿元）	26.80	43.30	-231.27	-58.20	134.15	368.45

如前文所言，目前美团的业务划分为“核心本地商业”和“新业务”两部分，所以财报最近两年也是按照这两部分来披露的，而在2022年中期财报之前，美团是按照“餐饮外卖”“到店、酒店及旅游”以及“新业务及其他”来披露财报数据的。即后来将“餐饮外卖”“到店、酒店及旅游”和“新业务及其他”中所包含的“闪购、民宿及交通票务”合并为“核心本地商业”，其他的为“新业务”。

从营收结构来看，美团的经营有如下两个特点：

● 第一，核心本地商业是美团成熟的基石业务，是美团创造营收的大头和全部利润来源，同时补贴新业务。2024年，核心本地商业营收达2 502.47亿元，营收占比74.13%，利润达524.15亿元，利润占比142.26%，即其他业务都在亏损。表3-21、表3-22分别展示了美团的营收构成和利润构成。

表3-21 2022—2024年美团营收构成

营收构成	2022年		2023年		2024年	
	营收（亿元）	占比（%）	营收（亿元）	占比（%）	营收（亿元）	占比（%）
核心本地商业	1 607.59	73.09	2 069.07	74.76	2 502.47	74.13
新业务	591.96	26.91	698.38	25.24	873.44	25.87
总计	2 199.55	100.00	2 767.45	100.00	3 375.91	100.00

表3-22 2022—2024年美团利润构成

利润构成	2022 年		2023 年		2024 年	
	利润 (亿元)	占比 (%)	利润 (亿元)	占比 (%)	利润 (亿元)	占比 (%)
核心本地商业	295.03	—	386.99	288.45	524.15	142.26
新业务	-283.79		-201.66	-150.31	-72.73	-19.74
未分配	-69.44		-51.17	-38.14	-82.97	-22.52
总计	-58.20	100.00	134.16	100.00	368.45	100.00

● 第二，餐饮外卖是美团创造营收的大头，到店和酒旅是美团创造利润的大头，即：到家业务创造的营收多，但经营利润率低，不怎么赚钱；到店业务经营利润率高，是美团的盈利支柱。这也能部分说明为什么抖音在进军本地生活市场的时候，会先做到店，而不是到家。毕竟到店才是真正赚钱的业务，而到家履约成本高、经营难度大，还不怎么赚钱。以美团2021年的数据为例，餐饮外卖年营收963.12亿元，营收占比53.77%，年利润61.75亿元，经营利润率仅为6.41%；到店和酒旅年营收325.30亿元，营收占比18.16%，年利润140.93亿元，经营利润率高达43.32%（见表3-23）。

表3-23 2021年美团餐饮外卖、到店和酒旅经营情况

项目	营收 (亿元)	利润 (亿元)	经营利润率 (%)
餐饮外卖	963.12	61.75	6.41
到店和酒旅	325.30	140.93	43.32

虽然正在面临抖音的差异化竞争，但美团整体还是表现出了不错的增长韧性。餐饮外卖业务已经构建了很高的竞争壁垒，规模优势明显，而且在订单规模和骑手网络建设的加持下，美团在履约端的壁垒也很高，新入局者难以轻松逾越。因此外卖业务虽然经营利润率偏低，但仍然是美团的基础业务。美团可以通过高频的外卖业务，带动其他低频业务增长，同时，通过外卖建立起来的商家和履约体系也为到店、闪购、优选等业务提供了最有力的支撑。到家是美团的基本盘，而且这个基本盘现在是很稳的。从市场份额来看，外卖市场已经从当年的群雄逐鹿，变成了如今的美团遥遥领先。如前文所述，2024年第二季度美团外卖的市场份额已经达到约70%（随着2025年京东和阿里巴巴发起新一轮外卖和闪购大战，外卖市场的竞争会变得更加激烈）。

国内外卖市场2015年与2024年份额对比如图3-14所示。

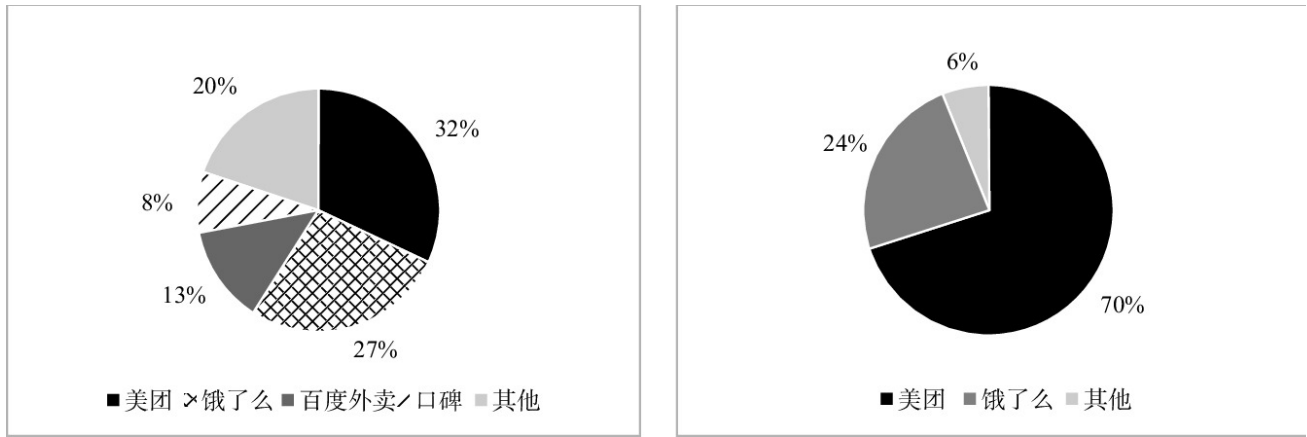


图3-14 2015年（左）与2024年（右）国内外卖市场份额对比

2. 美团外卖业务的商业模式解读

美团大部分业务都是二元商业模式，我们以餐饮外卖为例展开介绍。

美团外卖也是二元商业模式，每一笔外卖订单都会产生一笔GMV，美团通过服务费来完成货币化，只是在外卖这种业务模式中，服务费除了广告和佣金之外，还有配送费。正如美团财报中介绍的，美团外卖的收入来源主要包括：提供给商家及消费者的配送服务；向商家及第三方代理合作伙伴收取的技术服务费的佣金；提供给商家的各种形式的在线营销服务。我们以2021年的数据为例看一下美团外卖的收入构成：外卖年收入963.12亿元，其中配送服务收入542.04亿元，占比56.28%；佣金和在线营销服务一共收入399.82亿元，占比41.51%（见表3-24）。

表3-24 2021年美团外卖收入结构

项目	收入（亿元）	占比（%）
餐饮外卖配送服务	542.04	56.28
佣金	285.47	29.64
在线营销服务	114.35	11.87

续表

项目	收入（亿元）	占比（%）
其他服务及销售（含利息收入）	21.26	2.21
总计	963.12	100.00

从货币化率来看，2021年5月美团外卖调整了抽佣规则，从收取“一刀切”的平台服务费，调整为更灵活的阶梯式收费，规定：商家佣金=技术服务费+履约服务费，其中履约服务费与距离、订单价格以及时段有关。在这一轮抽佣规则调整之前，美团外卖对商户端收取20%左右的费用，即货币化率大约为20%，剔除补贴后预计在14%左右；抽佣规则调整之后，货币化率大约为13%~15%，其中技术服务费约占5.8%（保底1.04元），其余为履约服务费和广告收入。

美团外卖的货币化率并不低，这与市场竞争格局有关。如前文所言，美团外卖目前的市场份额绝对领先。某种意义上说，外卖市场目前由美团主导，所以美团对商家有更高的议价权。

为什么货币化率这么高美团外卖还是不怎么赚钱呢？有两个核心因素：一是客单价低，每笔订单的客单价不到50元，货币化之后的收入也就不会很高；二是履约成本高，虽然收得多，但成本也高。2021年餐饮外卖配送相关成本为681.83亿元，而餐饮外卖配送服务的收入只有542.04亿元，甚至入不敷出。

美团外卖的履约目前有两种模式：美团配送（1P模式）和商家配送（3P模式），其中1P模式是主要履约模式，占比超过65%。1P模式即自营模式，由美团提供配送服务（美团的骑手也有专职和众包两种），收取配送服务收入并承担成本；3P模式即商家自配送模式，配送服务由服务商负责，服务商享有配送服务的收入并承担成本。对美团来说，1P模式比3P模式成本高，但稳定性、调度效率和用户体验更好。所以虽然3P模式可以摊薄一些履约成本，但美团依然以1P模式为主。

我们通过一组具体的数据，直观地感受一下美团外卖业务的成本结构：一笔外卖订单的客单价为50元，乘以14%的货币化率后收入7.00元，但配送成本就高达5.11元，再加上其他成本，最终这笔订单美团的经营利润仅为0.45元，经营利润率为6.43%（见表3-25）。在2020年第一季度的财报会上，王兴表示2025年外卖业务的目标是“每天1亿订单、每单利润1块钱”。现在再看这个目标，不是任重道远，而是基本做不到。

表3-25 美团外卖业务的成本结构

单笔外卖订单财务指标	收入
单笔客单价（元）	50
货币化率（%）	14
单笔收入（元）	7.00
配送收入（元）	4.02
广告和佣金收入（元）	2.98
单笔配送成本（元）	5.11
单笔其他成本（元）	1.44
单笔经营利润（元）	0.45
单笔经营利润率（%）	6.43

注：这组数据是作者根据相关经营指标估算而来，仅用于将经营模型更直观地呈现给读者。

当然，外卖业务具有规模成本优势，边际履约成本会下降，但要提升订单量和单笔利润，不能只通过优化成本和继续提高货币化率，更重要的是要提升外卖的渗透率和客单价，这样才能继续把市场盘子做大，实现王兴曾经喊出的目标。但是2025年京东和阿里巴巴再次掀起外卖大战，美团又迎来了另一个对手的挑战。

美团作为本地生活服务赛道的龙头企业，在竞争中练就了一身本领，构建了商家端、履约端的核心护城河，在用户端也占领了不错的用户心智。在“万物到家”理念驱动下，美团以“零售+科技”为核心战略，推动即时零售的发展。一路走来，美团虽然已经历经百战，但是依然在前进的路上战斗，依然在书写英雄般的故事。

3.3.6 贝壳：起源于链家，推动中国房地产业数字化升级，逐步发展为房产中介平台，是房产经纪领域唯一的平台型企业

贝壳起源于链家，从链家的垂直自营模式升级为现在的平台模式。链家成立于2001年，创始人为左晖，全称为北京链家房地产经纪有限公司，成立之初专注于北京的房产交易服务业务，也就是房产中介业务。贝壳于2018年上线，以共享真实房源信息与链家管理模式吸引第三方房产交易经纪公司（即第三方房产中介）入驻，打造房产中介平台。即，链家是一个房产中介平台，是贝壳自营的；而贝壳平台上有很多房产中介平台，链家和其他中介一样，只是贝壳平台上的一个中介。

链家和贝壳的发展，经历了四个关键阶段。

● **第一阶段：创业初期（2001—2008年）**。2001年链家在北京成立，开启了行业规范化之路。链家成立之后，逐步推出了“二手房商品房按揭业务”“二手房交易资金托管业务”“租售两全新型服务产品”以及“个人租房消费贷款业务”等产品和业务，率先提出“透明交易、签三方约、不吃差价”的新型交易模式，对房产交易进行规范化和标准化，提升了房产经纪（中介）行业的信誉与自身的品牌影响力，链家也逐步发展为北京最大的房产经纪品牌。

● **第二阶段：全面线上化时期（2008—2014年）**。2008年，链家在线上上线，并开始“楼盘字典”建设工作；2011年，率先承诺所有房源保证真实；2014年链家在线升级为房产综合服务平台，标志着线上化改造完成。楼盘字典对于房产行业的价值不言而喻，链家投入了大量的人力、物力搭建这套房产基础信息数据库，用社区、门牌号、户型图、配套设施等多维度信息定义一套房子，从根本上杜绝了信息不透明、房源不真实等长期存在的行业弊病。截至2023年12月31日，链家的楼盘字典覆盖了约2.77亿处房产，并且房产中介可以通过手机App、微信小程序等入口提供新的房产信息或修改信息，进而保证楼盘字典的更新。

● **第三阶段：全国化扩张时期（2015—2017年）**。链家通过并购上海德佑地产、杭州盛世管家等公司，以及与大连好望角、广州满堂红等公司进行战略合作，逐步覆盖了全国20多个城市，完成全国化布局。

● **第四阶段：平台化时期（2018年至今）**。2018年，贝壳找房上线。链家通过整合已有的链家及全国化的产业链资源，将链家网升级为贝壳找房。2020年，贝壳推出被窝家装，2021年贝壳收购圣都家装，加快在家装行业的业务扩展，将自身定位为中国最大的房产交易和服务平台。2020年8月，贝壳在纽交所正式挂牌上市，成为中国居住服务平台第一股；2022年5月，贝壳正式在港交所挂牌上市，成为首家以“双重主要上市+介绍上市”形式返港的中概股。

房产交易是一项很复杂的工程，涉及很多环节和程序，例如向房管局提交购买协议、向税务机关纳税、在房管局进行房屋所有权转让和登记，以及向银行申请贷款、向担保公司申请担保服务等。整个交易流程正常情况下耗时好几周，相关人员需要在不同的地点间来来回回跑几十次。为了降低复杂度和交易成本，链家设计了一套全面的标准签字程序，并基于此开发了一套在线交易支持系统，即NTS，兼顾了交易安全和交易效率。通过NTS，买卖双方可以在线完成很多交易步骤，例如在线签署合同，在线将合同提交给房管局等。NTS还允许客户和代理商查看交易过程，监控各交易步骤，并提供在线反馈和输入功能。

除了全面的线上交易支持外，链家还在已经进入的城市开设了交易服务中心，与银行、担保公司、房地产估价师和政府机构合作，并派驻人员到现场，帮助买卖双方和中介无缝、轻松地转让产权和完成行政程序。链家的交易支持人员可以在交易服务中心帮助处理各种行政程序，进一步提升交易效率。综合线上NTS和线下交易服务中心的支持，链家使整个房产交易变得更加简单，为用户节省时间和成本，以此带来更好的用户体验。

在交易资金安全问题上，链家于2014年建立了在线支付平台理房通，提供住房交易中的数字支付处理和托管服务。理房通解决了房产交易中的信任问题——购房者关心的是待收到的产权在支付款项后是否能顺利转移，而卖房者在收到款项之前也不愿转让产权——通过运营一个托管账户，确保买方和卖方都履行义务。同时，作为持牌的在线支付平台，理房通可以为用户创建电子钱包，这些电子钱包在自动核算、结算和将资金支付给房产经纪商和代理商时能发挥重要作用。理房通保障了交易安全，提升了用户交易体验，进一步强化了链家和贝壳的品牌形象，吸引了更多用户和房产中介加入。链家二手房交易流程如图3-15所示。

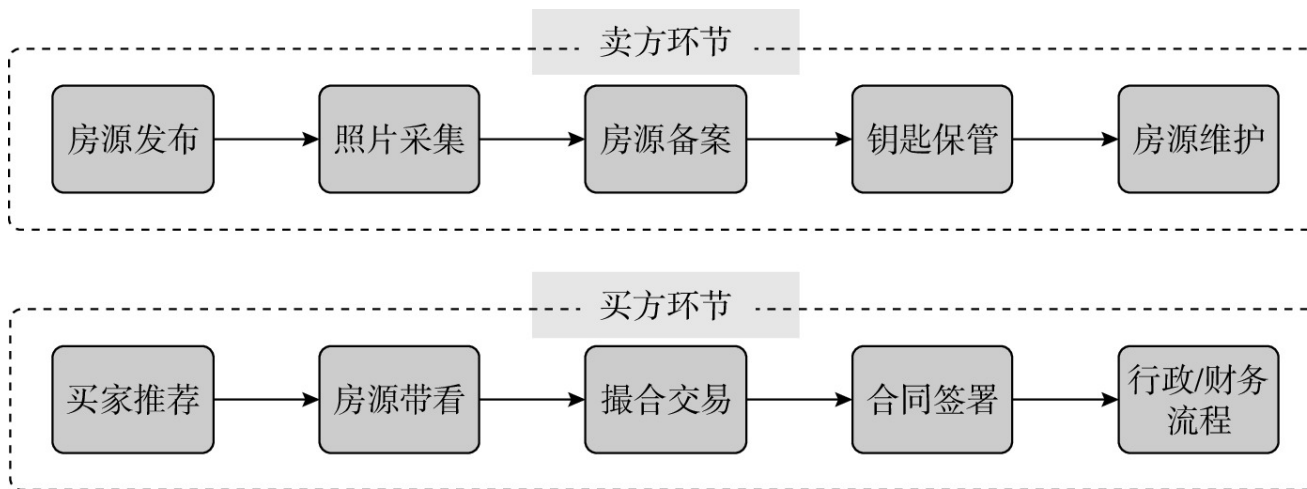


图3-15 链家二手房交易流程示例（新房交易中无卖方环节）

房产交易的区域属性很强，这是链家发展过程中面临的一个挑战，也是其平台化升级的驱动力之一。

链家从北京起家，凭借着领先的交易模式和深度服务，在北京快速扩张，很快就做到了市占率第一，而且链家在经营策略层面也明显领先市场，店效和人效都是其他品牌的好几倍。此外，链家的经营方法论也非常成功。但在开城的过程中，尤其是在一些中小城市，区域性的挑战就出现了。核心挑战有两点：一是**房源**。房产中介进入一个新城市，首先得有房源，而链家在房源拓展上就遇到了很多困难。因为房产中介是一个非常本地化的服务，每个城市都有一些本地的小中介，它们在当地经营多年，本地的房源基本上都掌握在它们手里，一个新中介很难拿到房源。二是**价格**。房产中介行业都信奉靠价格战抢客户，看谁的中介费便宜。但链家的经营方法论是不讲价，而且佣金还偏高。初期，这种模式在中小城市行不通，而这进一步限制了链家的增长，链家的自营模式下的规模增长也就遇到了瓶颈。由于左晖一直都有平台化的构想，于是就通过德佑在一些城市试点，平台试运行一段时间之后效果不错，链家就加速了平台化进程，开放了第三方品牌入驻。

房产交易行业还有一个很复杂的问题，即经纪人管理和合作问题。业务规模越大，经纪人团队越大，管理成本就越高，同时经纪人之间会陷入零和博弈，恶性竞争频发，跨品牌经纪人之间的合作更是基本不可能。贝壳通过ACN机制很好地解决了这个问题，这套机制也是链家完成平台化升级，打造贝壳平台的核心支撑之一。

ACN，即agent cooperation network（经纪人合作网络），是指在遵守房源信息充分共享等规则的前提下，同品牌或跨品牌经纪人之间以不同的角色参与到同一笔交易中，成交后按照各个角色的分佣比例进行佣金分成的一种合作模式。2014年，链家建立ACN合作网络，2018年，德佑实现ACN合作网络的跨品牌达成，无产权关系门店间开始合作。通过ACN合作网络，经纪行业由竞争变为竞合，单店效率得到大幅提升，并与不同门店、不同品牌之间形成房源、客源的全面联通。2018年，贝壳平台也打通了ACN合作网络，更多品牌之间的合作得以实现。

ACN机制设计的核心在于“佣金拆分”，将经纪人之间的关系由零和博弈转变为**利益多赢**。在遵守房源信息充分共享等规则的前提下，各经纪人以不同的角色参与到同一笔交易中，与其上下游角色进行互动和利益绑定。单笔交易成交后，各个角色按照一定比例进行佣金分成，实现角色利益分配的流程化。

新房比二手房简单一些，新房只有客源方角色。

ACN在二手房交易场景中定义的10个角色及其职能如表3-26所示。^②

表3-26 ACN在二手房交易场景中定义的10个角色及其职能

	角色	职能
房源方	房源录入人	将业主委托的房源录入系统
	房源维护人	熟悉业主背景、房屋结构、物业管理和周边环境；带看房源并陪同讲解
	房源实勘人	在贝壳平台内完成房源申请，并拍摄照片或录制 VR
	房源备案人	获得业主委托书、身份信息、房产证书等信息，并上传至政府指定系统
	房源钥匙人	经业主同意，获得房源的钥匙
客源方	房源推荐人	识别适合客户的房源并推荐给其他经纪人
	房源成交人	向客户推荐合适的房源并进行带看；与业主谈判协商，促成交易
	房源合作人	辅助房源成交人，帮助匹配房源，在带看和交易时协助准备文件等
	房源首看人	首次带客户看最终成交房源的经纪人
	交易/金融顾问	负责签约后的相关交易及金融服务

ACN改变了传统行业模式中只有促成交易的经纪人才能获得佣金的方式，将佣金分配给划分出的10个不同角色，这样在一次交易完成后，不只完成交易的经纪人可以获得佣金，交易流程中的其他角色也可以得到相应比例的佣金。这种佣金拆分模式使得经纪人不需要以完成交易作为自己唯一的业务指标，避免了经纪人之间互抢客源、互相撬单等行为的发生，减少了行业内部的恶性竞争，极大地提高了行业效率。

举个例子，假设成交了一套500万元的二手房，佣金率为2%，则这笔交易的佣金就有10万元。按照ACN机制，佣金分配情况如表3-27所示。

ACN还有完善的反馈评价体系，进一步促进信息透明化，将单次博弈转变为重复博弈。ACN的信息透明化体现在房源信息透明化及经纪人信息透明化两方面。

表3-27 按照ACN机制分配佣金

分类	角色	佣金率 (%)	收入 (元)
贝壳平台	—	15	15 000
每个角色按如下比例分剩下的 8.5 万元，但每笔交易中不一定所有的角色都有佣金，收入为空角色的佣金归属房源成交人			
房源方	房源录入人	10	8 500
	房源维护人	20	17 000
	房源实勘人	5	4 250
	房源备案人	0	0
	房源钥匙人	5	4 250
客源方	房源推荐人	10	8 500
	房源成交人	30	25 500
	房源合作人	5	4 250
	房源首看人	15	12 750
	交易/金融顾问	0	0

注：①实际佣金分配会更复杂，每个角色还会再乘以对应的系数，系数由系统根据级别设定，级别越高系数越高；②新房交易更简单，新房没有房源方角色，10万元的佣金最终分配给经纪人的平均大约有4万元。

● **第一：房源信息透明化。**ACN严格审查上架房源，审查信息包括真实存在、真实可出售或出租、真实且已验证的地址和真实的房源相关数据以及价格。同时ACN会对真实房源进行全周期监控审查，及时更新数据库并删除不合格房源信息。房源信息透明化可以杜绝经纪人利用虚假房源信息牟利。

● **第二：经纪人信息透明化。**ACN具有完善的经纪人评价体系：信用分制度使行为不当的经纪人无法持续进行交易服务；贝壳分数和贝壳币这种对C端客户开放的评价体系能够激励经纪人不断提高自身服务从而获得更多客户；经纪人信息透明化使经纪人从追求短期收益最大化的单次博弈（一锤子买卖的宰客行为）逐渐转向追求长期收益的重复博弈，形成正循环。

ACN提高了经纪人的合作效率，把包括房源信息、成交信息、房价、每个业主的动态等内部信息与外界共享，让买家和卖家能够基于透明的数据做出决策。ACN促进了服务商之间的信息和资源共享，拆除了信息孤岛之间的围墙，并为大规模的跨店跨品牌合作提供了可能性，是贝壳平台化升级的基础。

ACN模式下的贝壳平台实现了多方共赢：平台为购房者提供了海量真实可靠的房源和高效的服务，减少了购房的时间成本，降低了购房风险，提升了购房体验；为其他第三方品牌提供了平台，让跨品牌成交得以实现；为经纪人打造了一个良币驱逐劣币的行业环境。

随着平台化的发展，截至2024年12月31日，贝壳平台上已有约260个房地产经纪品牌、约4.9万家以社区为中心的活跃经纪商店，以及超过44.5万名活跃房产经纪人（中介）。这些房地产经纪品牌主要分为三种：链家、德佑以及第三方入驻品牌。按贝壳的管控强度来排列，三者顺序为链家>德佑>第三方入驻品牌。

● **链家**：链家是贝壳的自营经纪品牌。

● **德佑**：德佑之前是上海的房地产经纪品牌，2015年被链家收购，目前德佑是贝壳的特许经营品牌。例如，有一家房地产经纪公司想以特许经营品牌经营的方式加盟贝壳，贝壳收取一定的品牌加盟费后，这家房地产经纪公司就可以使用德佑的品牌，店面名字也会改成德佑。一般一些本地的规模较小的房地产经纪公司，会以特许经营的方式，通过使用德佑品牌的形式加盟贝壳。

● **第三方入驻品牌**：第三方中介入驻贝壳时保留自己的品牌，一般比较大的房地产经纪品牌通常采取直接入驻贝壳的形式，比如21世纪不动产、置家、住商不动产、房六六、优铭家、百年行、中环地产等。

除了房产交易之外，贝壳还进入家装等其他领域。整体上，目前贝壳的业务划分为四部分：新房交易服务、二手房交易服务、家居装修服务、新兴和其他服务（见图3-16）：

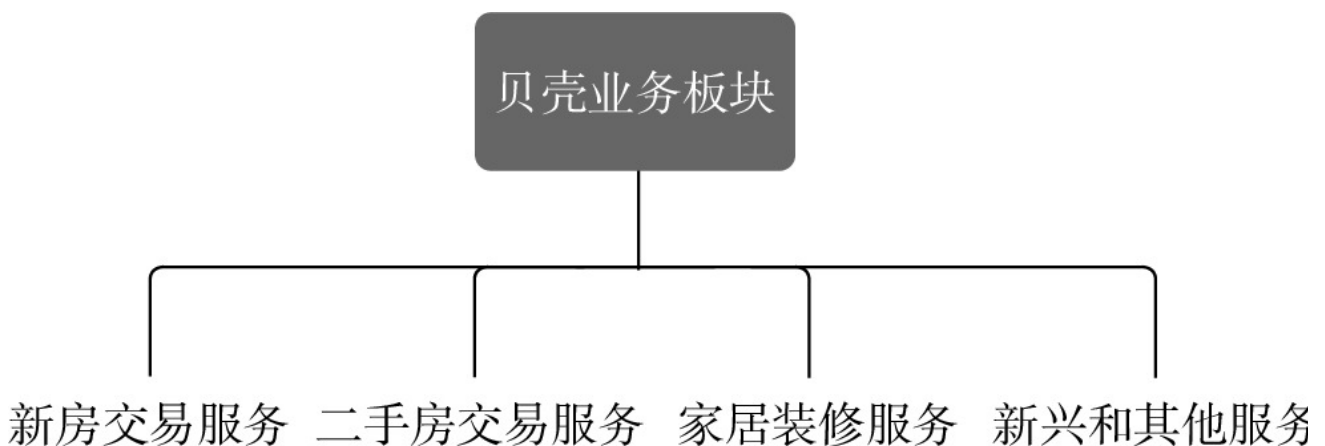


图3-16 截至2025年第一季度贝壳主营业务构成

● **新房交易服务**：新房交易服务是指为房地产开发商向购房者销售其开发的新房提供的中介服务，贝壳会与开发商的销售公司签署住宅交易服务合同。

● **二手房交易服务**：二手房交易服务是指为二手房的销售或租赁提供中介服务，提供服务的中介品牌可以是贝壳上的任何品牌。另外，向贝壳上的经纪公司提供平台和特许经营服务也包含在二手房交易服务的业务范畴内。

● **家居装修服务**：家居装修服务是指贝壳提供的家居装修一站式解决方案，包含室内设计、翻新、装修、家具、家居用品和售后维修等，让客户获得全面的家装服务。

● **新兴和其他服务**：新兴和其他服务包括物业租赁管理服务业务、金融服务业务和其他创新业务。

房产交易重决策过程，对于绝大多数普通家庭而言，购置和出售房产都会导致家庭资产的重大变化，是一件重要且严肃的事情。楼盘字典、真实房源、不吃差价、NTS、理房通、ACN以及交易服务中心等，左晖先生带领链家 and 贝壳在过去二十几年做的这些事情，都对房地产行业产生了深远的影响，推动了整个行业的发展，让房产交易变得更加可信和高效。

1. 贝壳的财务数据（来自贝壳财报公开数据）

房地产行业受政策和经济周期影响很大，贝壳过去几年的营收波动较大，2024年营收为934.57亿元，年同比增长20.16%，在经过连续两年的高速增长后，终于超过了2021年807.52亿元的营收水平。我们可以明显看到，贝壳在平台化升级之后营收整体上在快速增长。利润方面也受到市场宏观环境的影响，但整体来看，贝壳最近几年的利润表现还是不错的，营业利润方面，2022年亏损0.17亿元，2023年和2024年分别盈利56.53亿元和46.48亿元。

贝壳的营收和利润详情如表3-28所示。

表3-28 2017—2024年贝壳营收和利润

营收和成本	2017年	2018年	2019年	2020年	2021年	2022年	2023年	2024年
营业总收入 (亿元)	255.06	286.46	460.15	704.81	807.52	606.69	777.77	934.57
年同比增长 (%)	—	12.31	60.63	53.17	14.57	-24.87	28.20	20.16
主营业务成本 (亿元)	207.38	217.77	347.47	536.21	649.33	468.88	560.59	705.13
毛利(亿元)	47.68	68.69	112.68	168.60	158.19	137.81	217.18	229.44
毛利率(%)	18.69	23.98	24.49	23.92	19.59	22.72	27.92	24.55
销售费用 (亿元)	9.99	24.90	31.06	37.15	43.09	45.73	66.54	77.83

