

大数据大创新

阿里巴巴

云上数据中台之道

邓中华 / 著

- 联动阿里巴巴云上数据中台能力与阿里云云上服务
- 基于十年大数据亲身经历，精彩演绎数据中台之道

大数据大创新

阿里巴巴

云上数据中台之道

邓中华 / 著



电子工业出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

阿里巴巴云上数据中台正服务着阿里生态中的数十个业务板块、百余家公司、千万级客户，在帮助决策层看清甚至决定业态走向的同时，在上万个业务场景中应用并催生创新。

本书基于作者在阿里巴巴的十年大数据亲身经历，精彩演绎云上数据中台之道。全书基于大数据探索的大趋势，讲述阿里巴巴云上数据中台顶层设计，再以实际案例详述阿里巴巴云上数据中台建设及其业务模式的形成过程，总结云上数据中台沉淀的独特价值，并开诚布公地分享阿里巴巴以赋能为本质的大数据战略。

作者希望通过一名老数据人发自内心的分享，帮助尽可能多的有志于大数据构建、应用与创新者，构建自己的云上数据中台，从而拥有既“准”且“快”的“全”“统”“通”的智能大数据体系，以驱动各类业务数据化与数据业务化创新！

本书著作权归阿里巴巴（中国）网络技术有限公司所有，未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

大数据大创新：阿里巴巴云上数据中台之道 / 邓中华著. —北京：电子工业出版社，2018.11
ISBN 978-7-121-34902-7

I. ①大… II. ①邓… III. ①数据处理—研究 IV. ①TP274

中国版本图书馆 CIP 数据核字 (2018) 第 185839 号

策划编辑：张慧敏 王 静

责任编辑：王 静 石 倩

印 刷：北京富诚彩色印刷有限公司

装 订：北京富诚彩色印刷有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×980 1/16 印张：13.5 字数：262 千字

版 次：2018 年 11 月第 1 版

印 次：2018 年 11 月第 1 次印刷

印 数：5000 册 定价：99.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819, faq@phei.com.cn。

推荐序

从大数据的概念被正式提出，到马云老师预言人类正从 IT 时代走向 DT 时代，大数据浪潮迭起。大数据同仁共同认知的一点是，大数据会对社会创新、产业变革、业务创新及每个人的角色定位产生近乎决定性的影响。

马云老师早在 2014 年就提出了业务数据化和数据业务化，阿里巴巴因此付诸行动，通过全局数据监控、数据化运营、数据植入业务场景的各个链路等推进业务数据化；通过统一商家端数据产品、计量信用分，以及风险预测与防控等，推进数据业务化。但与此同时，我们也清楚地知道，在实现这个愿景的道路上存在着很多挑战，其中，如何提升大数据能力则是关键。

在 2014 年以前的阿里巴巴，我们的计算资源分散各处，我们的数据指标经常相互冲突，我们的数据应用大多从数据源头向上耗费大量资源进行垂直开发。这种状况无法可持续地推进业务数据化与数据业务化。于是，阿里巴巴数据人协同起来，从业务视角而非纯技术视角出发，建设既“准”且“快”的“全”“统”“通”的智能大数据体系。其间，在极致追求技术提升的同时，我们自成 OneData、OneEntity、OneService 三大体系，开发了致力于智能数据构建与管理的 Dataphin、高效数据分析与展现的 Quick BI 等产品，培育了一大批独具阿里特色的大数据人才，云上数据中台也就水到渠成了。

今天的阿里巴巴，几乎所有的业务都运行在大数据之上，几乎所有的小二都在用大数据改善工作甚至创新。云上数据中台正服务着阿里生态中的数十个业务板块、百余家公司、千万级客户，在上万个业务场景中应用并催生创新。而每一年的双十一都在上演着数据奇迹。以 2016 年的双十一为例，当天实时计算处理的数据量达到 9400 万条 / 秒，面向业务系统提供应用服务的单日数据调用约百亿次，而全链路完成数据采集、整合构建、服务展现仅需 2.5 秒，这些惊人的数字背后是因为有强大的云上数据中台大数据能力在支撑。

