



大数据的新方向：公开同享趋势下的开放数据产业



SupStat™

北京数博思达信息科技有限公司
数据科学专家组



概览



- 开放数据进行时
- 开放数据500例网站
- 公司案例： 美国
英国
上海大数据局
台湾开放数据联盟。



开放数据进行时



“如今开放数据领域的革新正带领我们快速进入一个全新的时代。”

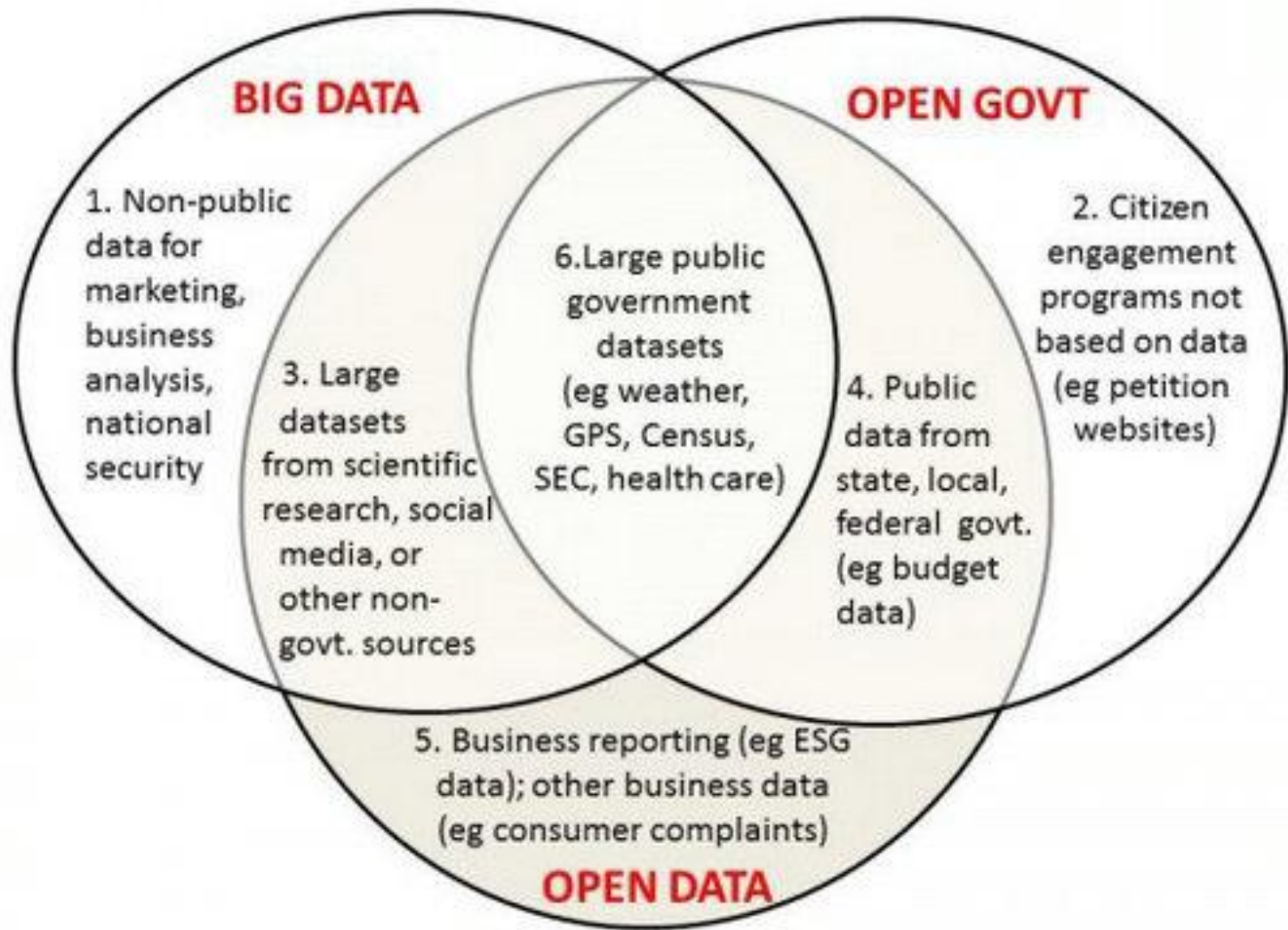
“开放数据正在逐渐成为全世界优秀商家的秘密武器。”



开放数据进行时



- “开放数据的最佳定义，即能够被个人、公司或其他组织自由获取，并用来发展新的业务，进行趋势分析，解决复杂问题的公共数据。”
- “虽然大数据与开放数据相似，但对大数据的介绍并不能完全涵盖开放数据的情形。”





	共享数据	公开数据	开放数据
是否确保「所有人」可获取	不是	是	是
是否「免费」使用	不一定	不一定	是
是否「无使用限制」	不是	不是	是
是否「机器可读」	不一定	不一定	是



开放数据进行时



- “本书借助多个公司的案例，介绍了开放数据在商业领域的应用。”
- 本书反映了该领域领袖们的观点，认知及建议。这些素材都源自作者曾经对他们的采访。
- 此外，本书也包含了“一些可帮助读者对开放数据进行探索的资源”。

+ 开放数据500例网站

- 网站源自一项针对美国公司和组织的调查研究。该调查囊括了那些以全新方式利用政府发布的开放数据来发展新业务的案例。
- 这项研究是由奈特基金会（Knight Foundation）支持，旨在促进媒体的创新，推动交流方式的革新，并支持艺术的发展。
- 调查由位于纽约大学的政府实验室（“GovLab”）主持。该实验室为来自不同领域的创新工作者提供了一个合作的平台，以期运用新技术提高政务水平。
- 开放数据500例网站则担负了发表研究结果的责任。