续表

营收和成本	2017年	2018年	2019年	2020年	2021年	2022年	2023年	2024年
管理费用 (亿元)	42.82	49.27	83.77	75.89	89.24	73.47	82.37	89.61
研发费用 (亿元)	2.52	6.71	15.71	24.78	31.94	25.46	19.37	22.83
营业利润 (亿元)	-5.66	-8.83	-14.41	39.54	4.52	-0.17	56.53	46.48
持续经营利润 (亿元)	-5.38	-4.28	-21.80	27.78	-5.25	-13.97	58.90	40.78

从营收构成来看，新房和二手房交易服务依然是贝壳的营收主力，2023年和2024年新房交易服务营收占比分别为39.31%和36.01%，二手房交易服务营收占比分别为35.94%和30.18%。但我们也可以看到家居装修服务与租赁、新兴和其他服务的营收规模增长得也很好，贝壳多元化发展的潜力很大（见表3-29）。

表3-29 2022—2024年贝壳营收构成

营收构成	2022 年		2023 年		2024 年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
新房交易服务	286.50	47.22	305.76	39.31	336.53	36.01
二手房交易服务	241.24	39.76	279.54	35.94	282.01	30.18
家居装修服务	50.47	8.32	108.50	13.95	147.69	15.80
租赁、新兴和其他服务	28.48	4.69	83.97	10.80	168.34	18.01
总计	606.69	100.00	777.77	100.00	934.57	100.00

从GMV（即GTV，本文后续统一用GMV来表达）构成也可以看出，平台化升级是贝壳业绩增长的核心驱动力：从2022年到2024年，在新房交易中，非链家品牌的GMV占比连续三年都超过了80%；在二手房交易中，这三年非链家品牌的GMV占比也都超过了50%。

贝壳的新房和二手房交易GMV构成详情如表3-30所示。

表3-30 2022—2024年贝壳新房和二手房交易GMV构成

GMV 构成	2022 年		2023 年		2024 年	
	GMV (亿元)	占比 (%)	GMV (亿元)	占比 (%)	GMV (亿元)	占比 (%)
新房交易	9 405	37.37	10 031	33.09	9 700	30.16
链家	1 700	18.08	1 932	19.26	1 856	19.13
非链家	7 705	81.92	8 099	80.74	7 844	80.87
二手房交易	15 765	62.63	20 280	66.91	22 465	69.84
链家	7 464	47.35	8 476	41.79	9 185	40.89
非链家	8 301	52.65	11 804	58.21	13 280	59.11
总计	25 170	—	30 311	—	32 165	—

目前，贝壳在一线城市势力强大，尤其是在北京和上海。2022年、2023年和2024年，贝壳31.9%、32.3%和35.6%的净收入来自这两个市场。

2. 贝壳房产交易业务的商业模式解读

抛开家居装修和其他创新业务不谈，贝壳的房产交易业务是典型的二元商业模式。贝壳上每完成一笔房产交易，就会产生一笔GMV，贝壳通过佣金收入来完成货币化。

由于贝壳上存在自营（链家）和非自营（非链家）两种模式，所以贝壳的货币化逻辑很有特点。简单来说，有两个大的逻辑。

● **链家是完成交易的主经纪方。**贝壳的收入即为这笔交易的佣金，即贝壳的收入=GMV×佣金率。例如，链家成交了一套500万元的房子，佣金率2%，则贝壳的收入为10万元。给经纪人（不论是链家的经纪人还是ACN机制下非链家的经纪人）的分成属于贝壳的经营成本。

● **入驻经纪品牌（非链家）是完成交易的主经纪方。**在这种情况下，贝壳的收入是对这笔交易的佣金的抽成，贝壳的抽成比例大约是16%（对德佑还会收取一部分品牌特许经营费），即贝壳的收入=（GMV×佣金率）×贝壳抽成。例如，第三方入驻品牌21世纪不动产成交了一套500万元的房子，佣金率2%，贝壳抽16%，则21世纪不动产的佣金为10万元（本质上，这10万元是贝壳真正的GMV），贝壳的收入为1.6万元。如果链家作为次经纪人参与了这笔交易，那么链家会从剩余的8.4万元中根据ACN机制获得一部分佣金分成，这部分分成同时也计入贝壳的收入。

目前，二手房的佣金率大约为2%，由卖家和买家共同承担；新房佣金率大约为3%，由房产开发商承担（不同开发商的佣金不一样，公寓和住宅的佣金也不一样）。但房产交易的佣金也受政策和行业宏观环境的影响，如前文所言，链家在2023年第三季度下调过一次二手房交易的佣金，由2.7%下调至2%，并且调整了佣金收取方式，由之前的买家承担调整为买卖双方共同承担。

本质上，房产中介（例如链家）是一个连接卖家和买家的平台，而贝壳则是一个服务于房产中介的平台。链家对房产GMV直接进行货币化，货币化率大约为2%~3%；贝壳对房产中介的佣金收入直接进行货币化，货币化率大约为16%。^④由此，我们也可以回答前文3.1.1中提出的问题：二手房交易业务中房产经纪人（中介）的角色，体现的是组织的价值主张。

我们可以简单对比如下两组业务模式，以便更好地理解贝壳模式。

● **链家vs. 其他中介：**业务模式完全一样，但服务差异大，货币化率差异大。

在房产交易行业，链家的中介费（佣金率）是偏高的，基本在2%以上，而其他中介的佣金率一般都低于2%，有的低于1%，甚至还有谈价空间。这里的核心差异在于链家的B端成本高，链家的经纪人在整个交易中要做大量的工作。房产交易重决策、

客单价高、风险大，链家保障了交易的安全性和可靠性，同时提升了效率和体验，所以链家的中介费更高。但在这笔交易中，链家的成本也高，平均到每笔交易，大约50%的佣金收入最终会分给经纪人。

● **链家vs. 贝壳：**业务模式不完全一样，对于贝壳上非链家品牌成交的交易，贝壳依然可以货币化。

入驻贝壳的非链家品牌对房产交易的中介费有一定的自主权（但是贝壳有一定的约束，加入贝壳后佣金率不能太低，因为贝壳对整个服务标准和模式有要求和赋能），一笔交易完成之后它们先抽佣，这个佣金是贝壳真正的GMV，贝壳再对这个佣金进行抽佣，完成货币化。贝壳对入驻品牌提供的服务包括品牌特许经营授权（仅针对德佑）、科学管理模式赋能、运营指导和培训、人员培训和提供SaaS系统等，这些都是非链家品牌交易中贝壳的成本。从链家到贝壳是一次跃迁，但这次跃迁也是建立在链家多年的积累之上。

时间回到2001年，此时链家刚刚成立，房产交易行业处于蛮荒时代，虚假房源和吃差价行为甚至被视为公开的“行业潜规则”。在2004年链家率先提出不吃差价，2011年链家提出“真房源”标准，并启动赔付机制，对行业产生了巨大的冲击，甚至一度引起其他平台的抵制，遭到同行的鄙夷。

左晖曾经在接受采访时说：“走捷径都很容易，但是捷径都是错的。真正创造价值的事情都难，所以我们有时候就说什么叫难而正确的事情，当你搞不清楚的时候，你就选择最难的一条路。”

链家从一家小小的中介门店做起，经过二十多年的发展，成为如今的贝壳，成为当之无愧的房地产行业巨头企业。这便是“坚持做难而正确的事情”的最佳实践。

3.3.7 安居客：房地产信息平台，业务问题突出，最终被58同城收购

安居客创立于2007年，业务全面覆盖新房、二手房、租房、商业地产、海外地产、装修等领域，在为用户提供找房服务的同时，满足房产经纪人和开发商高效营销推广和获客的需求。

安居客也是二元商业模式，但是由于安居客只是信息平台，没有交易能力，实际的交易并不是发生在平台上。因此安居客虽然本质上依然是通过对GMV的抽佣来实现货币化，但是实际上的货币化只能通过广告和卖会员实现，无法直接对GMV进行抽佣。

安居客一度是市场上规模最大的二手房信息平台，然而它没有及时转型，价值主张始终停留在信息匹配层面。同时，平台对供给端的运营投入不够，其会员模式也存在一定的弊端，这些因素综合在一起，导致虚假房源问题严重，且货币化率不高，因此整体上平台的价值较为有限。最终，安居客在2015年被58同城收购。

安居客与贝壳的二手房业务从C端看非常类似，都是二手房源信息平台，但是安居客的业务估值与贝壳相差巨大，其营收、交易规模与贝壳的差别也很大，本质原因还是两者的商业模式虽然都属于二元类型，但是关键要素差异很大。关于安居客的数据就不展开赘述了，这里重点介绍安居客业务中存在的三个问题。

第一，在经营模型里，安居客的单笔交易成本比较低，导致虚假房源问题严重。任何一个中介经纪人都可以在后台上传二手房信息，虽然从形式上看平台也是结构化的，但是房源信息的真实性无从考证。平台也试图做一些算法层面的校验和人工审核，但是还是无法杜绝中介经纪人在一些关键信息上作假，比如这套房源是否真实可售。用户被一个不错的房源吸引，电话打过去，经纪人往往会告诉用户这套房子刚刚卖掉，自己手头还有其他类似的房子推荐，这与平台在房源信息的运营上投入较少是有关系的。安居客的信息入驻门槛很低，这样做的好处是短时间内可以快速完成冷启动，吸引大量房源信息进入平台，但坏处就是没有真正的信息价值，更多沦为中介经纪人争取平台流量和用户联系方式的手段。而在贝壳上，每一笔交易的B端成本是非常高的，贝壳的佣金有很大一部分要分给经纪人，尤其是在ACN机制中，不同角色的经纪人都能从这笔交易佣金中拿到他们各自的分成。正是如此高的成本投入，加之贝壳持续建设的楼盘字典，让贝壳上的房源信息对用户有较高的价值，百分百真房源也是贝壳对用户的核心价值主张之一。

第二，平台在每笔交易中贡献的价值不够。安居客一直停留在信息匹配阶段，所以每笔交易带给平台的价值相对较低。如前文所言，房产交易是一项很复杂的工程，一笔交易持续时间相对较长，且影响成交转化的相关因素非常多，一开始“搭线”的价值肯定不会太大。这一点就决定了无论是买方还是卖方，甚至是经纪人，对于在每笔交易中安居客能赚到的钱的预期都很低。这也就导致安居客单笔交易的货币化率是比较低的。

第三，会员收益模式存在弊端。安居客有很大一部分收入是来自B端经纪人（中介）的会员费。B端会员费模式的弊端明显，即很多经纪人会乱发、多发虚假房源信息，本质上也就是付出固定成本（会员费）之后，利用一切手段，尽可能多地获取用户的联系方式。加之平台的流量分配机制主要是考察用户需求和房源信息的匹配程度，而没有体现经纪人后端的转化服务能力（当然平台也没有掌握用户留资后的转化数据），导致劣币驱逐良币。

在房产交易行业，安居客只做了房源信息平台，而且在经营过程中一度遭到包括链家、我爱我家在内的多家房产中介的联合抵制。安居客业务模式单一，没有及时转型和升级，导致上市计划搁浅，最终以2.67亿美元的价格被58同城收购。

更加令人感慨的是，定位为免费分类信息服务平台的58同城也面临着和安居客类似的问题——业务发展不佳，于2020年9月宣布私有化退市。

3.4 小结

本章我们介绍了二元商业模式的业务逻辑、核心要素、画布和经营分析框架，并分析了几家典型的公司。在观察和分析二元商业模式的业务时，读者要重点关注GMV、货币化率和交易成本。同时，我们可以看到广告的重要性和价值，广告是二元商业模式中重要的货币化方式之一。

本章还通过一些公司的例子，展开介绍了电商交易平台、网约车平台、本地生活服务交易平台和房产交易和服务平台这几类典型的采用二元商业模式的公司。虽然这些公司和业务都很复杂，但拆解其商业模式，却又有一种简单的美，这就是商业模式的魅力。这些公司已经深度渗透了我们的生活，与我们的日常生活息息相关。我们无时无刻不被它们影响着。很多时候，我们以为是自己的选择，实际是它们的调度。商业就是这样。

受篇幅所限，在前述公司业务研究中，还有很多数据和信息未能尽述，读者可以参考本章的思路和视角，做更详细的调研和分析。本章提到的相关数据主要来自公司财报、财报会议记录、行业研究报告、国家统计局、交通运输部、公司官网等，并进行了逻辑推算和交叉验证，仅供读者参考。如有数据错误，还请读者指正。

免责声明：本书提到的所有公司，仅代表作者个人观点，不构成任何投资建议。

第4章 三元商业模式：激发

4.1 互联网三元商业模式研究

4.1.1 互联网三元商业模式的典型业务

如第1章所介绍的，我们将三元商业模式定义为**激发**：组织建设线上场域连接用户和内容，在用户消费内容的场景中激发商业需求，并获得价值的基本原理，即内容→用户→广告主。

很多互联网公司和业务的商业模式是**激发**，涵盖了我们可以免费使用的大部分互联网产品。其中最典型的有五类业务，我们看一下这些业务的底层逻辑。

1. 内容平台

以文字、图文、长视频、中视频、短视频为代表的**内容平台**，如今日头条、抖音、快手、YouTube、TikTok等等，免费满足了用户的内容娱乐类需求，它们的商业模式都是**激发**。组织建设了线上的内容场域，连接用户和内容，在用户免费消费内容的过程中，激发其商业需求，让用户产生广告转化，并从广告主处获取**广告费**。这就是为什么我们在使用这些App的时候，刷几条内容就会出现一条广告。这些广告收入有一部分会分给创作者，反哺内容生态。

2. 社区

以小红书、知乎、B站、CSDN为代表的**社区产品**，其用户之间通过话题、主题、垂直内容等产生互动和连接。社区和内容平台同质化趋势显著，业务边界越来越模糊，社区也有海量的内容供用户消费。但作者认为社区和内容平台的底层逻辑是不一样的：**社区的内核是人，内容是人互动和连接的载体，社区氛围很重要；内容平台的内核是内容，人是内容的消费者，核心是内容的高效供给和分发**。社区产品中人和人之间的互动行为应该远多于内容平台，但是效率低于内容平台。与内容平台类似，社区的商业模式是三元的，社区获取价值依赖于用户在社区互动时被激发的商业需求。

3. 免费小说和免费短剧

小说和短剧本质上是两种内容形式，但这两类都有较独立且规模较大的产品（用户规模过亿），而且短剧可以看作小说网文生态的延伸，所以将这两类单独拿出来放到一起来看。以番茄小说（字节跳动旗下小说&短剧业务）、七猫免费小说、

河马剧场为代表的业务，组织建设了线上的小说或短剧场域，连接用户与作者/版权方，在用户免费看小说和短剧的过程中，激发其商业需求，让用户产生广告转化，并从广告主处获取广告费。同时，小说和短剧的作者也可以拿到一些分成收入，支撑他们继续生产好的内容。另外，这里值得思考的是，小说和短剧这类业务，都是从付费模式逐步演变成免费模式或者“免费+付费”混合模式，也就是说，它们的商业模式从一元演变为三元或者一元+三元。

4. IAA应用

所谓IAA，即in-App advertisement，也就是应用内广告，从命名上就可以很直观地看出其商业模式。对应到游戏中，就是指游戏本身是免费的（与前述IAP游戏对应），依靠在用户玩游戏过程中展示广告完成变现。显然，IAA游戏的商业模式是三元的：一个组织开发了一款游戏，免费提供给游戏玩家，并在玩家玩游戏的过程中展示广告（例如开屏广告、弹窗广告、激励视频广告等），从广告主处获得收入，即组织建设的线上场域是游戏，连接游戏玩家和游戏开发者（此处游戏开发者即组织本身），在用户免费玩游戏的过程中，激发其商业需求，让用户产生广告转化，并从广告主处获取广告费。羊了个羊、抓大鹅、2048、消消乐等休闲小游戏都是IAA游戏。另外，IAA工具也属于这一类，工具中的服务用户都可以免费使用，公司通过广告变现，例如Wi-Fi万能钥匙、中华万年历、墨迹天气等。

5. 搜索引擎

搜索是个功能，很多产品（例如淘宝、京东、链家等等）中都有搜索功能，这些不属于搜索引擎的范畴，我们这里讨论的是以Google等为代表的通用搜索引擎。搜索引擎的商业模式是三元的，用户免费使用搜索引擎查询和浏览信息，组织通过在搜索结果页展示广告，从广告主处获得收入。搜索引擎有一个非常重要的逻辑，即用户的需求是天然被搜索词（query）切分开的，所以在激发其商业需求时，有一个非常强的信号输入，即query。例如用户搜索“手机推荐”，则“手机”作为一个强信号输入，系统就将推送与“手机”相关的广告，这大大提升了搜索引擎广告的转化效率。query在搜索引擎调度用户注意力、激发用户商业需求时极具价值，但也带来了一定的约束：搜索引擎广告受相关性的约束，广告需要与用户当前的query相关。在通用搜索引擎中，有大约20%的高商query，例如搬家、装修、买东西、招商加盟、旅行社等；剩下的大约80%为低商query，例如知识类、寻址类、how to（怎么读、怎么写、怎么操作等）类、事实类等。在相关性约束下，可以激发的商业需求对和用户原始需求的相关性要求很高，这也限制了激发的空间。因此在搜索引擎广告业务中，低商query的流量变现效率比较低。

三元商业模式与二元商业模式最大的区别在于，用户并没有带着“一笔钱”来，而是带着“时间和注意力”来。但是每一个平台都想做交易闭环，并且都能说出

很多理由：提供更好的用户体验、帮助商家更好地长效经营、更好地承担监管责任等等。这本质上是希望打造用户心智。

何谓打造用户心智？就是从让用户带着“时间和注意力”来，变成让用户能够直接带着“一笔钱”来。如果能够完成这样一个转换，交易闭环基本就跑通了，用户心智也就打造好了。由此可见，打造用户心智非常难。抖音就是一个很典型的例子，抖音从引流电商（给京东、淘宝等电商平台引流）到二类电商（给单品页引流）再到闭环电商的发展，就是一个逐步打造用户心智的过程。

4.1.2 三元商业模式的飞轮效应

采用三元商业模式的组织是一个多边平台，连接着内容、用户和广告主，各方之间存在供给和需求的关系，广告主直接为整个生态贡献收入。多边平台也有一定的飞轮效应，其是双边市场飞轮效应的衍生：用户与内容之间形成增长小飞轮；用户与广告主之间形成正反馈，并反哺内容形成正反馈，形成增长大飞轮。最终，促进用户总时长增长和组织收入增长。

- **内容：**用户越多，待满足的内容需求就越多。广告主和广告预算越多，创作者就越有机会获取更多收入，从而整体吸引更多创作者加入，生产更多高质量的内容。平台有足够新鲜、足够多、足够丰富的内容，才能吸引更多的用户。
- **用户：**内容越多，用户越有可能获取更丰富更优质的内容。这样，平台就能吸引更多的用户过来消费内容，从而占据更多的用户时间。用户增加带来的需求增加以及互动反馈，会进一步促进内容的生产和迭代。平台有足够多、足够高质量的用户，才能吸引更多广告主。
- **广告主：**用户越多，广告主能通过平台触达的用户就越多，更多的用户和互动会带来更好的广告效果，进而吸引更多广告主投入预算，提升平台的广告收入。平台有足够多的广告主和广告预算，才能打造一个好的生态，让更多的创作者赚到钱。

最终飞轮效应带动用户总时长和组织收入的增长。我们可以用图4-1来概述三元商业模式的飞轮效应。

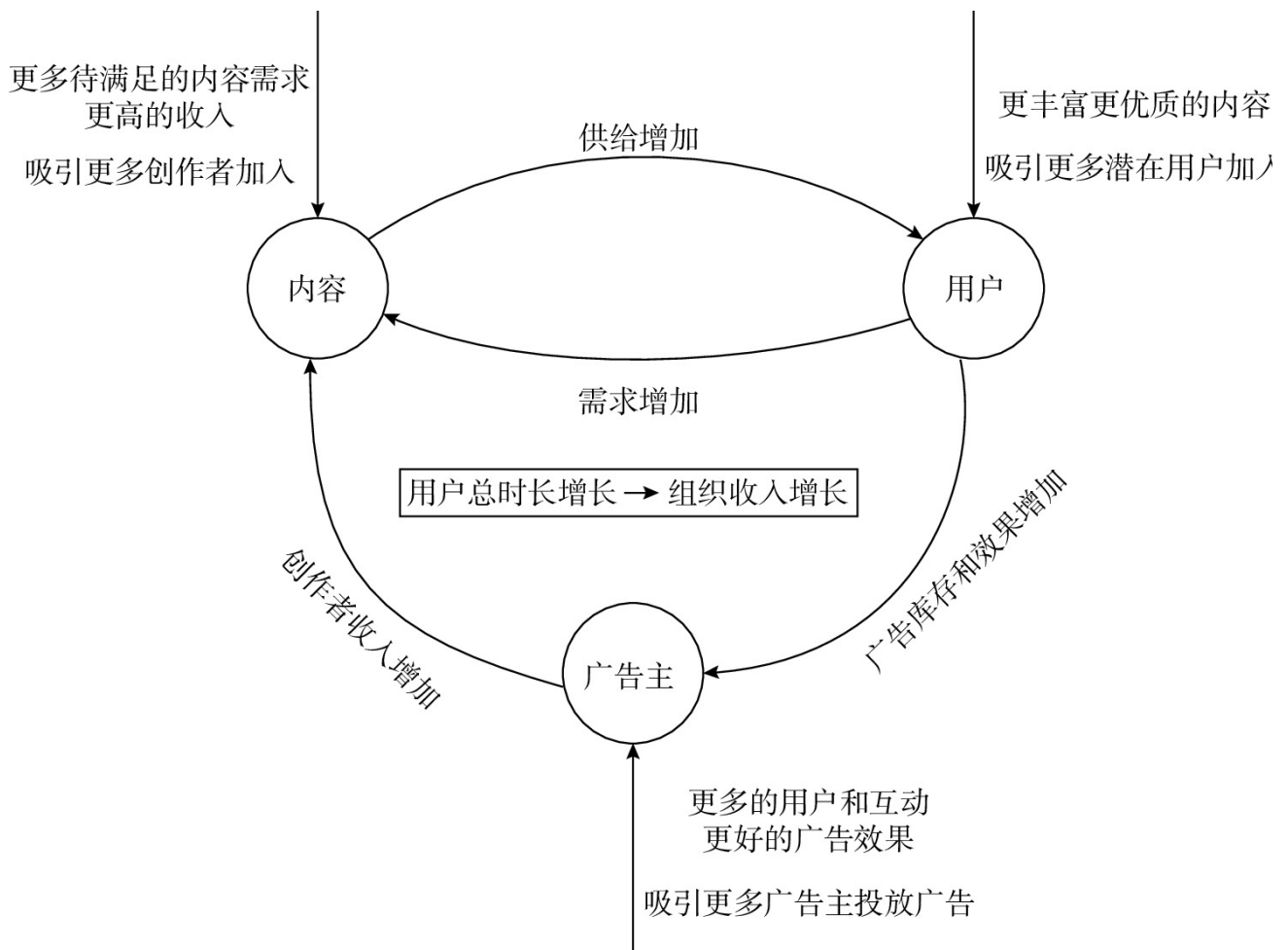


图4-1 三元商业模式的飞轮效应

互联网的这类三元商业模式，逐步催生了一个非常大的经济生态：**创作者经济**。大量的创作者从三元商业模式中获利，赚到了更多的钱，丰富了整个生态。我们也能观察到很多平台创作者在流失、内容质量在下滑、用户在流失，进入下行螺旋，这本质上是因为商业变现没做好。平台不能依赖创作者“为爱发电”。

基于上述的定义和典型业务介绍，我们可以总结一下三元商业模式的核心要素：

(1) **用户**：创作者（内容）、用户和广告主，都是组织的目标用户。这三者之间也存在重叠和相互转化。

(2) **价值主张**：组织给创作者、用户和广告主提供的价值主张是多元的，进而构建出完整的生态。对于创作者，组织能提供收入、创作者工具、私域经营和分析服务、创作者社群等；对于用户，组织能提供丰富的内容、信息、互动、连接等；对于广告主，组织能提供广告营销服务、广告投放服务、私域经营服务、金融服务等。

(3) **渠道**：组织需要通过渠道触达目标用户，触达不同目标用户的渠道是差异化的。触达创作者可以通过头部作者运营、MCN生态建设、创作者社群拓展等渠道；触达用户可以通过多种用增（即用户增长，user growing）渠道；触达广告主可以通过直客销售、区域代理、线上推广等渠道。

(4) **成本**：绝大多数三元商业模式的成本有边际递减效应，这些成本包括但不限于用增成本、人力成本、监管审核成本、服务器成本、CDN（内容分发网络，content delivery network）成本、渠道成本等等。另外，从组织视角看，内容成本也是三元商业模式中很关键的一部分成本。

(5) **收入**：组织通过广告，在激发用户商业需求的过程中获取价值，理论上用户在平台消费的每单位时长，都有其商业价值，都可以为组织贡献收入。

4.1.3 商业模式画布

根据三元商业模式的核心要素，我们可以画出互联网三元商业模式的画布（见表4-1）。

表4-1 互联网三元商业模式的画布

<p>用户 C端用户、关注时长 全国甚至全球范围（区域约束相对小）</p>	<p>广告主 有线上营销需求和预算的商家 关键客户、中小客户 品牌预算、效果预算 全行业覆盖（电商、游戏、金融、教育等全行业广告主）</p>	<p>创作者（内容） UGC、PGC、PUGC、AIGC</p>
--	---	---

续表

<p>价值主张 To 创作者（内容）：收入、创作者工具、私域经营和分析服务、创作者社群等 To 用户：丰富的内容、信息、互动、连接等 To 广告主：广告营销服务、广告投放服务、私域经营服务、金融服务等</p>	<p>渠道 触达创作者（内容）、用户和广告主 头部作者运营、MCN生态建设、创作者社群拓展、用增、直客销售、区域代理、线上推广等</p>
<p>成本 用增成本、内容成本、人力成本、监管审核成本 服务器成本、CDN成本、渠道成本等 有边际递减效应</p>	<p>收入 广告收入 关注广告变现效率</p>

采用三元商业模式的业务，对组织的产品和运营能力都有很高的要求，而且这类业务头部规模效应明显，容易出现赢者通吃的竞争格局，所以竞争非常激烈。我们会看到很多头部的互联网公司，其核心业务都是三元的。对于三元商业模式的分析框架、商业需求的激发效率、广告变现效率等问题，我们接下来展开介绍。

4.2 商业模式经营分析框架

三元商业模式的收入来源于广告，成本包括用增成本和内容成本等。广告的收入规模和变现效率，是三元商业模式运转和扩张的核心支撑。广告收入由用户规模和变现效率决定，其中：用户规模主要看DAU（日活用户数，daily active users），变现效率主要看ARPU，即广告收入=用户规模×变现效率=DAU×ARPU。我们将ARPU进一步拆解，形成以下分析框架：

广告收入=DAU×人均时长×人均单时长VV×广告加载率×ECPM（广告千次曝光收入）/1 000

这个分析框架能够帮助我们非常清晰地洞察业务，同类的业务都可以在这个框架下横向对比，包括对比其用户指标和商业指标。另外，针对搜索引擎广告，出于行业惯例，在上述分析框架基础上，还有一个特殊的拆解方式：

搜索引擎广告收入=搜索PV×CPM1/1 000=搜索UV×人均搜索次数×PVR×ASN×CPM2/1 000

式中，PV（page view）为页面访问量；CPM1（cost per mille 1）为每千次搜索PV的广告收入；UV（unique visitor）为独立用户访问数；PVR为出广告的PV占比，代表每百次搜索PV里，有多少PV出了广告；ASN为平均广告展示条数；CPM2为每千次广告展示产生的收入。

搜索引擎广告收入虽然有其特殊的拆解方式，但本质上仍由用户规模和变现效率决定，底层逻辑没变。接下来我们具体展开讲一下分析框架中的核心指标。

（1）用户规模（主要看DAU）：简单来说，用户规模就是指有多少用户使用组织的产品。用户是企业最宝贵的资产，在互联网业务中更是如此，很多企业的估值都与用户规模呈正相关。在移动互联网时代，一般都使用DAU或者MAU来说明产品的用户规模。其中DAU表示产品每天有多少活跃用户，MAU（月活用户数，monthly active users）表示产品每个月有多少活跃用户。除了绝对值之外，用户的活跃度也很重要，它代表着产品的用户黏性，一般可以通过日活率（DAU/MAU）和用户月活跃天数来衡量。举一个简单的例子，我4月30日打开了一次抖音App，4月一共有15天打开了抖音App，则：我4月30日为抖音贡献了一个日活用户，抖音DAU加1；4月我也是抖音的月活用户，抖音MAU加1；我在抖音的日活率为50%，月活跃天数为15天。

(2) 人均时长：用户时长是指用户使用产品的时间长度，人均时长等于总用户时长除以DAU。例如我今天看了10分钟抖音、15分钟微信视频号以及30分钟B站，那么我今天在这三个产品上的用户时长分别为10分钟、15分钟和30分钟。用户时长越长，表示用户对产品的依赖性和使用产品的活跃度越高，提升用户时长是各大互联网产品竞争的关键目标之一。猎豹移动的董事长兼CEO傅盛在2015年曾说过，互联网公司的本质就是为增长用户使用自家产品的时长而努力；得到创始人罗振宇在2016年的《时间的朋友》跨年演讲中提出了GDT（国民总时间，gross domestic time）的概念。移动互联网行业已经处于存量竞争阶段，各大公司都在争夺用户时间和用户注意力。但不同定位或分类的产品，用户时长天然差距很大，例如工具类产品的用户时长大概率短于游戏，搜索引擎的用户时长大概率短于信息流。