如今，中国正处于数字化转型阶段，政府的各个部门及各行各业越来越相信大数据的

力量。我们认为，这套在阿里生态内检验过的云上数据中台大数据能力及其推进业务数据化、数据业务化的云上数据中台业务模式，可以在阿里生态之外推而广之，赋能全社会！2016年9月，我们不再将云上数据中台深藏于阿里生态内，开始以亲身经历的各种积淀对社会各界有志于大数据战略者伸出“合作之手”，帮助诸如零售、旅游、环保、地产、传媒、运营商、文教、政府部门等领域的客户构建自己的大数据能力，并在数字化转型之路上逐步走向成功！

一群有情有义、有梦想、有担当又有极强战斗力的人共同缔造了今天的阿里巴巴。但如何驱动大数据让世界更加美好呢？我们希望携手同道者，开拓，进而提升大数据能力，共同在大数据实践之路上走向成功！大数据能力本应无边界，越多地参与，才越有可能真正实现无边界。

永不停歇地奋斗，正因乐在于志！我们坚信，一切美好，都会因此而开始得刚刚好！

胡晓明（花名孙权）

阿里巴巴合伙人，阿里云总裁

前言

正说阿里巴巴云上数据中台及云上数据中台业务模式

今天的阿里巴巴，几乎所有业务都运行在大数据之上，几乎所有小二都在用大数据改善工作、进行创新。阿里巴巴云上数据中台正服务着阿里生态中的数十个业务板块、百余家企业、千万级客户，在帮助决策层看清甚至决定业态走向的同时，在上万个业务场景中推进业务数据化，尝试实现数据业务化并催生创新。而这背后则是因为有强大的云上数据中台大数据能力的支撑。

阿里巴巴的大数据观

在阿里巴巴，我们自己的大数据观。我们坚信，“大数据拥有超能力”，大数据所具备的数据计算能力、智能数据能力和数据智能能力，就像人的经络、血液和大脑，缺一不可！围绕着“统一数据建设与数据资产化管理能力”“统一实体连接识别与标签画像高效生产能力”“统一数据服务能力”（在书中分别代称为“OneData”“OneEntity”和“OneService”^[1]）三大体系，阿里巴巴实践着云上数据中台，进行着一系列自我命题与自我解题。

阿里巴巴云上数据中台建设之路

从2012年开始，阿里巴巴云上数据中台建设经历了不断革新理念和实战、不断量变和质变的过程。其中，2014年4月至2015年11月的阿里巴巴数据公共层建设和2016年9月正式开始实施的阿里巴巴大数据能力赋能社会战略，是阿里巴巴在大数据领域的两次关键质变。第一次质变确定了阿里巴巴云上数据中台及数据中台团队，第二次质变确定了阿里巴巴云上数据中台业务模式的社会赋能战略。

2014年，适逢阿里巴巴数据登月元年，首批登月预算数亿元且很快面临耗尽的局面，这引起了时任阿里巴巴CTO姜鹏（花名三丰）的特别关注，OneData体系特别是其方法论也因此进入高层管理者的视线。

[1] OneData、OneEntity和OneService代表着阿里巴巴大数据能力的三大体系，包括方法论、原则、规范、数据产品、数据技术、数据与业务融合的能力等。所以书中提到OneData、OneEntity和OneService体系时，指的是包括方法论、数据产品、数据技术等综合能力；提到OneData、OneEntity和OneService体系方法论、规范、数据产品等时，则是指三大体系中具体对应的方法论、规范、数据产品等。

经过一段时间的方案细化和多轮评审及沟通，2014年4月8日，阿里巴巴数据公共层建设项目正式启动。在保障平稳支持日常业务的前提下，一期启动全局架构，二期启动18个子项目，三期启动9个子项目，并启动6大数据技术领域。一年后，阿里巴巴数据公共层建设项目即取得了阶段性成果，除深度参与的淘系、B2B等BU外，涉及或影响小微金服^[1]、阿里云等10多个BU^[2]。其数据服务20多个BU，主打小二端统一的数据产品平台——阿里数据，统一商家端数据产品平台为生意参谋，并推出数据大屏助力双十一；同时深入业务，协助业务创新及探索数据自主创新。与此同时，以阿里巴巴数据公共层建设为切入点繁荣发展起来的数据构建、管理和服务自成体系，其特别之处在于OneData体系的升级、OneEntity体系方法论的提出、OneService体系数据产品的升级。这些不仅在技术上降低成本、提高效率，在业务上验证着数据价值化的可能性，也同时促进了组织优化。此时，阿里巴巴数据中台已经在事实上存在了。