+ 开放数据500例网站



■ 研究目的：

1. 评估政府发布的开放数据的经济价值；
2. 促进新开放数据公司的发展；
3. 建设一个政府为商业发展提供更有效开放数据信息的交流平台。

+ 开放数据500例网站

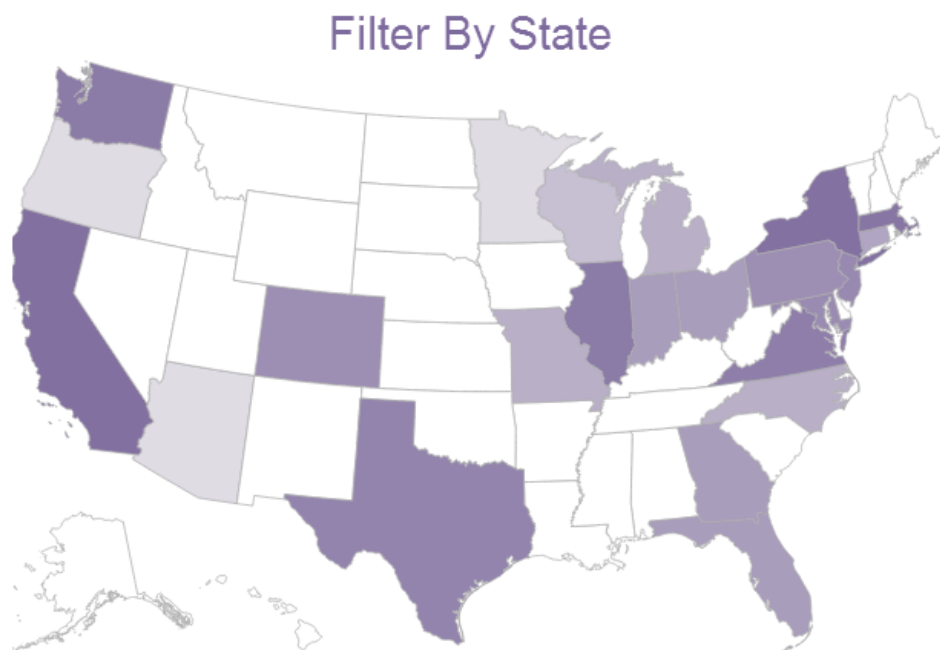


■ 项目蓝图:

1. 在美国境内，建设一系列为政府、商业和其他组织、以及开放数据提供者 and 使用者之间提供发展合作关系的交流平台，以期更好地利用开放数据资源；
2. 在国际环境中，与其他政府和国际组织合作，将美国的开放数据模式推向全球。

+ 开放数据500例网站

- 开放数据500例网站（OpenData500.com）提供了筛选功能，能够根据行业、地址、数据来源对调查所涉及的公司进行查找。



Category

Clear Filter

Business & Legal Services

Data/Technology

Education

Energy

Environment & Weather

Finance & Investment

Food & Agriculture

Geospatial/Mapping

Governance

Healthcare

Housing/Real Estate

Insurance

Lifestyle & Consumer

Research & Consulting

Scientific Research

Transportation

Federal Agency

Clear Filter

Dept of Commerce

HHS

SEC

Dept of Labor

Dept of Energy

Dept of Education

EPA

Dept of the Treasury

Dept of Agriculture

Dept of Transportation

Dept of the Interior

Dept of Defense

Dept of Homeland Security

Dept of Justice

Federal Reserve Board

NASA



研究案例信息



- 公司背景介绍
- 数据来源
- 数据处理方式
- 为客户提供的服务
- 盈利模式
- 声誉及现状



公司目录



纽约

[Enigma](#)

[ZocDoc](#)

[Honest Buildings](#)

[Capial Cube](#)

[Bloomberg](#)

[Aidin](#)

[Calcbench](#)

[Consumer Reports](#)

[GetRaised](#)

[Palantir](#)

美国其它城市

[Archimedes Inc.](#)

[SigFig](#)

[PsychSignal](#)

[CARFAX](#)

[Zillow](#)

[Brightscope](#)

英国

[Mastodon C](#)

[Locatable](#)

[Open Corporates](#)

[DueDil](#)

[CarbonCulture](#)

中国

[上海大数据局](#)

(Big Data Bureau)

[台湾开放数据联盟](#)

(Open Data Alliance)



基础设施



Enigma.io

公司背景介绍



■ 公司名字的意义：

Enigma：难以理解的事物；谜；谜一般的人。

“为了纪念由计算机先锋Alan Turing在二战期间发明的密码破译机；此外，公司的创始人认为，公共数据本身并不像它目前呈现的那样庞大而令人迷惑。”

■ 公司规模：

由10余个现居纽约的成员组成的，



Enigma.io

数据来源

- 他们的数据包括政府数据和客户提供的数据，比如 **Nike** 公开的供应商。
- 他们的数据量现已达到了**10万个数据集**和**200亿个**独立的数据点。





Enigma.io

数据处理方式

- 获取有价值的公共数据并增强其可用性；
- 使在全部数据集中进行快速搜索成为可能；
- “开发了一个强健的数据源和一系列有效工具”。





Enigma.io

为客户提供的服务



- 帮助客户找到与其公司发展相关并可开放获取的信息；
- 收录各样的数据并转换成可用性更强的数据集；
- 把公共信息转换成开放数据。



Enigma.io

盈利模式



- 目前，该公司的利润主要来源于对使用其数据服务的服务费。然而，他们对学术机构、非营利组织和政府机构的收费低于对对冲基金公司的收费。
- 该公司最终希望能将他们的数据完全开放免费地提供给公众，而通过分析数据或其他更优质的服务而获利。



Enigma.io

声誉





医疗领域开放数据

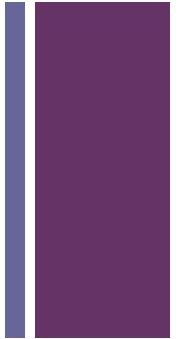


ZocDoc

公司背景介绍

ZocDoc

Get well sooner.



