

价值密度下沉的挖掘者

- 大数据思考之一 -

核心观点

程杲 执业证书编号: S1100513040003
互联网 (021) 68416988-222
研究员 chenggao@cczq.com

- ❖ **数据价值下沉：驱动高价值数据向低价值数据迁移。**过去大数据的主战场是有能力产生高价值数据的行业，如银行、运营商、石油开采等，IOE 是这一时期主要技术解决方案。2011 年以来大数据应用主战场转移到数据价值密度中等的行业，例如电商（淘宝京东）、社交（腾讯）、搜索（百度），影视（优酷土豆）等。我们预计，价值密度的下降将持续，因而未来的低价值数据将成为主战场，例如视频资料、人的行为模式、语音轨迹、天气数据等。虽然价值密度较低，但低价值数据的体量要比高价值数据高几个数量级，对其挖掘的市场价值将远高于高价值数据市场。
- ❖ **计算成本下降正在快速驱动大数据应用场景的拓宽。**纵观过去 40 年 IT 的发展史，每一次大的变革浪潮都是由成本的下降驱动；此次大数据浪潮的主要驱动因素在于分布式计算逐渐取代集中式计算架构，这导致大规模数据处理对服务器的要求大幅降低，数据存储计算成本大幅降低，数据对象的价值密度不断下沉。
- ❖ **大数据应用行业的横向和纵向拓展。**大数据的市场空间来自于先行行业的纵向深化应用，以及新行业的横向拓展。预计随着物联网的普及，各种传感器将为视野之外的各大传统行业带来大量数据和应用场景。其中，比较重点的八个行业包括零售、物流仓储、汽车制造、教育、电子科技行业、制造行业、医疗卫生、交通运输。
- ❖ **政策支持形成持续催化剂。**在经济增速下滑的背景下大数据板块是为数不多的亮点行业；9 月初的《促进大数据发展行动纲要》是产业发展的一个标志性事件，然而市场对此反应并不充分。接下来市场将能看到《纲要》在地方层面的落实，成为下一阶段的催化剂。
- ❖ **六个选股评价维度：**1) 上市公司 B2B 业务较为聚焦，且 B 端有能力提供丰富 C 端数据的；2) 大数据分析能力沉淀程度强；3) 数据可接触性强；4) 数据量级及广度宽阔；5) 数据单位价值相对较高；6) 数据可脱敏等。我们看好在数据分析能力方面较为突出的东方国信、拓尔思，在特定行业数据方面积淀深厚的华宇软件、千方科技，以及业务已经全部云化的全通教育。

川财证券研究所

成都：
高新区交子大道 177 号中海国际中心 B 座 17 楼，610041
总机：8628-86583000
传真：8628-86583002

大数据下沉：驱动高价值数据向低价值数据迁移

大数据并不是一个新的行业，这个行业已经存在多年。如果将大数据定义为“一块硬盘装不下”的级别，那么大数据的必要条件就是数据量达到 TB 之上。这个量级在诸如银行等行业应用中早已触及这一标准，只是之前是以高性能计算和存储的形式出现的。

因此我们可以看到大数据并非凭空冒出来的行业，其实际应用已经有了数十年的时间；过去大数据的主战场是业务中会产生海量高价值数据的行业，如银行、证券交易所、运营商、政府、石油开采等，而这些行业的生产经营也高度依赖于数据的处理。IOE 是这一时期主要技术解决方案。

而 2011 年以来大数据概念的提出、发展，实质上是大数据应用的横向和纵向拓展。在这个过程中，呈现出数据的价值低密度化。近几年的大数据浪潮主要以 x86 集群云计算为代表，其应用主战场转移到数据价值密度中等的行业，例如电商（淘宝京东）、社交（腾讯）、搜索（百度），影视（优酷土豆）等。比较消费者在银行产生的真实交易数据价值，其在互联网上的商品选购、聊天记录、搜索偏好显然价值密度要低很多。

我们预计，价值密度的下降将持续成为未来的低价值数据将成为主战场。这些低价值数据可能在目前来看存储成本小于存储收益的数据，例如视频资料、人的行为模式、语音轨迹、天气数据等。

