

a mindstyle business 与思想有关

BIG DATA

THE BUSINESS
REVOLUTION

车品觉◎著

大数据实践的先行者，红杉资本中国基金专家合伙人
阿里巴巴集团前副总裁，首任阿里数据委员会会长
车品觉 首部个人专著

决战大数据

升级版

大数据的关键思考



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE



BIG DATA
THE BUSINESS
REVOLUTION



决战大数据

大数据的关键思考

车品觉◎著



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

图书在版编目 (CIP) 数据

决战大数据 (升级版) : 大数据的关键思考 / 车品觉著 . —杭州 : 浙江人民出版社, 2016.4

ISBN 978-7-213-07227-7

I. ① 决… II. ① 车… III. ① 企业管理—数据管理 IV. ① F270.7

中国版本图书馆 CIP 数据核字 (2016) 第 053493 号

上架指导: 经济趋势 / 大数据

版权所有, 侵权必究

本书法律顾问 北京市盈科律师事务所 崔爽律师
张雅琴律师

决战大数据 (升级版) : 大数据的关键思考

作 者: 车品觉 著

出版发行: 浙江人民出版社 (杭州体育场路347号 邮编 310006)

市场部电话: (0571) 85061682 85176516

集团网址: 浙江出版联合集团 <http://www.zjcb.com>

责任编辑: 朱丽芳

责任校对: 陈 春

印 刷: 北京鹏润伟业印刷有限公司

开 本: 720mm × 965 mm 1/16

印 张: 19.5

字 数: 19.8万

插 页: 1

版 次: 2016年4月第1版

印 次: 2016年4月第1次印刷

书 号: ISBN 978-7-213-07227-7

定 价: 56.90元

如发现印装质量问题, 影响阅读, 请与市场部联系调换。

BIG DATA

THE Business Revolution

—— 重磅赞誉 ——

在我眼中，品觉一直是我非常尊敬的国内数据领域的大师级人物，这是他的第一本著作，也是汇聚了他十几年数据行业经验，呕心沥血的精髓之作。他一直强调数据是尊重人性的，只有有商业实效性的数据才是有质量的数据；一直强调不能只是纯粹地看数据，要用数据还原真实的现实与场景。这些见解不仅对每个企业的未来发展具有极大的指导意义，而且对每个普通人建立数据化思维和进行个人数据管理都有很好的启发作用。这本书贵在坦诚和实用，每个人都能从书中收获颇多。

沈 亚

唯品会创始人，董事长兼 CEO

我跟品觉相识多年。在大数据还没有流行以前，他就开始关注这个趋势，研究数据应用的意义以及怎样真正读懂数据和利用数据。这本书深入浅出，把深奥的大数据讲得生动有趣，非常好读；书里面集中了品觉个人的很多真知灼见，非常值得一读。阅读此书是一种享受，正如我享受与品觉的友谊一样。

林奕彰

eBay 大中华区 CEO

在中国，可能没有人比品觉玩的数据多，更没有人像他一样玩转了大数

据。在这本书里，品觉以他特有的通俗易懂的语言风格，讲活了大数据，并结合他在阿里巴巴的实践，揭示了数据化运营和运营大数据的实战秘笈，非常赞！

刘 星

红杉中国合伙人

作为一个企业经营者，说起大数据，最怕只见数据不见价值；作为一个互联网产品老兵，最焦虑于数据挖掘和分析脱离用户场景。本书好就好在实践了作者书中提出的观点，立足应用场景，聚焦如何让数据产生实际价值。因其实战性和创新的角度，值得所有关心大数据实战者认真一读。

