

## 文字游戏



埃雷兹·艾登、简-巴普蒂斯特·米歇尔

**未知**

将大数据作为人类文化的透视镜

河源出版社，纽约，2013年版，288页，27.95美元（精装）。

本书中包含了許多未知事物。但是“未知”并不意味着埃雷兹·艾登（Erez Aiden）和简-巴普蒂斯特·米歇尔（Jean-Baptiste Michel）不重视这些事物——事实上，“图表示意”或许是一个更加恰当的标题。

本书讲述的是搜集世界上所有图书中数亿个单词的故事。先前在文本意义中由于“未知”而丢失的单词，却被本书编制为核心内容。作者希望整个制图过程会带领我们探索文化中许多有趣的方面，他们称之为“文化经济学”。艾登和米歇尔利用谷歌制作一个强大的网络工具，但是有关其实用性的言论也许有些夸大。

这并不是说本书毫无乐趣可言。从本书一开始的鸣谢中你就能够发现这是一本有趣的书：在鸣谢中艾登对他的三个孩子致以谢意，他一个女儿的中间名字为“香蕉”（Banana）（他一贯这样古怪精灵；他的儿子名叫伽利略）。

目前，我急于寻求乐趣。但是艾登和米歇尔正在撰写重要的科技

著作，在书中举一些“有趣的”例子并不会有助于他们的工作。无需大量的数据说服我们，卓柏卡布拉（chupacabra）一词（据报道，1995年人们在波多黎各亲眼目睹了这种吸血生物）远比北美野人或尼斯湖水怪罕见。在20世纪40年代（难于详述本书中再现图表的起始日期）至2000年间发行的图书中，图表标注“argh”和“aargh”的使用变化是一件看似愚蠢的事情。书封皮引用了《琼斯母亲》杂志的一句话，将Ngram Viewer称之为“互联网有史以来最伟大的时间浪费者”。出版商作此引用实属气魄之举。

通过机器人阅读每本发行图书上的每一个单词，以记录文化历史是一个大胆的尝试。那么，我所说的这项工作有所欠缺是什么意思呢？艾登和米歇尔承认他们正通过单词细微的例子来探寻文化历史，尽管他们说谷歌迄今为止已浏览约3000万本图书（截止到目前也许数量更多），但仍然还有约一亿本图书需要浏览。

此外，倘若一个文字的使用是我们窥探文化历史的线索，本书中忽略了许多其他资料来源：报纸和杂志文章、信件、电影、电视和电台访问、改写本和讲座——事实上，忽略了所有除已发表图书以外的以书面或口头形式存在的文字。此外，著作撰写完成后，时常进行语法和拼写的编辑和修订，更别说翻成其他语言。每一位作者都了解编辑会根据他们所在出版社的风格修改文章。我不知道图书（甚至3000万本图书）中所使用的语言是否是语言使用变化的可靠来源。

我猜想艾登和米歇尔会说谷歌所浏览的图书全部是他们必须搜集的素材，但是考虑到有所遗漏，他们所提出的“人类文化透视镜”理论或许是一个太过大胆的言论。

至于书中的图表：图表数量很多，但被精简过；以黑、白、灰三色印刷；并且从数据资料中直接生

成。简单的图表并没有什么不妥，但是缺少相对的标签和网格线以及标注在区分度很小的灰色阴影里的图表分度线（有时多达6条），使其难于阅读。本书作者告知我们用网络浏览，这样所有这些问题都将

**本书讲述的是搜集世界上所有图书中数亿个单词的故事。**

被解决：颜色区分线条，并且点击分度线上的任意一点都会有标签和日期显示。这是一个关于印刷图像和网络图像之间差距的例子。

但从乐观的角度来说，作者并没有规避在浏览众多图书时所引发的版权问题，或者用看似不道德的“阴暗”方式处理这些问题。书中生动的描述了有关拉里·佩奇（Larry Page）和玛丽莎·梅耶尔（Marissa Mayer）在2002年所进行的实验，她们计算出浏览完世界上所有的图书需要花费多长时间。显然，这将需要花费“几千年，甚至更久”。因此，作者们是如何处理这个问题的？阅读本书，其乐无穷！

奈杰尔·霍姆斯  
(Nigel Holmes)

Explanation Graphics 负责人，  
最新出版的《Wordless  
Diagrams》和《The Book of  
Everything》的作者

# 了解每件事物的价格，却不了解其价值



戴安娜·科伊尔

## GDP

简短而充满深情的历史

普林斯顿大学出版社，普林斯顿，新泽西州，2014年，168页，19.95美元（精装）。

既然“经济”一词起源于希腊文 oikos（意为家庭管理），那么为什么在2000年前没有像亚里士多德这样的聪明人提出国内生产总值（GDP）的概念呢？这是因为亚里士多德潜心研究运动的物体，如月亮和其他行星。