图 1：数据的价值密度下降，数据量提升



资料来源：川财证券研究所

虽然价值密度较低，但低价值数据的体量要比高价值数据大几个数量级，随着存储和计算成本的下降，对其挖掘利用的市场正在形成。

第四次成本下降成就大数据应用拓展

成本下降正在快速驱动大数据应用场景的拓宽。纵观过去 40 年 IT 的发展史，每一次大

的变革浪潮都是由成本的下降驱动的，我们正在面临的大数据浪潮也不例外。

1、个人计算机和互联网浪潮

“全世界只需要 5 台电脑就足够了！”IBM 的创立者托马斯·沃森曾做出这样的预测，但 PC 出货量与单个 PC 价格之间的动态均衡很快让计算能力增长和价格快速下跌成为现实，于是 PC 在很快的速度内成为每个人都可以配备的电子产品。

2、云计算浪潮

带宽成本下降驱动了云计算的应运而生。与运营数据存储在本机上不同，云计算将存储和大部分计算都转移到云端，让移动处理的多端同步查看变得更容易。这一切应该归功于光纤的大规模应用，让网络传输成本快速下降，刺激了网络带宽需求爆发。

3、移动互联浪潮

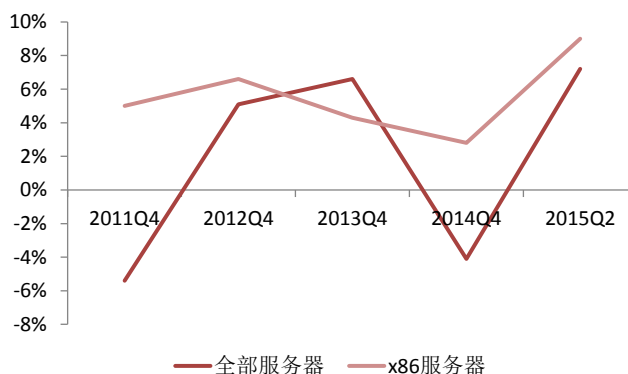
智能手机的普及，供应链的大规模铺开让手机成本快速下降，人手一部手机带动了整个手机生态圈在 2010 年之后的大繁荣，包括移动 App、手游，以及各种 o2o、LBS 等多种移动互联网业态。

4、大数据浪潮

与前面几次浪潮发端多于 C 端不同，大数据浪潮的主要驱动因素在于 B 端。互联网公司受益于其自身业务对数据一致性的弱要求（CAP 定律之中舍 C 而取 AP），分布式计算逐渐取代集中式计算架构，这导致大规模数据处理对服务器的要求大幅降低，廉价的 x86 得以取代大型机及小型机成为服务器市场的主要组成部分，让大数据存储成本大幅降低，低价值数据存储第一次变得可行。

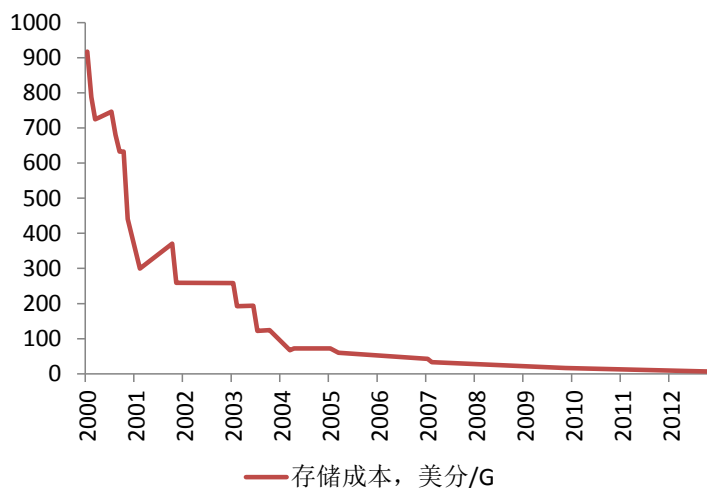
此外，革命性的技术在近 10 年相继出现，加速了大数据的应用基础，其中包括虚拟化技术，Map-Reduce & Bigtable，NoSQL 数据库，Deep Learning 技术等，这些技术帮助我们解决了海量数据的存储和计算问题。

图 2：x86 服务器市场增速远超其他类型服务器（营收增速%）



资料来源：Gartner，川财证券研究所

图 3： 数据存储成本下降容许数据存储价值降低



资料来源：Gartner，川财证券研究所

我们对大数据的应用场景的未来发展缺乏想象能力，这是因为对于以眼前的知识去推测未知的事物非常困难。但我们可以从大数据发展的趋势做一窥测。目前银行早已是大数据应用的受益者，可远在没有装备电子计算机的年代，客户查询一次对账单都是难以办到的业务。彼时业务的处理仅能以手工簿记完成，以凭证为中心的业务流需要客户拿着各种凭证来驱动，于是交易明细在过高的成本下没有开发的價值。电子化让银行第一次可以以客户为中心搭建交易系统，客户也第一次得以享受到以个人账户为核心的查询服务。