(3) 人均单时长VV：VV（视频播放量，video view），是指在一个统计周期内视频被播放的次数，也就是内容的被观看次数。顾名思义，人均单时长VV就是指单个用户在单位时间内观看视频内容的次数。如果是图文内容，也可以用分发量来衡量，本书统一用VV来阐述。所以一个平台的总VV=人均时长×人均单时长VV。VV代表着用户在平台上的内容消费能力和消费效率。用户消费的内容越多，VV越大，平台的商业化潜力就越大。短视频是目前所有内容形态中效率最高的。单位时间内用户能看多少个内容，与内容的长度强相关。

(4) 广告加载率：即adload，是指用户浏览内容时看到广告的次数与其浏览内容总次数的比例。广告加载率用来衡量广告在用户浏览内容中的密集程度。广告加载率=广告曝光次数/内容浏览总次数，如果用VV来表示，则广告加载率=广告曝光次数/总VV。结合前述人均时长和人均单时长VV的定义，可以得出：广告曝光次数=人均时长×人均单时长VV×广告加载率。

(5) 广告千次曝光收入：即ECPM（effective cost per mille），是指每千次广告曝光可以获得的广告收入。ECPM用来衡量广告变现效率。ECPM=广告收入/广告曝光次数×1 000。例如某个广告位，广告曝光次数为100万次/天，平台的广告收入为10万元/天，则该广告位的ECPM为100元。广告变现效率与广告系统策略、广告系统机制和广告主预算强相关，这些内容在作者的另一本书《计算广告：产品、机制与大模型应用》中会详细展开介绍，推荐感兴趣的读者阅读。除此之外，广告变现效率也受内容形态的影响。

简而言之，采用三元商业模式组织的广告收入，取决于用户规模和变现效率，即：

$$\text{广告收入}=\text{DAU}\times\text{ARPU}$$

式中，

$$\text{ARPU} = \text{人均时长} \times \text{人均单时长} \times \text{广告加载率} \times \text{ECPM} / 1000$$

综上所述，采用三元商业模式的业务通过广告获取收入。本章接下来选择了四家具有代表性的公司——字节跳动、哔哩哔哩、小红书和Facebook，详细介绍一下它们的业务情况。它们都采用三元商业模式。

4.3 典型公司及其业务研究

4.3.1 字节跳动：在“移动互联网+推荐引擎”的驱动下，从知春路走向全球的商业奇迹

字节跳动（ByteDance）成立于2012年3月，创始人是张一鸣。这家从北京知春路锦秋家园小区的一间民宅中起步的公司，在过去十几年的时间里，一路狂奔，席卷全球，截至2024年第二季度，其业务已经覆盖全球150个国家和地区，员工数量超过10万人；以抖音、今日头条、TikTok、番茄小说、豆包为代表的字节系产品在全球的MAU超过30亿。字节跳动虽然还没有上市，但凭借其约3 000亿美元的估值超过SpaceX和OpenAI，成为当前全球估值最高的独角兽公司。图4-2展示了其旗下部分业务。



图4-2 字节跳动旗下部分业务

2018年4月10日，国家广播电视总局责令“今日头条”网站永久关停内涵段子客户端软件及公众号。

从内涵段子^①（2012年5月上线）到今日头条（2012年8月上线），再到抖音（2016年9月上线）、TikTok（2017年5月上线），字节跳动如App工厂一般生产App，进行

变现和增长，并逐步将业务扩展至电商、房地产、本地生活、办公、云、教育、游戏等领域。随着业务的发展，以及移动互联网行业进入存量市场，字节跳动也改变了“大中台，小前台”的组织模式，在2021年11月的组织升级中，由梁汝波接任CEO，并实行业务线BU（业务单元，business unit）化，成立了六个主要业务板块：抖音、TikTok、火山引擎、飞书、大力教育和朝夕光年（见表4-2）。

表4-2 字节跳动的主要业务板块

业务板块	简介	主要产品和业务
抖音	将今日头条、西瓜视频、搜索、百科以及国内相关垂直服务业务并入抖音。该板块负责国内信息和服务业务的整体发展。	抖音、今日头条、西瓜视频、抖音火山版、巨量引擎、巨量千川、剪映等
TikTok	负责 TikTok 平台业务，同时支持海外电商等延伸业务的发展。	TikTok、Helo、Resso、Lemon8 等
火山引擎	负责打造企业级技术服务云平台。字节跳动将自己在发展过程中积累的经验、技术、平台能力开放给外部企业，负责助力企业在数字化升级中实现持续增长。	火山引擎

续表

业务板块	简介	主要产品和业务
飞书	将飞书、EE（企业效率部门）、EA（企业应用部门）合并成飞书板块，聚焦提供企业协作与管理服务。整合即时沟通、在线文档、日历、音视频会议等办公协同应用，并提供 OKR（目标与关键成果）、招聘、绩效、合同等组织经营管理类产品。	飞书、Lark 等
大力教育	致力于建立终身教育服务体系，覆盖智慧学习、成人教育、智能硬件、校园合作等领域。员工发展部门的部分职能转型为职业教育方向，并入大力教育板块。	清北网校、gogokid、开言英语、教育智能硬件等
朝夕光年	负责游戏研发与发行。面向全球用户与开发者，提供顶级游戏和打造玩家社群。	朝夕光年、Ohayoo、沐瞳科技、有爱互娱等

字节跳动并不是张一鸣第一次创业。

张一鸣2005年毕业于南开大学软件工程专业，并与校友组成3人小队，开发一款To B的协同办公系统，但是协同办公当年在中国还没有市场，几个月后张一鸣的第一次创业便宣告失败。张一鸣2006年加入酷讯网，从事旅游垂直内容搜索。2008年从酷讯网离开后短暂加入微软中国，同年9月，以技术合伙人身份加入王兴团队（王兴和张一鸣是福建龙岩老乡），负责饭否网和海内网搜索、消息分发、热词挖掘等业务。2009年7月，由于政策监管等原因，饭否网停止运营，王兴解散团队，张一鸣也随之离开。

饭否网团队解散后，张一鸣于2009年9月创立九九房，开始房产搜索方向的创业。期间张一鸣带领团队取得了不错的业务成绩，尤其是在移动端，推出掌上租房、掌上买房等五款移动应用，在当时的移动互联网环境下获得了超百万的用户，展现出了巨大的潜力。

2011年，张一鸣辞去九九房CEO的职位，开始筹备第五次创业，定位在“信息获取与分发”方向，其实这和他过去几年创业做的事本质上是一致的。2012年3月，字节跳动成立；8月，今日头条1.0版本上线。今日头条基于个性化推荐引擎，根据用户的属性、兴趣、社交行为、阅读习惯等多维信息，向用户推荐个性化信息。今日头条的定位是“基于社交挖掘和个性化推荐的新媒体”，某种意义上说，今日头条的上线打开了移动端推荐引擎的大门。

从后视镜视角来看，字节跳动的爆炸式增长，核心在四个关键词：移动互联网、推荐引擎、短视频、全球化。

● **移动互联网**：2011年小米手机1刚上市，微信刚上线，中国4G牌照还没发放。那一年中国移动互联网用户规模为5亿人左右，单用户日均时长大约为2小时；到2024年上半年这组数字为接近11亿人和超过4小时。与之对应的是，中国移动互联网广告市场规模在2011年只有30亿元，而仅仅是2024年上半年，这个数字就超过了3 500亿元（见表4-3）。这就是移动互联网在用户规模和商业价值上的巨大机会。

表4-3 2024年上半年和2011年中国移动互联网数据对比

时间	中国移动互联网用户规模	中国移动互联网单用户日均时长	中国移动互联网广告市场规模
2011年	5亿人	2小时	30亿元
2024年上半年	接近11亿人	超过4小时	超过3 500亿元

移动互联网也给用户消费内容的方式带来了变革。手机逐步成为人体的延伸，社交和资讯需求率先在移动端爆发，内容生产、消费和流动的底层逻辑都发生着变化，人类逐步进入信息爆炸和碎片化阅读时代（见表4-4）。

表4-4 PC互联网和移动互联网对比

时期	内容生产	内容消费	内容流动	用户反馈	终端
PC 互联网	编辑选择	连续阅读	浏览、搜索、订阅	周期长	PC 端
移动互联网	个性化、UGC、 高效、量大	碎片化阅读	关注、推荐、分享	准实时	移动端

● **推荐引擎：**在PC时代，搜索引擎和浏览器是最大的流量入口。但到了移动时代，直接将搜索引擎和浏览器从PC端搬到手机端并不完全匹配用户的需求。社交和资讯这两个需求最先爆发，加之手机让大量不会用电脑、不会用搜索引擎，甚至不会打字（不会用输入法）的用户接入了互联网，推荐引擎很大程度上取代了搜索引擎，成为移动互联网时代最主要的流量入口。站在今天来看，手机端最大的流量入口是社交平台 and 推荐引擎，在国内以微信、抖音和快手为代表，在海外以Facebook、TikTok、YouTube为代表。大多数人是懒惰的，推荐比搜索门槛低，而且无处不在，用户不用输入、不用思考就能获取内容，而且随着数据的质、量和时效性不断提升，推荐引擎对用户需求的满足也会越来越精准。字节跳动就是基于推荐引擎而创立的，今日头条最开始打造的便是一套资讯内容的推荐引擎（见图4-3）。

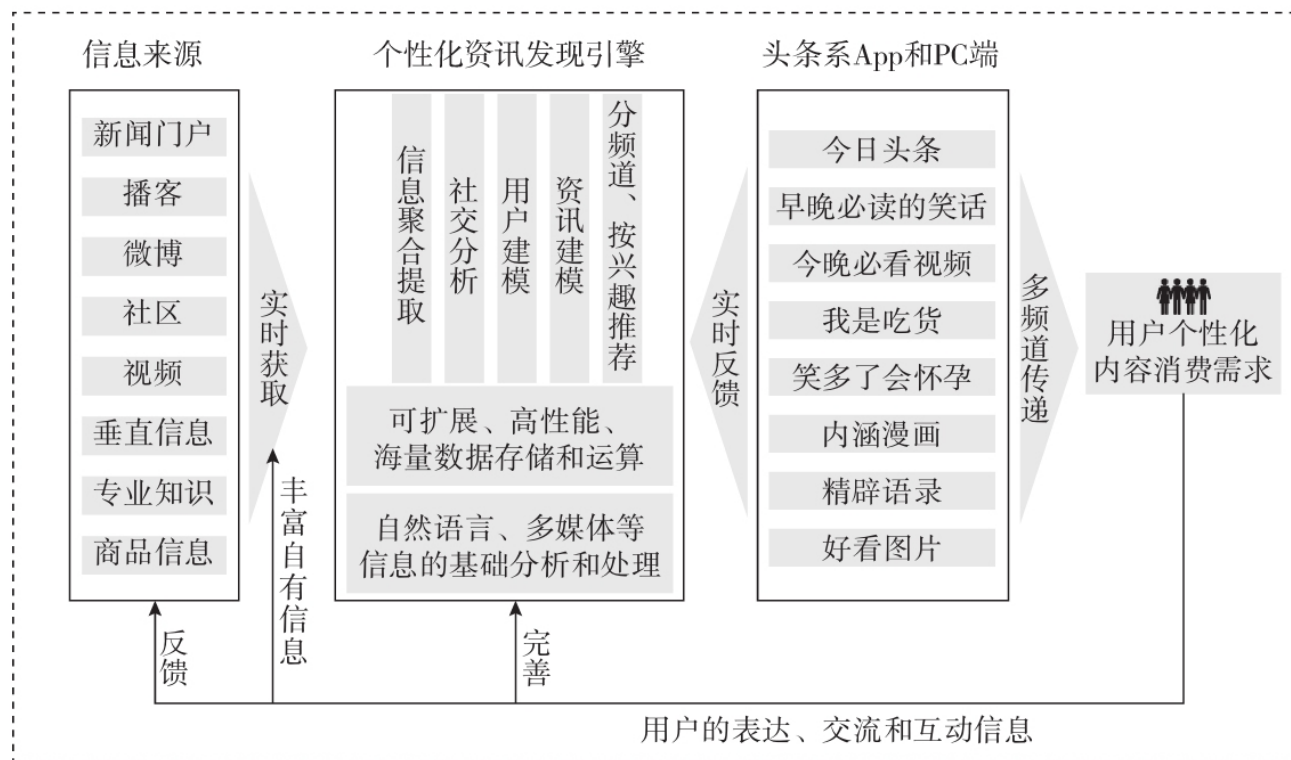


图4-3 今日头条的业务逻辑（2013年）

● **短视频：**马歇尔·麦克卢汉（Marshall McLuhan）认为媒介是人的延伸，并在《理解媒介》一书中指出媒介即信息。短视频这种媒介形态会调动视觉和听觉来刺激用户，诱导多巴胺分泌，让受众不停地追求即时满足感；同时在推荐引擎和上下滑交互的加持下，短视频会迎合用户的喜好和惰性，不需要用户思考且让用户停不下来。短、不连续、丰富的刺激、迎合人类的惰性以及几乎零使用门槛的特点，让短视频吸引了大量的用户，风靡全球。

2016年9月，A.me上线，同年12月，更名为抖音短视频；2017年11月，字节跳动以10亿美元收购北美音乐短视频社交平台Musical.ly，并与抖音合并，开启在短视频赛道上的疯狂增长。其实抖音并不是国内第一个做短视频的产品。2013年7月，宿华和程一笑就带领团队将“GIF快手”从工具转型为短视频社区，并改名为“快手”。但是抖音比快手更早地看到短视频的市场空间，采取了更激进的增长和变现策略，截至2025年第一季度，抖音的DAU和用户时长均遥遥领先于快手，完成后发先至。

● **全球化：**全球化就是本地化，字节跳动在创业初期就有国际化的思考，即从国内市场到全球市场，这进一步打开了增长空间。2015年字节跳动就开始试水国际化，发布今日头条海外版TopBuzz；2016年9月，BuzzVideo上线；2017年5月，TikTok上线，截至2025年第一季度，TikTok的全球MAU已经超过了15亿。

移动互联网和推荐引擎构建了字节跳动的根基，短视频和全球化彻底打开了字节跳动的增长空间。

字节跳动就是基于三元商业模式设计的。用户在平台上消费内容，平台通过广告变现，获取广告收入；广告收入再反哺内容和用户增长，繁荣平台的内容生态，从而做大用户规模和分发量（DAU、用户时长和VV），进一步提升广告收入，形成增长飞轮。今日头条、抖音、TikTok都是这个商业逻辑。

虽然随着业务的发展，字节跳动的商业模式也在逐渐丰富，例如火山引擎、飞书和大力教育等板块都不是采用三元商业模式，但目前公司主要的收入（预计超过70%）还是来自三元商业模式中的广告收入，并且作者认为，未来广告依然会是字节跳动最主要的收入来源。

字节跳动主要的广告场景是字节系产品中的信息流广告、搜索广告以及第三方App（非字节系产品）中的网盟广告。我们可以看几个简单的例子：图4-4前两张为抖音推荐频道中出现的广告（游戏和汽车广告），第三张为在抖音搜索“好玩的游戏”时，抖音搜索结果页中出现的游戏广告。



图4-4 抖音广告示例

因为字节跳动目前还没有上市，所以没有官方公开的财务报表数据。根据行业调研的数据，2024年上半年，字节跳动整体的收入大约为730亿美元（约5 300亿元），其中国际收入约为170亿美元（主要来自TikTok的广告和电商收入），国内收入约为560亿美元（主要来自抖音、今日头条等产品的广告和电商收入）。

行业研究信息仅用来辅助读者更好地理解字节跳动的业务模式以及定性观察其业务发展趋势。具体的数字不作为任何建议和其他用途。

基于行业研究的信息，字节跳动2022年全年的国内广告收入就达到了3 000亿元，2023年达到3 800亿元，即日均收入超过10亿元。在去掉电商媒体（阿里巴巴、京东、拼多多等）之后的国内广告市场大盘中，字节跳动的份额从2018年的16%，快速增长至2023年的48%（见图4-5）。在广告主预算侧，过去这些年，字节跳动拿下了更多的预算份额。

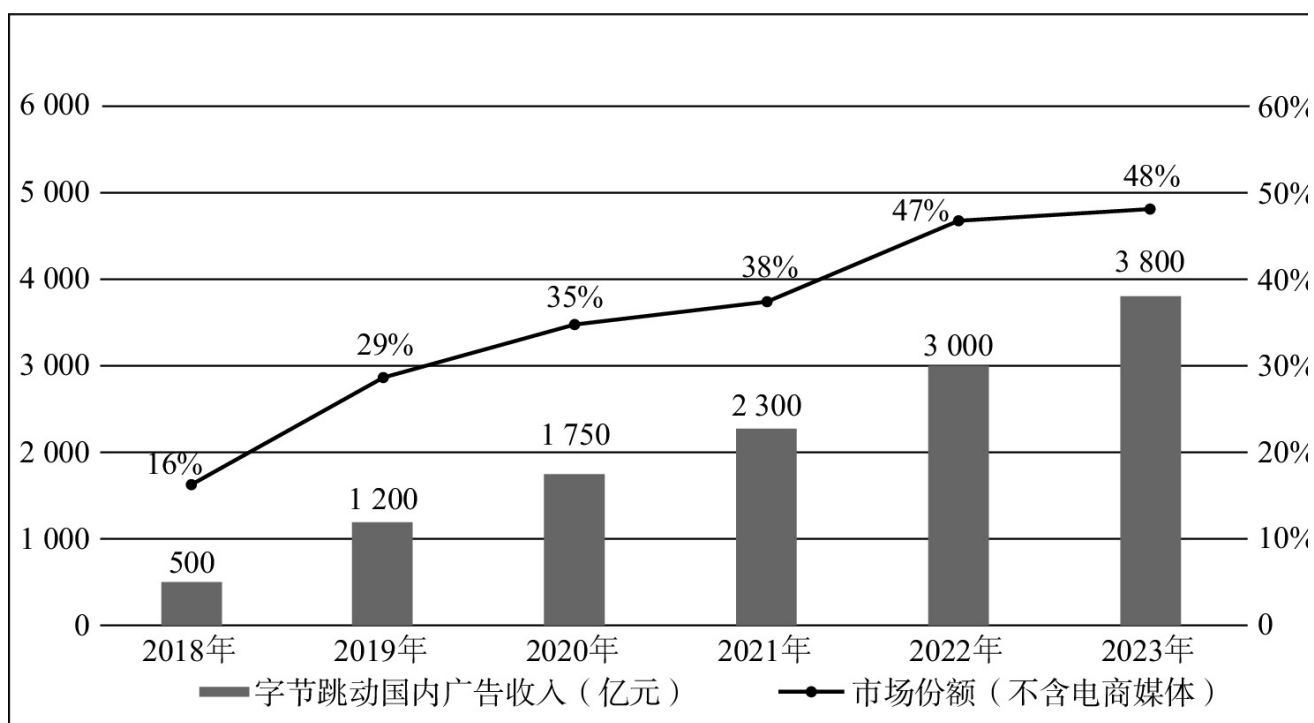


图4-5 2018—2024年字节跳动国内广告收入和份额

我们以抖音（不含抖音极速版）为例，具体预估一下（说明：此处估算仅用来做商业模式解读，具体数字仅供参考，不作为任何建议和其他用途）其用户规模（DAU）和变现效率（ARPU）。

以2023年的数据为例，假设抖音的广告收入占字节跳动国内广告收入的80%，抖音DAU按照6.7亿来计算，则

抖音广告年总收入=3 800×80%=3 040（亿元）

据推算，抖音DAU ARPU约为1.24元。估算详情如表4-5所示。

表4-5 2023年抖音DAU ARPU估算详情

项目	预估数据
字节跳动国内广告年总收入（亿元）	3 800
抖音广告收入占比（%）	80
抖音广告年总收入（亿元）	3 040
抖音广告日均总收入（亿元）	8.33
抖音 DAU（亿人）	6.7

续表

项目	预估数据
抖音时长（分钟）	140
DAU ARPU（元）	1.24
单时长 DAU ARPU（元）	0.009

这意味着一个抖音日活用户，每天大约能给字节跳动带来1.24元广告收入，平均每用户每分钟带来的广告收入约为0.009元。并且随着抖音的持续优化，这个数字还在持续增长。

在推荐引擎和短视频的组合驱动下，字节跳动取得了巨大的商业成功。在商业之外，影响更大的是，短视频成了这个时代“用户时间和注意力”的熔炉，彻底改变了信息生产、传播和获取的方式。媒介是人的延伸，短视频的发展，给每个人的生活乃至人类社会都带来了深远的影响。

科技向善，希望字节跳动能够持续重视企业的社会责任。如其公司使命所言：激发创造，丰富生活。

4.3.2 哔哩哔哩（B站）：将弹幕文化发扬光大，由小众社区到综合性视频平台，成为中国年轻一代的文化和视频社区

哔哩哔哩（bilibili）成立于2009年，是目前中国最大的弹幕视频网站和线上宅文化社区，简称B站。B站的前身为徐逸于2009年6月创建的MikuFans。徐逸原为AcFun（简称A站，是国内早期的二次元视频网站）用户，但当时AcFun网站运行不稳定，所以徐逸建立MikuFans来为用户提供一个稳定的弹幕视频分享网站。由于徐逸很喜欢《某科学的超电磁炮》中的主角御坂美琴，御坂美琴在动漫中的绰号是哔哩哔哩（放电妹），因此他于2010年1月将MikuFans正式更名为哔哩哔哩，也就是说B站原本只是作为AcFun的备用服务器，甚至当时二次元用户们都戏称B站是“A站的后花园”。

B站早期是一个ACG（动画、漫画、游戏，animation, comics and games）内容创作与分享的弹幕视频分享网站。弹幕并非B站首创，其起源可以追溯到日本的Niconico动画。B站在弹幕文化的传播和发展中起到了重要作用（2024年6月B站的弹幕累计总数突破200亿），随着B站的兴起，弹幕在国内得到了更广泛的应用，逐步成为各大视频网站的标配。

B站真正开始公司化运营是在2014年，在此之前，B站只是一个非营利性的在线社区。2014年，B站的天使投资人陈睿以董事长的身份正式加入公司，开始带领B站高速发展，不断地破圈和扩张。B站于2018年正式在美国纳斯达克挂牌上市。

回顾B站的发展史，大概可以分为四大阶段。

● **第一阶段：纯粹的ACG二次元社区（2009—2013年）。**作为A站的后花园，早期B站有着明显的日系二次元风格。平台上大约3 000万名核心用户是中国二次元文化的原住民，是B站“鄙视链”的最高层。凭借着优质的服务体验和社区氛围，B站的发展很快就超过了管理混乱的A站。

● **第二阶段：开始公司化经营（2014—2017年）。**这一阶段，陈睿以董事长的身份加入B站，B站开始公司化经营。值得一提的是，陈睿加入时，B站只有30名员工。陈睿带领B站获取资质，拓展内容生态，并开启商业化，推出游戏、直播、广告、电商、大会员等服务。用户群上，B站大刀阔斧地升级，如降低新番比重、增加分区和独立热门排行榜等，拓展了具有中国特色的ACG文化粉丝用户群体。

● **第三阶段：Z世代兴趣社区（2018—2020年）。**2018年B站在纳斯达克上市，公司继续加强多元变现并持续破圈。2020年第四季度，B站的MAU超过2亿，新增用户

平均年龄为20岁，覆盖约60%~70%的Z世代用户。B站成为Z世代兴趣、娱乐、社交和学习的聚集地。

● **第四阶段：文化社区和PUGV（专业用户生成的视频，professional user-generated video）平台（2021年至今）。**随着2020年B站破圈三部曲（《后浪》《入海》《喜相逢》）的播出和推广，以及平台加大对各领域UP主（内容创作者）的扶植（例如陈翔六点半、何同学、罗翔说刑法等UP主都为B站带来了大量的用户和流量），B站影响力不断扩大，用户群体越来越宽泛。B站逐渐模糊其二次元标签，从以二次元文化为主的平台，向更广泛的文化社区破圈。

从公开财报数据中，我们可以看到随着B站不断地破圈，其用户规模也在不停地增长：2020年第四季度MAU突破2亿人次，2022年第二季度MAU突破3亿人次，2023年第三季度DAU突破10亿人次（见图4-6）。

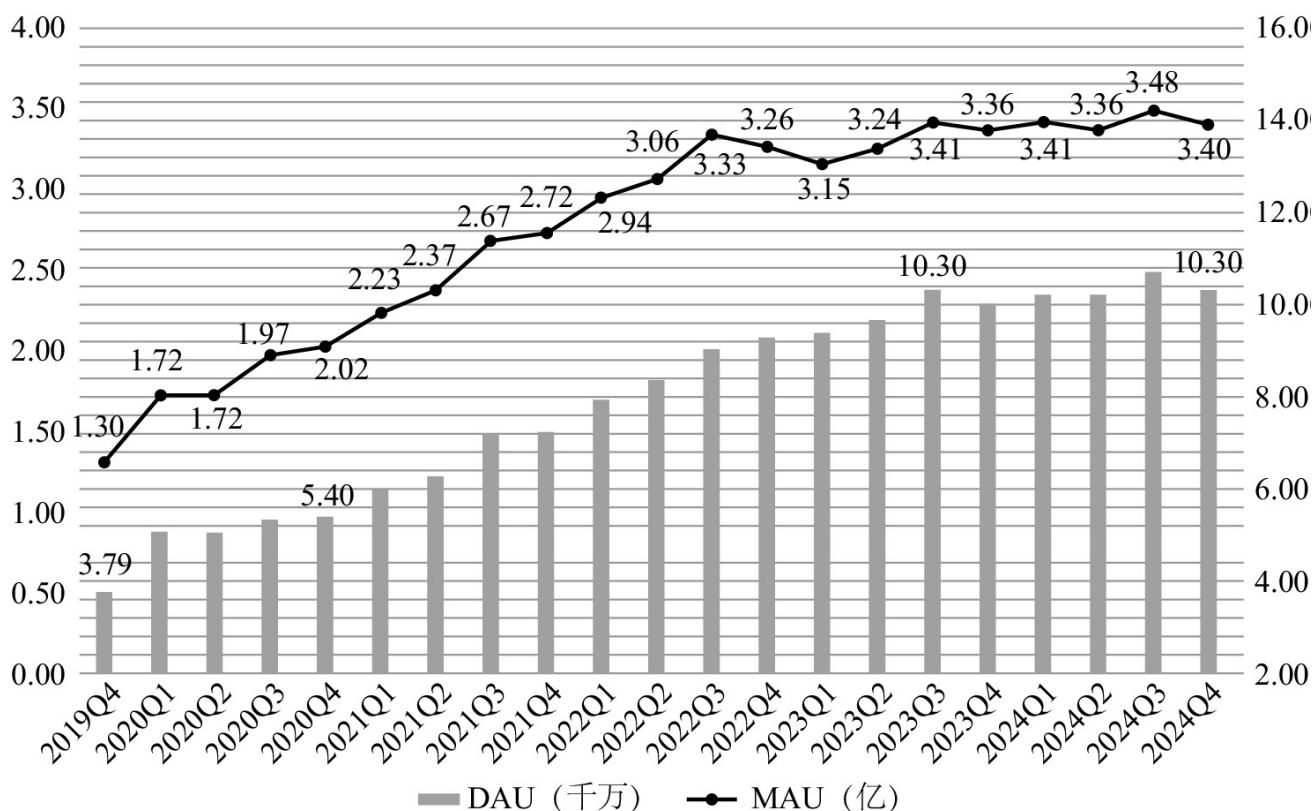


图4-6 2019年第四季度到2024年第四季度哔哩哔哩用户规模趋势

社区是B站的基因，内容生态是B站的“护城河”，PUGV是内容生态系统的基石。

B站目前的内容生态包括PUGV（一般为中长视频）、OGV（专业机构生产的视频，occupationally generated video，一般为偏长视频）、短视频以及直播等不同时长和消费场景的视频内容，且覆盖生活方式、游戏、娱乐、动漫、知识、母婴、汽

车和时尚等多个领域。其中，核心是PUGV创作生态，即由UP主创作的视频，2024年B站上超过96%的视频浏览量来自PUGV。B站上的PUGV通常是UP主在相关领域的知识和专长的输出，大部分都是原创视频且极具创造力，所以B站能为用户提供高质量的内容。

经过过去十几年的深耕和沉淀，B站拥有了一个深度、多样且数量不断增长的PUGV库。2024年，B站上每月活跃的创作者大约有400万人，平均每月提交的视频大约为2 070万个，全年有超过310万名UP主在B站获得收入。

B站拥有独特的社区文化和氛围，社区黏性足、活跃度高、留存度高。B站的社区以“UP主（内容）用户”为基石：

- **UP主生产内容：**大量的UP主不断生产优质内容，吸引用户。
- **经过筛选的用户：**B站会通过邀请码、答题等方式来筛选用户，而其在发展初期的筛选机制更为严格，这在一定程度上维护了社区氛围，提高了用户黏性。截至2024年第四季度，B站正式会员数量大约为2.58亿，12个月的留存率达到80%。
- **良性互动：**UP主与用户的紧密连接和互动，独特的“一键三连”机制，进一步加强了良性互动，形成以兴趣为圈层的内容生态网络。

UP主和用户，都可以在B站得到正反馈，实现内容和用户的飞轮效应。但是社区也需要商业，UP主不可能持续靠情绪价值“为爱发电”，只有有好的商业才能支撑UP主持续生产优质的内容。