[1] 今天的蚂蚁金服在2014年时称为小微金服。

[2] BU, Business Unit, 中文意思为业务单元。在阿里巴巴，通常会把相关或相近的业务或技术团队组合为一个业务单元，以便组织管理、业务推进与协同等。

2015年12月7日，张勇（花名逍遥子）在给阿里巴巴全体小二的邮件《阿里巴巴集团全面组织升级，启动中台战略》中正式宣布：“构建符合DT时代的更创新、更灵活的‘大中台小前台’组织机制和业务机制。”自此，数据中台成为阿里巴巴中台事业群的重要组成部分，阿里巴巴数据中台团队正式成立！

在此过程中，阿里巴巴独特的大数据观越来越清晰明确，也累积了独特价值，包括云上数据中台大数据技术、云上数据中台建设方法论、云上数据中台产品化服务、云上数据中台业务体感、OneTeam协同作战思维、特色大数据人六大方面。

阿里巴巴云上数据中台业务模式赋能全社会

在阿里巴巴，有一句众所周知的话：“今天最好的表现是明天最低的要求！”数据中台团队的正式成立预示着其将要面临更高的要求 and 更多的挑战。与此同时，阿里生态内越来越多的全资或投资子公司提出了数据诉求，其中既有同时输入数据和消费数据的，也有只消费数据的，还有只输入数据的。于是，从2015年12月开始，整个阿里巴巴数据中台团队开始面向阿里生态内建设智能大数据体系！智能大数据体系的建设极大地丰富和完善了阿里巴巴大数据中心，OneData、OneEntity、OneService渐趋成熟并成为上至CEO、下至一线员工共识的三大体系，云上数据中台深入业务、赋能业务，也让数据价值及赋能业务的潜在价值越来越被认可和期待。

同时，我也深深地反省将其放大到阿里生态内时的不足之处，以及在阿里生态之外，这套大数据能力是否可以推而广之，赋能全社会呢？

于是，2016年9月，为了使命，为了达成愿景，我们不再将云上数据中台深藏于阿里生态内，开始将阿里巴巴大数据能力同时赋能阿里生态内外！

历经了阿里生态内各种业态挑剔的实战历练，云上数据中台除自身具备的内核能力外，还向上与“赋能业务前台”连接、向下与“统一计算后台”连接，并与之融为一体，形成云上数据中台业务模式，具备了对全社会赋能的可能。关于云上数据中台与云上数据中台业务模式的关系，如图0-1所示。

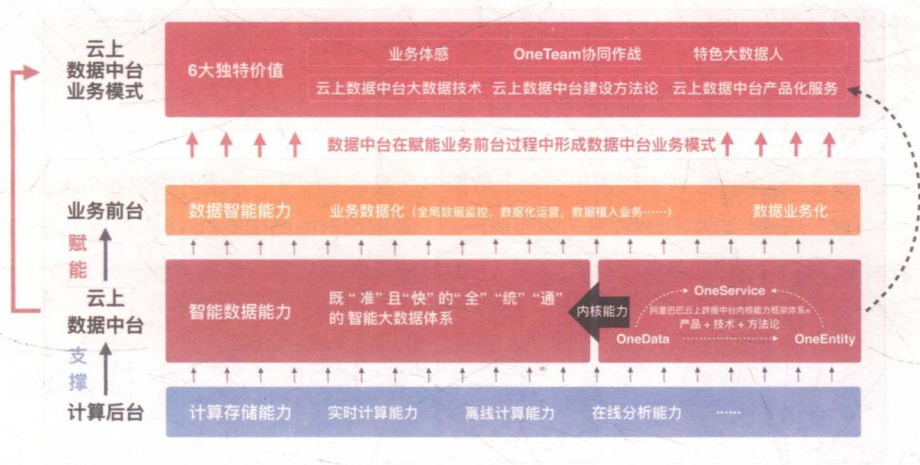


图 0-1 阿里巴巴云上数据中台与云上数据中台业务模式的关系

云上数据中台定位于计算后台和业务前台之间，其内核能力是以业务视角而非纯技术视角，智能化构建数据、管理数据资产并提供数据调用、数据监控、数据分析与数据展现等多种服务；承技术启业务，是建设智能数据和催生数据智能的引擎。而指导云上数据中台内核能力不断积累和沉淀的正是 OneData、OneEntity、OneService 三大体系的方法论。