张 涛

大众点评网创始人

大数据像铺天盖地的洪水一样涌来！由于近年来数字媒体、网络和移动技术的迅猛发展，数据的积累速度已对数据的存储、管理、分析和决策应用提出了前所未有的挑战。很多企业、政府机构、学校和研究机构为了能在大数据时代继续生存和发展，都在重新定位和寻找新的方向。让我们高兴的是，作为一个先行者，品觉在阿里巴巴已经走出了一条大数据运营的路，并又在运营大数据方面总结了很多成功经验。他的《决战大数据》一书尤其为读者提供了“数据化思考”的模式和框架。通过讲故事、作比喻，品觉打破了大数据的神秘，然而又能使读者产生丰富的联想，开动脑筋，真正理解大数据成功运用的要诀。感谢品觉为“大数据金矿”的探索开发作出的贡献。

程 杰

Acxiom（安客诚）前全球副总裁，数据科学家

“大数据”和绝大多数“新概念”一样，已经成为许多闭门造车者宣扬的主题，而品觉——一个被我视为数据科学家的朋友，将他在 eBay 和阿里巴巴近十几年的实战经验进行梳理和总结，使“大数据”不再是一个人云亦云的概念，而是一个可以用心领会和使用的科学方法。

吴海

桔子水晶酒店集团创始人

品觉的书主要讲的是决战！那么咱们就从兵法的角度来看大数据！只讲三点：第一，对“道”的理解：即分析师对人和事物基本规律的诠释。分析问题千万不要从分析大数据开始，而是要从对人、世界、产品或者商业行为最基本的认知着手！第二，对“计”的理解：计就是计谋！交战之前用“商业智能”的一个重要作用就是要造成信息情报不对等，然后进一步造成了战略优势的不对称，从而造就取得优胜的“势态”。第三，对“胜”的理解：决战的目的是要胜利，兵法上取胜的一些基本的要领比如以快打慢、以少胜多、以众击寡，在大数据分析上完全适用，要做到分析得快速而精准、大规模部署以及产品化等。这次为品觉的新书做推荐，既感到非常感激和荣幸，又感到诚惶诚恐。我的水平非常有限，完全是抛砖引玉，希望读者们细细品味《决战大数据》！

张溪梦

GrowingIO 联合创始人兼 CEO

未来几年将是大数据技术及其应用发展的黄金时期，本书是车品觉老师根据自身在蚂蚁金服以及阿里集团的多年数据实践，并结合理论对数据产品规律进行了深入阐述。本书构思流畅，由术入道，深入浅出，既适合初学者以及业务人员了解大数据的相关知识，也适合专家级别的读者

细细品味，无疑是大数据领域的必读书籍。

崔晓波

TalkingData 创始人兼 CEO

大数据领域一直以来都是理论和实作两大山头并立：如果说谈理论是文、卷起袖子实作是武，那品觉兄肯定是这个大数据时代文武兼备的一代宗师，因为他不但在实务界有着丰富的操盘经验与彪炳战功，又能透析数据核心、自创理论心法。最难得可贵的是，他一直以来都秉持着利他无我的气度，对提携数据后进始终不遗余力。听品觉谈大数据，除了“激动”与“热血沸腾”之外，我再想不到更确切的形容词了！

陈杰豪

MIGO 集团 CEO

老车是“手艺人”，数据时代解读数据世界奥秘的手艺人；对我来说，亦师亦友、亦火亦水。我们曾若干次品茗促膝、围炉夜话，人生的起伏、数据的理解与应用、佛法修行，每个话题都始于数据终于数据，获益匪浅。难能可贵的是，我们对数据的理解与应用，观点出奇的一致，相互平添好感。老车是数据的修行者，我十分期待这位无疆行者的感悟与升华之作。一杯清茶、一本好书、一位知己好友，人生幸事。

徐 雷

京东集团高级副总裁

我在阿里的6年

回顾我在阿里的6年，从经历来说，是先负责支付宝，继而负责淘宝，最终负责整个阿里集团的数据工作。实际上，这个过程同样也是我对数据应用的一个历程：第一阶段用数据，第二阶段养数据，第三阶段从看数据到用数据。