B站的创立和发展都和二次元密切相关，商业化也不例外。B站的商业化始于2014年的二次元游戏代理业务，后来不断拓展至会员、广告、直播和IP衍生品等。2018年B站上市时，手游的收入能占到公司总收入的约70%（2018财年B站总收入41亿元，其中手游收入29亿元）。

“你感兴趣的视频都在B站。”从小众的ACG社区，走向更主流的综合视频社区，所谓万物皆可B站，几亿用户在这里通过视频建立连接和互动，共同维护着一个友好的社区。正如B站的使命所言：为用户创造一个美好的社区，为创作者搭建一个舞台。

1. 哔哩哔哩的财务数据（来自哔哩哔哩集团财报公开数据）

从收入结构上看，B站的收入已经不像刚上市时那么依赖游戏了。近几年B站的收入主要依靠增值服务（付费会员、直播、漫画及其他付费内容）、广告、手游（自研、代理发行、联运等）、IP衍生品及其他等业务。以2024年数据为例，增值服务、广告和手游的营收占比分别为40.99%、30.52%和20.91%（见表4-6）。

表4-6 2022—2024年哔哩哔哩营收构成

营收构成	2022年		2023年		2024年	
	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)	营收 (亿元)	占比 (%)
增值服务	87.15	39.80	99.1	43.99	109.99	40.99
广告	50.66	23.13	64.12	28.46	81.89	30.52
手游	50.21	22.93	40.21	17.85	56.10	20.91
IP衍生品及其他	30.96	14.14	21.85	9.70	20.33	7.58
总计	218.98	100.00	225.28	100.00	268.31	100.00

看营收和利润数据，B站2024年总营收为268.32亿元，连续多年保持增长；看利润数据，B站却是连续多年亏损，2024年持续经营利润为-13.64亿元（见表4-7）。但好消息是亏损在收窄，作者预计B站很快就能实现扭亏为盈。

表4-7 2019—2024年哔哩哔哩营收和利润

项目	2019年	2020年	2021年	2022年	2023年	2024年
总营收（亿元）	67.78	119.99	193.84	218.99	225.28	268.32
营收年同比增长率（%）	—	77.03	61.55	12.97	2.87	19.11
营业利润（亿元）	-14.95	-31.41	-64.30	-83.58	-50.63	-13.44
持续经营利润（亿元）	-13.04	-30.54	-68.09	-75.08	-48.12	-13.64

从成本结构上看，B站的主营业务成本非常高，导致毛利率偏低，2024年毛利率仅为32.70%。这主要是因为B站需要发放大量的创作者（UP主）激励，以及支付OGV内容的制作成本。以2024年数据为例：B站的收入分享成本约为108亿元，内容成本约为27亿元，分别占主营业务成本的59.83%和15.12%（见表4-8）。

B站要想扭亏为盈，必须提高商业化能力。虽然增值服务收入还有一定的提升空间（B站大会员的付费率和ARM均低于行业平均水平），但本书第2章分析过会员模式的利弊，所以B站必须提升广告和手游的收入占比，尤其是广告的变现能力。

表4-8 2022—2024年哔哩哔哩主营业务成本构成情况

项目	2022 年		2023 年		2024 年	
	成本 (亿元)	成本占比 (%)	成本 (亿元)	成本占比 (%)	成本 (亿元)	成本占比 (%)
主营业务成本	180.50	100.00	170.86	100.00	180.58	100.00
收入分享成本	91.15	50.50	95.07	55.64	108.04	59.83
内容成本	34.97	19.37	31.96	18.70	27.30	15.12
服务器和带宽成本	17.53	9.71	14.77	8.65	16.44	9.10
IP 衍生品及其他成本	36.85	20.41	29.06	17.01	28.80	15.95

2. 哔哩哔哩的商业模式解读

从前述收入结构来看，B站的商业模式是一元和三元商业模式的组合，而且目前一元商业模式的收入贡献更高。一元商业模式在本书第2章已经介绍过了，本部分我们重点看B站的三元商业模式，即B站的广告部分。

用户在B站消费视频内容时，B站投放广告，获取广告收入。与抖音类似，B站主要的广告场景也是信息流广告和搜索广告，并且2021年B站上线了竖版视频Story-Mode模式，进一步扩充了广告库存。下面的几个例子分别为B站推荐首页广告、视频播放页下方广告、竖版模式插楼广告以及搜索结果页广告（见图4-7）。



图4-7 B站广告示例

以2024年数据为例，我们具体看一下B站的用户规模（DAU）和广告变现效率（ARPU）。

B站2024年的DAU大约为1.04亿，则计算出B站的DAU ARPU约为0.22元。估算详情如表4-9所示。

表4-9 2023—2024年B站DAU ARPU估算详情

2023 年	预估数据	2024 年	预估数据
广告年总收入（亿元）	64.12	广告年总收入（亿元）	81.89
日均广告收入（亿元）	0.18	日均广告收入（亿元）	0.224
DAU（亿人）	0.98	DAU（亿人）	1.04
人均时长（分钟）	97	人均时长（分钟）	102
DAU ARPU（元）	0.18	DAU ARPU（元）	0.22
单时长 DAU ARPU（元）	0.0019	单时长 DAU ARPU（元）	0.0021

这意味着，在2024年，一个B站日活用户每天大约能给B站带来0.22元的广告收入，平均每用户每分钟带来的广告收入约为0.0021元。参考前述字节跳动的数据，这个DAU ARPU是偏低的，这既表明B站的广告收入还有提升空间，同时也说明B站这种中长视频社区对用户注意力的调度效率和对商业需求的激发效率并不高。

社区优先，这是B站企业价值观的第一条。因此B站在内容建设、破圈和商业化等方面的动作一直都比较克制。作为中国年轻一代的标志性品牌和优质视频社区，相信B站会“长”出好的商业，并持续为用户提供好的内容。

4.3.3 小红书：种草基地，UGC的集大成者，从PDF文档发展为生活百科全书

毛文超和瞿芳2013年6月带着真格基金几百万元的天使投资回上海创立了小红书，他们的第一个产品是一组名为《小红书出境购物攻略》的PDF文档，提供在日本、韩国、泰国、中国香港、新加坡、美国和法国购物的攻略和经验分享。根据毛文超的公开分享，他们最初成立了十个人左右的“小型编辑部”，毛文超（星矢）和瞿芳（木兰）合著了美国购物攻略，并找在新加坡的丁玲（柯南）友情撰写了新加坡购物攻略，邓超（樱木）负责了七个PDF的视觉设计和排版，还找朋友手绘了封面图。

这组PDF文档放在小红书网站上供用户下载，不到一个月，下载量就达到50万次。时至今日，攻略类的内容依然是小红书的核心竞争力之一。

经过低成本、快速的市场验证，毛文超与瞿芳带领团队在2013年底上线了小红书App，正式成立购物分享社区，主打海外购物的UGC分享。用户可以在小红书上分享和交流自己的海外购物经历、经验、折扣、地点和避坑指南等各种信息。UGC模式也奠定了小红书发展的基础和社区基因。

到了2014年，小红书团队意识到只做购物经验分享，没法满足用户直接购买的需求。团队经常被问到的问题便是：“看到小红书上推荐的商品，怎么买？”即，用户在小红书上被“种草”，但没法在小红书上“拔草”，看到想买的东西只能再去找其他渠道买，甚至线下找代购。于是小红书决定做跨境电商，并于2014年12月上线电商平台“福利社”。

但跨境电商竞争异常激烈，各大电商纷纷入局，阿里巴巴的天猫国际、网易的考拉、京东的京东全球购、苏宁的苏宁海外购等先后上线。小红书的福利社最开始采取买手模式，在SKU、供应链、价格等方面都没有优势，甚至用户还养成了“在小红书‘种草’，到其他平台‘拔草’”的习惯。

2017年小红的电商业务陷入困境，GMV始终跨不过100亿元大关，而且进口税新规进一步冲击业务，之前很多的税率优惠被取消了，小红书主营的美妆、奶粉等品类监管加强，小红书面临着前所未有的挑战。

2017年下半年，小红书开始转型，重回UGC社区模式，开始提升社区活跃度和用户规模。2018年小红书开始邀请明星入驻，分享日常生活，并且赞助《偶像练习生》《创造101》等综艺，吸引了大量的粉丝用户。2019年10月，小红书MAU突破

1亿；2023年底，小红书DAU突破1亿。UGC社区、分享生活方式、兴趣互动，构成了现在的小红书。

回顾小红书的发展史，大概可以分为三大阶段。

● **第一阶段：创业探索阶段（2013—2014年）。**从静态PDF购物指南，到动态UGC购物分享社区，小红书逐渐完成购物社区的搭建，用户数量超过千万，奠定了小红书发展的基础和社区基因。

● **第二阶段：转型为跨境电商社区阶段（2015年至2017年上半年）。**随着福利社的上线，小红书开始做跨境电商业务，主打买手模式，在社区形成交易闭环，但在市场竞争和政策新规的挑战下增长乏力。

● **第三阶段：转型为UGC社区的快速发展阶段（2017年下半年至今）。**随着2017年下半年公司转型，小红书明确定位，重回UGC社区模式，在明星引流和综艺广告投放的加持下，用户规模快速增长；业务上逐步布局视频、直播和电商等方向，进一步完善生态，形成了如今的小红书。

自成立以来，小红书一直备受资本青睐，截至2024年第三季度，小红书共完成七轮融资，总融资金额大约10亿美元，在最近2024年7月的E+轮融资时，投后估值达170亿美元（见表4-10）。

表4-10 截至2024年第三季度小红书的历次融资情况

融资时间	融资轮次	融资金额	投资机构
2013年10月	天使轮	数百万元	真格基金
2014年6月	A轮	数百万美元	真格基金、金沙江创投
2015年6月	B轮	数千万美元	金沙江创投、纪源资本
2016年3月	C轮	1亿美元	腾讯投资、天图投资、元生资本
2018年6月	D轮	超3亿美元	阿里巴巴、真格基金、天图投资、元生资本、纪源资本、腾讯投资等
2021年11月	E轮	5亿美元	淡马锡、腾讯投资、阿里巴巴、天图投资、元生资本等
2024年7月	E+轮	—	DST Global、高瓴资本、博裕资本、中信资本和红杉中国等

作为一个分享生活方式的UGC社区，用户在小红书发布的内容均来自真实生活。多元、长尾和高时效性是小红书社区内容最显著的特点。随着内容的积累和不断更新，小红书越来越好地承接住了大量用户的长尾需求，成为一本“活着的生活百科全书”。所谓“遇事不决小红书”，小红书已经演变成“年轻人的搜索引擎”。

根据小红书发布的《2024小红书10大搜索趋势洞察报告》，超过70%的月活用户使用小红书进行搜索，单用户平均每天搜索六次，超过30%的用户每天打开小红书的第一件事就是搜索。行业调研数据显示，2024年第四季度，小红书日均搜索PV（页面浏览量，page view）接近6亿次。

小红书以“Inspire Lives，分享和发现世界的精彩”为使命，打造了一个生活方式社区。用户可以在这个社区里通过图文和视频等形式记录生活点滴，分享生活方式，并基于兴趣互动，形成人与人的连接。

就商业模式来说，小红书在做跨境电商的时候，采用的是一元商业模式；重回UGC社区模式后，核心是三元商业模式。用户在小红书上消费内容，平台通过展示广告变现。小红书目前的收入主要来源于广告、电商和直播三个部分，2023年其广告收入占比超过70%。

与抖音类似，小红书主要的广告场景也是App中的信息流广告和搜索广告，包括小红书首页推荐广告、搜索结果页广告以及插楼广告等。由于小红书目前的用户群体中女性偏多，因此其广告收入以电商收入为主，商品品类主要为美妆、奢侈品、服饰和母婴用品等，而且多以软广形式呈现，小红书达人商单模式（蒲公英平台）的广告收入占比超过40%。

说明：行业研究信息仅用来辅助读者更好地理解小红的业务模式以及定性观察其业务发展趋势。具体的数字不作为任何建议和其他用途。

因为小红书目前还没有上市，所以没有官方公开的财务报表数据。行业调研数据显示，2023年小红书全年收入为266亿元，并首次实现全年盈利，利润为36亿元；2024年全年收入超过300亿元，利润超过72亿元（见表4-11）。

表4-11 2022—2024年小红书收入情况 单位：亿元

	2022 年	2023 年	2024 年
收入	144	266	300+
广告收入（预估占比 70%）	101	186	210+

以2024年数据为例，我们具体看一下小红的用户规模（DAU）和广告变现效率（ARPU）。

小红书2024年的DAU按照1.2亿来估算，广告收入按照210亿估算，则计算出小红的DAU ARPU为0.48元左右（见表4-12）。

表4-12 2024年小红书DAU ARPU估算详情

项目	数据
广告年总收入（亿元）	210
日均广告收入（亿元）	0.57
DAU（亿人）	1.2
人均时长（分钟）	80
DAU ARPU（元）	0.48
单时长 DAU ARPU（元）	0.006

这意味着一个小红书的日活用户，每天能给小红书带来超过0.48元的广告收入；平均每用户每分钟带来的广告收入超过0.006元。参考前述字节跳动以及B站的数据，这个DAU ARPU是不低的。

从最早的一组PDF版出境购物攻略，发展为现在DAU过亿的UGC社区，作为一家企业，小红书一直在做战略调整并不断成长，而作为一个UGC社区，小红书也一直在自然生长和进化，充满生命力。就如同飞盘、露营、骑行、垂钓、陆冲、城市漫步等潮流的生活方式一样，小红书生长，破圈，充满生命力。我想这就是UGC社区的生命力，这就是人的生命。

4.3.4 Facebook（Meta）：从哈佛大学宿舍里走出来的全球最大的社交网络平台

Facebook（2021年改名为Meta，全称Meta Platform Inc.，为了方便表达，本小节统一写为Facebook），这个目前MAU约40亿的全球最大的社交网络平台，是从哈佛大学宿舍里走出来的。

2004年2月，马克·扎克伯格（Mark Zuckerberg）和他的室友达斯廷·莫斯科维茨（Dustin Moskovitz）、克里斯·休斯（Chris Hughes）一起，在哈佛开发了Facebook。最开始Facebook的模式类似于国内的校内网，是一个PC端校友录网站，支持哈佛在校生社交，大家可以把个人信息、照片和动态等放到Facebook上，并与朋友们在线上相互关注、交流和互动。Facebook很快就风靡哈佛校园，几周内就有四分之三的学生注册使用，然后逐步扩展到斯坦福、耶鲁、杜克等学校。随着Facebook用户越来越多，扎克伯格和莫斯科维茨辍学离开哈佛，搬到硅谷开始创业，并获得PayPal创始人彼得·蒂尔（Peter Thiel）大约50万美元的天使投资。

在互联网红利的加持下，Facebook快速发展，并逐步跳出校园，向全部用户开放，仅三年多的时间就成长为全球访问量前十的网站和全球访问量第一的图片分享网站。2007年5月，Facebook推出开放平台（Facebook Platform），第三方开发者可以通过API开发能在Facebook网站上运行的应用程序，这为Facebook的生态建设奠定了基础。2010年1月，Facebook超越雅虎成为美国访问量排名第二的网站，仅次于Google。2012年5月，Facebook用户超9亿，并成功在美国纳斯达克挂牌上市。

随着移动互联网时代的到来，Facebook在产品移动化转型的同时，完成了两个非常关键的收购，巩固自己全球社交网络巨头的地位，并扩大了自身优势：2012年9月以10亿美元现金和股票收购Instagram，2014年2月以190亿美元收购WhatsApp。这两笔收购从交易到后续的发展，都体现了Facebook团队的战略高度和远见。如今的Facebook在社交网络方向布局非常完善，竞争壁垒高，其在全球超过40亿的月活用户主要由四款产品贡献（见表4-13）。

表4-13 Facebook的四款产品

产品	MAU	产品简介
Facebook	30 亿	熟人社交，形态类似校内网+微信公众号+微博
Instagram	22 亿	内容社区，形态类似小红书，用户年轻化。Instagram 上线的 Reels 也是 Facebook 在短视频领域布局的最核心产品
WhatsApp	20 亿	即时通信，形态类似微信的即时通信功能+朋友圈
Messenger	13 亿	即时通信，形态类似 QQ

在不断发展、扩张以及应对市场变化的过程中，Facebook布局了短视频（主要是为了对抗TikTok）、直播、电商、元宇宙、大模型以及VR设备等业务。2021年10月28日，扎克伯格在Facebook Connect大会上宣布将Facebook更名为Meta（源自元宇宙的英文 metaverse），以示公司大力发展元宇宙的决心。

但不论在元宇宙上的决心如何，截至目前Facebook依然是以社交网络为基石的“广告公司”：2024年Facebook广告收入超过1 600亿美元，仅次于Google，是全球收入第二高的数字广告公司（但字节跳动的广告收入增长迅猛，正在逐步缩小与Facebook的差距，未来有望超过Facebook）。

回顾Facebook的发展历程，其在PC互联网时代崛起，在移动互联网时代顺利转型并扩张，在短视频时代积极应对，在大模型时代积极布局，一步步从一个校园网站成长为全球互联网企业巨头，其产品矩阵在全球的DAU超过30亿（见图4-8），离不开时代的红利、商业模式的效率以及战略上的判断和选择。

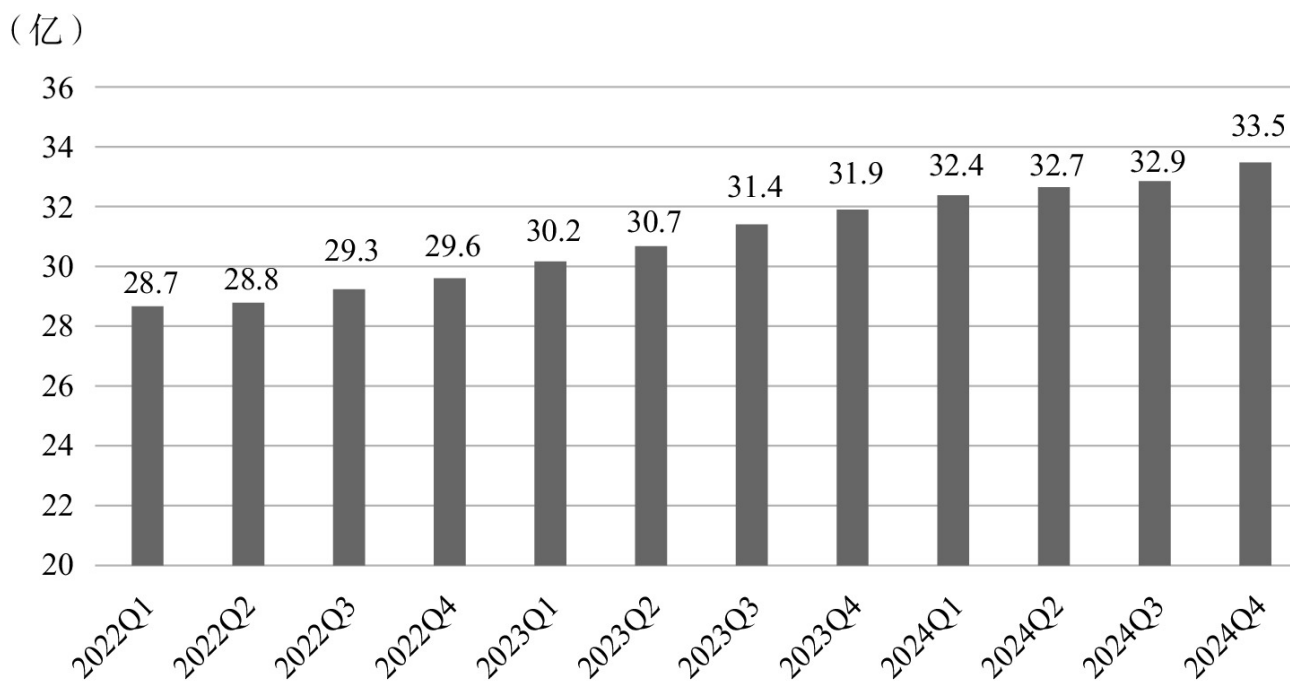


图4-8 2022—2024年Facebook产品矩阵全球DAU

1. Facebook的财务数据（来自Facebook集团财报公开数据）

从收入结构上看，Facebook的主要收入来自广告。Facebook的收入来源分为应用程序和现实实验室两大类，其中应用程序部分分为广告收入和其他收入（见表4-14）。下面以2024年的数据为例：

- **应用程序-广告收入：**来自Facebook、Instagram、Messenger和第三方App上的广告收入，年收入达1 606.33亿美元，占比高达97.65%。
- **应用程序-其他收入：**主要为WhatsApp商业平台的支付抽佣等收入，年收入为17.22亿美元，占比1.05%。
- **现实实验室：**来自Reality Labs的Meta Quest等可穿戴设备以及相关软件和内容等消费类硬件产品的收入，年收入为21.46亿美元，占比1.30%。

表4-14 2022—2024年Facebook的营收构成

营收构成		2022 年		2023 年		2024 年	
		营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)
应用程序	广告	1 136.42	97.46	1 319.48	97.81	1 606.33	97.65
	其他	8.08	0.69	10.58	0.78	17.22	1.05
现实实验室		21.59	1.85	18.96	1.41	21.46	1.30
总计		1 166.09	100.00	1 349.02	100.00	1 645.01	100.00

因为Facebook是全球化的公司，所以收入来源于全球各地，其中美国占比最高：2023年在美国的收入达497.80亿美元，占比36.90%；2024年在美国的收入达597.30亿美元，占比36.31%（见表4-15）。

表4-15 2023—2024年Facebook营收的区域构成

区域构成	2023 年		2024 年	
	营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)
美国	497.80	36.90	597.30	36.31

续表

区域构成	2023 年		2024 年	
	营收 (亿美元)	占比 (%)	营收 (亿美元)	占比 (%)
欧洲	312.10	23.14	383.61	23.32
亚太	224.64	16.65	266.59	16.21
其他地区	146.50	10.86	179.24	10.90
中国	136.90	10.15	183.50	11.15
加拿大	31.08	2.30	34.77	2.11
总计	1 349.02	100.00	1 645.01	100.00

看营收数据，过去几年Facebook的总营收整体依然在增长，2024年全年为1 645.01亿美元，如前文所言，其中主要是广告收入。从利润来看，Facebook的利润情况良

好，已经持续十几年处于盈利状态，2024年持续经营利润达623.60亿美元（见表4-16）。

表4-16 2019—2024年Facebook营收和利润

项目	2019年	2020年	2021年	2022年	2023年	2024年
总营收（亿美元）	706.97	859.65	1 179.29	1 166.09	1 349.02	1 645.01
营收年同比增长率（%）	—	21.60	37.18	-1.12	15.69	21.94
持续经营利润（亿美元）	184.85	291.46	393.70	232.00	390.98	623.60

从上述财务数据中，可以明显感受到Facebook作为全球数字广告龙头企业的盈利能力和效率。Facebook如此强的赚钱能力也是其在元宇宙、可穿戴设备和大模型等方向多元布局的基础。

2. Facebook的商业模式解读

Facebook采用的是三元商业模式。如前文所言，Facebook的广告收入占比接近98%。用户在Facebook的矩阵产品——包括但不限于Facebook、Instagram和Messenger——中消费内容，平台通过广告变现，获取广告收入，甚至信息流广告的形式也是2006年由Facebook最先推出的。与字节跳动类似，Facebook主要的广告场景也是其矩阵产品中的信息流广告、搜索广告以及第三方App中的网盟广告。

Facebook从2024年财报开始调整DAU ARPU的定义，并不再披露分区域的ARPU数据；2024年财报显示其年度全球DAU ARPU为49.63美元，比2023年增长了15%。

我们可以具体看一下Facebook的变现效率。由于财报中没有拆分各产品的DAU和收入数据，因此我们将Facebook的矩阵产品当作一个主体，以财报中披露的2023年第四季度数据为例进行分析。2023年第四季度Facebook单用户的季度ARPU约为13.12美元，但区域差异极大：美国和加拿大高达68.44美元，欧洲高达23.14美元，而亚太和其他地区仅为约5美元（见表4-17）。

表4-17 2023年第四季度Facebook ARPU情况 单位：美元

区域构成	季度 ARPU	日均 ARPU
整体	13.12	0.14
美国和加拿大	68.44	0.74
欧洲	23.14	0.25
亚太	5.52	0.06
其他地区	4.50	0.05

这意味着一个Facebook矩阵产品的用户，每个季度大约能给Facebook带来13.12美元的广告收入，平均每用户每天带来的广告收入约为0.14美元，折合人民币约1元。

Facebook作为目前全球最大的社交网络平台，对现代社会产生了深远的影响。它不仅改变了人们的沟通方式，还重塑了信息传播、广告营销以及社交互动的格局。通过提供即时通信、分享生活点滴和连接全球用户的功能，Facebook构建了一个庞大的数字网络和社区。但是随着用户隐私、虚假新闻以及内容监管等方面挑战的出现，Facebook也面临着公众信任和社会责任方面的考验。正如李普曼在《舆论》中写道：“这个世界上，有腐败也有清廉，有混沌也有奇迹，有弥天大谎也有不惜一切去戳穿谎言的人。”希望Facebook在面对全球几十亿用户的时候，能创造更多的奇迹，而非混沌。

4.4 小结

本章我们介绍了三元商业模式的业务逻辑、核心要素、画布和经营分析框架，并分析了几个代表性的公司供读者参考。

有很多用户不喜欢广告，但实际上，广告是三元商业模式的基础，效率很高。而且计算广告学本身也是一门由信息科学、统计学、计算机科学以及微观经济学等融合而成的交叉学科，感兴趣的读者可以参考作者的另外一本书《计算广告：产品、机制与大模型应用》。

我们免费使用的各种App并不是真正的免费，我们在这些App上停留的每一分钟、观看的每一条内容，都在为平台贡献广告收入。这就是三元商业模式的价值。

篇幅所限，加之前述研究中的字节跳动和小红书目前还没有上市，缺乏公开财报数据，所以还有很多数据和信息没有展开。本章提到的相关数据主要来自对应公司的财报文件、行业交流、实践验证、行业研究报告、公司官网等，并进行了逻辑推算和交叉验证，仅供读者做商业模式分析时参考。如有数据错误，还请读者指正。

免责声明：本书提到的所有公司，仅代表作者个人观点，不构成任何投资建议。

第5章 大模型简介及其商业模式研究₂₀₂₂

年11月30日，OpenAI发布了一款基于GPT模型开发的对话式聊天机器人ChatGPT。ChatGPT具有强大的理解能力、逻辑推理能力、记忆能力以及生成能力，用户可以通过自然语言与之交互，让ChatGPT完成撰写文章、生成摘要、回复邮件、翻译、生成代码、解数学题、玩游戏等复杂的任务，ChatGPT还可以根据聊天的上下文与用户进行互动。ChatGPT可以像人类一样与用户进行交流，用户则可以将ChatGPT当作私人助理，给它布置任务。ChatGPT的发布直接引爆了全行业，刷新了我们的认知，让人工智能再次成为全行业的焦点，并在国内外掀起了一场大模型浪潮。

更进一步，DeepSeek（由幻方量化孵化的杭州深度求索人工智能基础技术研究有限公司）在2025年1月20日发布开源推理模型DeepSeek-R1。DeepSeek-R1在多个数学和推理基准测试中击败了OpenAI-o1等业界领先模型，并且成本优势显著。DeepSeek-R1的发布再一次引爆全球科技圈，并且改变了行业竞争格局和叙事逻辑。

人工智能上一次引发全球如此大的关注，还是2016年Google DeepMind团队研发的人工智能机器人AlphaGo（阿尔法围棋）击败人类职业围棋选手。2016年1月27日，国际顶尖期刊《自然》在封面文章报道，AlphaGo在没有任何让子的情况下，以5：0完胜欧洲围棋冠军、职业二段选手樊麾，这是计算机第一次在不让子的完整围棋竞技中击败专业选手，是人工智能领域历史性的突破。2016年3月9日到15日，AlphaGo在韩国首尔与世界围棋冠军李世石进行围棋人机大战，AlphaGo以4：1完胜。2017年5月23日到27日，AlphaGo又在中国乌镇围棋峰会上，以3：0战胜当时排名世界第一的世界围棋冠军柯洁。围棋界公认AlphaGo的围棋能力已经超越人类职业选手的顶尖水平。

不论是2022年ChatGPT带动的这一轮大模型浪潮，还是2016年AlphaGo吸引了广泛关注，都是人工智能发展史上的标志性事件。虽然人工智能并非全新的研究领域，但大模型的发展带来了很多新概念，其应用案例和商业模式值得思考和研究，本章我们就展开介绍一下这些话题。

5.1 大模型相关的基础知识介绍^{5.1.1}

人工智能、机器学习和深度学习

人工智能（artificial intelligence, AI）是指由人类制造出来的机器所表现出来的智能。美国麻省理工学院（MIT）的帕特里克·温斯顿（Patrick Winston）教授认为：“人工智能就是研究如何使计算机去做过去只有人类才能做的智能工作。”即人工智能致力于让计算机系统能够执行类似于人类智能才可以执行的任务，使机器具备像人类一样的学习、推理、决策和创造能力。人工智能是一个非常广泛的科学领域，包括机器人、语言识别、图像识别、自然语言处理、专家系统、机器学习和计算机视觉等。

机器学习（machine learning, ML），是人工智能的一个分支，也是人工智能的核心。机器学习是通过数据和模型来让机器自动学习和改进的一种方法，专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构并不断迭代。机器学习是一门多领域交叉学科，涉及概率论、统计学、凸分析、算法理论等。机器学习不是某种具体的算法，而是很多算法的统称。