以云上数据中台自身内核能力为中坚力量的云上数据中台业务模式，同时关注着与大数据能力相关的技术、数据、业务等，其基于技术而又深入业务，是包括数据产品+数据技术+方法论+场景价值实现等在内的综合性能力输出。它为智能化数据、极致提升技术和数据智能化业务负责，即一方面专注于从业务视角建设既“准”且“快”的“全”“统”“通”

的智能大数据体系，并且极致化追求技术上的成本降低及效率提高；另一方面致力于智能大数据与业务场景深度融合中的各类应用与智能化价值创新。

我们的追求与努力

为何而来？将往何处？这是我们一直思考的问题！

云上数据中台业务模式是活性的，是有血有肉、有情感的，其信念是“大数据拥有超能力”，其使命是“让大数据催生大创新”，其愿景是实现“大数据创新无处不在，大数据人才无处不在”。所以，未来，我们将全力以赴地分享云上数据中台业务模式，不遗余力地将云上数据中台业务模式中的方法论、数据产品、数据技术、数据与业务融合的价值化经验等付诸帮助国内、国际的云上客户的实战中！希望可以帮助越来越多的志同道合者一起“华山论剑”！

从2016年9月开始，我们结合一些客户的实际诉求，分享了不少大数据领域相关经验，而本书则是应多方期望，对于阿里巴巴云上数据中台及云上数据中台业务模式的系统、全面、深入的分享。后续，我们计划推出《阿里巴巴云上数据中台的赋能实战》《大数据产品经理》《大数据大设计》《深入大数据产品与技术》《大数据大管理》等一系列大数据相关图书。对于分享，我们是认真的！

邓中华（花名宗华）

阿里巴巴大数据人，资深产品专家

特别说明，特别感谢

本书所有内容均基于笔者十年来在阿里巴巴大数据领域实战中的亲身经历，所有图片均来自于笔者的 100 余份 PPT 和数本手稿，以及笔者所在阿里巴巴数据中台团队的战友积累的 20 余份 PPT。所以，这本书代表着真正意义上的云上数据中台实战。十年寒来暑往，十年风刀霜剑，相信读者可以从中或多或少感受到阿里巴巴大数据人用初心和信仰、求变与务实、不懈与努力铸就的云上数据中台及其业务模式！

特别感谢相信我、支持我、陪伴我一起战斗的阿里巴巴数据公共层建设若干期项目中的数百名伙伴、生意参谋平台中的 80 名队友、新能源实验室与新行业赋能团队的 120 名热血同仁！书中图片中引用的部分图标来自 iconfont 平台，在此一并感谢！因人数众多，在此不一一列出姓名。

谨以此书，献给这么多年来不离不弃陪我奔跑在大数据之路上的梦想战友！献给一直包容我、疼惜我并且支持我投身大数据事业的家人！献给在本书写作过程中无欲无求给予我帮助的好友！

生命因为有你们而美好，征程因为有你们而无悔，因为相信，所以坚持，终于看见！

目录

第1章 笔者自述 001

一千个人眼中有一千个哈姆雷特。对于阿里巴巴,大家给出的评价有“年轻又独特”“神秘又很土”“天马行空又脚踏实地”“强大又懂得分享”“源于电商又超越电商”“中国的、国际的”“了不起且有社会责任感”“小人物成就大梦想”“大数据驱动的互联网创新公司”“务实且重价值观”……而在我眼中,感触最深的却是“情义”“梦想”“担当”。