因此，随着技术让成本进一步降低，我们一定可以看到更多此前难以想象的客户需求被开发出来并被满足。

数据应用行业的横向和纵向拓展

我们认为大数据的市场空间来自于原有已被覆盖的行业进一步深化应用，以及新行业的拓展。

我们整理了大数据的部分潜在应用行业。一些行业数十年前已经开始应用广义上的大数据，例如银行、运营商等，政府也已将技术用于交通、司法的政务处理。

以淘宝、百度等为代表的互联网公司是目前狭义上的大数据应用代表。这些行业的大数

据应用大部分还处于初级阶段，纵向拓展深度仍比较大：更多的数据被加入分析，更多的分析工具被开发出来。

图 4：大数据的已有应用场景



资料来源：川财证券研究所

我们预计随着物联网的普及，各种传感器将为尚未应用大数据工具的其他各大传统行业带来大量数据，传统行业也将产生大量的应用场景。其中，比较重点的八个行业包括零售、物流仓储、汽车制造、教育、电子科技行业、制造行业、医疗卫生、交通运输。这些行业进入应用范围将带来横向拓展。

图 5：大数据的拓展应用场景



资料来源：川财证券研究所

政策支持形成持续催化剂

政策的支持让大数据板块存在持续性高弹性投资机会。在经济增速下滑的背景下大数据板块是为数不多的亮点行业，是国家产业结构调整升级的重要受益部分。

刺激将来源于需求和供给两个方面，从需求来看，国家将从中央到地方加大对大数据的

预算投入，来自政府的需求将迎来持续的高速增长；从供给来看，政府是大数据的主要沉淀地，政府数据的打通、公开，将在供给层面激发行业资源。

9月初的《促进大数据发展行动纲要》是产业发展的一个标志性事件，然而市场对此反应并不充分。接下来市场将能看到《纲要》在地方层面的落实。随着地方文件的出台，相关上市公司也将与地方政府陆续形成合作，成为下一阶段的催化剂。

并购驱动价值外延增长

受制于数据可得性、技术人才短缺等限制，国内大数据产品多数处于萌芽阶段，多数公司还处于 A、B 轮的成长期。部分上市公司认识到了这一点，将并购作为公司大数据战略发展的方向。

在 A 股上市公司估值与一级市场存在较大差距的情况下，上市公司有质量的外延并购将显著提升公司内含价值，从而相较竞争对手获得更多的赢面，因此我们认为明确建立外延收购战略的上市公司应该享有溢价。下图中，我们整理了我们看好的 5 家大数据相关公司近期的并购情况，可以看到，相关上市公司对并购活动持非常积极的态度。

图 6：部分涉及上市公司大数据相关并购活动

上市公司	并购方式	金额	目标公司	简述
东方国信	定增	17.96 亿	na	投资于大数据分析服务平台、互联网银行平台、工业大数据智能互联平台、城市智能运营中心、分布式大数据处理平台等五个项目。
拓而思	现金收购	共计 2 亿	耐特康赛、花儿绽放等四家	通过并购进而拥有全国领先的搜索引擎优化（SEO）技术服务提供商、经验丰富的互联网营销团队和大量互联网营销客户资源；有能力为双方目标客户群提供网站体验优化、网络流量提升、互联网营销等高增值服务。
华宇软件	现金收购	共计 6 亿	浦东中软、万户网络等三家公司	浦东中软主要从事法院、非银行金融机构、食品药品安全等业务领域的软件开发、运维服务与系统建设，其法院业务在华东地区的优势明显。万户网络主要从事协同办公软件的研发、销售及相关技术服务。
千方科技	现金收购	共计 3 亿	一二三零八、远航通等 6 家公司	12308 为出票量最大的全国性客车售票平台 www.12308.com 及 APP。远航通为民航用户进行航后分析、生产主数据管理提供数据整合和应用服务平台，市场占有率达到 80%。
全通教育	股权+现金	共计 14 亿	广西慧谷、湖北音信等 8 家公司	广西慧谷、湖北音信等公司为电信增值服务型公司，可以弥补公司在广东等优势地区之外渠道覆盖不足的短板。继教网、智园控股、西安习悦为产品性公司，对公司全课云产品线形成补充。