- 该公司于2007年成立，旨在为公众获得医疗服务提供更便捷的途径。ZocDoc免费帮助病人寻找附近的医生。

- 公司总裁Cyrus Massoumi说：

“我曾经在飞机上鼓膜破裂，但是却在之后的四天内都无法找到一个医生。我相信应该有一个更便捷的途径来帮助病人找到医生。这就是我创建ZocDoc的初衷。”

Reason for Visit ▼ Language ▼ Doctor Gender ▼ ☐ Only doctors that see children

	8:40 PM	8:40 PM		4:30 PM			5:15 PM
	more...	more...		more...			more...

☒ In-Network



Dr. Angelina Lauchangco MD

Primary Care Doctor
155 East 47th Street
New York, NY 10017
0.23 miles



It was very informative. A great visit.

aetna

Book Online

View Profile

Mon 05-19-14	Tue 05-20-14	Wed 05-21-14
	10:00 am	9:45 am
	10:45 am	10:45 am
	11:00 am	11:00 am
	11:15 am	11:15 am
	more...	more...



Dr. Luis Correa MD

Primary Care Doctor
46 Essex St
Jersey City, NJ 07302
4.39 miles



Great doc, great staff, great experience.

aetna

Book Online

View Profile

No availability these days.
Next: Thu, May 22



Dr. Michael Aziz MD

Primary Care Doctor
509 Madison Avenue
New York, NY 10022
0.58 miles



I was not in his office very long but he seemed knowledgeable and personable.

aetna

Book Online

View Profile

9:00 am

9:30 am

10:00 am

10:30 am

more...



Dr. Eugene Edynak MD

Primary Care Doctor
155 East 34th Street
New York, NY 10016
0.61 miles



Great! Quick and convenient.

aetna

Book Online

View Profile

9:00 am

9:20 am

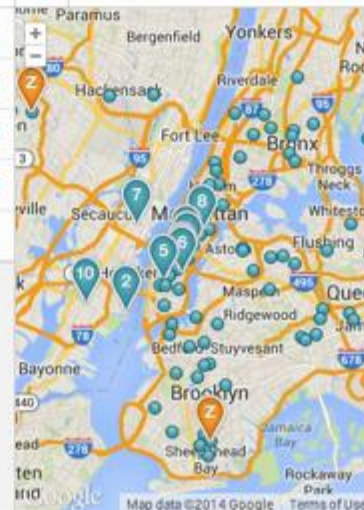
9:40 am

10:00 am

more...

9:45 am

9:00 am



Doctors in New York, NY

Whether you're basking in the bright lights of Broadway or soaking up the serenity of Central Park, use ZocDoc to instantly book a nearby doctor's

[Read More...](#)



Dr. In-Hei Hahn MD



ZocDoc

数据来源

- 首先，他们从病人那里获取信息并提供给医疗人员和组织；
- 另外，他们从医生处获取诸如地址、专科、所接受的保险和病人的评价等信息，并将其提供给病人作为选择的参考。





ZocDoc

数据处理方式



- 使用银行安全级别的保密方法将信息加密；
- 不断分析数据以了解谁在使用ZocDoc及如何提高服务质量。



ZocDoc

为客户提供的服务



- 帮助病人找到附近接受相应保险的医生，查看他们的时间安排并通过手机APP或者ZocDoc的网站进行预约；保证病人可以在24-72小时之内得到治疗。
- 帮助医生进行有效的门诊以及治疗的时间管理。降低办公的管理成本。



ZocDoc

盈利模式

- 在ZocDoc注册的医生、诊所及其他医疗机构，需支付一定的费用。这就是ZocDoc的利润来源。





ZocDoc

声誉



2012年度时代杂志评出的50个最佳网站之一；

2013年世界最佳APPs中商业领域前100名之一；

2012、2013年财富杂志最佳中小型公司之一；

2010至2013年连续4年被Crain杂志评为纽约市最佳工作场所；

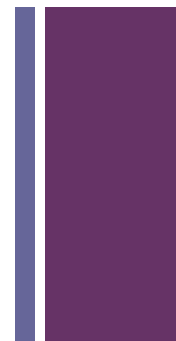
2011至2013年连续3年被现代医疗杂志（Modern Healthcare）评为最佳工作场所；

2013年度Arizona商业杂志最受欢迎公司。



Aidin

公司背景介绍



- 该公司于2011年创立，成员目前不超过10人。
- 该公司的创立是启发于其创始人的亲属在离开医院时接受了大量繁琐的离开医院的程序，并无法得到疗养院信息。
- 该公司创立的目标是，提高医疗机构的透明度，并且为病人提供他们需要的医疗信息。