1.1 我的阿里九年:相信一切都是最好的安排 002

1.2 我眼中的阿里:有情有义,有梦想,有担当 005

第2章 阿里巴巴云上数据中台之顶层设计 007

如今的阿里巴巴,几乎所有的业务都运行在大数据之上,几乎所有的小二都在用大数据改善工作甚至促进创新。云上数据中台正服务着阿里生态中的数十个业务板块、百余家公司、千万级客户,在帮助决策层看清甚至决定业态走向的同时,正在上万个业务场景中被应用并催生创新。

2.1 大数据的发展历程与价值探索 008

2.1.1 国内外大数据发展研究 008

2.1.2 同仁如何用大数据 013

2.2 阿里巴巴的大数据主张 017

2.2.1 阿里巴巴的大数据主张概述 017

2.2.2 大数据催生大创新 018

2.2.3 大数据拥有超能力 019

2.2.4 云上数据中台是“承技术启业务”的关键存在 021

2.3 水到渠成的云上数据中台	022
2.3.1 云上数据中台赋能业务全景图	023
2.3.2 云上数据中台的顶层设计	025
2.3.3 云上数据中台业务模式的愿景与使命	031

第3章 阿里巴巴云上数据中台之建设过程 035

阿里巴巴云上数据中台的发展主要经历了四个阶段：阶段一，2012.2—2014.3，初探期；阶段二，2014.4—2015.11，质变期；阶段三，2015.12—2016.6，升华期；阶段四，2016.7至今，新一轮质变期。

3.1 云上数据中台执行计划	036
3.1.1 现状梳理和未来展望	036
3.1.2 项目计划与组织协同	041
3.1.3 实际落地与项目总结	048
3.2 从零散的数据到统一的数据	062
3.2.1 烟囱式开发造成业务困扰和技术浪费	062
3.2.2 数据公共层力求让业务和技术都满意	067
3.3 从成本中心到资产中心	083
3.3.1 大数据容易陷入成本中心的深渊	083
3.3.2 大数据应该且必须回归资产本质	086
3.4 从孤岛数据到融通的数据	097
3.4.1 孤岛数据存在的必然性	098
3.4.2 数据只有融通才能真正产生价值	099
3.5 从“授人以鱼”到“授人以渔”	106
3.5.1 给数式服务——“授人以鱼”	106
3.5.2 业务主题式服务——“授人以渔”	107

3.6 从“数据可有可无”到“无数据不智能” 109

3.6.1 野蛮增长时期：数据是可有可无的 109

3.6.2 智能化增长时期：无数据则不智能 111

3.6.3 大数据引爆“双十一全球盛典” 120

第4章 阿里巴巴云上数据中台业务模式之独特价值 141

阿里巴巴经过较长时间的大数据探索、量变积累，进而达成质变，形成云上数据中台，并在云上数据中台服务阿里生态业务过程中形成云上数据中台业务模式。

为何阿里巴巴云上数据中台能够支撑整个阿里生态的业务发展？为何阿里巴巴云上数据中台能够从思想意识到决策行为上引起从“数据可有可无”到“无数据不智能”的改变？这些都与云上数据中台业务模式的独特价值密不可分！

4.1 云上数据中台业务模式的独特价值 142

4.2 云上数据中台大数据技术 144

4.3 云上数据中台建设方法论 146

4.4 云上数据中台产品化服务 149

4.4.1 云上数据中台产品化服务概览 150

4.4.2 云上数据中台核心产品 Dataphin 151

4.4.3 云上数据中台核心产品 Quick BI 154

4.5 业务体感 156

4.6 OneTeam 协同作战思维 157

4.7 特色大数据人 158

第5章 走向大数据成功之路 161

阿里巴巴在智能大数据体系建设与数据智能化各类应用与创新中开拓了一条成功之路。我们不愿独享这些经历，正在向社会各行业中有志于大数据战略者伸出“合作之手”。我们希望携手越来越多的各行业中有志于大数据战略者，开拓、提升大数据能力，共同在大数据探索之路上走向成功！大数据所具有的能力本应无边界，越多地参与，才越有可能真正实现无边界。

5.1 如果可以选择，你会走怎样的大数据之路	162
5.2 阿里巴巴以赋能为本质的大数据战略	163
5.2.1 基于云上数据中台业务模式的解决方案	163
5.2.2 阿里巴巴真正在意的和能够给予的	166
5.2.3 循序渐进的云上数据中台业务模式	166
5.3 你的成功依赖于你的远见和执行力	168