我在支付宝

加入支付宝之前，业内都认可我是一个比较懂得用数据去设计、优化、迭代产品的首席产品官。彼时我在支付宝，大胆地把产品与数据分析部门结合，积累了不少经验。

• 数据产品化，这个理念随我加入支付宝

说起加入支付宝，其实有些偶然。当时，支付宝邀请我去分享如何用数据做产品。分享会之后，我去了华星大厦的一个会议室，里面坐着 Lucy（彭蕾，时任支付宝 CEO）、Eric（井贤栋，时任蚂蚁金融 COO）。在那个会议室里，他们问了我几个如何能用好数据的问题。当时，我反问了 Lucy 一个问题：“你每天大概会看多少数据？”

“大概几十个吧。”

“不对，你不应该一天看几十个数据那么多，

BIG DATA

The Business Revolution

前言

我觉得你应该只看几个数据就够了。”

或许，他们已经不记得我提出的这个问题了，但是我却记得非常清楚。因为在支付宝的两年里，这是我做数据产品时一直遵循的原则。

正式加入支付宝后，我面对的第一件事情就是着手组建数据分析师团队。2010年的支付宝，公司大部分业务团队对数据团队都不太满意，而且高层对数据能产生的价值也没有太多认识。而数据负责人的调职，也导致了支付宝的数据分析和数据技术部门被分开管理。

面对数据分析和数据技术团队的目标不一致，加上业务部门对为什么要使用数据及如何用好数据充满疑问，甚至有人认为数据分析仅仅是门面功夫，所以技术团队有时候会跟我的团队“对着干”，因为两个团队的老板不是同一个人，他们也没有义务一定要支持我，所以甚至有时候会撇开我们直接跟业务部门对接。

但是最后，我坚持的数据分析产品化路线最终得到了双方的认同，技术团队甚至还主动过来问我：“我们怎么合作比较好？”所以那时，双方的关系从“对着干”变成了“我们会全力支持你”。

终于有一天，Lucy走进我办公室跟我说：“你知道吗？你们做的产品让我觉得很爽，我很喜欢这个产品。”而当我在支付宝的第二个产品“地动仪”出来时，她对我说：“这就是我梦寐以求的数据产品。”

同事和老板对结果的支持，让我肯定自己在支付宝走的这条路是对的。如果让我总结这其中的精髓，我的秘诀就是：**数据分析也要讲求用户体验。**

简单来说就是，别说是数据产品，哪怕一个小小的数据报告，我们

都必须要知道以下 5 个问题：

- 这是什么问题？
- 这是谁的问题？
- 这个问题现在必须马上解决吗？
- 这个问题我们能解决吗？
- 我们有足够的数据来解决吗？

当我“苛刻”地提出“把月度经营报告当成产品来做，目的是让这份报告像 iPhone 一样好用”的要求时，整个团队都傻了：报告怎么能像产品一样呢？！那么，为什么数据报告不可以做到行云流水，让人看起来很爽呢？

所以，我是这么要求的：**首先，报告的前 3 页一定要吸引人，让人有欲望继续看下去；其次，当别人看到一个数据，心里在猜测数据背后的各种原因时，我的下一页报告就要解答他心中的疑问。**

因此，如果要用产品管理的理念来运行分析团队的话，必须要问“问题是什么”，并以此作为切入点。接着，解决用户的问题，得到用户的信任，然后就可以做更多的好产品。这是一套做产品的理念，因为产品需要不断迭代，并非一劳永逸。

那么，我怎么评估报告是否获得成功呢？简单来说，如果一个月度经营报告仅仅 5 分钟就被 pass 掉了，那它肯定是不成功的。

我在支付宝，几乎每一个月度报告都会让管理层讨论三四个小时，一份数据报告居然会成为管理层讨论焦点，每次都会有很多高层管理者关注我讲什么，这时我就知道自己肯定成功了。