Aidin

数据来源

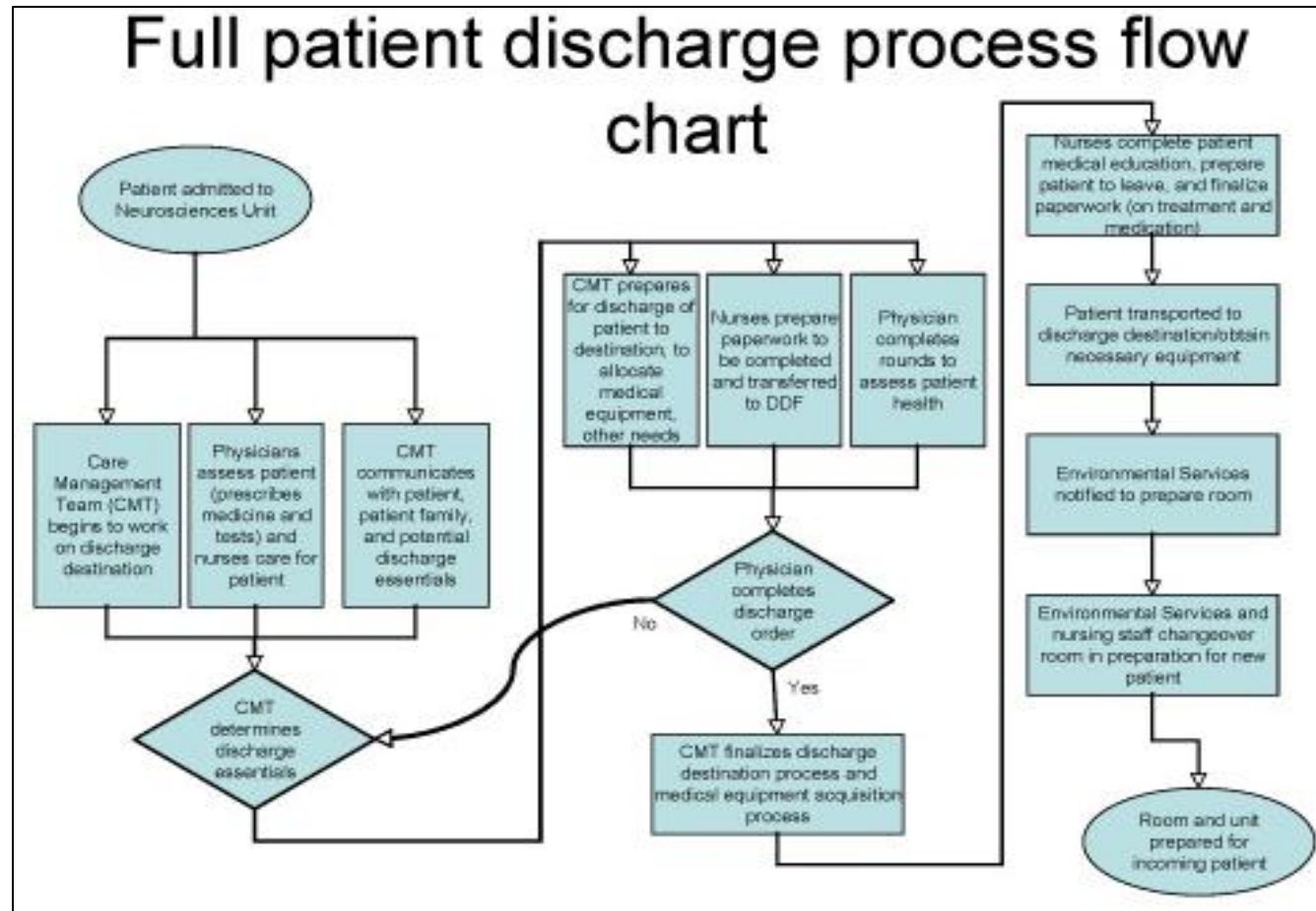
- 政府开放数据；
- 病人对医院的评价。

Provider Name	Address	Phone
JEFFERSON COUNTY HOME CARE	530 BEACON PARKWAY WEST, SUITE 300	205-916-9500
ALACARE HOME HEALTH & HOSPICE	2970 LORNA ROAD	205-979-2619
GENTIVA HEALTH SERVICES	557 GLOVER STREET, SUITE 5	334-347-0234
AMEDISYS HOME HEALTH OF BLOUNTSVILLE	1106 2ND AVENUE E, SUITE E	205-274-4919
SOUTHEAST ALABAMA HOMECARE	806 GLOVER AVENUE, SUITE A	334-794-0591
SOUTHERN RURAL HEALTH CARE	508 ST CLAIR STREET SE	256-332-1631
GENTIVA HEALTH SERVICES	3225 RAINBOW STREET	256-442-1138
AMEDISYS HEALTH CARE	820 SOUTH UNIVERSITY BOULEVARD, SUITE 3F	251-380-0492
SOUTHEAST ALABAMA HOMECARE	810 HEDSTROM DRIVE, SUITE ONE	334-794-0591
SAADS HEALTHCARE SERVICES	1515 UNIVERSITY BLVD, SOUTH	251-343-9600
GENTIVA HEALTH SERVICES	2200 RIVERCHASE PARKWAY, BUILDING 700	205-739-7800
HGA HOMECARE OF DECATUR	2708 HIGHWAY 31 SOUTH, SUITE B	256-350-4182
CARETENDERS	420 10TH STREET SW	205-783-7900
ST VINCENT'S EAST HOME HEALTH	1400 URBAN CENTER DRIVE, SUITE 240	205-838-5795
MERCY MEDICAL HOME HEALTH	101 VILLA DRIVE	251-621-4431
INFIRMARY HOMECARE OF MOBILE	4358 MIDMOST DRIVE, SUITE A	251-450-3300
BAPTIST HOME HEALTH	300 INTERSTATE PARK DRIVE, SUITE 312	334-395-5100
AMEDISYS HOME HEALTH OF HUNTSVILLE	7047 OLD MADISON PIKE, SUITE 305	256-922-9919
ABLE HOMECARE OF BIRMINGHAM	2090 COLUMBIANA ROAD, SUITE 1950	205-979-3180
GENTIVA HEALTH SERVICES	1800 HIGHWAY 84 WEST	334-493-4555
SPECTRUM HOME HEALTH AGENCY	3125 UNIVERSITY DRIVE SUITE C,	256-539-4457
GENTIVA HEALTH SERVICES	1390 NORTH MCKENZIE STREET	251-943-3002
CARESOUTH HOMECARE PROFESSIONALS	3700 ROSS CLARK CIRCLE, SUITE 1	334-793-5758
GENTIVA HEALTH SERVICES	1108 EAST PARK DRIVE	205-836-0777