后记 一位老数据人的心路与情怀	169
-----------------	-----

附录 A 云上数据中台演进格局选图	177
-------------------	-----

A.1 笔者见证的阿里巴巴大数据团队演进	178
A.2 一张图说尽云上数据中台顶层设计演进	179
A.3 阿里巴巴云上数据中台的成长历程	180
A.4 阿里巴巴云上数据中台赋能业务运行图	184
A.5 阿里巴巴云上数据中台赋能业务全景图	185

附录 B 云上数据中台核心产品选图	187
-------------------	-----

B.1 阿里巴巴云上数据中台核心产品 Dataphin 图示	188
B.2 阿里巴巴云上数据中台核心产品 Quick BI 图示	192

附录 C 插图索引	197
-----------	-----

第 1 章

笔者自述

我自南京大学毕业，在懵懂无知中加入阿里巴巴，在不知“大数据之所谓”中，误闯入大数据这片汪洋大海，却在历经无数风浪后深深爱上了这片海洋，并且遨游不休。

从 RA 与 ETL^[1] 到 PD 与 PM，再到 TL，从初期连接业务与技术的需求分析工作开始，到从个性化推荐项目开始产品之路，规划并落地生意参谋，再到设计出 OneData、OneEntity、OneService 体系并推进阿里巴巴数据公共层建设，进而探索 DT 上云和云上 DT^[2] 赋能阿里生态内外的云上数据中台业务模式；其间，从 B2B 走进淘系，从阿里生态内联动阿里生态外，我一直致力于大数据技术及产品领域的平台化及商业化。

我不否认其间有过悲观、不满，但每每回想初心，总会坚信一切都是最好的安排；我不否认其间有过想要放弃，但“战友”们的有情有义、有梦想、有担当，给了我坚持下去的力量。

[1] ETL 即 Extract-Transform-Load，用来描述将数据从来源端经过抽取（extract）、转换（transform）、加载（load）至目的端的过程。ETL 一词比较常用在数据仓库，但又不限于数据仓库。ETL 是构建数据仓库的重要一环，用户从数据源抽取所需数据，将数据进行清洗转换后，按照预先定义好的数据仓库模型加载到数据仓库中。

[2] 2016 年 12 月，经过几个月的讨论，我所在的数据中台团队调整了业务方向，一方面继续服务阿里生态内，另一方面将服务阿里生态内的大数据能力上云，进而在云上用大数据能力赋能阿里生态外社会各界，简称 DT 上云和云上 DT；相应地，我们的组织架构也进行了调整，直接成立新能源实验室，负责 DT 上云和云上 DT 的赋能之路的探索。

1.1 我的阿里九年：相信一切都是最好的安排



2009年4月，还有两个月才正式毕业的我，怀着梦想与忐忑的心情，加入了阿里巴巴，从此便与大数据结下不解之缘。一路走来，不知不觉已近十年！而事实上，在阿里巴巴这样一家互联网科技创新型公司中工作一年，差不多相当于在一般公司中工作五年，相信很多了解这个圈子的人都不会觉得这样比喻夸张。

我入职于当时阿里巴巴 B2B 技术部下的数据仓库部门，没过多久就成为阿里巴巴“百阿 99 期”的学员。也是在那时，我第一次对大数据产生了懵懂的认识，这源于一次简单的自我介绍。当我介绍自己来自数据仓库部门时，同学们似乎都不知道这个部门是做什么的。突然有一位通过社会招聘进来的同学说：“我知道，数据仓库就是给老板和业务人员做报表的！”也许他是无意的，但当时的我感觉到的是满满的不屑，却无力反驳，瞬间涨红了脸，心想：不该是这样的，不该只是这样的！

今天回头再看，我们可以说，从大数据的概念被正式提出，到马云老师预言人类正从 IT 时代走向 DT 时代，再到后续的各种日新月异的变化，大数据时代终于扑面而来！

身处其中而不忘初心、矢志不渝的人，必然经历过刻骨铭心的变化，在体力、脑力和心力的经受过各种锻炼。于我而言，在阿里巴巴这么多年一直坚持在大数据领域中探索，每当想要放弃或有一丝松懈时，脑海中总会浮现当时的情景，耳边总会响起当时的心声，或者这正是我在大数据的道路上一路狂奔的原动力吧！