资料来源：川财证券研究所



六个选股评价维度

我们从下面几个维度出发制定大数据相关上市公司的选股思路。

1、B2B 业务是转型 B2B2C 的护城河

历史原因，A 股上市现有计算机公司基本上都是 B2B 的盈利模式，软件产品所处理的数据一般是 B 端即企业端的数据。B2B 模式软件企业的优点是客户付费意愿强，盈利可靠性较高；但多为项目制收费，很难摆脱传统软件行业 ROE 受制于人头费用的劣势。

大数据的盈利模式必然需要企业做出转型，从 B 端转向 C 端，这主要有两个原因：第一，转型 2C 后将能实现多种商业模式创新，摆脱项目制的依赖；第二，转型后产品使用人数将成倍增长，这使得 2C 的市场空间更大。

但在这之前，大数据公司要成为大数据公司，必须借助 B2B 业务接触到 C 端数据；这意味着转型过程必须依托原有 2B 业务，这样的公司有着天然的 B 端行业进入壁垒，才能在竞争中拥有独特优势。否则市场就会问一个自然而然的问题：“Why you?”（为什么是你？）

因此，我们更看好上市公司中在当前 B 端业务较为聚焦，且 B 端有能力提供丰富 C 端数据的公司。从这个角度我们看好在特定行业数据方面积淀深厚的华宇软件和千方科技。

2、大数据分析能力需要长期沉淀

大数据分析主要包括 3 个方面的内容，分别是大数据存储、大数据计算、价值输出，其中，前两个环节侧重于计算机技术和数据可得性，对公司的技术能力要求较高。价值输出环节则侧重于行业理解和数据变现能力，这对行业客户认知、变现模式成熟度、公司团队执行能力有着较高要求。这些都需要多年的行业积累和耕耘。

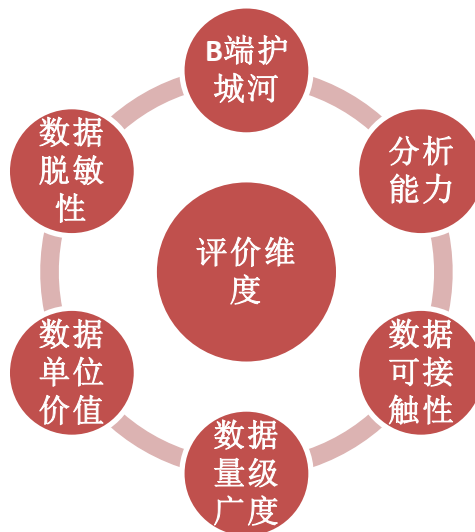
从这个角度我们看好在大数据分析能力方面较为突出的东方国信、拓尔思。前者与联通合作，从事运营商数据分析多年；后者是国内文本检索和语义分析的领军公司。

3、数据可接触性

传统的软件主要以单机形式运作，这种离线版对软件公司来说数据可接触性差，在线版软件比较好，云计算方式最好。这个角度我们看好业务已经全部迁移至全课云的全通教育。

其他的标准还包括涉及数据量级及广度、数据单位价值、数据脱敏性等。

图 7：大数据上市公司的选股维度



资料来源：川财证券研究所

此外，在资本市场层面，对上市公司专注程度、安全边际、市值、并购动力、政策推动等方面也应给予一定考虑。



投资评级说明

证券投资评级:

以研究员预测的报告发布之日起 6 个月内证券的绝对收益为分类标准。

买入: 20%以上;

增持: 5%-20%;

中性: -5%-5%;

减持: -5%以下。

行业投资评级:

以研究员预测的报告发布之日起 6 个月内行业相对市场基准指数的收益为分类标准。

超配: 高于 5%;

标配: 介于-5%到 5%;

低配: 低于-5%。

免责声明

本报告由川财证券有限责任公司（已具备中国证监会批复的证券投资咨询业务资格）制作。本报告仅供川财证券有限责任公司（以下简称“本公司”）客户使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司对该等信息的准确性及完整性不作任何保证。本报告所载的意见、评估及预测仅为本报告出具日的观点和判断，该等意见、评估及预测无需通知即可随时更改。在不同时期，本公司可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。本公司不保证本报告所含信息保持在最新状态。本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供投资者参考之用，不构成所述证券的买卖出价或征价。该等观点、建议并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素，必要时就法律、商业、财务、税收等方面咨询专业财务顾问的意见。对依据或者使用本报告所造成的一切后果，本公司及作者均不承担任何法律责任。本公司及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，也可能为之提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。本公司的投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“川财证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权力。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

本公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：J19651000。

©版权所有 2015 年川财证券有限责任公司