Aidin

数据处理方式

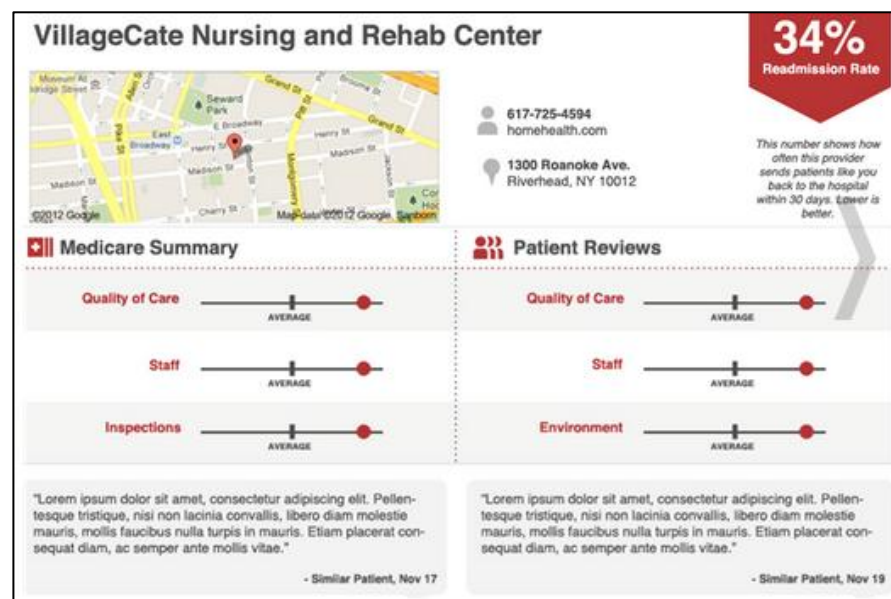




Aidin

为客户提供的服务

- 该公司给医疗机构提供他们的数据并帮助其提高服务质量；
- 病人也可以利用该公司提供的信息对医院进行选择。





Aidin

盈利模式

- 通过他们的产品获利，包括：The full Aidin Solution, Aidin Lite, 以及Aidin Provider。



Reduce Length of Stay



Automate Medicare
Compliance

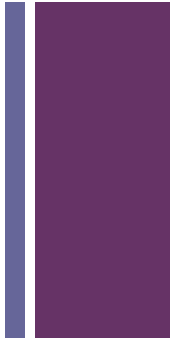


Reduce Marketing &
Admin Expenses



Aidin

声誉



- GE Quest: 医院方面一等奖获得者。





Archimedes Inc.

ARCHIMEDES
Quantifying Healthcare

公司背景介绍

- 自1993年起作为Kaiser Permanente公司的一部分而成立，后于2006年成为一个独立的公司。
- 该公司具有数十年的开发针对医疗建立算法和预测模型的经验。

Healthcare Modeling
and Analytics



+ Archimedes Inc.

数据来源

- 医疗保险与救助中心（Centers for Medicare & Medicaid Services, CMS);
- 医疗试验数据库;
- 国家健康与营养调查（National Health and Nutrition Examination Survey）数据。

Facts, Figures,
and Free
Resource Center





Archimedes Inc.

数据处理方式



- 他们使用阿基米德模型（Archimedes Model）作为他们的核心数据分析工具。
- 他们的科研队伍分析了多样的相关数据并推导出了百余公式，来代表不同疾病、体检和治疗的效果。
- 这些公式通过面向对象编程，而被整合到一个单独的、大规模的虚拟模型中。



Archimedes Inc.

为客户提供的服务



- 他们提供一套在线的医疗模拟和分析工具（**ARCHeS**），来分析不同治疗方法在医疗和经济方面的效果。
- 他们的另一个产品（**IndiGo**）针对个人提供健康指导，并对不同预防性治疗进行优先级排序。



Archimedes Inc.

盈利模式

- 他们通过销售产品（ARChES和IndiGo）来获取利润。
- 同时，他们也提供建模和咨询服务，包括针对疾病和新的医疗手段的建模，和对不同预防措施的分析。



+ Archimedes Inc.

声誉

- 2012年6月6日获得由医疗数据倡导行动颁发的最佳医疗软件奖。

Better Information
Better Adherence
Better Outcomes

IndiGO





金融领域



Capital Cube

 analytixinsight

公司背景介绍

- Analytixinsight于2010年创立，现有约10名员工。
- 作为该公司的门户网站，CapitalCube.com为一个全球投资者提供全面的公司分析。
- 按需对4万多家全球企业进行基本面分析，投资组合评估，以及筛选服务。

Google Inc. Class C (NASDAQ:GOOG) May 19, 2014

Summary Peer Analysis Trends Analysis Reports Alerts Chart News Financials

Summary

Company scores, performance ratings, and analysis. All at a glance.
Register to get the full analysis. [Learn More](#)

Stock Price

[View Comparative Chart →](#)

520.63

USD
as of May 16 2014 52 week range (USD): 502.80 - 604.83

3 Month Price - GOOGUS



Peers

	close	52 week range	
Apple Inc.	597.51	388.87 - 604.41	USD
Microsoft Corporation	39.83	30.84 - 41.66	USD
Amazon.com, Inc.	297.70	258.34 - 408.06	USD
eBay Inc.	51.95	48.06 - 59.70	USD
Yahoo! Inc.	33.41	23.82 - 41.72	USD
Adobe Systems Incorporated	61.64	41.91 - 71.11	USD
Akamai Technologies, Inc.	53.33	40.08 - 63.15	USD

[Register To Customize Peers →](#)

Analysis

[Unlock The Complete Analysis →](#)