当然，我也同时会提醒分析师们，千万别以为靠一份报告就可以解决所有事情，你一定要让业务部门认为这份业务报告是大家一起探索的，这是我管理团队的一个基本理念。

● 用产品化理念管理数据团队，这是推进的秘诀

总的来说，数据产品化的理念最终得到了 Lucy 及其他高层管理者，以及团队和同事们的认可。那么针对这个理念，我是如何在团队内一步步推动的呢？

首先，我让团队做好基础服务——“白米饭”。为什么叫白米饭？因为对人来说，如果“饭”不够就会出大问题。对于公司来说，如果数据基础服务满足不了我们的用户，那么他们就会跳起来，那时候再高大上的报告也没用。所以，我要求分析团队一定要跟业务部门达成一致，知道什么东西如果我不提供给他们，结果会很严重。

但是，长期做“白米饭”的分析部门是没有前途的。所以其次，我抽出团队一部分人成立了“突击队”。每当我听到管理层对数据的需求时，就会把信息带到团队，让团队成立两三支“突击队”，让管理层用到有用的、合适的以及具有带探索性的内容。把自己最精锐的部队放在“突击队”里，这个决定非常重要。

但是值得注意的事，“突击队”有一个前提，必须要有人能接触到最高管理层的信息通道，否则无法得到高层的信息。而这个角色就是我，我接触最高管理团队，并且把信息输送给“突击队”。

基于“白米饭”和“突击队”，最后一步，我们开始沉淀数据分析的框架。为什么数据分析框架如此重要？因为它是数据产品的养分。我

用建立数据产品为目标的理念来运营一个数据团队，所以数据分析框架才是给到数据产品养分的关键点。

在支付宝时，基于 Lucy 和 Eric 给我的支持，数据产品都是自上而下推行。我先做了“观星台”，这是一个高管的数据仪表盘工具；然后做了“地动仪”，一个给中层管理看的数据产品，接着还有客服 360 等。

所有这些数据产品，都是从数据分析框中提炼出来的。**把分析的理念和框架变成数据产品，本质上是一个数据泛化的过程。**这个过程非常重要，因为数据报告的需求会越来越多，如果没有泛化数据给使用数据的人，分析团队将永远被冗杂和重复的工作所困。

在这期间，我们也做过一个改变。因为“观星台”和“地动仪”还是传统的分析工具，只不过让你有更好的“眼睛”看到更好的数据，用我一直强调的理念来讲，这属于描述性的报告，用于描述一个公司今天做的怎么样。但由于我们有了分析框架，可能稍微会带一点诊断性。所以，这种产品主要是描述性产品，偶尔会做一些诊断，但还未达到探索预测的状态。

在这个看法的基础上我有了个新想法，想法的灵感源于 Facebook。当时 Facebook 已经积极开展探索性的数据产品，可视化及交互功能已经非常灵活。探索和静态报告的区别在于，用户可以选择自己的角度灵活地去看数据，真正做到“逛”数据。

因此，我们创造了一个产品叫“黄金策”。“黄金策”可以让用户很轻松地在 3 秒钟之内就看到不同标签下用户购物行为的异同。比如，30~35 岁的上海女性，有一张信用卡和多张信用卡的两种不同群体，看她们的购买行为有什么不一样。或者，看广东和上海的女性在购物表现

上有什么不一样。用户可以非常轻松地比较不同用户群的购物表现。所以，我认为好的数据产品应该是可以激发用户灵感的。

在做“黄金策”这个产品的过程中，我深有感触的是：**数据行业的人会很容易让技术制约产品的发展**。当时我提出做“黄金策”时，大部分技术工程师都断言，3秒钟出结果是不可能的。因为以前类似的灵活配置的报告，一般都不能做到即时查询，如果要快速产出的话只能找分析师。即便如此，也可能因为资源排期等原因等待很久。

当时只有一个人站出来说“我可以帮你搞定3秒钟”，这个人就是蒋杰（现任腾讯数据中心总经理）。于是很自然地，我就跟蒋杰一起合作了这个项目，并且成了好朋友。后来，蒋杰很出色地帮我完成了这个项目。所以，我觉得做技术的人不应该让当前的技术和思维限制未来的新产品。