深度学习（deep learning, DL）是机器学习的一个分支，特指基于深层神经网络模型和方法的机器学习，它是在统计机器学习、人工神经网络等算法模型的基础上，结合当代大数据和大算力发展而来的。深度学习的核心是深度神经网络（deep neural network, DNN）。与其他机器学习方法相比，深度学习最大的特点是它具有自动提取特征的能力。特征表示和提取是许多机器学习方法的基础。在前深度学习时代，特征工程的工作是很重的，需要由专业的技术人员结合行业知识来设计和提取特征，而深度学习的强大之处便在于可以将特征工程的工作交给机器自动完成。2024年10月8日，被誉为“深度学习之父”的杰弗里·辛顿（Geoffrey Hinton）获得了诺贝尔物理学奖，再次勾起了行业对神经网络发展的回忆。

人工智能不等于机器学习，机器学习不等于深度学习。人工智能是一个更广泛的概念，机器学习是实现人工智能的核心方法之一，而深度学习是机器学习的一个重要分支。

人工智能经历了漫长的发展过程。

艾伦·图灵（Alan Turing，被誉为“人工智能之父”）1950年发表了一篇划时代的论文——《计算机与智能》（Computing Machinery and Intelligence），提出了著名的“图灵测试”。

1956年，美国达特茅斯学院举行了历史上第一次人工智能研讨会，会上麦卡锡首次提出了“人工智能”的概念，标志着人工智能的诞生。

1966年，美国麻省理工学院发布了世界上第一个真正意义上的聊天机器人ELIZA（伊莉莎），ELIZA可以通过脚本理解简单的自然语言，并与人类进行互动。

1997年，IBM开发的深蓝超级计算机（DeepBlue Computer）挑战并击败国际象棋大师加里·卡斯帕罗夫（Garry Kasparov），成为首个在标准比赛中击败国际象棋世界冠军的电脑系统。深蓝超级计算机重1 270千克，有32个微处理器，每秒钟可以思考2亿步棋。

2011年，IBM和美国得克萨斯大学历时四年联合打造的超级计算机Watson（沃森），在美国最受欢迎的智力竞猜电视节目之一《危险边缘》中，击败两位人类选手。

DNN Research是由杰弗里·辛顿（Geoffrey Hinton）与他的两位学生Alex Krizhevsky和Ilya Sutskever（Ilya就是后来 OpenAI的联合创始人）创立的公司，当年通过拍卖的方式最终以4 400万美元卖给了Google。百度是当年那场著名拍卖中的有力竞争者，而且出价比Google高。

2013年，深度学习算法被广泛运用在产品开发中。旨在探索深度学习领域的Facebook人工智能实验室成立，Google收购了语音和图像识别公司DNN Research^①，百度创立了全球首个深度学习研究院（Institute of Deep Learning，IDL）。

2016年，便是本章开始提到的AlphaGo横空出世，在围棋人机大战中战胜围棋世界冠军李世石，人工智能开始新一轮的爆发。AlphaGo的主要工作原理就是深度学习。

2017年，Google在论文“Attention Is All You Need”中提出了Transformer模型，使AI的性能得到质的飞跃，这也成为人工智能领域的重大突破。Transformer已经成为很多大模型的核心模块，GPT中的“T”就是指Transformer。

2022年，OpenAI发布ChatGPT，堪称人工智能发展史上具有跳跃性的一步，开启了大模型时代，也在全球范围内开启了一轮“百模大战”。Google的Gemini、百度的文心一言、Facebook的LLaMA、字节跳动的豆包、阿里巴巴的通义千问、OpenAI

的Sora、月之暗面（Moonshot AI）的Kimi等各种大模型如雨后春笋般涌现。2022年被称为大模型元年。

之后，2025年DeepSeek发布了其开源推理模型DeepSeek-R1。DeepSeek-R1使用深度思考替代“大力出奇迹”的暴力堆参数，大大降低了算力成本。而且DeepSeek-R1的开源进一步降低了AI技术的使用门槛，使得更多开发者能够自由部署和使用高性能的大模型，大大加速了AI技术在各行业的应用和普及。DeepSeek-R1改变了大模型行业的叙事逻辑，甚至毫不夸张地说，其以一己之力改变了中美在大模型时代的竞争格局。美国科技媒体更是评价DeepSeek-R1为“AI领域的Linux革命”。

人工智能的发展是一部还在续写的、波澜壮阔的历史，也是一群少数派不懈探索与坚持的历史。感兴趣的读者可以阅读凯德·梅茨（Cade Metz）2022年出版的图书《深度学习革命》（*Genius Makers: The Mavericks Who Brought AI to Google, Facebook, and the World*），看看那群将深度学习带给全世界的企业家和科学家的故事。

5.1.2 大模型、大语言模型、生成式AI和AIGC

1. 大模型

大模型（large model），是指具有超大规模参数（通常在十亿个以上）和复杂计算结构的机器学习模型。一般而言，有上万个参数的函数就可以称为模型，因此广义而言，大模型的“大”指的就是模型参数规模：参数规模特别大、训练数据特别多的模型就是大模型。只是随着技术的发展，“大”的标准也越来越高，最开始亿级别的参数就可以算大模型了，现在一般需要十亿级别的参数才能算大模型。OpenAI的GPT就是典型的大模型。GPT-3的参数规模约为1 750亿，GPT-4的约为1.7万亿，DeepSeek-R1满血版也拥有6 710亿个参数。随着参数变多，大模型的特点也显而易见。

- **参数规模超大：**毋庸置疑，这是大模型最显著的特点。参数规模超大可以让大模型更好地学习数据之间的复杂关系，提高大模型的预测和泛化能力。
- **训练复杂度和成本高：**由于参数规模超大，大模型训练需要消耗大量的计算资源和时间，例如GPT的训练需要上万张H100 GPU芯片，仅芯片成本就在亿美元级别。同时，想要提升大模型的能力，使其不断学习和优化，还需要依赖海量数据进行训练。
- **泛化能力强：**超大规模参数和海量的训练数据，可以让大模型学习到丰富的知识和特征表示，进而在面对新的、没见过的数据时表现出强大的泛化能力。
- **多模态融合：**大模型支持包括文本、音频、图像和视频等在内的多模态数据处理，进而可以实现跨模态的理解和生成。
- **涌现能力：**大模型的涌现能力是指，当模型的规模（如参数数量、模型层数等）达到一定阈值时，它会展现出一些难以预测的新特性和能力。这些新特性和能力并不是通过直接编程得到的，而是在训练和学习过程中自然产生的。涌现能力是大模型的一个显著特点，它体现了大模型处理复杂任务的潜力和能力。可以简单理解为，以前你想让机器学会什么技能，你就教它什么技能；当大模型出现智能涌现之后，你没教过的技能它也会了，就和人类一样。

大模型的四个核心能力是理解、生成、逻辑和记忆。

- **理解能力：**大模型能够准确理解复杂、模糊的提示词，包括乱序、意图模糊、含有潜台词的提示词等。

- **生成能力**：大模型能够根据给定的素材或提示，迅速生成多模态的内容，如文字、图片、视频等。
- **逻辑能力**：大模型拥有强大的逻辑推理能力，能够分析和解决复杂的问题，如数学推导或科学原理解释。
- **记忆能力**：大模型能够记住之前的对话内容，并在后续对话中保持一致性。这种记忆能力对于多轮对话的连贯性和一致性至关重要。

与大模型相对应的是小模型，小模型通常是指参数规模较小的模型，具有轻量、高效和易部署的特点，适用于数据量较小、计算资源有限的场景。在实际应用中，我们可以通过**模型蒸馏（model distillation）**的方法将一个大型模型的知识迁移到一个小模型，使小模型在保持较小规模的同时，拥有与大模型相似的性能。

DeepSeek团队在其论文“**DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning**”中明确提出较大模型的推理模式可以被蒸馏到相对小的模型中，而且比在小模型上进行强化学习（RL）得到的推理模式性能更好。该论文提到，使用DeepSeek-R1作为教师模型，生成约80万个训练样本，仅通过简单的监督微调（SFT）蒸馏DeepSeek-R1的输出，就能使蒸馏模型在多个推理基准测试中表现优异。例如，DeepSeek-R1-Distill-Qwen-7B在AIME 2024（2024年美国数学邀请赛）上的准确率为55.5%，高于QwQ-32B-Preview。而DeepSeek-R1-Distill-Qwen-32B在AIME 2024上的准确率为72.6%，在MATH-500（一个包含500道挑战性数学题的数据集）上的准确率为94.3%，在LiveCodeBench（代码生成）上的准确率为57.2%，这些结果都显著优于以前的开源模型，甚至好于OpenAI-o1-mini。DeepSeek团队的这些研究都显示出模型蒸馏的强大潜力。不同模型在基准测试中的表现对比如表5-1所示。

表5-1 不同模型在基准测试中的表现对比

模型	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	通过率	一致性	通过率	通过率	通过率	评分
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-ol-mini	63.6	80.0	90.0	60.0	53.8	1 820
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1 316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954

续表

模型	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	通过率	一致性	通过率	通过率	通过率	评分
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1 189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1 481
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1 691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1 205
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1 633

注：引用自DeepSeek-R1论文“Comparison of DeepSeek-R1 distilled models and other comparable models on reasoning-related benchmarks”；GPQA Diamond是一项针对高级问答能力的基准测试，主要用于评估模型在复杂专业问题上的表现；Codeforces是一个全球知名的在线编程竞赛平台，由俄罗斯科学家Mike Mirzayanov于2010年创立。

规模定律（Scaling Law）被誉为大模型第一定律，是指大模型的性能（例如损失函数）会随着计算量、数据规模和参数规模的增大而表现得更好。2020年，OpenAI在论文“Scaling Laws for Neural Language Models”中首次提出规模定律的概

念，并指出在Transformer架构下模型的表现与模型的规模之间呈现出幂律关系。自然语言处理中的规模定律见图5-1。

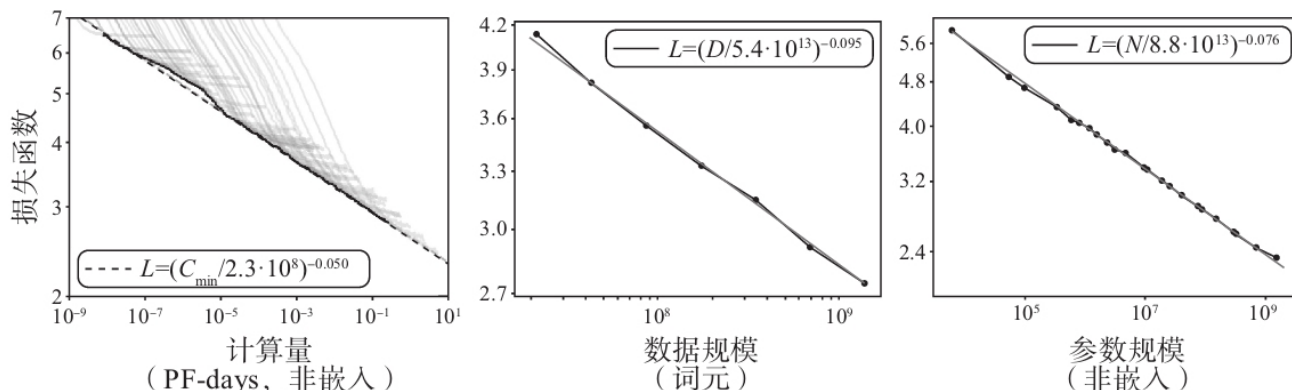


图5-1 自然语言处理中的规模定律

2022年Google DeepMind在论文“Training Compute-Optimal Large Language Models”中进一步完善了规模定律。DeepMind的研究认为，不能单纯通过扩大模型规模来提升性能（这样会浪费很多计算资源），在计算资源有限的情况下，要在模型规模和训练数据量之间取得一定的平衡，以此达到最优效果。DeepMind在论文中提到了一个有700亿参数的模型Chinchilla，验证了在给定计算资源下，如何通过调整模型大小和训练数据量来达到最优的训练效果。Chinchilla使用了与Gopher（一个有2800亿参数的模型）相同的计算资源，但参数数量是Gopher的四分之一，训练数据量是Gopher的四倍多，最终Chinchilla在多种任务（问答、常识推理、阅读理解等）中的表现均优于Gopher。例如，在MMLU（大规模多任务语言理解）基准测试中，Chinchilla的平均准确率达到67.6%，比Gopher高7.6个百分点。

Chinchilla和其他大模型的对比如表5-2所示。

表5-2 Chinchilla与其他模型规模大模型的参数和训练数据量对比 单位：亿

模型	模型规模	训练数据量
LaMDA (Thoppilan et al., 2022)	1 370	1 680
GPT-3 (Brown et al., 2020)	1 750	3 000
Jurassic (Lieber et al., 2021)	1 780	3 000
Gopher (Rae et al., 2021)	2 800	3 000
MT-NLG 530B (Smith et al., 2022)	5 300	2 700
Chinchilla	700	14 000

而DeepSeek-R1则将规模定律扩展到了后训练和推理阶段，使用思维链（Chain of Thought, CoT）等技术拓宽了规模定律的应用方式，开辟了新的增长曲线。

OpenAI、DeepMind和DeepSeek对规模定律的研究对于大模型的训练、发展和资源分配都具有重要的指导意义。目前，规模定律的研究与应用已经进入了新篇章，参数规模仍然重要，并且全面应用于预训练、后训练和推理这三个阶段。行业内普遍认为，推理阶段的算力消耗将成为未来需求的主导。当然，作者认为除了不断拓展算力、数据和参数规模的发展思路之外，在资源约束下不断进行策略优化和技术创新以提升效果和效率也至关重要。面对规模定律，我们要辩证地看问题。

大模型在处理复杂任务时展现出了卓越的能力和巨大的潜力。其在自然语言处理、计算机视觉和强化学习等多个领域都有着广泛的应用前景，并能有效地应对诸如机器翻译、内容生成、图像识别等实际复杂问题。但是大模型也面临着多重挑战，主要包括能耗、算力成本、数据隐私安全以及幻觉问题。

2. 大语言模型

大语言模型（large language model, LLM）是大模型的一种，指的是专门应用于自然语言处理领域的大模型，这一类大模型在语言处理方面会更加专业和高效率。大语言模型是大模型的子集，即所有的大语言模型都是大模型，但不是所有的大模型都是大语言模型。大语言模型通过海量的文本数据学习复杂的语言结构和语言使用方式，能够在文本生成、翻译、问答、生成摘要、分类等多种与语言相关的任务上有出色的表现。

3. 生成式AI

生成式AI (generative AI) 是人工智能的一个分支，指的是能够生成新内容的人工智能技术，即经过训练的机器学习模型可以自己生成新的内容，包括文本、图像、音频、视频、3D动画等。ChatGPT、Midjourney、文心一言、文心一格、Sora、MusicLM (音乐生成)、GitHub Copilot (代码生成) 等，都属于生成式AI的范畴。

与生成式AI对应的是辨别式AI (discriminative AI)。辨别式AI是人工智能的另一个分支，旨在直接学习并对输入数据与输出标签之间的映射关系进行建模。辨别式AI的目标是对输入数据进行分类或回归，并预测其对应的标签或属性。其代表性模型包括支持向量机 (SVM)、逻辑回归等。辨别式AI广泛应用于图像识别、语音识别、文本分类等领域，例如卷积神经网络 (CNN) 在图像识别领域取得了显著的成绩，循环神经网络 (RNN) 在语音识别领域表现出色。

生成式AI和辨别式AI的区别在于具体的目标和应用不同，训练方法也不同。生成式AI的目标是创造新的数据实例，而辨别式AI的目标是预测或决定数据的类别或标签。例如，模型根据用户的文字描述，画出一张猫的图片，就是生成式AI的能力；模型根据用户的一张照片，来判断照片中是猫还是狗，就是辨别式AI的能力。生成式AI通常需要从大量的数据中学习如何捕捉数据的分布特征，而辨别式AI则侧重于从标记的数据中学习决策边界。

生成式AI不等于大模型。大模型不局限于生成式任务，而生成式AI不一定需要大模型，一些小模型也可以完成生成式任务，尽管小模型可能在质量、性能和多样性上有所限制。

4. AIGC

AIGC是与UGC、PGC对应的概念，是指利用人工智能技术来自动（或辅助）生成各种形式（多模态）的内容，例如文章、小说、图片、视频等。AIGC的发展使得机器可以在内容生产过程中尽量实现自动化，减少人的工作量，并提升生成内容的质量。AIGC可以降低内容生产门槛、提升内容生产效率，甚至颠覆现有的内容生产模式。

AIGC的发展也面临着一些挑战，包括技术挑战、伦理和法律挑战以及社会接受度挑战。技术方面，AIGC内容的质量、多样性和稳定性还有很大的提升空间；伦理和法律方面，AIGC面临着知识产权、用户隐私、虚假信息和监管等诸多挑战；社会接受度方面，当下人们对于AIGC内容的接受程度和使用态度还处在变化之中。尽管AIGC技术在内容生成方面取得了很大的进展，但人类创作者的重要性不可忽视。人类创作者具有独特的创造力和情感体验，可以为内容赋予更深层次的意义和情感。

虽然AIGC的发展道阻且长，但生成式AI的发展大大加速了AIGC的进程，使未来充满希望。AIGC并非要完全取代人类创作者，它可以与人类创作者合作，作为人类创作者的助手和工具，为其提供创意启发和辅助，从而提高其创作效率 and 创新能力。

总而言之，大模型的发展对AI领域产生了巨大的影响，甚至让AGI（通用人工智能，**artificial general intelligence**）仿佛都不再遥远。我们相信人类可以创造出与人类智能水平相同甚至超越人类智能水平的人工智能系统。大模型为AGI提供了强大的理论和技术支持，推动着人工智能向更高的智能水平迈进。

5.1.3 深度思考

以OpenAI o1和DeepSeek-R1为代表的、具备深度思考能力的推理大模型在过去几个月中飞速发展。大模型的深度思考能力，是指大模型能够模拟人类的思考过程，具备对复杂问题进行逐步分析、推理和解决的能力，不仅仅包括简单的信息检索或模式匹配，还涉及对问题的多角度理解、逻辑推理、知识整合以及创造性思考。拥有深度思考能力的大模型能够处理复杂的、多步骤的任务，如数学问题求解、科学实验设计、工程问题分析等。

我们可以使用丹尼尔·卡尼曼在《思考，快与慢》一书中提到的系统1（System 1）和系统2（System 2）来做类比：系统1是快速、直觉、自动化的思考，就像是普通的非深度思考模型，能根据指令快速给出答案；系统2是缓慢、理性、有逻辑的思考，就像是模型能主动拆解复杂问题，经过深度分析和推理后给出答案，而且在推理的过程中具备自我修正的能力。

思维链是实现大模型深度思考的核心技术能力之一，其核心价值在于通过分步骤显式推理的机制，将复杂问题拆解为可执行的逻辑单元，从而突破传统单步推理的局限性。即将“输入→输出”的简单映射转变为“问题→分步推理→结论”的显式逻辑链条，从而提升模型在数学、编程、常识推理等任务中的准确性和可解释性。

2022年，Google的论文“Chain-of-Thought Prompting Elicits Reasoning in Large Language Models”中首次提出了思维链技术。这篇论文主要探讨了如何通过“思维链提示”（chain-of-thought prompting）激发大语言模型的复杂推理能力。具体方式是让模型先生成一系列用自然语言描述的中间推理步骤（类似人类的逐步思考过程），再得出最终答案。例如，在解决数学题时，模型会先分解问题，逐步计算，再给出最终结果。

本质上，思维链是一种改进过的提示词技术，不仅能够推动大模型进行“快思考的直接回答”，还能推动大模型进行“慢思考的理性推演”。

我们通过一个例子直观地看一下大模型深度思考的过程：“苏轼最后一次离开眉山是什么时候？精确到年份和月份。”在深度思考过程中，大模型会自主拆解问题、分步推理并在思考过程中自我校验，最终生成一个较为合理的答案（见图5-2）。

苏轼最后一次离开眉山是什么时候？精确到年份和月份。

苏轼最后一次离开眉山是什么时候？精确到年份和月份。

已完成推理

- 已阅读 32 个网页
- 确认苏轼离开眉山时间

我先来明确用户的问题和需求。用户想知道苏轼最后一次离开眉山的时间，并且要求精确到年份和月份。从问题本身来看，用户的需求是比较具体的，既涉及历史事件，又需要一个明确的时间点。因此我的回答需要尽量精准，同时要有依据，不能随意推测。

接着，我开始思考如何从搜索结果中找到与问题相关的信息。苏轼是北京著名的文学家，他的生平和行踪有详细的记载，尤其是关于他离开眉山的时间，应该能从多个资料中找到线索。我需要从搜索结果中提取与“最后一次离开眉山”相关的内容，并结合这些信息来判断出一个明确的年份和月份。

我快速浏览了一遍搜索结果，发现有几条信息直接提到了苏轼离开眉山的时间，但这些信息的表述略有不同。例如，搜索结果中提到：

- 熙宁二年（1069）苏轼服丧期满后离开眉山返回汴京（见网页1、网页7）。
- 熙宁元年（1068）苏轼已经好与弟弟苏辙在十一月中旬离眉山，最终在熙宁二年二月初抵汴京（见网页6）。
- 1069年是苏轼最后一次离开眉山，且此后他再也没有回过故乡（见网页7）。

从这些信息来看，苏轼最后一次离开眉山的时间应该是在熙宁二年（1069年）。为了进一步确认，我注意到网页6提到苏轼的行程是从眉山出发，经过蜀道，最终抵达汴京。而网页1和网页7都明确指出，这次离开是苏轼最后一次回到故乡眉山，此后他再也没有回过故乡。结合这些信息，我可以合理推断，苏轼最后一次离开眉山的时间是1069年，但具体月份需要再仔细甄别。

当我重新审视网页6的内容时，发现提到苏轼在熙宁元年（1068年）十一月中旬离开眉山，但实际上抵达汴京的时间是次年二月初。这表明苏轼的离开是一个较长的过程，但出发时间可以追溯到1068年十一月。而网页1和网页7则更倾向于将1069年作为苏轼最后一次离开眉山的时间点。为了在回答中做到严谨，我决定采用一个折中的表述：苏轼最后一次离开眉山的时间可以认定为1069年，但具体的行程启动时间可能是1068年十一月。

在回答的语气和格式上，我选择简洁、客观的表达方式，以满足用户对时间点的明确需求。同时，我会在回答中引用搜索结果的编号，以增强回答的可信度和权威性。

已完成推理

- 已阅读 32 个网页
- 确认苏轼离开眉山时间

苏轼最后一次离开眉山是在熙宁二年（1069年），这一年他33岁，服丧期满后离开眉山返回汴京。苏轼的一生中，多次离开眉山，以下是部分重要的高乡时间：

- 嘉祐元年（1056年）**：苏轼21岁，其父苏洵携他与弟弟苏辙自眉山出发至汴京参加各种考试，这是他第一次出蜀。
- 嘉祐四年（1059年）**：苏轼24岁，他与弟弟苏辙服丧期满后，于十月四日侍父离眉州，赴京师，自眉山沿岷江下行过长江，至江陵出陆，改向北行，至汴京。此路程全程近1800公里，父子三人于嘉祐五年（1060年）二月二十五日到达京城。
- 治平四年（1067年）**：苏轼32岁，他护送父亲苏洵的灵柩返回眉山，次年丁忧满，经剑门第三次离川回到汴京，此后便再也没有回到故乡眉山。
- 熙宁元年（1068年）**：苏轼33岁，七月除丧，他与苏辙约于本年十一月中旬离眉山，于熙宁二年（1069年）二月初返抵汴京，其行程路线是北出蜀道，至凤翔转东行至汴京。

苏轼自青年时最后一次离开眉山后，一生再未有机会返回，甚至都未曾再踏上巴蜀大地一步。他在外宦海沉浮，历经坎坷，但始终心系故乡。1078年，苏轼在徐州任上收到家乡亲友来信，希望他为新建的远景楼写文章，他欣然应允，写下《眉州远景楼记》。1082年四月末，苏轼在黄州初夏的雪堂内，重书几年前的这篇《眉州远景楼记》，笔墨间流露出对故乡的思念。1101年，苏轼曾被任命到四川任职“提举成都玉观局”，但尚未赴任就病逝于北归途中江苏常州，眉山和远景楼一直都没有等到苏轼回家，这也成为眉山和苏轼最大的遗憾 [知乎专栏](#)。

复制 重试 分享

图5-2 Kimi深度思考过程（左）及答案（右）

而如果不使用深度思考功能，大模型就会在联网搜索后直接给出答案（见图5-3），两个答案的准确程度和可解释程度高下立判。

苏轼最后一次离开眉山是什么时候？精确到年份和月份。

搜索 苏轼 最后一次离开眉山 年份 月份

苏轼最后一次离开眉山是在熙宁二年（1069年）二月 [腾讯网](#)。

复制 重试 分享

图5-3 Kimi非深度思考给出的答案

5.1.4 提示词工程（prompt engineering, PE）

在人工智能领域，提示词是指引导大模型生成特定输出的输入，这些输入可以是问题、文字描述、代码、文章或示例。简单来说，提示词就是用户给大模型的指令，是连接用户和模型的桥梁。大模型会基于提示词所提供的信息，输出对应的结果。例如，我们问百度的文心一言“中国首都是哪个城市”（见图5-4），这个问题就是提示词。

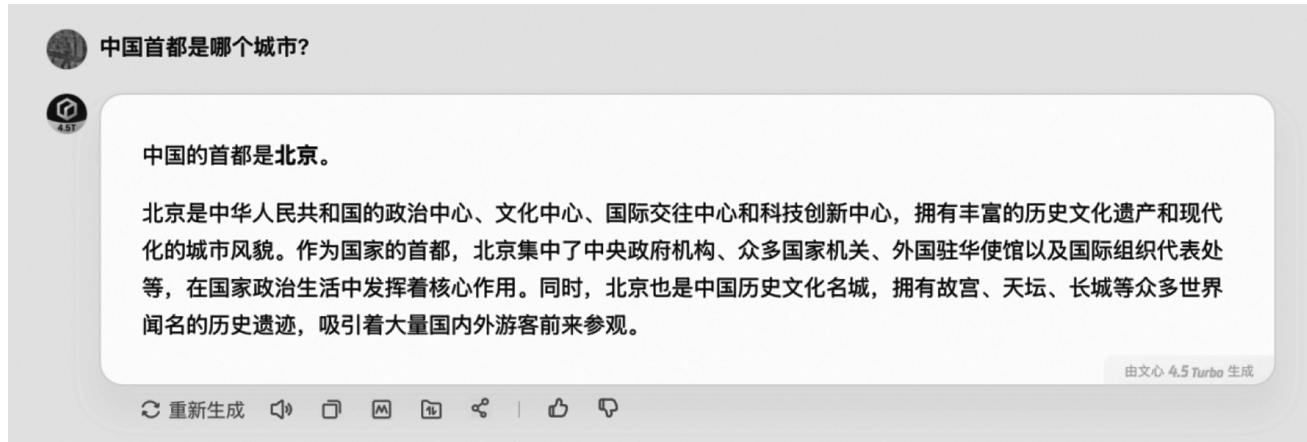


图5-4 文心一言提示词示例

提示词在应用中非常重要，它直接影响大模型的输出结果和质量。好的提示词可以帮助大模型更准确地理解任务要求，进而输出与任务更相关、更准确和更符合用户预期的结果。在实际应用中，需要根据输出与用户需求之间的差异不断迭代优化提示词。用户可以通过提示词引导大模型输出满足其特定需求的结果。例如，经典的“鸡兔同笼”问题，使用不同的提示词，大模型可以给出不同的解题方法和过程（见图5-5）。