Fundamental Analysis

Relative Valuation

Valuation Drivers

Operations Diagnostic

Earnings Leverage

Sustainability of Returns

Drivers of Margin



Closing Stock Price
USD 520.63

As of Date
May 16, 2014

One Year Performance
N/A%

One Month Performance
-6.45%

Industry Group
Internet Software/Services

52 wk. Range
USD 502.80 - USD 604.83

Market Cap (Millions)
USD 351,162.00

Total Shares Outstanding
N/A

Latest Filing Dates
Annual

2013-12-31

Semi Annual
N/A

Quarterly
2014-03-31

Next Release Date
2014-07-17

Filing Currency
USD

Listed Exchange
NASDAQ

Peers

✓ **Apple Inc.**

✓ **Microsoft Corporation**



Capital Cube

数据来源



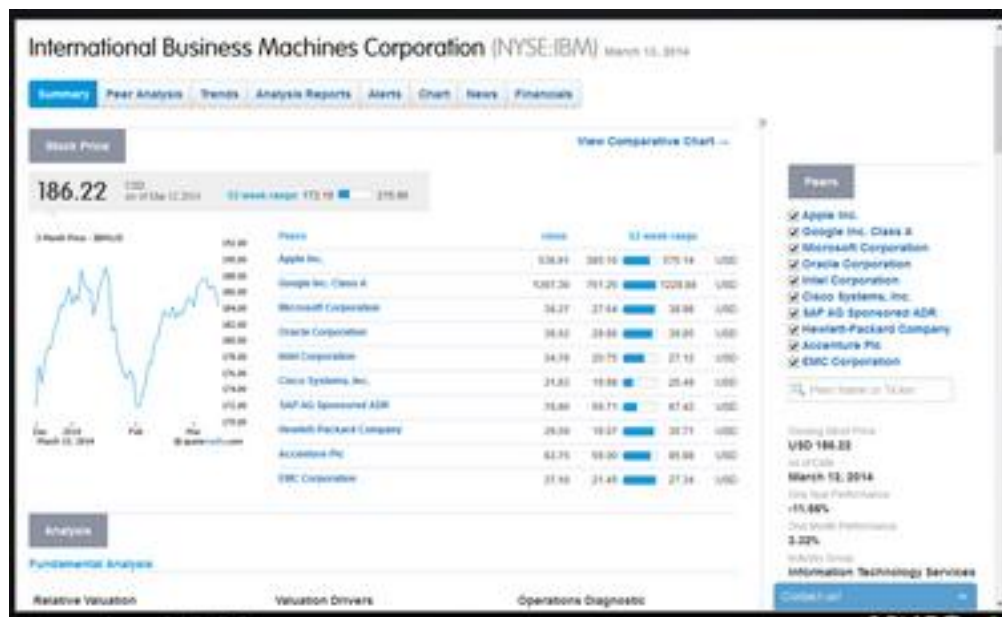
- 美国证券交易管理委员会（Securities and Exchange Commission, SEC）
- 第三方数据：
 1. FactSet
 2. Thomson Reuters
 3. Capital IQ



Capital Cube

数据处理方式

- 开发了一个软件，可每日获取来自4万家国际化运营的上市公司的数据；
- 将数据转换成投资者可用于投资决策的文字报告和图表。





Capital Cube

为客户提供的服务

- 投资工具

- 分析服务：

1. 股利分析

2. 盈利质量分析

- 优势：

1. 覆盖面广（4万多个公司）

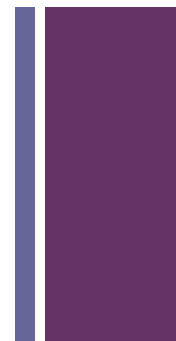
2. 时效性（每日更新）





Capital Cube

盈利模式



- 为经销商提供能够使用其数据的权限。
- 两大主要经销商是：
 1. 三星（Samsung）：每次三星的用户下载了CapitalCub引用的数据，AnalytixInsight就会获得分红；
 2. 道琼斯（Dow Jones）：AnalytixInsight能够从每一次其挂于道琼斯网站上的内容得到浏览中获利。



Capital Cube

声誉



■ 2013年10月28日:

该公司被白宫科学技术政策局 (the Office of Science and Technology Policy, OSTP)和科学技术政策研究所 (the Science and Technology Policy Institute, STPI)评为以创新并激动人心的方式成功开放联邦数据的公司之一。





Calcbench

公司背景介绍



- 创立于2011年的Calcbench是第一家完全利用了政府强制的扩展的商业数据语言XBRL（Extensible Business Reporting Language）的公司。
- 目前该公司拥有约10名成员。



Calcbench

数据来源

- XBRL：一个开放的、政府强制的、用于商业信息交换的标准数据格式；
- XBRL报告。





Calcbench

数据处理方式

- 经过这个公司的处理之后，开放的但可用性差的XBRL数据被转换成更详细具体的数据，因此数据的价值得以提高。

The screenshot displays the XBRL Viewer application window. The main area shows a table titled 'Disclosure of exploration and evaluation assets' for the period 2009. The table has multiple columns, including 'Identifier', 'DAEMT', and 'Period'. The table content is organized into sections like 'Disclosure of exploration and evaluation assets', 'Exploration and evaluation assets', and 'Changes in exploration and evaluation assets'. The bottom left pane shows the 'Properties, plant and equipment' section with a list of items and their values. The bottom right pane shows the 'Fact properties' section with a list of properties and their values.

Identifier	DAEMT	Period	2009
Disclosure of exploration and evaluation assets			
Disclosure of exploration and evaluation assets			
Exploration and evaluation assets			
Changes in exploration and evaluation assets			



Calcbench

为客户提供的服务



- 该公司把政府强制收集的XBRL原始数据转换成可用性强的金融数据。
- 该公司同样使用其技术帮助SEC提高数据质量，如查找数据中的错误并改正。

How It Works



Calcbench Engine processes new filings in seconds:

- Error correction
- Normalizing
- Analytics & Reporting
- Proprietary database

← artificial intelligence

← machine learning

← big data

Cloud based data and analytics platform

Investors

CFOs

Auditors

Academics



Calcbench

盈利模式

- Calcbench为他们的客户提供数据并收取一定费用。
- 它同样提供其他分析服务来获取回报。





Calcbench

声誉

- 在2012年2月29日召开的XBRL与金融分析技术会议上，该公司获得XBRL挑战一等奖。



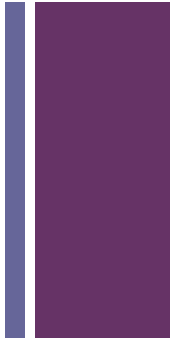


SigFig

公司背景介绍



SigFig



We're an investment company that puts people before profits.