从 2009 年到 2011 年 6 月，当时的数据仓库部门里有一个岗位叫 RA（即需求分析师）。这个岗位是业务人员与数据技术人员之间的桥梁，负责将业务线上的产品经理或运营人员提出的报表等各类需求，转化为数据研发人员能理解的需求文档。然后，数据研发人员基于此需求文档完成 ETL 开发并生成数据，需求分析师将数据通过报表或者其他方式反馈给产品经理或运营人员。这个岗位从侧面反映了当时数据不是业务部门所熟悉的和必需的现状，以及数据仓库团队努力往前服务业务人员的态度。

当时，我通过担任 RA，既了解了业务，又学习了数据技术。2009 年 11 月，我以当时还无足轻重的个性化推荐项目“一见钟情”为切入点，开始走上数据产品经理之路。2011 年 6 月公司取消了 RA 岗位，我顺利地转型为数据产品经理，并将打造个性化推荐产品作为阶段性目标。到 2012 年年初，我们协同一众团队成员，将个性化推荐产品打造

成为当时在 B2B 领域备受瞩目的大数据应用场景，并沉淀出个性化推荐引擎 iRecom。

在这个过程中，我如饥似渴地自修了数据产品能力、平台化思想和项目管理能力，同时也主动地向身边的小伙伴虚心求教各种业务知识和算法，以及 ETL、测试、Java、前端等相关技术。那段日子里的我，真像是一个饥饿的人扑在面包上。

2012 年 2 月，我接到一个新的任务——梳理支持 B2B 业务的 370 多个 API 并设法将其统一。这是我第一次从一个相对成功的产品中走出来，试着从零开始自我命题与自我解题，没想到从此走上了建设阿里巴巴云上数据中台之路，这是后话了。当时的我来不及多想就开始剖析每一个 API 及研究 API 调用的每一条数据，我发现，如果 API 背后的数据没有实现标准化和规范化，那么盲目地统一 API 只能治标，而不能治本，业务人员用数据难、技术响应业务慢等问题无法被根本解决。于是，我与我的主管商量，先从数据的标准化和规范化着手，进而解决 API 统一的问题。同时，我也收获了未来很多年一路相伴的第一个“梦想战友”。于是，我们便激情满满地投入战斗。

2012 年 6 月，OneData 体系方法论首次被提出。不久之后，OneService 体系的前身——OpenAPI 被提出。我们在 B2B 领域的数字应用中尝试建设 1688 数据公共层，并向上连接 OpenAPI 提供服务，这样既实现了数据的标准化和规范化，又实现了统一服务，同时深入业务服务而不仅仅是支持业务。我们首次推出傻瓜数据平台用于服务小二，首次推出生意参谋用于服务 1688 商家。

2013 年 4 月，B2B 领域的 3 个数据团队与淘系数据团队融合。也许是前一场小胜的激励，也许是初生牛犊不怕虎，团队融合后，我们快速推出了生意参谋的淘宝版和天猫版，并以此反推，以 OneData 体系方法论建设淘宝和天猫商家数据公共层。虽然在意料之中，但还是让我们感到惊喜的是——OneData 体系方法论竟具有快速扩展的能力。在这个过程中，我收获了未来很多年一路相伴的第二个和第三个“梦想战友”。

2014 年是阿里巴巴数据登月元年，首批登月预算达数亿元但很快被消耗完了，这引起了时任阿里巴巴 CTO 姜鹏（花名三丰）的高度关注，OneData 体系特别是其方法论也因此进入阿里巴巴高层管理者的视线。

经过一段时间的方案细化和多轮评审沟通，2014 年 4 月 8 日，阿里巴巴数据公共层建设项目正式启动。在保障平稳支持日常业务的前提下，一期启动全局架构，二期启动 18 个子项目，三期启动 9 个子项目，并启动 6 大数据技术领域^[1]。一年后，阿里巴巴数据公