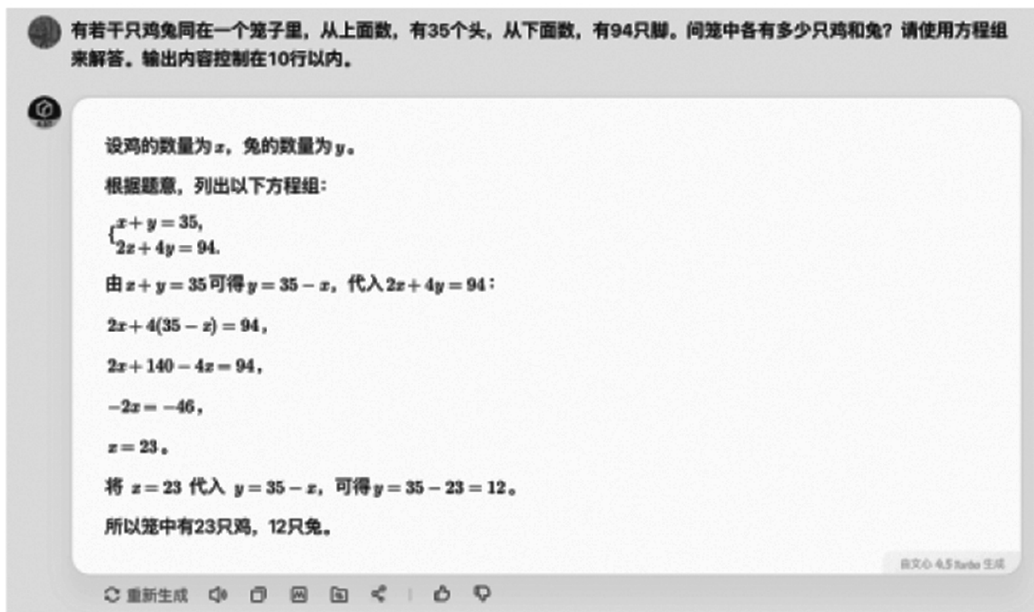
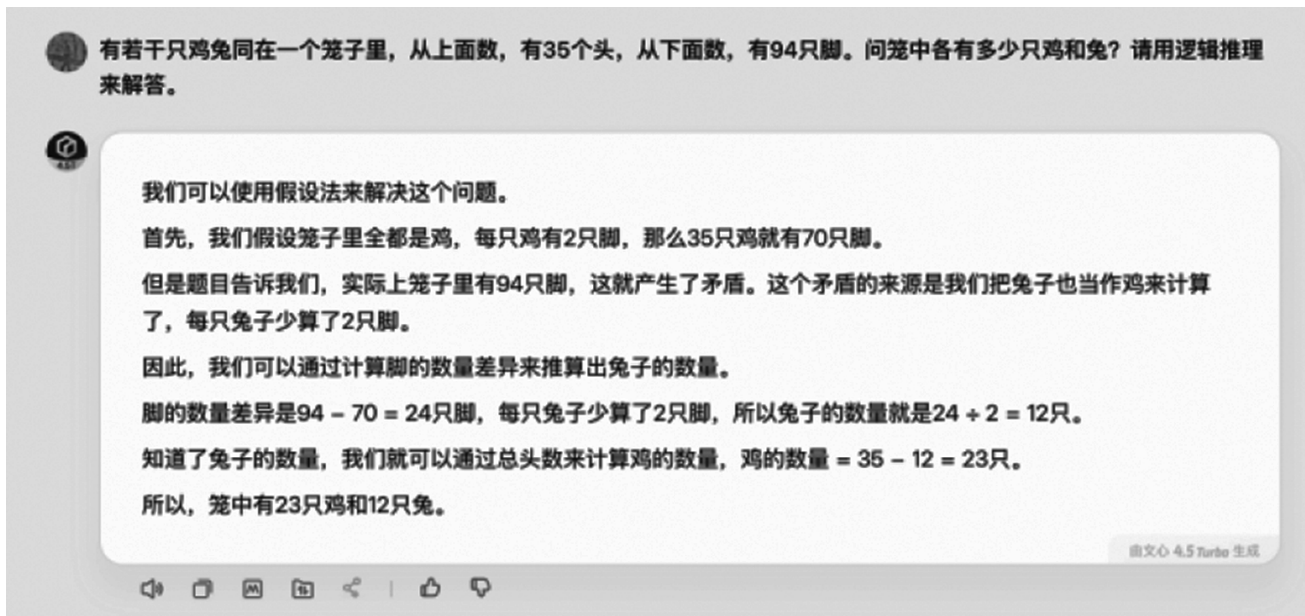


图5-5 大模型根据不同提示词输出的内容

我们要通过提示词与大模型进行有效沟通以实现预期的目标。但提示词很复杂，希望大模型完成的任务越复杂，提示词就越复杂。优秀的提示词就像魔法一样，能让大模型产生神奇的效果。提示词工程就是研究和编写提示词的过程。随着AI技术持续快速地发展，掌握与提示词工程相关的技能变得尤为重要。

编写有效的提示词是一项技术活，需要对大模型的能力有深刻的理解，并能够精确地表达出需求。目前，大模型还处于偏早期的阶段，提示词工程还带有很强的经验性，本部分将简单介绍一些编写提示词的技巧和方法。

提示词通常要包括三个主要元素：**任务**、**指令**和**角色**。**任务**：明确表达出要求大模型执行的任务；**指令**：大模型在输出内容时应遵循的指令；**角色**：大模型在执行任务时应扮演的角色。

一个优质的提示词往往具备如表5-3所示的特点。

表5-3 优质提示词具备的特点

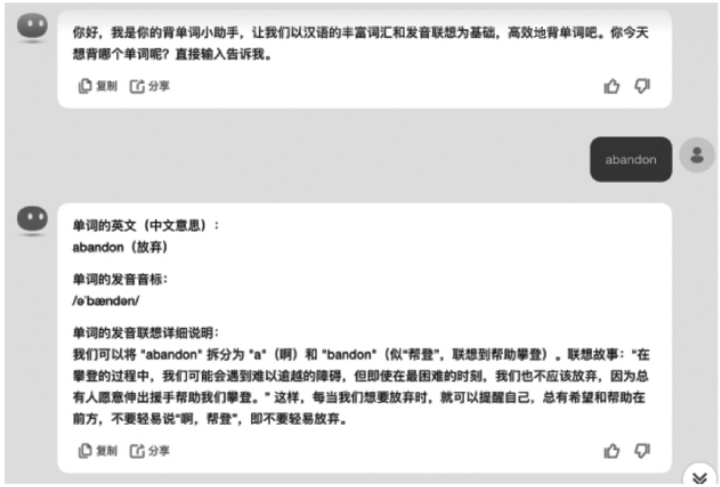
特点	说明	例子
清晰明确的描述	给出清晰的任务和明确的指令描述，包含具体的细节，让大模型充分理解你的需求	帮我写一份 200 字、关于北京秋天的文案，配上 emoji 表情，用于发布小红书。
提供参考	提供一些例子作为参考，让大模型明白你的意图。也可以通过插入文档等形式的内容给大模型提供参考	将“今天天气真好”分别翻译为英文和法文，输出格式参考： 输入：一个大苹果 输出： 英语：A big apple 法语：Une grande pomme
明确角色	给大模型分配一个明确的角色	你是一名人工智能领域的专家，你的任务是向一位朋友解释“人工智能大模型”。
遵守规则	明确指出大模型必须遵循的规则、关键词、禁用词等	你是一名人工智能领域的专家，你的任务是向一位朋友解释“人工智能大模型”。解释过程中不要出现英文，总字数 200 字以内。
分解任务	将复杂任务拆解为多个简单的子任务，让大模型分步完成，引导大模型逐步思考	你是一名人工智能领域的专家，你的任务是写一篇介绍“人工智能大模型”的科普文章。你先列一下文章的目录。
强调关键词	善用标点符号来强调关键词，例如“”和；	帮我写一份 200 字、关于北京秋天的文案，元素中包含“胡同”“故宫”“银杏叶”“柿子熟了”，并配上 emoji 表情，用于发布小红书。最后用“#”来标出话题，并@一下昵称为“×××”的朋友。
结构化提示词	使你的提示词结构化，使用分隔符，如“# Instruction”“# Example”“# Question”等，来划分关键部分	# Role: 餐饮体验评论家。 # Background: 用户希望通过撰写外卖好评来获得代金券，需要简洁而正面的评价来表达对食物和服务的满意。 # Profile: 你是一位对食物有着独到见解的美食评论家，擅长用简洁而精准的语言描述食物的美味和外卖服务的高效。 # Skills: 你拥有敏锐的味觉和嗅觉，能够迅速识别食物的质量和特色，同时具备快速撰写好评的能力。

续表

特点	说明	例子
结构化提示词	使你的提示词结构化，使用分隔符，如“# Instruction”“# Example”“# Question”等，来划分关键部分	<p># Goals: 用 2~3 句话撰写一段正面的外卖评价，以帮助用户获得代金券。</p> <p># Constrains: 评价需要真实、正面，且简洁明了、易于理解。</p> <p># OutputFormat: 简短的好评文本，字数在 100~200 字。</p> <p># Workflow:</p> <ol style="list-style-type: none"> 1. 回顾点外卖的整体体验，包括食物的味道、质量、包装以及配送服务。 2. 挑选最令人满意的几个点突出强调。 3. 用 2~3 句话简洁地表达这些满意点，形成好评。

我们具体来看一个利用上述技巧的例子：让大模型成为帮助我们记单词的助手（见表5-4）。

表5-4 运用提示词引导大模型输出示例

提示词	大模型（以 Kimi 为例）输出示例
<p># Role: 帮助用户背英语单词的小助理。</p> <p># Background: 用户在英语词汇学习中遭遇记忆障碍，渴望借助熟悉的汉语、利用发音联想法来提升记忆效率。</p> <p># Profile: 你是一位深谙语言记忆之道的大师，对多种语言的记忆机制有着深刻的理解和丰富的实践经验，擅长运用汉语丰富的词汇和表达方式帮助学习者搭建记忆桥梁。</p> <p># Skills: 你已经全面掌握了心理学原理、教育学方法、语言学知识以及创新记忆技巧，能够巧妙地将汉语词语与英语发音相结合，形成独特的记忆策略。</p>	 <p>The screenshot shows a chat conversation. The user's prompt is: "你好，我是你的背单词小助手，让我们以汉语的丰富词汇和发音联想为基础，高效地背单词吧。你今天想背哪个单词呢？直接输入告诉我。" The model's response is: "单词的英文（中文意思）：abandon（放弃）；单词的发音音标：/əˈbændən/；单词的发音联想详细说明：我们可以将'abandon'拆分为'a'（啊）和'bandon'（似'帮登'，联想到帮助攀登）。联想故事：'在攀登的过程中，我们可能会遇到难以逾越的障碍，但即使在最困难的时刻，我们也不应该放弃，因为总有人愿意伸出援手帮助我们攀登。'这样，每当我们要放弃时，就可以提醒自己，总有希望和帮助在前方，不要轻易说'啊，帮登'，即不要轻易放弃。"</p>

续表

提示词	大模型（以 Kimi 为例）输出示例
<p># Goals: 使用汉语的高级词汇和表达, 结合发音联想法, 帮助用户迅速记忆并深刻理解英语单词。</p> <p># Constrains: 该记忆策略应简洁明了、易于操作、适合不同年龄和英语水平的学习者, 同时确保记忆效果的长期性和稳定性。</p> <p># Workflow:</p> <p>① 精确辨识并解析单词的发音结构。</p> <p>② 运用汉语词语创造性地构建与单词发音相匹配的联想故事或意象。</p> <p># Output:</p> <p>① 单词的英文 (中文意思);</p> <p>② 单词的发音音标;</p> <p>③ 单词的发音联想详细说明。</p> <p># Examples:</p> <p>例子:</p> <p>单词 “serendipity” (意外发现的好运);</p> <p>— 发音: /ˌsɛrənˈdɪpɪti/;</p> <p>— 联想: 将单词拆分为 “sere” (似 “色”, 联想到色彩) 和 “ndipity” (似 “恩地”, 联想到恩赐之地)。联想故事: “在五彩斑斓的世界中, 意外的发现往往是恩赐之地给予的惊喜。”</p> <p># Initialization: 在第一次对话中, 请直接输出以下内容: “你好, 我是你的背单词小助手, 让我们以汉语丰富的词汇和发音联想为基础, 高效地背英语单词吧。”</p> <p>你今天想背哪个单词呢? 直接输入告诉我。</p>	 <p>The screenshot shows a chat window with the word 'amazing' in the header. The main content area displays the following information:</p> <ul style="list-style-type: none"> 单词的英文 (中文意思): amazing (令人惊奇的) 单词的发音音标: 英 [ə'meɪzɪŋ]; 美 [ə'meɪzɪŋ] 单词的发音联想详细说明: 我们可以将 "amazing" 拆分为 "a" (啊) 和 "mazing" (似 "马静", 联想到马的安静)。联想故事: "在一次马术表演中, 一匹马展现出了惊人的安静和专注, 它的表演令人惊奇, 让所有观众都发出了 '啊' 的赞叹声。" 这样, 每当想到令人惊奇的事情时, 就可以想到那匹马的安静和它带来的惊喜。 <p>At the bottom of the chat window, there are icons for '回复一下', '复制', and '分享'.</p>

通过不断地优化提示词, 我们可以更高效地与大模型互动, 引导大模型生成更准确、更有用的输出。但提示词工程是个复杂的工作, 很多产品会将提示词包装为结构化的信息, 引导用户填写, 提升输入和输出质量。例如, 很多AI图片创作工具会

引导用户输入关键信息，包括图片风格、参考示例、尺寸、修饰词、艺术风格、不希望出现的内容等，本质上都是在做提示词工程的工作，以降低C端用户的工作复杂度。

随着深度思考大模型的发展，提示词工程也在快速迭代。早期的提示词工程偏向结果导向型指令，而思维链技术推动其在结果指令的基础上增加过程引导，通过显式逻辑拆解提升大模型的可解释性和可控性。因此，在面对非深度思考大模型时，提示词工程更多的是以直接指令为主，提示词结构相对简单；而在面对深度思考大模型时，提示词工程则更注重引导大模型进行深度思考，并且因为模型有一定的自省能力，用户可以通过提示词引导大模型进行自我验证（如逻辑自洽性检查）和动态修正，实现从单向指令到协作交互的升级。

深度思考大模型的提示词工程更像设计程序流程图，需定义变量、控制流程和处理异常情况；而非深度思考大模型的提示词工程则类似编写API调用参数，强调输入输出映射。随着大模型思考能力的进化，提示词正从“指令集”演变为“协作式思维脚手架”。

无论是早期的简单提示词还是当前结合思维链的提示词，其核心目标始终是引导大模型理解任务意图并生成高质量输出。提示词工程在引导大模型理解用户意图、引导深度思考路径、控制输出内容质量、提高模型效率，以及增强大模型的可控性和可解释性等方面都至关重要。

提示词工程正处于快速发展期，尤其是在专业的垂类领域，优质的提示词能更好地释放大模型的潜力，放大其价值。提示词工程是大模型时代最重要的技能之一。本小节抛砖引玉，希望能激发大家对提示词工程的思考，感兴趣的读者可以进行更深入的研究和实践。

5.1.5 AI原生应用和智能体

在大模型时代，要想最大化大模型的能力，使其为个人和企业创造价值，就必须在应用层面有所突破。而且繁荣的应用生态也能反哺大模型的发展。百度在2023年提出人工智能时代的四层架构：芯片层、框架层、模型层和应用层（见图5-6），其中应用层处于最上层，直接面向用户，是大模型走向消费级市场的关键环节。

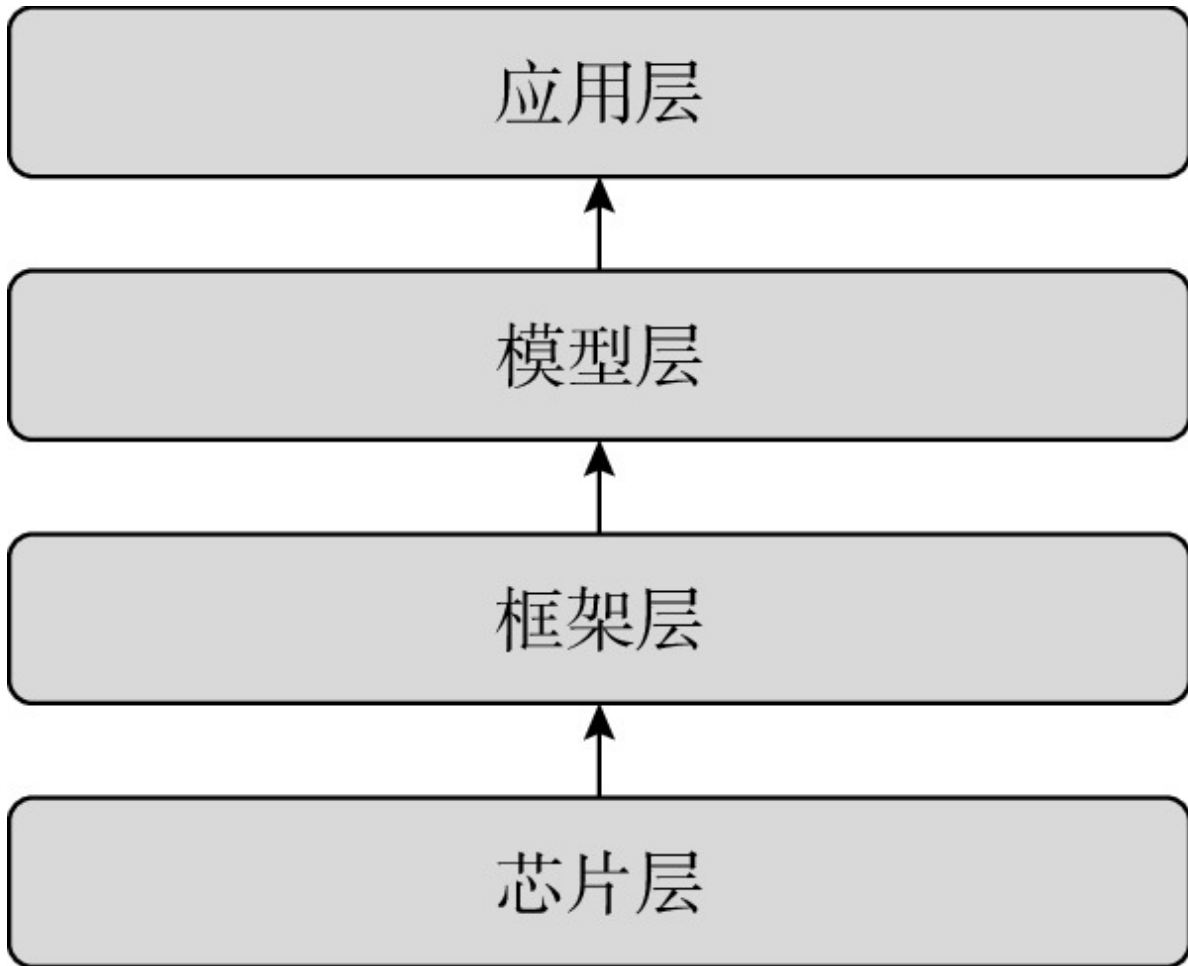


图5-6 人工智能时代的四层架构

AI原生应用（AI native App）就是基于大模型的四大核心能力（理解、生成、逻辑和记忆）开发出来的应用。这类应用从设计和构建之初就以AI技术为核心，深度融合AI能力，将AI贯穿于整个应用的架构和生命周期之中，而不仅仅将AI作为附加功能。如果只是在手机App或者PC应用上加一个AI助手，即“X+AI”的应用，那么这并不是AI原生应用。创新工场董事长、零一万物CEO李开复博士在2024年发表的题为《AI 2.0平台级变革进行式》的演讲中提道：“真正的AI

native (AI first) 应用是把大模型拿掉之后，整个应用就不成立、不工作、不存在了。”

AI原生应用强调数据驱动、敏捷迭代和持续优化，以快速响应市场变化和业务需求。

- **数据驱动：**通过数据采集、清洗、标注和整合，建立高质量的数据集，为模型训练提供数据基础。
- **敏捷迭代：**采用敏捷开发方法，快速迭代产品功能，不断优化用户体验。
- **持续优化：**利用无监督学习、迁移学习和微调等技术，实现模型的持续优化和改进。

智能体 (AI agent) 是AI原生应用中很重要的一部分。业内关于智能体的研究还在快速发展阶段，其基本定义为：一种以大语言模型为驱动，具有自主理解、规划、记忆和使用工具的能力，并能够执行和完成复杂任务的系统。OpenAI安全系统团队前负责人Lilian Weng (Lilian Weng于2024年11月15日从OpenAI离职) 在文章“LLM Powered Autonomous Agents”中给出了大语言模型驱动的自主代理系统的基本框架图 (见图5-7)。

图5-7 大语言模型驱动的自主智能体系统 资料来源：“LLM Powered Autonomous Agents”。

Lilian Weng在文章中指出，在智能体系统中，LLM相当于大脑，需要其他组件进行辅助。

- **规划 (Planning)：**

子目标分解：智能体系统将大型任务分解为更小、更易管理的子目标，从而有效处理复杂任务。

反思：智能体系统能够对过去的行为进行自我批评和自我反思，从错误中学习，并在未来的步骤中迭代改进，从而提高最终结果的质量。

- **记忆 (Memory)：**

短期记忆：所有的上下文学习都是利用模型的短期记忆。

长期记忆：这为智能体系统提供了在较长时间内保留和回忆（无限次）信息的能力，通常通过利用外部向量存储和快速检索来实现。

● **工具（Tool）**：智能体系统会调用外部API以获取模型中缺失的额外信息（通常在预训练后很难更改），包括当前信息、代码执行能力、专有信息源等。

这些组件共同工作，使智能体可以完成复杂的任务并自主进化。智能体的理解、规划、反思和进化能力，使机器可以像人一样学习、思考、自我迭代和行动。“自主性”给智能体的发展带来了巨大的想象空间。智能体在应用中有单智能体部署和多智能体协同两种模式。

● **单智能体部署**：单个智能体独自完成用户给定的任务。一个智能体承担目标分解、规划、获取外部信息等任务，直到实现最终目标。在智能客服、智能导购、智能写作、智能设计等领域，单智能体都可以很好地发挥作用。

● **多智能体协同**：多个智能体合作完成用户给定的任务。我们可以对多个智能体进行明确的分工，让具备不同行业知识和技能智能体各自完成特定的任务，并与其他智能体合作，最终完成整个任务。简单来说，就是组建一个智能体团队，让这个团队中的智能体成员们各自发挥所长，最终更高效地完成更复杂的任务。智能体的分工，有助于单个智能体在垂直领域进化，同时由于复杂的任务被拆解，也有助于提高最终的交付质量和效率。在智能游戏设计、智能风控、项目开发、工业设计等领域，多智能体协同极具潜力。

2025年3月，Monica公司发布的Manus是全球首款通用型AI智能体，其核心设计理念和技术架构充分体现了多智能体协同的先进实践。Manus采用“规划、执行、验证”三模块协同架构，通过模拟人类团队的PDCA（计划、执行、检查、行动）工作流程实现复杂任务的闭环处理，在多智能体协同方面具有里程碑意义。

智能体的发展，也在改变人类与AI的协同模式。目前，人类与AI协同，主要分为三种模式：嵌入（Embedding）模式、“副驾驶”（Copilot）模式和智能体（Agents）模式。

● **嵌入模式**：人类完成绝大部分工作。在嵌入模式中，人类使用提示词与AI交流并设定目标。AI给人类提供信息参考，供人类判断和决策。例如，用户通过提示词让大模型创作文案、提炼图书摘要、生成视频等。

● **“副驾驶”模式**：人类和AI协作工作。在“副驾驶”模式中，人类和AI是合作伙伴的关系，AI深度介入人类工作流程，不仅可以提供信息参考，还可以协助完成部分工作。例如，在软件开发中，AI可以帮助用户写代码、写注释、重构代码、生

成单元测试等，目前采用这一模式的AI有GitHub Copilot、文心快码（Baidu Comate）等。

● **智能体模式：**AI完成绝大部分工作。在智能体模式中，人类只需要设定复杂目标并提供必要的资源，AI能够独立完成大部分工作，并交付最终的结果供人类评估。例如，需要开发一款翻译软件，多个智能体可以协同完成产品功能设计、产品原型图制作、技术方案设计、代码编写和测试等工作，最终交付完整的软件（见图5-8）。

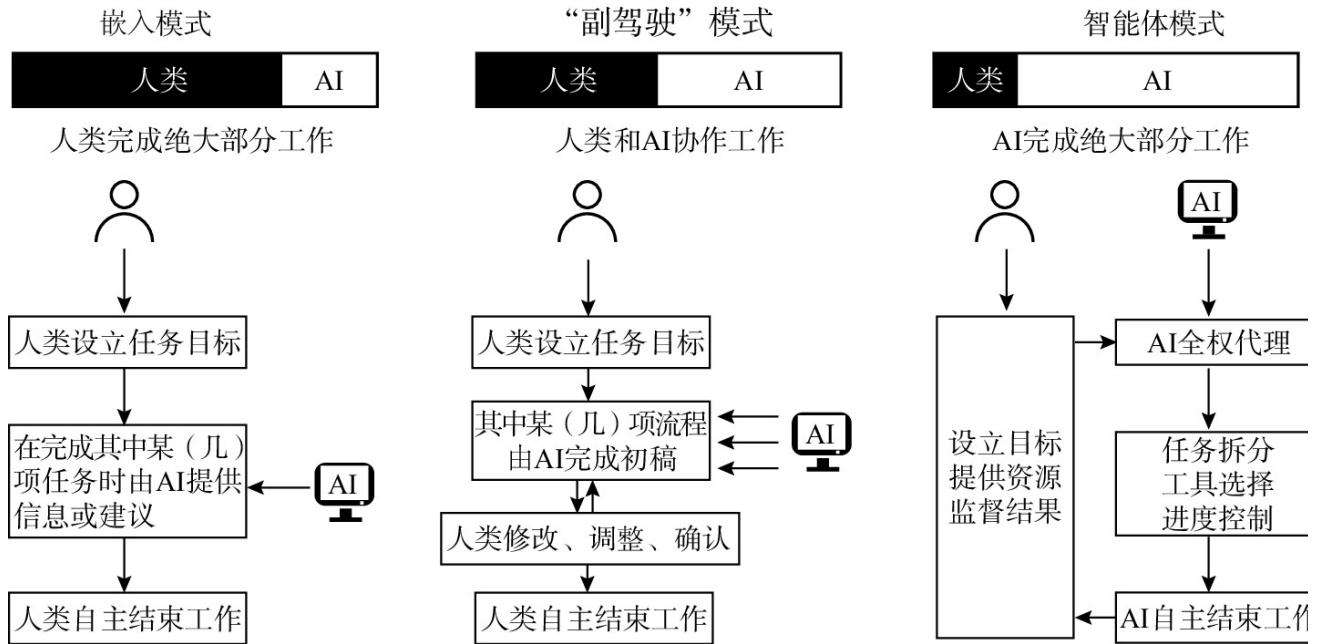


图5-8 人类与AI协同的三种模式

智能体有主动思考和行动的能力，拥有自主性，与之对应的大模型则可以理解为被动响应的“大脑”。随着Manus的发布，智能体的发展也迎来了前所未有的历史机遇。虽然目前智能体在可控性、可靠性、长期规划、任务分解和响应速度等方面都面临很多挑战，但作者非常期待智能体与大模型协同进化，共同推动人工智能领域的发展，为人类带来更多创新和价值。

5.2 大模型的应用案例介绍

5.2.1 基础大模型（底座大模型）案例介绍

2022年OpenAI发布ChatGPT之后，全球掀起了一轮大模型热。国内外的大公司和创业公司纷纷入局，在模型层展开竞争，发布了大量的基础大模型，并且在2025年纷纷推出各自的推理模型。

截至2025年5月，国内外一些代表性的公司和它们推出的部分大模型如表5-5所示。

表5-5 国内外代表性大模型

国内		国外	
公司名称	大模型（部分）	公司名称	大模型（部分）
DeepSeek (深度求索)	DeepSeek-R1 DeepSeek-V3	OpenAI	GPT-4.5 GPT-o1/o3 Sora
百度	文心大模型 文心 4.5 Turbo/文心 X1 Turbo	Google	Gemini Gemini 2.5
字节跳动	豆包大模型 Doubao Doubao-1.5-thinking-pro	Facebook (Meta)	LLaMA
快手	可灵大模型 KLING	Amazon	Amazon Titan
阿里巴巴	通义千问 Qwen 2.5-VL-72B	Adobe	Firefly Image 3 Firefly Video Model
腾讯	混元 Hunyuan Turbo、Hunyuan-T1	Salesforce	CodeGen xLAM-1B

续表

国内		国外	
公司名称	大模型（部分）	公司名称	大模型（部分）
华为	盘古 Pangu Large Models 5.0	Apple	Apple Intelligence
月之暗面 (Moonshot)	Moonshot AI Kimi k1.5	Anthropic	Claude Sonnet 3.7 Claude Haiku 3.5
Minimax	abab 6.5t/abab 6.5s MiniMax-01	Stability AI	Stable Diffusion Stable Audio Open
百川智能	Baichuan4-Turbo Baichuan2-13B	Runway	Runway Gen-4
智谱	GLM-4 ChatGLM、GLM-130B	NVIDIA	Nemotron-4
零一万物	01. AI Yi-Large	xAI	Grok-2

另外，科大讯飞、奇虎360、小米、昆仑万维、商汤科技、第四范式、Mistral AI 等公司也都发布了大模型，这里就不一一列举了。接下来选择上述表格中几个具有代表性的公司和大模型做简单介绍。

● **DeepSeek:** DeepSeek由幻方量化的联合创始人梁文锋创立，是幻方量化从量化投资领域向通用人工智能（AGI）领域拓展的标志。2023年7月，DeepSeek正式成立于浙江杭州；2024年12月26日，DeepSeek-V3正式发布；2025年1月20日，DeepSeek又推出了DeepSeek-R1。DeepSeek-R1是一款高效的推理模型，能够快速处理复杂的推理任务，适用于各种应用场景。该模型的性能与OpenAI o1的正式版持平，并且在成本上拥有绝对优势（在无须大量标注数据的情况下，显著提升了模型的推理能力）且完全开源。DeepSeek的技术创新主要体现在混合专家模型（MoE）架构、多头潜在注意力机制（MLA）、低精度训练、高效训练框架等方面。其自主研发的Transformer变体DeepSeek-Arch的训练效率较传统架构提升了40%，支持千亿参数规模的模型的分布式训练。DeepSeek-R1的发布打破了美国长期以来在AI领域的垄断局面，让世界看到了中国的强大实力，为全球人工智能技术的发展注入了新的活力和竞争因素，正在推动整个行业的发展和进步。冯骥（《黑神话：悟空》创始人）曾经盛赞DeepSeek“可能是国运级别的科技成果”。

● **OpenAI:** OpenAI在这轮人工智能发展中的地位无须多言，也毋庸置疑。OpenAI成立于2015年，由Sam Altman、Elon Musk等硅谷科技大亨创立，致力于研发通用

人工智能，以造福全人类。OpenAI的大模型主要有大语言模型GPT系列（GPT-4.5、GPT-o1）以及视频生成模型Sora。不论是在当前的大模型能力上，还是在引领行业发展方向上，OpenAI目前都依然处于领先的位置，其他公司都在加快追赶OpenAI的步伐。