Our mission is to build a new kind of investment company: one that uses science and data to help all investors, large and small, invest better. We're a company that treasures innovation and values beautiful design; one whose DNA is not Wall Street; one which values simplicity and whose team revels in the opportunities created by technology. We are a company that creates transparency in an industry mired with conflicts of interest and hidden fees.



SigFig

数据来源

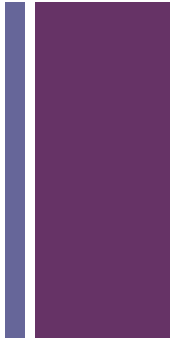


- 基本面数据和价格数据的历史
- 一些第三方数据源如Lipper、Thomson Reuters、Interactive Data和Xignite



SigFig

数据处理方式



The Journey of a SigFig Recommendation



We Diagnose Your Portfolio

Every week, we run 100+ tests on your investments to find ways for you to save and make more.

We Gauge Your Situation

Next, we personalize and tailor your advice to your unique financial situation.

We Send You Advice

Finally, we deliver clear, actionable investment advice right to your virtual doorstep.



SigFig

为客户提供的服务



■ 一个简明的信息表

BALANCE				TODAY'S GAIN/LOSS			
\$129,932				\$1,226 0.9%			
SYNCED ACCOUNTS				MANUAL PORTFOLIOS			
Fidelity	\$121,778	\$23,478	0.9%	Watchlist 01	\$0	\$0	0.0%
401K (Acct***5426)	\$101,924	\$1,169	1.2%				
IRA (Acct***1358)	\$21,024	\$22	0.1%				
Charles Schwab	\$6,933	\$49	0.7%				
Acct *****1234	\$6,933	\$49	0.7%				

Accounts

Account Total **\$234,567**
Today's Change **\$3,456**
Percent Change **1.6%**

E*Trade **1.1%**
\$91,827 **\$992**

Scottrade **1.8%**
\$128,239 **\$2,618**

Fidelity **-0.5%**
\$21,023 **-\$168**

U.S. Markets 3:45 PM EST

S&P-500 **1.0%**
1,300.00 13.00

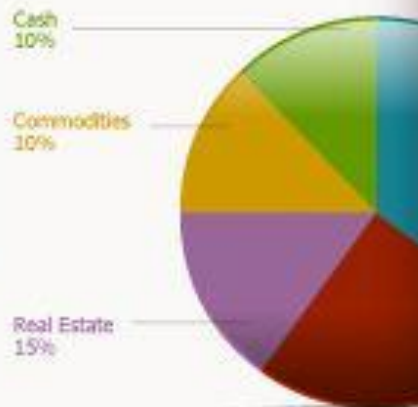
DJIA **1.0%**
13,000.00 130.00

Nasdaq **-1.0%**
3,000.00 -30.00

Performance



Asset Allocation



Geographic Allocation



\$234,567 **\$3,456** **1.6%**

Investment Advice

BAD INVESTMENT

4 out of 5 funds perform better than your fund.

\$5,345

HIDDEN BROKER FEE

You were charged \$120 in late settlement fees.

\$120

EXPENSIVE ADVISOR

Your advisor charges more than 75% of other advisor

\$5,345



SigFig

盈利模式



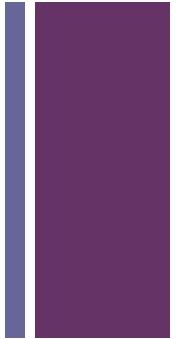
- 该公司不收取管理费用，他们只通过发布与其他网站的合约以及将用户介绍给其他经纪人来获取收益





SigFig

声誉



The Trophy Case



Best Apps of 2013



*10 Most Innovative
Companies in
Finance*



*5 Best Investing
Apps*



*8 Most Useful
Financial Apps*





Brightscope

公司背景介绍



- BrightScope®, Inc是一家金融信息公司。
- 他们使用数据给独立投资人、公司、资产经理人等提供信息，帮助他们进行决策。
- 他们主要有两个板块的服务：退休金计划（Retirement Plan）和财富管理（Wealth Management）。



Brightscope

数据来源



- 美国劳工部的5500表格（Form 5500）中数据；
- 公司的创始人，Mike和Ryan Alfred，成功劝说劳工部将表5500的信息收集、整理并发布到网上。



Brightscope

数据处理方式

- 他们收集并重新发布处理后的可用性更强的数据，并且使这些数据更易被公众获取。





Brightscope

为客户提供的服务



- 给401k计划和403b计划评分排名；
- 发起了第一个全面公开的金融分析师的目录；
- 提供免费的关于固定缴款计划（**Defined Contribution, DC**）养老金的白皮书。



Brightscope

盈利模式



- 通过出售他们的市场情报产品（**Beacon**）和针对养老金计划的销售情报产品（**Spyglass**）来获利；
- 针对5万个固定缴款（**Defined Contribution, DC**）养老金计划，提供包含细节和数据的研究报告，并以此获利。



Brightscope

声誉



- 第一家说服美国劳工部收集、整理并在网上发布 5500 表格中信息，并以此获利的公司；
- 提供第一个全面、公开的金融分析师目录。

Data. Diligence. Decisions.™



PsychSignal

公司背景介绍



- 由另一个分析大数据的创业公司SmogFarm发起。

Who We Are

ENGINEERING & PRODUCT

James Crane-Baker



James had +10 years of trading experience while actively leading the revolution in electronic trading before turning to startups. [Learn more about him.](#)

Ramon is a Visual Designer and Front-end Engineer.
[Learn more about him.](#)