就用户体验来说，用户之所以不用我们的产品，其中一个很重要的原因是，没有人愿意耗时太久等一个报告，因为当报告出来之后，他的问题可能早已经解决了。试想，我们拍下脑袋决定只需两分钟，但等一个报告要两个星期，而且报告的结果还不一定能帮上他，你说他会选拍脑袋还是等报告？所以，这也是我们产品人的理念：到底怎样才能解决用户的问题？这也是产品人和数据人不一样的地方。

• 让用户喜欢数据，就会让团队喜欢自己

除了以上这些有关数据的工作外，那时我还做了另外一件事。我启动了一个项目叫“西湖品学”，每年请公司外的数据牛人来分享经验，这是数据人的一场盛会，当时很受欢迎。我做“西湖品学”有两个原因：首先，我认为人们不使用数据不是他真的不想使用；其次，没有人推广

使用数据的好处。

那时，我常在支付宝说一个比喻：“我右手用得好好的，你突然让我用左手，那你就要告诉我用左手的好处，不然我无法改变我的习惯。”

用数据的习惯也是一样，当他知道数据好用的时候，就肯定会用。所以，每一年的“西湖品学”，我都请外面很多人来讲数据到底有什么好处，让公司内部的数据人知道数据应用的商业场景是怎样的，从而激起更多人对数据的渴望和期待。“西湖品学”就是这样一个全新的尝试。

在数据分析人员中，我们大部分分析员都被戏称为“表哥”“表弟”（因为他们每天都要做很多表）。但是，每天仅仅做表的人会有幸福感吗？每天做这么多表，也不知道别人怎么用，没有结果肯定就没有幸福感可言。而“西湖品学”让他们知道，他们的责任在哪里、怎样跟业务对接，他们不仅仅是“表哥”“表弟”。所以我的目的就是，不让他们抱怨自己的才华被埋没了，而让他们知道自己努力的价值所在。

当我的分析师看到管理层很激烈地讨论他们做出的报告时，我能看见他们脸上的笑容：“不枉我们两个星期不眠不休地把这个报告做出来。”其实，大多数数据人都不怕辛苦，怕的就是做出来的东西没人看。

总的来说，我对分析师们要求有些不同。我要求他们对商业有一定理解，这个本身就是我在这本书中提到的“混、通、晒”——**混在商业场景中，把数据与商业场景打通，将数据产品当成载体，把这个东西晒出来，让更多人使用。**当你看到你的用户不能没有你，当你看到你的用户满意的时候，你自然就会感到兴奋，整个闭环就会让人感觉很兴奋。

总结起来，数据产品化的理念是我在支付宝对整个数据分析部门的

顶层设计，我一直用这个理念去运行一个部门，决定资源如何配置。怎样让我的数据使用者喜欢用我的产品，取悦我们用户的办法就是帮他解决问题。

然后就是让数据更泛化。我们必须宣传数据的好处，要教育我们的用户如何使用数据。我们不能假定用户最终会知道数据的好处，而是要一边宣传理念，一边教授技巧，让用户知道怎么用数据、怎么用好数据，知道数据的价值。

最后一点就是：让我们部门的员工喜欢自己，喜欢自己的工作。

我在支付宝的成功，也离不开我的坚持。首先我得益于自己坚持的整个顶层设计的理念，我的老板 Lucy 和 Eric 也十分支持我，给我机会，对我很支持和包容；最后就是我很棒的下属，能够把我的理念付诸实践。

记得当我晋升到副总裁，成了当时阿里巴巴数据领域职位最高的人的时候，Lucy 曾经说：“在中国，品觉是数一数二的分析师。”这句话当然是对我个人的偏爱，但如果她指的是电商平台数一数二的最好的分析师，我觉得能勉强接受。