在大约5万年的时间里，人类的GDP并不会上下浮动。在亚里士多德那个时代，并没有美国全国广播公司财经频道通过广播GDP的上下波动的方式影响市场。也没有人希望过上比他们父母更好的生活。当然，人们也关心资金和债务。在16世纪，亨利八世命令其财政大臣密切关注他在酒馆消费和在与法国的战争期间的账单，但他绝不会想到去过问一下人均GDP是否比去年有所增长。

戴安娜·科伊尔（Diane Coyle）的新作《GDP：简短而充满深情的历史》巧妙易懂，讲述了20世纪GDP产生的故事，这一数值的创造使得各国首相在之后的每一季度倍感头疼。在本书的开头部分，科伊尔向我们提及了GDP对现代雅典而非古希腊的重要性，雅典国家统计局的负责人将其工作称之为“格斗

运动”。她还分享了她的朋友经济学家宝拉·苏拉奇（Paola Subacchi）的经历。宝拉·苏拉奇在皇家国际事务研究所工作，曾拜访希腊相关机构，期待在那里看到超级计算机或者至少看到一个算盘。然而，当她爬上一座建于20世纪50年代的住宅楼时却发现“三五个人待在一个满是灰尘的屋子里”，而屋内并没有电脑。

但是国家统计局必须想出一些对策，经常制造虚假数据以达到出售债券的目的并哄骗他人提供援助。科伊尔称中国官员有时大肆鼓吹其雄厚的GDP，而有些时候却减少GDP以使其有资格获取援助。苏联解体以后，我曾拜访过圣彼得斯堡。陈旧的经济学教科书曾说苏维埃社会主义共和国联盟在共产主义的领导下国力昌盛。甚至像保罗·萨缪尔森这样的诺贝尔奖获得者也曾发表过这样模棱两可的数字。但是我需要做的仅是在这个日益衰败的冬宫博物馆里呼吸着刺鼻的空气并意识到：共产主义所面临的问题不是跟不上西方的发展步伐，而是它还赶不上1917年西方的生活标准！

科伊尔完成了一项重要任务。她提醒我们，对GDP精确计算（ $C + I + G + [X - M]$  或消费 + 投资 + 政府支出 + 净出口）刺激政府领导人支出更多。原因何在？因为加强政府支出能力可以重复地增加GDP。因此，国家领导人只需放开支出限制，就可以依靠其统计专家增加GDP。而且，政府支出的价值是基于政府公职人员的工资进行计算的，而非其产出价值。我的一名哈佛学生曾经说过，鉴于这种无意义的重复，偶尔考虑预算紧缩的国家领导人值得特别的尊重。

科伊尔在将GDP问题进行拆分方面同样出色，包括一些悖论，如娶了女管家的鳏夫会因此降低GDP，因为他不用再发给女管家任何工资。合理评估信息经济服务尤其困难。举一个当下现实生活中的例子，我设计了一个新的数字矩阵，以帮助孩子学习算术。当孩子们通

过这种叫数字箭头的矩阵图学习加法时，他们会增加，比如说，10万美元的收入潜力。但是其应用成本仅为4.99美元。这个矩阵的每一次下载都会创造10万美元的价值？还是仅仅一点而已？

将GDP的问题进行拆分以后，科伊尔质疑我们是否能做得更好，通过竞争者的指数列表，包括人类发展指数（HDI），经济福利标准（MEW）以及各种各样的幸福指数。科伊尔的质疑是正确的，尤其是质疑那些被幸福大师和煽动者所鼓吹的言论。乌戈·查韦斯将GDP称为一个“资本主义的阴谋”。但是其他选择更易被扭曲，就像太妃糖一样。例如，2009年幸福星球指数显示，在众多国家当中哥斯达黎加的幸福指数位于榜首，古巴紧随其后，排名第七。同样，还发现巴勒斯坦自治政府的子民比以色列人民更幸福、更健康。如果一个支持犹太复国主义者发言人称巴勒斯坦人民经济状况更优裕，他将会受到嘲笑或者是被扔石子。呵呵，美国在该份名单中排名第114位。有趣的是，我从未见过一位居民乘伐子离开迈阿密前往古巴。所以，科伊尔剖析GDP的缺陷和对其临近的风险提出警告都是再正确不过的了。

在她短小精悍的论述中，我仅发现一处遗漏。当我想知道一个国家的生活标准的时候，我经常会同这样一个简单的问题：一个普通工人需要工作几个小时才能买到一只鸡？在20世纪20年代，赫伯特·胡弗的竞选宣言承诺“每家每户都能吃到鸡”。在那些日子里，人们花两个或一个半小时的时间就能挣到购买一只鸡的钱，而如今只需不到15分钟。对于我而言，这似乎是进步。当然，除非你是家禽。

托德·布赫霍尔茨  
(Todd G. Buchholz)

前任美国白宫经济政策主任，  
《New Ideas from Dead Economists》的作者，  
教育软件公司 Sprogit LLC 的首席执行官