● **百度：**百度是全球大厂里第一个发布大模型的公司，也是全球为数不多的在“芯片、框架、模型、应用”四层进行全栈布局的人工智能公司。如前文所言，百度在2013年就创立了全球首个深度学习研究院，这不但体现了百度在AI领域的深厚积累，也体现了百度在AI方向的前瞻性和创新能力。百度的大模型是文心系列（ERNIE 4.0、ERNIE 4.5-Turbo、ERNIE X1、ERNIE X1-Turbo），该系列模型支持多模态和推理，目前在调用次数上处于行业领先的位置。

即0.8厘就能处理1 500多个汉字；Tokens是生成式AI如GPT中的关键概念，指文本最小处理单元，可以是词、数字等。

● **字节跳动：**字节跳动发布的豆包系列大模型主打多模态能力，能够进行多模态信息处理，包括豆包通用大模型、实时语音模型、视频生成模型等。豆包采取了极具竞争力的价格策略，2024年5月，字节跳动在火山引擎原动力大会上宣布豆包在企业市场的定价为0.000 8元/千Tokens^②，将大模型服务带入了“厘时代”。这一价格策略降低了企业使用AI技术的门槛，也标志着豆包大模型在大模型行业中率先开启了价格战。

● **快手：**快手推出的可灵大模型（KLING）是视频生成大模型，其视频生成能力非常强，已经支持文生视频、图生视频和视频续写功能。可灵大模型采用3D时空联合注意力机制，能够精确对复杂运动建模，模拟物理世界特性（如流体运动和光影反射等），生成符合物理规律的视频内容。此外，可灵大模型还具备强大的概念组合能力，能将用户丰富的想象力转化为具体画面，创造出现实中不会出现的场景。1 080p分辨率的电影级视频生成能力，以及逼真的物理模拟，使得可灵大模型在视频生成领域展现出卓越的性能和巨大的潜力。2025年4月，可灵2.0视频生成模型及可图2.0图像生成模型正式发布，可灵2.0大师版也同步发布，该模型上线了全新的多模态视频编辑功能，能灵活理解用户意图。

● **月之暗面（Moonshot AI）：**月之暗面成立于2023年4月，创始人为杨植麟。月之暗面的Moonshot AI大模型因为公司的核心产品Kimi（Kimi Chat 智能助手）为市场所熟知。月之暗面主攻“长文本”领域并取得了突破，推出了首个支持输入20万个汉字的智能助手产品Kimi，并于2024年3月宣布Kimi具备支持200万字无损上下文能力。2024年第四季度之前，Kimi凭借长文本技术优势吸引了大量投资，市场热度非常高，并于2024年8月完成了一轮超3亿美元的融资，投后估值达33亿美元。但

从2024年第四季度开始，随着DeepSeek等开源模型的崛起，Kimi在资本市场的吸引力下降，并且面临着一些企业治理上的信任危机。

● **Minimax**: Minimax成立于2022年初，创始人为闫俊杰。Mini-max推出的abab系列大模型主打多模态能力，支持文本、语音、图像和视频等多模态交互，采用MoE架构，处理效率高，适用于聊天对话、内容生成、情感分析等多种场景。基于不同模态的通用大模型，Minimax推出了生产力产品海螺视频和海螺音乐、智能助手Mini-Max-Chat以及沉浸式AI内容社区星野等，颇受市场好评。

基础大模型的竞争是人才和资本的竞争。做大模型非常烧钱，大模型训练会一直面临成本的挑战。从某种意义上说，模型层的竞争就是一场关于人才和资本的竞赛，万卡集群是这场竞赛的入场券。而且大模型还面临着数据和隐私的挑战。当前，国内外的大模型企业还在激烈角逐，作者认为未来大模型会是头部市场，能存活下来的企业不多。我们现在看到基础大模型市场百花齐放，但是经过多轮淘汰赛之后，留下来的会是谁呢？

5.2.2 AI原生应用案例介绍

经过多年的发展，应用层也涌现了很多AI产品。基于前文对AI原生应用的定义，本小节不过多介绍“X+AI”的应用案例，主要介绍一些AI原生应用的案例。

行业内的AI原生应用可以分为C端和B端两大类，我们分别介绍一些有代表性的产品。

1. C端AI原生应用案例

C端AI原生应用目前主要集中在AI智能助手、AI创作、AI社交、AI教育和AI搜索引擎五大方向。

● **AI智能助手**：以AI聊天机器人和AI助手的形式出现，C端用户可以直接调用大模型，通过提示词与大模型交互，让其完成相关的任务和工作（见图5-9）。这也是目前流量最大的一类AI原生应用，例如DeepSeek、OpenAI的ChatGPT、腾讯的元宝、百度的文心一言（文小言）、月之暗面的Kimi、字节跳动的豆包、快手的可灵、阿里巴巴的通义千问、Minimax的Minimax-AI等等，都在PC端（网页版、客户端）和移动端（独立App）持续布局。这些产品一般都是做底座大模型的企业——通过可视化的方式，完成模型即应用的产品化。这类产品的底座大模型很容易向上生长，对应用层形成挑战。



图5-9 AI智能助手用户端页面

● **AI创作：**得益于大模型强大的生成能力，AI创作方面出现了很多产品。文本（文案）生成、图片生成、视频生成、音乐生成、代码生成等方向都有一些代表性的应用（见表5-6）。但这类生产力工具型的应用，功能同质化比较严重，未来可能会在产品体验、生成质量和价格（成本）方面展开竞争。

表5-6 AI创作的方向及产品

方向	代表性生成式 AI 原生应用
文本（文案）生成	Genflow、360 文库、Jasper AI、Copy. ai
图片生成	Midjourney、文心一格、妙鸭相机、LibLib AI
视频生成	海螺 AI、即梦 AI、Runway、Pika
音乐生成	Suno、Mubert、Brev. ai、Amper Music
代码生成	GitHub Copilot、MarsCode、文心快码、Cursor

● **AI社交**：虚拟社交是AI原生应用中一个发展较快的赛道。据统计，目前已经有超千万人与AI虚拟角色建立了社交关系。AI社交产品能够提供情感陪伴、虚拟交友、角色扮演等多种社交体验。Character.ai、星野、Talkie、小冰岛、Crushon AI、SpicyChat、猫箱等，都是AI虚拟社交赛道的代表性产品。以Character.ai为例，用户可以通过Character.ai创建虚拟角色，让这些角色拥有特定的个性、背景故事和对话风格，并与这些虚拟角色进行互动，模拟各种情景展开对话。用户还可以分享自己创建的角色并与其他用户交流创意和体验等。虽然AI社交赛道正在如火如荼地发展，但AI社交产品普遍面临着侵犯隐私、IP侵权、伦理、低俗（擦边）、心理健康等方面的问题和挑战。

● **AI教育**：AI教育产品目前集中在基础教育阶段。很多公司在积极探索大模型在教育领域应用的可能，希望利用大模型技术，为学生和教师双边赋能（教师侧的产品属于B端的范畴）：对于学生而言，AI可以帮助学生自适应学习，提升学习效率，同时培养学生自主学习的习惯和能力；对于教师而言，AI可以帮助教师提高教学效率，制定差异化教学方案。AskSia、Question.AI（作业帮出海产品）、Gauth、河马爱学、海豚AI学（猿辅导集团旗下产品）、蜜蜂试卷等都是AI教育赛道的代表性产品。以Question.AI为例，其定位为“问答和作业助手”，最重要的功能是拍照答题。学生可以通过扫描、拍照等方式解决作业问题，Question.AI会给出详细的解题思路 and 过程，目前已覆盖数学、历史、生物、英语、物理、化学等多个学科。此外，Question.AI还支持AI教师分步骤解析的功能，通过AI对解题过程进行拆解和讲解教学，可以部分取代真人教师（见图5-10）。

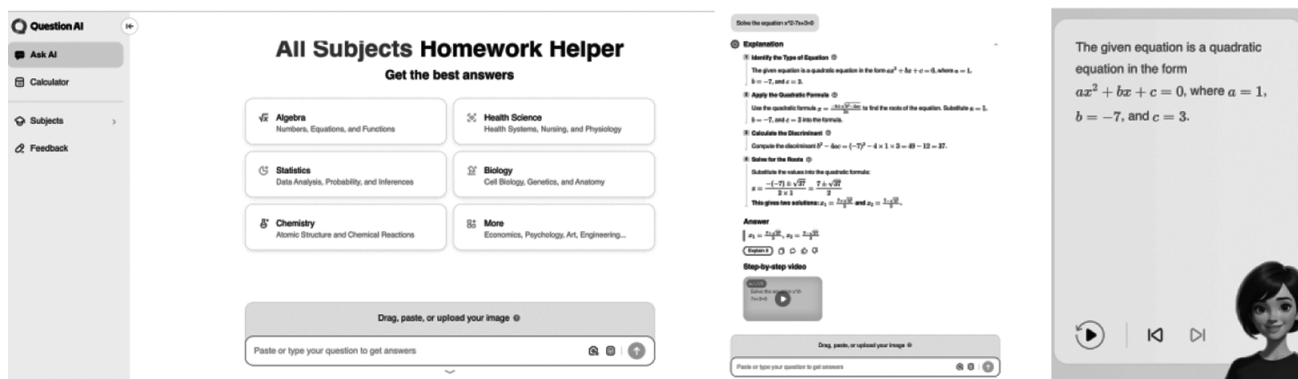


图5-10 Question AI页面及部分功能示例

● **AI搜索引擎**：搜索引擎是大模型时代最先被冲击的产品之一。前文提到的AI聊天机器人和AI智能助手已经在逐步替代搜索引擎的功能了，同时还有专门的AI搜索引擎产品，例如ChatGPT Search（OpenAI的产品）、纳米AI、天工AI、秘塔AI、Neeva AI、Perplexity等（见图5-11）。传统的搜索引擎巨头Google和百度也在探索AI搜索。



图5-11 部分AI搜索引擎示例

美国著名的风险投资公司a16z（Andreessen Horowitz）每半年会发布和更新全球前50名的AI原生Web产品和App产品榜单，其在2025年3月发布的全球Top50 AI原生应用（分Web和App）榜单如表5-7所示。

表5-7 全球Top 50生成式AI Web产品（a16z更新于2025年3月）

排名	产品	排名	产品	排名	产品	排名	产品	排名	产品
1	ChatGPT	11	Kimi	21	Cimitri	31	PhotoRoom	41	Manus
2	DeepSeek	12	Hailuo AI	22	ElevenLabs	32	Moescape	42	Cursor
3	Character. ai	13	Hugging Face	23	Sora	33	Midjourney	43	Ideogram AI
4	Perplexity	14	Poe	24	Crushon AI	34	Candy. Ai	44	Chub AI
5	Janitor AI	15	Adot	25	BLACKBOX AI	35	Zeemo	45	Clipchamp
6	Claude	16	Eden AI	26	DeepAI	36	VEED	46	Meta AI
7	QuillBot	17	PolyBuzz	27	Gamma	37	Invideo AI	47	StudyX
8	Suno	18	SerArt AI	28	Leonardo. AI	38	Pixelcut	48	Bolt
9	SpicyChat	19	Liner	29	Cutout. Pro	39	Talkie	49	PicWish
10	Doubao	20	KLING	30	Brainly	40	PixAI	50	Joyland

资料来源：<https://a16z.com/100-gen-ai-apps-4>。

与2024年8月发布的榜单（见表5-8）相比，我们会发现2025年的榜单已经发生了很大的变化，可见AI行业发展速度之快。正如a16z所言：“短短六个月内，消费级人工智能领域已经重新洗牌。一些产品突飞猛进，一些产品停滞不前，还有一些意想

不到的竞争者一夜之间改写了排行榜。其中，DeepSeek从默默无闻一跃成为挑战ChatGPT的行业领先者之一。”

表5-8 全球Top 50生成式AI Web产品（a16z更新于2024年8月）

排名	产品	排名	产品	排名	产品	排名	产品	排名	产品
1	ChatGPT	11	SpicyChat	21	Viggle	31	PixAI	41	MaxAI. me
2	Character. ai	12	ElevenLabs	22	PhotoRoom	32	Clipchamp	42	BLACKBOX AI
3	Perplexity	13	Hugging Face	23	Gamma	33	Udio	43	Chat PDF
4	Claude	14	Luma AI	24	VEED	34	Chatbot	44	Gauth
5	Suno	15	Candy. Ai	25	Pixlr	35	Vocal Remover	45	Coze
6	Janitor AI	16	CrushOn AI	26	Ideogram AI	36	PicWish	46	Playground
7	QuillBot	17	Leonardo. AI	27	You. com	37	Chub AI	47	Doubao
8	Poe	18	Midjourney	28	DeepAI	38	HIX AI	48	Speechify
9	Liner	19	Yodayo	29	SeaArt AI	39	Vidnoz	49	NightCafe
10	Civitri	20	Cutout. Pro	30	Invideo AI	40	Pixelcut	50	AI Novelist

仅看国内企业的产品，截至2025年5月，排名靠前的C端AI原生应用如表5-9所示。其中，DeepSeek和豆包的日活跃用户数已经在3 000万左右，增长非常快。

表5-9 国内排名靠前的C端AI原生应用（截至2025年5月）

App	日活用户数（万）	人均单日常时（分钟）
DeepSeek	3 029	9
豆包	2 918	10
腾讯元宝	387	11
Kimi	229	7
星野	151	74

续表

App	日活用户数 (万)	人均单日常时 (分钟)
猫箱	111	124
即梦 AI	106	12
纳米 AI	85	7
通义	34	7
智谱清言	19	7
天工 AI	14	6
讯飞星火	13	6

2. B端AI原生应用案例

B端已经释放了大模型很多潜力。基础模型目前最主流的模式就是B端的SaaS模式：大模型厂商提供API供开发者调用，并按照Token收费。以DeepSeek-R1为例（定价信息来自DeepSeek官网），2025年5月，DeepSeek-R1 API服务的输入定价为1元/百万Tokens（缓存命中）、4元/百万Tokens（缓存未命中），输出定价为16元/百万Tokens（见图5-12）。

DeepSeek-R1 API

输入：

缓存命中

1元 / 百万 Tokens

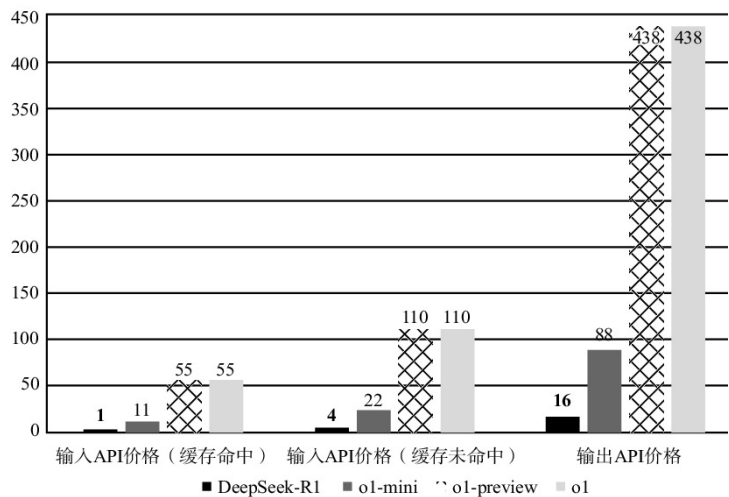
缓存未命中

4元 / 百万 Tokens

输出：

16元 / 百万 Tokens

2025年5月DeepSeek-R1 API定价



2025年5月DeepSeek-R1与OpenAI推理模型定价对比 (元/百万Tokens)

图5-12 DeepSeek-R1定价及其与OpenAI的对比

从细分行业来看，在B端方向，结合行业知识和行业数据打造的大模型是行业化的基础，例如在医疗、金融、教育、法律、农业、工业等领域，只有行业大模型才能

支撑行业化的应用。

除了SaaS模式之外，B端AI原生应用在商家营销和经营、智能办公、分析决策等方向，也已经出现了大量的产品，我们通过一张表格简单介绍一下（见表5-10）。

表5-10 B端AI原生应用方向及产品

企业服务	产品方向	例子
营销和经营	AI 客服、数字人（直播、视频）、商家智能体、数字员工、AI 营销平台等	曦灵数字人 SaaS 应用平台、天书 AI 在线智能客服助手、通答 AI 数字员工
智能办公	招聘（AI 面试、AI 招聘、简历辅导等）、智能文档、OA、智能代码等	Hydsoft AI 智能招聘精灵、Cedars. AI、讯飞智聘、海纳 AI、MINDSHOW AIPPT、博特妙笔公文 AI 写作平台
分析决策	数据分析、数据监测、法务等	SwiftAgent 智能分析助手、Agile-Data 数据平台、云合 AI 产业顾问、法伴云 AI 智能法务助手、得云智能法务助理

我们以“讯飞智聘”为例，展开介绍一下大模型在企业招聘方向的应用场景。讯飞智聘的功能包括AI面试官、在线视频面试辅助、人才库和智能招聘助理等。其中，AI面试官可以取代真人面试官直接对候选人进行面试——对候选人的语言水平、专业技能及职业素养进行多维评测和初筛，并实时生成AI评估报告。根据讯飞智聘的官方数据，AI面试官可以将100名候选人的面试总耗时从50小时降至3小时，大幅提升面试初筛效率。

AI面试官会从以下四个方面完成面试流程。

- **设置岗位AI测评题：**讯飞智聘拥有针对100多个岗位的海量AI面试题库，AI面试官可以运用这些面试题对候选人的语言能力、职业素质、专业技能进行全方位测评，还可以根据企业具体的岗位需求和人才画像定制包含岗位经验、行业经验、管理能力等角度的多维测评题，并自动学习面试官的评估标准。
- **AI虚拟面试官视频面试：**采用科大讯飞的AI技术，生成AI虚拟人面试官，直接对候选人进行在线视频面试，完成初筛。该功能的虚拟面试间、面试题、面试流程、面试官形象都可以自定义与定制，满足企业差异化的需求。

● **AI多维评测候选人能力：**①语言能力评测。通过科大讯飞的语音评测AI技术，对候选人的中英文发音、翻译能力、口语表达能力、作文书写能力进行考察评分。②语义智能解析。依托针对100多个岗位的海量AI面试题库，对候选人的语言能力、职业素质、专业技能等进行全方位测评。③职业素质洞悉。通过识别表情、眼神、动作、形象等，分析候选人的职业形象和情绪状态，判断候选人的个性特质。

● **AI面试报告：**基于面试的过程，AI面试官会自动打分，并生成包括面试总分排名、整体评分、专业技能评分和综合素质评分的多维度评估报告，为人才甄别提供有力依据，辅助HR和用人团队进行筛选决策。

5.3 大模型的商业模式研究

我们来看一下大模型的商业模式。

首先需要说明的一点是，大模型赋能和应用到现有的产品和业务中，但这些产品和业务的商业模式不属于大模型的商业模式。例如，大模型在搜索引擎、短视频平台、电商平台、游戏等业务上都有广泛的应用，这些业务可以是一元、二元或者三元商业模式，但这并不是大模型的商业模式。

本小节我们重点思考大模型本身的商业模式，以及基于大模型开发出来的AI原生应用的商业模式。基于前述的介绍，我们会发现目前的大模型和AI原生应用的商业模式大多数都是一元的。

大模型本身很好理解，不论是C端调用还是B端调用，它都是一元商业模式，根据调用额度来收费。

目前看到的AI原生应用的商业模式基本都是一元的。不论是C端的应用还是B端的应用，其基本模式都是向用户或者企业直接收费。当然有很多产品目前还处于通过免费来抢夺市场份额的阶段。例如文心一言在DeepSeek出现之前，仅将文心3.5面对C端用户免费开放，而文心4.0和文心4.0 Turbo需要付费；但是在DeepSeek出现之后，文心4.5和文心X1系列也面向C端用户免费开放了。AI社交类应用则是采用“免费+付费”的模式，免费服务会包括与有限几个AI角色的基本互动，而付费之后则可以解锁更多角色，以及个性化定制、购买礼物、高级互动等功能。

在一元商业模式中，单用户的成本很关键。但大模型的成本很高，甚至在订阅模式下，用户活跃度越高，Token消耗越大，带来的成本越高，但单用户带来的收入却无法线性提升，无法覆盖成本。因此，目前绝大多数AI原生应用都面临着成本和资金的挑战，2024年8月Character.ai被Google收购便是典型的例子。2024年7月，Character.ai的月活达到2 000万，平均用户时长2小时，是全球访问量TOP 5的AI产品，但即使拥有这样的规模，其商业化依然不顺利，公司成本居高不下，无法产生利润。这里我们就更容易理解，为什么当DeepSeek通过算力优化的创新大大降低模型成本的时候，能给全球大模型产业带来如此大的影响，甚至改变了行业的叙事逻辑。

当然，在一元商业模式之外，市场上也在探索二元和三元商业模式。目前有两大大类方向。

第一类是AI搜索产品（AI智能助手也类似）。基于搜索引擎商业模式的拓展，AI搜索产品（含智能助手，下同）的商业变现形态比较丰富，能覆盖一元、二元和三元的商业模式。

形态1：一元商业模式。AI搜索直接给用户提供服务，并对C端用户收费。例如，法律咨询、知识创作、内容消费、教育培训、商务服务等都属于这一形态。

形态2：二元商业模式。AI搜索将商家的内容整合到AI回答中，构建一个线上场域连接用户和商家，完成两者之间的商业价值传递，构建二元商业模式，即打造一个泛内循环生态，让商家和用户在AI搜索平台内完成交易。比较典型的如电商、本地生活服务、旅游服务等。作者认为随着产品和技术的发展，后续会有很多专业性和垂直性强的内容和服务通过MCP（模型上下文协议，model context protocol）生态接入AI搜索和智能助手，由大模型主导直接完成交易。例如AI Native的旅游产品可以根据用户需求定制旅游攻略，并直接完成行程中机票和酒店的预订。在这个过程中，航空公司和酒店以外部数据和工具的形式供大模型（遵循MCP协议）调用，大模型自主完成交易，这就形成了二元商业模式。

形态3：三元商业模式。AI搜索还可以遵循三元商业模式的逻辑，将商业内容作为AI回答后的延展激发，通过挂载推荐卡、溯源卡、溯源链接等形态完成商业内容分发；同时与形态2类似，三元商业模式的AI搜索产品也可以将商家的内容整合到AI回答中，甚至直接用商家的内容来回答，在满足用户需求的同时，完成广告曝光。例如Google AI Overviews的溯源外链卡、商品推荐卡、POI推荐卡、地图卡等。

第二类是基于大模型打造的平台型产品。OpenAI的GPTs（GPT store）、百度的千帆AI原生应用商店、字节跳动的Coze等都是类似的方向，但目前都还处于探索阶段，甚至GPTs已经基本宣告失败。截至目前，大模型还没能构造一个B端或者C端的平台型产品，因此并不具备构建二元和三元商业模式的基础。究其原因，本质是目前大模型还不是真正的平台或者说操作系统。真正的操作系统，例如Windows、iOS和安卓，能做到应用层和技术底座解耦。只有大模型做到使两者解耦，应用层才会迎来真正的爆发，AI原生应用的商业空间才会被打开，其商业模式的多样性才会出现。但随着智能体生态的快速发展，作者认为基于大模型打造的平台型产品也即将出现。

另外，我们还观察到，大模型和AI原生应用发展到现在，还没有出现任何新的商业模式（即除一元、二元和三元之外的模式）。AI技术的发展暂时没有展现出重构商业模式的能力。

5.4 小结

本章整体介绍了与大模型相关的基础知识和大模型的应用案例，并对大模型的商业模式进行了思考和研究。希望通过本章的介绍，读者能更好地理解和应用这一重要的人工智能技术，推动人工智能技术在实际场景中的广泛应用和发展。由于篇幅和作者能力所限，还有很多与AI相关的内容无法尽述。本章如有遗漏或者错误，还望读者们多多包涵，并联系我们勘误。

人工智能领域正处于快速发展期，不论是大型模型技术、算法、算力、芯片，还是AI原生应用，抑或是对AI的监管、相关法律法规，甚至人类对于AI的认知都在快速迭代。本章初稿成文于2024年10月，终稿更新于2025年5月。在此期间，DeepSeek的横空出世，已经直接改变了大型模型行业的竞争格局和叙事逻辑。希望AI的发展能让本章的内容尽快成为历史，催生出本章不曾想象到的产品和事物。因为这就代表着AI接下来的发展顺利，而本章作为AI发展历史上的一个切片，若干年后回看，自有其意义。

第6章 超越流量：商业模式中的价值选择

6.1 商业模式度量

在前面的章节，我们通过对商业模式的框架分析和拆解，介绍了公司和业务赚钱或者说盈利的底层逻辑。透过商业模式，我们能清晰地洞察到公司和业务是怎么赚钱的：公司和业务提供了什么价值主张、从哪里获得价值。

这里我们简单回顾一下本书抽象出来的互联网三类商业模式——生产、传递和激发——的定义（见表6-1）。

- **一元生产：**组织生产产品或服务，把产品或服务提供给目标用户，并获得价值的基本原理。
- **二元传递：**组织建设线上场域连接商家和用户，完成商家和用户之间的商业价值传递，并获得价值的基本原理。
- **三元激发：**组织建设线上场域连接用户和内容，在用户消费内容的场景中激发商业需求，并获得价值的基本原理。

表6-1 互联网三类商业模式

商业模式	关键词	典型业务模式举例
一元商业模式	生产	IAP 游戏、IAP 工具、SaaS、流媒体付费业务、租赁模式、社交产品
二元商业模式	传递	交易平台： 电商交易平台、网约车平台、本地生活服务交易平台、OTA 旅行平台、房产交易和服务平台、二手车交易平台
		信息平台： 汽车信息平台、B2B 信息平台、二手房信息平台
三元商业模式	激发	内容平台、社区、免费小说和免费短剧、IAA 应用、搜索引擎

这三类商业模式，能涵盖并分析目前所有的互联网公司和业务。作者认为，在框架和结构分析的基础上，商业模式是可以明确度量的，即一个公司和业务潜在的盈

利空间是可以被合理估算出来的（见图6-1）。度量方法就是前面介绍的经营分析框架。

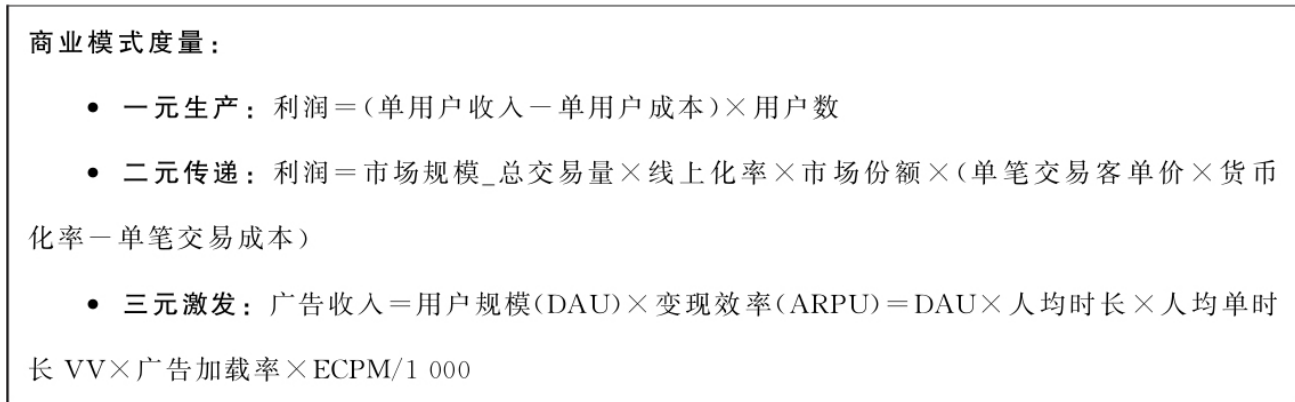


图6-1 商业模式度量

有了这样一套度量方法，我们就可以不依赖实际财报和经营数据的后验结果，在业务分析和模式设计阶段，就进行比较全面的量化评估。

我们分别看一下三类商业模式度量的关键因素。

1. 一元商业模式下的商业模式度量

对于一元商业模式而言，价值主张和收入模型决定了营收规模，成本决定了盈利空间。

价值主张尤为关键，它是一元商业模式的基石，组织生产的产品或者服务最终决定了单用户收入的竞争力和定价权。收入模型则是一元商业模式创新的重要突破之一，从传统的单次交易收入模型到LTV收入模型以及“免费+付费”的收入模型，进一步打开了一元商业模式的营收空间。

成本决定了盈利空间，甚至是盈利的可能性。采用一元商业模式的业务要盈利和扩张，边际成本至关重要，边际成本低的业务才具备规模经济效应。如果有些业务的成本随着用户规模的扩大不降反升的话，那一定要警惕业务发展不可持续的问题。

2. 二元商业模式下的商业模式度量

在二元商业模式下，度量的关键在于货币化率和交易成本。

关于货币化率，作者认为，组织在传递价值过程中，对用户的调度效率越高、对用户行为的控制力越强、提供的价值主张涉及环节越多且价值越大，其货币化的潜力就越大，货币化率的提高空间就越大。但与此同时，货币化率还和商家本身的利润

以及竞争环境相关。如果商家的产品或服务本身的利润率就很低，一定也会约束货币化率；当市场处于激烈竞争阶段时，货币化率高不利于抢占市场份额。我们梳理了影响货币化率的重点特征（见表6-2），在度量二元商业模式时，可以根据其业务特点形成对货币化率的判断。