每每回想起我在支付宝的两年，甚至离开支付宝到了淘宝之后很久，公司内外依然有很多人对我说支付宝的数据分析团队是最棒的。每当遇到这样的赞美，我都觉得非常欣慰。所以，那段经历让我感觉非常幸福。

我在淘宝

有一天，Lucy 把我叫过去说：“现在淘宝很需要你，淘宝的数据体系需要你去整理。”那是 2012 年年底，也是我在支付宝的黄金时代，就

这样我来到了淘宝，开始了我在阿里后4年的时期。

淘宝前两年的经历非常丰富，每每想起来都像过去了10年，我的白发基本上也都是在这段时间长出来的。

• 一样的开始，不一样的数据分析框架

从支付宝来到淘宝，意味着我要重新打一仗。由于我已经在支付宝打了一场胜仗，所以可以照搬很多顶层设计到淘宝使用。但我也不能完全照搬，因为淘宝使用的顶层设计有些跟支付宝一样，有些却完全不一样。

淘宝跟支付宝其相比其实要复杂很多，淘宝本身是个自下而上的组织，是个极其复杂的生态圈。有趣的是，当我来到淘宝时，居然面对的情况恍如回到了入职支付宝时的情形：团队的人差不多都走光了，我感觉就像踏上了一块杳无人烟的新大陆——分析师只剩下几个人，而且我被告知这几个人还会陆续离开。

面对这一切，我做了一个今天想起来仍然感觉很勇敢的决定：我说我不会带支付宝的一兵一卒过来。所以说，我是在支付宝建了一个完整的团队，然后在淘宝又重新组建了一支全新的团队。这段经历其实让我很痛苦，但也是极为珍贵的人生经历。和支付宝时代面对的团队不一样的是，我需要同时管理分析及技术团队，所以在淘宝不用经历支付宝前期那段业务分析和数据技术目标不一致的痛苦。

当时，淘宝的数据技术团队很乱，而且是累积多年的乱，所以要整理淘宝的数据体系，只用顶层设计和数据产品的方法来做，肯定是行不通的。

但还好此时我已经有很多经验，加上当时的总裁三丰跟我的感情非常好，比起我在支付宝时，其实办事更容易。比如在支付宝我还不能直接问 Lucy 到底你要什么，但在淘宝我可以随时到三丰办公室问他要什么，在支付宝我只是一步步尝试满足 Lucy 的需求，而在淘宝我有管理层的绝对支持。所以，我遇到的问题都是来自内部，没有外在问题。

在淘宝的分析团队中，我也同样分别推出了“白米饭”“突击队”和分析框架。不一样的是，我们很难给淘宝定 KPI 去说达到某个程度就是做得好。如果你说这样就好了，你会发现另一个指标可能就不好了。

也正因如此，当我们要建数据产品的时候，我犯了一个错误，我想用一个自上而下的方法去做，最终发现这在淘宝行不通。尽管得益于我在 eBay 的经验，对淘宝这样的双边市场比较熟悉，但后来我依然觉得淘宝的生态圈非常复杂，其分析框架和支付宝非常不一样。

当我发现自上而下的方法不行之后，我就逐步去做更多能发现问题的数据产品，总体来说还是：**坚持数据泛化的方向**。在阿里淘系这样自下而上的组织来说，因为权力已经下放给每个“小二”。泛化一个产品相对来说比较容易，每个“小二”需要数据帮他们做决策的时候，如果让他们每个人都用数据产品，这相比在支付宝时期其实更容易。反而因为数据非常零散，我们要决定到底如何做好一个数据产品，知道如何化繁为简更重要，因为虽然有资源，但是要用资源的人也很多。

关于在淘宝所做的数据产品我这里不再一一赘述，值得一提的是，我们当时做了一个非常重要的决定，就是移动终端。那时候，很多淘宝的消费者都开始使用智能手机，我就决定我们一定要做移动版的阿里数据 App。淘数据这个 PC 产品在我来淘宝之前已经存在，只是一直不温