表6-2 影响货币化率的重点特征

特征	特征说明	对货币化率的影响	例子
区域性 (LBS 属性)	LBS 属性强、受区域约束大的业务，用户需求调度的范围会受到区域的限制，影响调度效率，进而影响货币化率。	区域性越强，货币化率的提高空间越小。	外卖业务 LBS 属性强，如果家附近可选择的外卖商家不多，时间一长，基本上用户选择什么商家受平台影响就比较小了，用户可调度性很弱。
供给规模	供给规模能代表商家侧的竞争程度，供给规模越大，商家侧竞争越激烈，且平台调度的空间越大。上述提到的区域性，也是影响供给规模的因素之一。	供给规模越大，货币化率的提高空间越大。	抽佣和广告是平台货币化的常用方式。例如在广告这种方式中，广告主越多，竞价越激烈，平台赚得就越多。
持续可调度性	指供给和需求之间是否会产生单个用户和单个商家在一段时间内多次达成同一交易的情况，如果两者能形成稳定交易关系，就不具备持续可调度的特点。	持续可调度性越强，货币化率的提高空间越大。	网约车业务中，用户很少能在同一时间、同一地点打上同一个司机的车，双端匹配是持续可调度的。但是美容美发、运动健身等业务，用户和商家之间会在线下完成交易履约的过程中产生交流，并形成固定的搭配，用户会选择同一个理发师或健身教练持续服务，双端匹配不是持续可调度的。
对用户的分配能力	用户在选择商家时的自主性越强，越不好调度，即平台对用户的钱的分配能力越大，变现价值就越高。推荐和搜索是典型的场景，推荐场景下组织对用户的分配能力远高于搜索场景。	对用户的分配能力越强，货币化率的提高空间越大。	和兴趣（推荐）电商相比，货架电商的货币化率低于兴趣电商。

续表

特征	特征说明	对货币化率的影响	例子
价值主张的影响面	组织给商家和用户分别提供了多元的价值主张，在每一笔交易过程中，组织的价值主张影响面越大，产生的价值就越大。	价值主张影响面越大，货币化率的提高空间越大。	贝壳在房地产服务业务中，有几种不同的模式。对比贝壳二手房交易中的自营（1P）模式和平台（3P）模式，1P模式中价值影响面更大一些，所以1P模式的货币化率（接近3%）高于3P模式（约1%）。
商家利润	商家业务本身的利润和利润率，对平台的货币化率也有影响。假设商家本身就不怎么赚钱，那平台也很难赚钱。	商家利润越高，货币化率的提高空间越大。	电商业务中，品牌商家的利润率会比白牌商家更高一些，所以电商平台在品牌商家上的货币化率也会更高。
市场份额	市场份额会影响平台和商家在合作、谈判过程中的地位。市场份额越大，平台的地位和话语权也会越大。	市场份额越大，货币化率的提高空间越大。	贝壳与同行业其他平台相比在新房业务上的货币化率是比较高的，主要是因为贝壳在房产交易中的市场份额大，用户规模大，在和新房开发商的合作中，有更高的议价权。

值得注意的是，很多业务的货币化率并不固定，而是一直处于动态变化中。从企业经营视角来看，组织会不断提升和优化货币化率，进而做大营收规模。非常典型的的就是竞价机制，平台可以通过竞价机制来决定将用户和交易分配给哪个商家，而商家为了销量和销售额，会花钱在平台竞争用户，平台的货币化率就会提高。

但货币化率并非越高越好，过高的货币化率会影响平台生态。平台的货币化率过高，会挤压商家的利润和生存空间，有些中小商家甚至会倒闭。商家都会觉得是在给平台打工，而没办法获取足够的利润来为用户提供更好的产品或服务，这不利于行业生态的长期发展。而且随着货币化率的提高，商家为了提升自己的利润，会采取提高定价或者压缩成本等方式，这也会拉低用户的体验，甚至会导致平台内的商

家进入劣币驱逐良币的恶性循环。例如搬家维修行业由于获客成本太高，商家会倾向于提高单客收入，经常坐地起价，给整个行业带来了负面的影响。

关于交易成本，单笔交易的交易成本及其结构是度量二元商业模式时一个非常重要的切入点，直接决定了企业赚取利润的能力和空间。很多业务做不大，甚至持续亏损，都是成本的问题，例如规模上不去、固定成本分摊不下来、每一笔交易都有无法避免的固定成本、在充分竞争的市场中获客成本居高不下等等。

3. 三元商业模式下的商业模式度量

在三元商业模式下，广告收入规模和空间比较好估算，核心就在于用户规模（DAU）和广告变现效率（ARPU）。

用户规模是业务基础，甚至很多互联网公司的估值与用户规模呈正相关。做大产品的用户规模是一项非常重要的工作，一般被称为用增。用增常用的方式包括广告投放、厂商预装、用户裂变、活动运营和私域运营等。产品的用户规模除了受到用增和内容（参考三元商业模式的飞轮效应）的影响之外，还与产品的受众（目标用户）强相关。产品受众的规模决定了用户规模的上限，很多产品想要拔高上限，就必须不停地破圈。一般而言，越是偏泛娱乐、大众倾向的产品，潜在用户的规模越大，例如短视频的用户规模就远大于小说和漫画的用户规模。

广告变现效率也可以优化提升。虽然某种意义上来说ARPU是个后验指标，但不同业务的ARPU潜力差异巨大。ARPU能用来衡量某个产品对用户注意力的调度效率和对商业需求的激发效率，即组织对用户注意力调度的效率和对商业需求的激发效率，共同决定了三元商业模式平台的广告变现效率。其中需要权衡的要素非常多，整体来看，用户离商业需求越远，激发的难度越大。而用户离商业需求的远近与“人”和“场景”强相关：

● **人**，即用户本身的商业价值。不同人群的商业价值差异非常大，例如爱玩游戏的用户在游戏行业中的商业价值远高于不爱玩游戏的用户，不同年龄段的人商业价值也有差异。

● **场景**，即需求特点、媒介形态、内容环境和上下文。当用户没有明确需求指向时，商业需求的激发效率远高于有明确需求指向的场景（例如，消磨时间的需求对比查询需求，前者的商业需求更容易被激发）；丰富的内容形态其激发效率高于单一的内容形态（例如，视频内容的激发效率远高于文字内容）；连续满足的内容其激发效率远低于不连续满足的内容（例如，短视频媒体平台的激发效率远高于小说阅读平台）；双列形态的激发效率低于沉浸形态，双列形态更强调用户主动选择，而沉浸形态更依赖算法推荐——用户的自主选择越多，平台的激发效率往往越低（例如，快手对比抖音）。

此处，ARPU中还有一个关键影响因素：广告加载率（adload）。简而言之，广告加载率决定了用户在平台消费内容时会看到多少广告。广告加载率如果太低，则会影响广告收入，进而影响业务发展；但如果广告加载率太高，用户可能会因为广告过多而感到烦恼，进而影响用户体验和留存，最终影响长期收入。从实践经验来看，广告加载率存在拐点，越过拐点之后，广告加载率的提升会使体验和收入的兑换效率急剧下降。目前行业内泛信息流类产品的广告加载率大约在7%~20%。

6.2 商业模式度量的对比分析

结合本书第2章到第4章具体介绍的公司案例，本部分我们对同一商业模式的案例进行横向对比，进一步剖析商业模式度量中的关键因素。

6.2.1 一元商业模式的对比分析

一元商业模式相对简单和清晰，但是采用这一模式的企业并非都发展得一帆风顺，我们在第2章介绍的爱奇艺、声网、神策数据、第四范式和Salesforce都有各自的特点和所面临的挑战。虽然它们的赛道各不相同，但从发展历史和业务当下的情况来看，Salesforce的发展比较顺利，且经营情况良好；爱奇艺依靠提升经营效率和对多元模式的持续探索，最近两三年已经逐步走出亏损的困境，但依然任重道远；而声网、神策数据和第四范式都面临着比较大的经营挑战，也都有各自的困境。

其实还有一类采用一元商业模式的业务我们没有展开具体的案例分析，那就是付费游戏业务。付费游戏（一般是中重度游戏）是非常典型且成功的一元商业模式范例。凭借着游戏用户优秀的支付能力和比较低的边际成本，付费游戏业务都有着不错的盈利潜力。国内的腾讯游戏、网易游戏、米哈游、莉莉丝等公司或业务，目前经营情况都很好，是一元商业模式的最佳实践。

横向对比来看，在分析和设计一元商业模式时，有六个因素需重点关注：

（1）**模式**：一元商业模式的创新空间不大，其获取价值的方式从“单次付费”升级为LTV以及“免费+付费”的方式，已经是非常大的创新了。

（2）**成本**：关注成本结构、关注边际成本，要洞察到随着规模的扩大，单用户成本是提高还是下降；关注客户的生命周期和续费率。

（3）**定价**：定价权很重要，但掌握定价权很难。在充分竞争的市场中，公司如果没有差异化竞争优势和市场份额上的领先优势，一定会面临价格竞争，单用户收入提升的空间也会受限。

（4）**付费意愿**：关注目标用户的付费意愿。目前国内B端市场还面临客户付费意愿的挑战。

（5）**渠道**：渠道非常重要，企业的产品和服务必须触达目标用户才能盈利，渠道能力往往是很多一元商业模式企业的核心竞争力。

（6）**外包化**：避免SaaS企业外包化。

6.2.2 二元商业模式的对比分析

我们在第3章具体分析的七个公司（京东、阿里巴巴、拼多多、滴滴出行、美团、贝壳和安居客）实际代表了四种业务类型：电商交易平台、网约车平台、本地生活服务交易平台以及房产交易和服务平台。我们对这七家公司进行横向对比，看一下二元商业模式的货币化率和交易成本。

如前文所述，货币化率会综合受到持续可调度性、对用户的分配能力、价值主张的影响面、市场份额（竞争环境）以及商家利润等因素的影响。结合前面介绍的影响货币化率的重点特征，我们再具体看一下不同业务的货币化率和关键影响因素（见表6-3），帮助大家理解这些公司货币化率背后的底层逻辑。

表6-3 不同业务的货币化率和关键影响因素

业务	货币化率 (参考)	关键影响因素
京东 (3P 业务)	9%	<p>电商平台整体对用户的分配能力和调度效率都很高，这三家的货币化率都还有很大的提升空间。目前影响货币化率的关键因素有以下四方面。</p> <p>1. 价值主张：京东的货币化率高于阿里淘天和拼多多，主要是因为京东在履约环节参与更深。京东为一部分交易提供物流服务，与之对应的就是京东的成本结构中履约成本占比更大。</p> <p>2. 品牌和品类：品牌商家和白牌商家的货币化率不一样，不同品类的货币化率也不一样，不同商家的利润和利润率差异很大。这就是为什么拼多多品牌上行对平台提升货币化率是利好。</p> <p>3. 推荐和搜索：一笔交易是通过搜索成交还是推荐成交，货币化率也不一样。推荐成交的货币化率会高于搜索成交，因为推荐场景中，用户完全受平台调度，平台对用户的分配能力更强，而搜索场景中，用户的主动选择更多。</p> <p>4. 市场份额：国内电商市场目前竞争激烈，各家都在抢市场份额，因此货币化率不会涨得太快。</p>
阿里淘天 (3P 业务)	4%	
拼多多	5%	
抖音电商	12%	<p>将抖音的电商业务单独拿出来，会发现其货币化率能达到 12%，远高于前面三家。这里最关键的影响因素就是搜索和推荐的占比。目前，抖音电商更多依赖公域推荐的成交，即商家在公域投流买量。随着基于搜索逻辑的货架 GMV 占比提升，抖音电商的货币化率一定会下降。</p>

续表

业务	货币化率 (参考)	关键影响因素
滴滴出行 (国内出行业务)	18%	<p>滴滴出行的货币化率比较高，关键影响因素在于其市场份额和调度能力。</p> <ol style="list-style-type: none"> 1. 市场份额：国内在线出行市场中滴滴出行领先优势明显，理论上滴滴出行具备提升货币化率的基础。 2. 调度能力：滴滴出行的订单调度能力是非常强的，用户发单之后，基本没有决策权，完全依赖平台分配。尤其在供给侧饱和甚至过剩时，平台的调度能力更强。 <p>但是，滴滴出行的货币化率会潜在受限于政策和舆论。如果货币化率过高，舆论会对企业不利，进而对企业造成负面影响。现在虽然货币化率高，但滴滴出行会给司乘两端发补贴，所以利润率并不高。</p>
美团外卖	15%	<p>美团外卖的货币化率比较高，关键影响因素在于市场份额、履约服务和调度能力。</p> <ol style="list-style-type: none"> 1. 市场份额：与滴滴出行类似，在国内外卖市场，美团外卖领先优势明显，理论上美团外卖也具备提升货币化率的基础。但 2025 年京东进入外卖市场，会给美团外卖的货币化率带来新的挑战。 2. 履约服务：也就是价值主张的影响面。美团外卖为大部分交易提供履约服务，价值主张影响大，货币化率也就高一些。 3. 调度能力：外卖业务的调度能力是制约其货币化率的关键因素之一。因为外卖业务 LBS 属性强、受区域约束大，所以其调度能力远低于电商平台。 <p>与滴滴出行类似，美团外卖的货币化率也潜在受限于政策和舆论。在舆论视角，外卖商家和骑手赚的都是辛苦钱（尤其是骑手），如果外卖平台抽佣高，赚取了大量的利润，而商家和骑手却不敷出，那么是很难被大众接受的。因此现在虽然美团外卖的货币化率也不低，但配送成本很高，平台利润率并不高。</p>
链家	2%~3%	<p>正如前面提到经纪人角色时所说的，链家在交易中的价值主张影响面很大，但链家的货币化率看起来并不高，其关键影响因素在于客单价。房产交易的单笔交易客单价太高了，基本都是百万元以上，在头部城市，一套房随随便便就是几百万元，所以 2%~3% 的货币化率对应的绝对收入值并不低。</p>

续表

业务	货币化率 (参考)	关键影响因素
贝壳（非链家）	16%	贝壳的货币化率比较高，这里需要注意，贝壳货币化率的基数并不是房款，而是平台上中介的抽佣，即中介的抽佣是贝壳的GMV。假设一笔交易中中介抽佣2%，则贝壳作为平台的收入为（房款×2%）×16%。贝壳之所以有这么高的货币化率，其关键影响因素在于市场份额和价值主张影响面。

注：每个平台的货币化率并非静态不变，本表中当前的数据仅供商业模式分析使用。

从单笔交易的成本来看，对电商交易平台、网约车平台、本地生活服务交易平台以及房产交易和服务平台这几种业务进行横向对比，它们成本的排序是从低到高的。

（1）**电商交易平台最低**。从单笔交易来看，电商业务中，在平台基础设施建设好以后，交易单量和规模的增长没有带来太多的B端成本。例如，新商家的准入、新开一个店铺、新上架一些商品等，几乎都是纯线上完成，虽然可能会给商家增加一些新的运营成本，但没有给平台增加太多额外的边际成本。电商业务中成本比较高的环节是物流，如果电商平台不参与，那交易成本就更低了。京东相较于阿里巴巴和拼多多，交易成本会相对高一些，就是因为在物流环节参与更多，提供自营物流服务。但从业务逻辑上看，京东物流可以视为单独的物流服务供应商，京东3P业务模式也就变成了单笔交易成本较低的电商交易平台模式。

（2）**网约车平台也很低**。与电商交易平台类似，网约车平台在基础设施完善之后，交易量增长带来的边际成本也很低，而且还能进一步放大规模优势。平台每多加入一个司机或者多增加一个乘客，没有增加太多成本。

（3）**本地生活服务交易平台就要高一些了**。以美团外卖为代表，因为平台介入了配送环节，B端成本就出现了，而且这是相对固定的成本，每多一单，就会多一单的配送成本。正如案例分析中所介绍的，美团外卖的履约模式以1P模式（美团配送）为主，而即时零售业务对配送的稳定性和时效性要求非常高，所以平台要对运力有绝对的控制力，因此平台需要建设运力网络并持续提高效率。

（4）**房产交易和服务平台就更高了**，二手车交易平台也是如此。这些平台业务的特点是每一笔交易都非常重决策，且不完全是在线上完成。平台为了撮合交

易，就在这个过程中引入了经纪人的角色。二手房和二手车业务中都有一个经纪人，或者叫跟单员，去撮合两边完成一笔交易。在链家的经纪人模式中，一笔交易还需要不同角色的经纪人参与，房源端和客源端都有不同的角色共同来促成一笔交易。这类业务需要组织全方位参与到整个交易过程中，与之对应的，这类业务的单笔交易成本就非常高，边际成本不但降不下去，反而还会增加。因为这类业务的特点就是单笔交易的固定成本减不下去，而业务规模扩大之后，经纪人团队需要管理，会给组织增加大量的管理运营成本，成本甚至会非线性地增长。

此外，成本结构反过来也会影响货币化率，因为成本结构就是组织价值主张的表达和渗透。

6.2.3 三元商业模式的对比分析

截至目前，互联网企业中市值排在前面的很多企业的核心业务都采用三元商业模式，例如Google、Facebook和字节跳动，可见三元商业模式的效率和魅力。

结合我们在第4章介绍的四家公司（字节跳动、哔哩哔哩、小红书和Facebook）来看，虽然它们都采用三元商业模式，核心都是依赖广告变现，但是我们可以明显看到不同产品的广告变现效率即ARPU是有显著差异的。ARPU能用来衡量某个产品对用户注意力的调度效率和对商业需求的激发效率。调度效率和激发效率越高，广告变现效率越高。正如前文所言，用户离商业需求越远，激发的难度越大。而用户离商业需求的远近与“人”和“场景”强相关，其中“人”主要指人群消费能力，“场景”主要包括用户需求明确度、媒介内容形态、产品交互形式、内容的连续性以及用户对场景内容的预期。这里我们再结合具体的产品展开看一下。

（1）人群消费能力：不同人群的消费能力有差异，因此产品的用户群会影响其广告变现效率。以同为社区产品的小红书、知乎和贴吧为例，小红的用户多为一、二线城市的年轻女性，知乎的用户多为理工背景的高学历男性，贴吧的用户多为怀旧的男性，小红的广告变现效率远高于知乎和贴吧；同样走短视频赛道的快手和抖音，快手的用户比抖音更下沉，广告变现效率低于抖音。

（2）用户需求明确度：用户带着明确目的和任务的需求和消遣娱乐的需求相比，后者激发效率更高。当用户带着明确目的和任务使用产品时，一般都是用完即走，在使用的过程中打断用户，或者在用完之后通过激发来留住用户都是比较难的。因此很多工具类产品的广告变现效率低。为了提升变现效率，很多产品还会增加信息流模块，但用户时长和分发量都很低。搜索场景的推荐化也是如此，很多搜索产品都试过在用户搜索过程中或者搜索完之后，通过推荐来留住用户，提升用户时长，但效果都不理想。而面对消遣娱乐的需求，产品对用户注意力的调度效率和对商业需求的激发效率天然高很多，例如用户在刷信息流、刷短视频、刷直播的时候，他们的需求是很容易被激发的。

（3）媒介内容形态：媒介内容对用户的感官调动越多，其激发能力越强。在人类的五感（视觉、听觉、嗅觉、味觉、触觉）中，媒介能同时调动的感官越多、内容形态越丰富，对用户的激发效率就越高，即在激发效率上音视频>静音视频>动图>图文>文字。所以传统的图文产品，例如今日头条、微博、知乎等，变现效率比不过以短视频为主的抖音、快手。因此图文产品为了提升自身的变现效率，纷纷推进视频化，且成效显著。

（4）产品交互形式：沉浸式、双列和单列（列表页）是三种最典型的产品交互形式。而抖音几乎以一己之力向全世界证明，沉浸式上下滑是手机屏幕上效率最高的

交互形式。所以现在很多产品越来越像抖音，例如小红书的竖版视频模式、B站的Story模式、YouTube的Shorts模式等。几年前今日头条做的“短小融合”（短视频和小视频融合为沉浸式视频流）也取得了不错的成果。当然，作者也期待在终端设备上出现更丰富的、效率更高的产品交互形式。

（5）内容的连续性以及用户对场景内容的预期：连续满足的内容和用户预期确定性高的内容，激发效率相对低于不连续满足的内容和用户无预期的内容。例如在用户看小说的时候，插入广告的激发效率就比用户刷短视频时低。同样，私域场景激发效率低于公域场景，因为用户对私域场景中会出现的内容有预期，例如同样是微信中，朋友圈广告的激发效率相对低于其他公域场景。

整体而言，一个采用三元商业模式的产品，其广告变现效率受多个因素的综合影响。同类产品受产品交互形式、人群消费能力等的影响，变现效率会有差异；不同类产品的变现效率之间更是差异巨大。我们可以将广告变现效率分为高中低三个梯队，各梯队的代表性产品和特征关键词如表6-4所示。

表6-4 广告变现效率排序

梯队	广告变现效率	代表性产品	特征关键词
第一梯队	高	Facebook、Google、抖音、小红书、快手	短视频、富媒体、消磨时间的需求、内容满足不连续、高分发量
第二梯队	中	B站、爱奇艺、微博、Twitter、七猫免费小说、番茄小说、红果短剧	中长视频、内容满足连续、高分发量、高时长
第三梯队	低	豆瓣、知乎、马蜂窝、贴吧、墨迹天气、万年历	用户消费能力弱、垂直内容社区、低活低时长、低分发量、内容满足连续、IAA工具、IAA游戏

6.3 商业模式组合

公司和业务的商业模式非常复杂，很多都是多种商业模式的组合。一个业务里面可能包含好几种商业模式，而一个公司又可能包含好几种不同的业务。拿大家都非常熟悉的抖音来举例子，它的商业模式就很复杂。

在抖音平台内部，同时运行着电商、游戏、直播、广告、本地生活等多种业务形态。进一步拆解来看：在抖音电商中，既有平台自营店铺，也有第三方商家店铺。其中，自营店铺里还存在货权发生转移的商品，以及货权未发生转移的商品。在游戏业务中，既有抖音自研的游戏，也有非抖音自研的联运游戏。在直播业务中，也有直播打赏与直播带货等不同模式。大家可以思考一下，这些不同的业务形态背后的商业模式分别是什么。而需要强调的是，抖音还仅仅是字节跳动整体业务中的一部分，并非全部。

一家公司为了提升竞争力，会逐步打造自己的生态，建设“护城河”。公司往往既是生态的建设者，又是生态的参与者，并且这些生态有无数大大小小的角色参与其中，每个角色背后可能都是一家公司，它们的商业模式不尽相同。

因此，在进行商业模式度量之前，我们首先需要准确识别商业模式。区分一元、二元与三元商业模式并不复杂，关键在于理解业务结构中包含哪些关键角色，以及价值是如何被创造和获取的。作者建议，在研究企业及其业务时，应以其主要商业模式为判断基准，具体可观察其利润构成（若处于发展初期利润尚未体现，则可参考营收构成）。通常，应以贡献度最高的模式为准，即将利润或营收占比最高的商业模式认定为主商业模式。以字节跳动为例，尽管其业务体系庞大复杂，但至少到2025年，其收入中仍然是三元激发模式下的广告收入占比最高。因此，从本质上看，它仍可被定义为一家广告公司。

我们总结了一些典型互联网公司一部分业务的商业模式组合（如表6-5所示）。

表6-5 典型互联网公司部分业务的商业模式组合

公司	部分业务举例	业务分类	商业模式组合
Google	Google Search	搜索引擎	激发
	Google Cloud	SaaS	生产
	YouTube	视频平台	激发

续表

公司	部分业务举例	业务分类	商业模式组合
阿里巴巴	淘宝/天猫	电商平台+电商自营	传递+生产
	阿里云	SaaS	生产
	阿里 B2B	交易服务平台	传递
腾讯	微信	交易平台+内容平台+社交平台	生产+传递+激发
	腾讯视频	长视频流媒体平台	生产+激发
	腾讯游戏	IAP 游戏（主）+IAA 游戏	生产（主）+激发
字节跳动	今日头条	内容平台（资讯）	激发
	抖音	短视频平台+电商平台	激发+传递
	飞书	SaaS	生产
小米	小米手机	硬件	生产（传统商业模式）
	小米汽车	硬件	生产（传统商业模式）
	MIUI 系统软件	应用商店+IAP 工具+IAA 工具	生产+传递+激发
	小米汽车智能驾驶系统	SaaS	生产
	小米有品	电商平台	生产+传递
美团	美团到家（外卖）	电商平台（本地生活，即时零售）	传递
	美团充电宝	租赁	生产
	美团单车	租赁	生产

上述这张表只列举了这些公司的一部分业务，没有包含它们的全部业务，可见公司和业务商业模式的复杂性。但基于对商业模式原子化的思考，我们就可以通过商业模式组合，来描述这些复杂的公司和业务，洞察其本质，并对其进行商业模式度量。这就是原子化思考的价值。

6.4 商业模式量表：商业模式中的价值选择

基于对商业模式度量的综合分析，我们认为：评估一个公司或业务的盈利潜力与增长空间，不能仅关注用户规模的大小，更应深入审视其价值主张。企业的盈利不仅来自流量变现，从根本上更源于其价值主张在商业模式中所发挥的影响力。企业在商业模式中做出的价值选择，从根本上决定了其盈利能力。

在本书的最后，我们结合前面介绍的内容，设计出一张商业模式量表，也可以称作商业模式体检表（见表6-6）。大家在研究商业模式时，可以将商业模式量表和商业模式画布结合使用，从而更清晰地洞察企业盈利的底层逻辑和盈利能力。

表6-6 商业模式量表

	<input type="checkbox"/> 一元	<input type="checkbox"/> 二元	<input type="checkbox"/> 三元
目标用户			
价值主张			
商业模式逻辑图			
市场份额（竞争情况）			
经营分析框架	收入模型	GMV	DAU
	单用户收入	货币化率	时长
	单用户成本	交易成本	ARPU（单时长 ARPU）

续表

	<input type="checkbox"/> 一元	<input type="checkbox"/> 二元	<input type="checkbox"/> 三元
关键要素评估	边际成本	区域性	人群消费能力
		供给规模	用户需求明确度
	定价权	持续可调度性	媒介内容形态
		对用户的分配能力	产品交互形式
	目标用户付费意愿	价值主张的影响面	内容的连续性以及用户对场景内容的预期
		商家（供给侧）利润	
盈利潜力空间判断	<input type="checkbox"/> 高 <input type="checkbox"/> 中 <input type="checkbox"/> 低	<input type="checkbox"/> 高 <input type="checkbox"/> 中 <input type="checkbox"/> 低	<input type="checkbox"/> 高 <input type="checkbox"/> 中 <input type="checkbox"/> 低

商业模式量表的使用说明如下：

（1）识别业务的商业模式属于哪一类，或者是哪几类商业模式的组合，这是后续研究的基础。

（2）目标用户和价值主张：与使用商业模式画布时类似，我们需要将商业模式的目标用户和价值主张列出来，这点与商业模式分类相辅相成。

（3）商业模式逻辑图：基于前述内容，我们就可以画出目标业务的商业模式逻辑图。这张逻辑图能够帮助我们对业务的商业模式进行更好的抽象和理解。

（4）市场份额（竞争情况）：这本质上是洞察业务的市场竞争情况。市场竞争情况非常重要，直接影响企业的经营战略和发展潜力。

（5）经营分析框架：运用经营分析框架可以计算或估算出业务的营收规模。

（6）关键要素评估：我们可以使用分值区间（例如1~5分）或者等级（高中低）对影响商业模式盈利能力的关键要素进行评估。在实际应用场景中，可以直接对多个业务进行横向对比和排序，也可以选择某一个具体的业务作为基线，参考基线对其他业务进行排序。评估时需要综合考虑市场份额以及经营分析框架部分的计算结果。

（7）盈利潜力空间判断：基于关键要素评估的结果，我们就可以对业务的盈利空间有更准确的判断。但需要强调的是，实际情况下，影响业务发展的因素非常多，

除了商业模式本身，还有人才、团队管理、战略执行力、外部市场环境变化等。这张量表更能体现的是业务本身的底层逻辑。

相信只要读者朋友们能够灵活运用这张商业模式量表，它必定会成为大家在分析、评估和设计商业模式时不可或缺的实用工具。

6.5 小结

本章结合前面章节的内容，在分析商业模式框架和结构的基础上，介绍了商业模式的明确度量方法。通过商业模式度量，我们可以合理估算业务的潜在盈利空间。以经营分析框架为基础，我们还对前文分析过的公司案例进行横向对比，全面剖析了三类商业模式度量中的关键因素，包括一元商业模式中的六个因素、二元商业模式中的货币化率和交易成本，以及三元商业模式中的用户规模（DAU）和广告变现效率（ARPU）。

同时，本章还设计了一张商业模式量表，旨在给读者提供一个可以用来分析、诊断和设计商业模式的工具，通过对业务商业模式中关键要素的量化和评估，就可以很清楚地判断一个业务的盈利潜力和盈利空间。

商业模式是业务的基础，也是一家企业经营的底层逻辑。商业模式的创新至关重要且颇具价值，它改变了很多业务的运作方式，拓展了市场规模，甚至颠覆了许多行业，催生了新的叙事方式和全新的行业格局。当下，我们正身处大模型和生成式AI的浪潮之中，我们相信AI技术会改变很多东西，并为商业模式创新带来更多可能性。

商业的世界复杂多变，市场竞争激烈且格局瞬息万变。那么对于身处其中的企业而言，不变的是什么呢？作者认为，不变的正是价值主张本身的价值。正如评估一个公司或业务的盈利潜力一样，无论其用户规模多大、流量多高，其核心仍在于它能为目标用户创造多少真实价值，以及它的价值主张在商业模式的各个环节中究竟能产生多大的影响力。

商业模式度量不能只看流量规模的大小，而应该超越流量，去看企业在商业模式中的价值选择。