Ramon Bispo

Leonardo Almeida



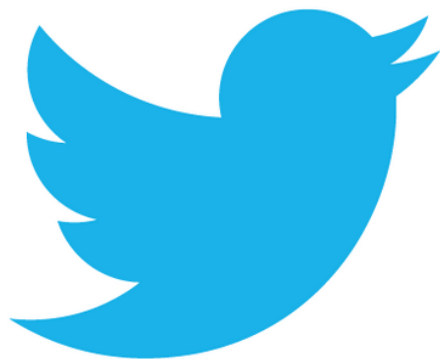
Leo is an information retrieval expert and Software Engineer. [Learn more about him.](#)



PsychSignal

数据来源

- 从社交媒体中获取对话信息。



TM





PsychSignal

数据处理方式

- 从在线对话中寻找表达心理、情感、态度的语句；
- 集结了数以百万的情感语句用以描述群众的情感；
- 建立了一个先进的情感搜索引擎；
- 使用由NASA开发的信号处理工具。





PsychSignal

为客户提供的服务



The most granular and accurate real time financial sentiment indices in the world.

We set out to quantify real world crowd psychology and in doing so we assembled leading experts in the fields of psychology, engineering and data science. Other sentiment products are generally built upon pre-existing technology or hastily built based on flawed theories. At PsychSignal we have the confidence to [compare our live results](#) to real world sentiment surveys and measurements.



PsychSignal

盈利模式

- 使用NASA最新开放的算法，每日分析在线对话。从噪声很强的信息中提取人们的情感，并以此获利。





PsychSignal

声誉



We do one thing exceptionally well.

Our skill set lies in the discovery and processing of sentiment signal.

What people are saying about us



"model results may be ground for cautious optimism about creating actionable model based on sentiment indicators."

— [Sergey Yurgenson](#)

Kaggle Top 10 ranked data scientist



"we believe there is sufficient evidence of alpha generation potential in the PsychSignal data feed for it to warrant additional research effort [...]"

— [Stuart Farr](#)

President, Deltix



建筑和房地产

+ **Honest Buildings**

公司背景介绍



- 创立于2011年，目前由约30名员工。
- 作为一家商业地产中介，它将建筑商、地产商、决策者和项目管理人员联系起来。
- 它是地产领域的社交媒介。

+ Honest Buildings

数据来源

- 该公司收集由建筑商、地产商、业主和其他相关人员提供的数据。





Honest Buildings

数据处理方式



- “大数据虽然是很好的资源，但是需要一定的处理。目前对将数据整体化的平台需求很大。”

—— Honest Building.com

- 该公司收集并发布各类相关数据，这些数据往往是用户很难找到的，例如能源花费，交通便利程度等。



SEARCH

OPEN PROJECTS

COMMUNITY

How It Works

List Your Business

Log In

START A PROJECT

MAY
29

GC for Restaurant/Bar Add-On Fit Out- Brooklyn

Brooklyn, NY

PROJECT CODE: N14079

APPLY NOW

PROJECT INFORMATION

Scope

Existing Restaurant/Bar seeks contractor for conversion of adjacent 500 sq ft former coffee shop into additional space. Project involves replacement of existing coffee shop storefront with brick overlay and interior game room

Approx Project Value

Building Description

- *Type:* Retail
- *Approx Size:* 500 sq. feet
- *Num. Stories:* 1

Building Location

Brooklyn, NY

Project Status

Soliciting Bids

Candidate Requirements

- Experience with Retail/Restaurant Conversions
- Prior Building Projects in Brooklyn
- Expedition of Permits through Brooklyn DOB

APPLICATION DUE

May 29th, 2014

APPLY NOW



About Honest Buildings

Honest Buildings helps real estate decision-makers identify trusted consultants and vendors who have an established track record of relevant projects. We are the client's partner as they find and evaluate candidates, solicit bids, and make decisions.

START A PROJECT

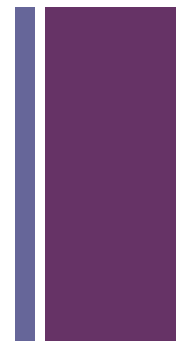
+ **Honest Buildings**

为客户提供的服务

- 为客户提供匹配服务，帮助其一步步找到合适的投标：
- 客户可在这个平台
 1. 注册公司；
 2. 建立过往案例的专案展示；
 3. 添加标签和联系人；
 4. 获得投标。

+ Honest Buildings

盈利模式



- 该公司对其提供的匹配服务提供一定的费用；
- 它收取建筑商等投标人一定费用为其匹配合适的开发商。



+ Honest Buildings

声誉

2014年纽约
黑客马拉松
获奖者





Zillow

公司背景介绍



- 创立于2005年，该公司总部在西雅图。在纽约、旧金山、芝加哥，加州欧文，内布拉斯加林肯等市均有分公司。
- 它给消费者提供关于居家、房地产、贷款的信息和工具，帮助他们做出决定。
- “在Zillow，我们把以前需要数小时才能获得的房地产信息收集起来，将它们转换、发布、免费提供给消费者。



Zillow

数据来源



■ 政府开放数据源，包括：

1. 劳动统计局（The Bureau of Labor Statistics）；
2. 联邦住房金融局（Federal Housing Finance Agency）；
3. 人口普查局（The Census Bureau）。



Zillow

数据处理方式



- 将政府的开放数据收集起来，并建立一个包含了1.1亿美国家庭的实时更新的数据库。
- 将这些数据在他们的网站上重新发布，给用户关于住房、地产、信贷的不同信息。



Zillow

为客户提供的服务



- 提供在售房、待租房、和即将上市的房产等信息，并给买家和房东提供信息；
- 给借款人和贷款人提供一个开放的、匿名的、自由的中介；
- 给不同的房屋提供一个估价，名为**Zestimate**，作为确定房屋价值的一个开始。

For Rent ▾

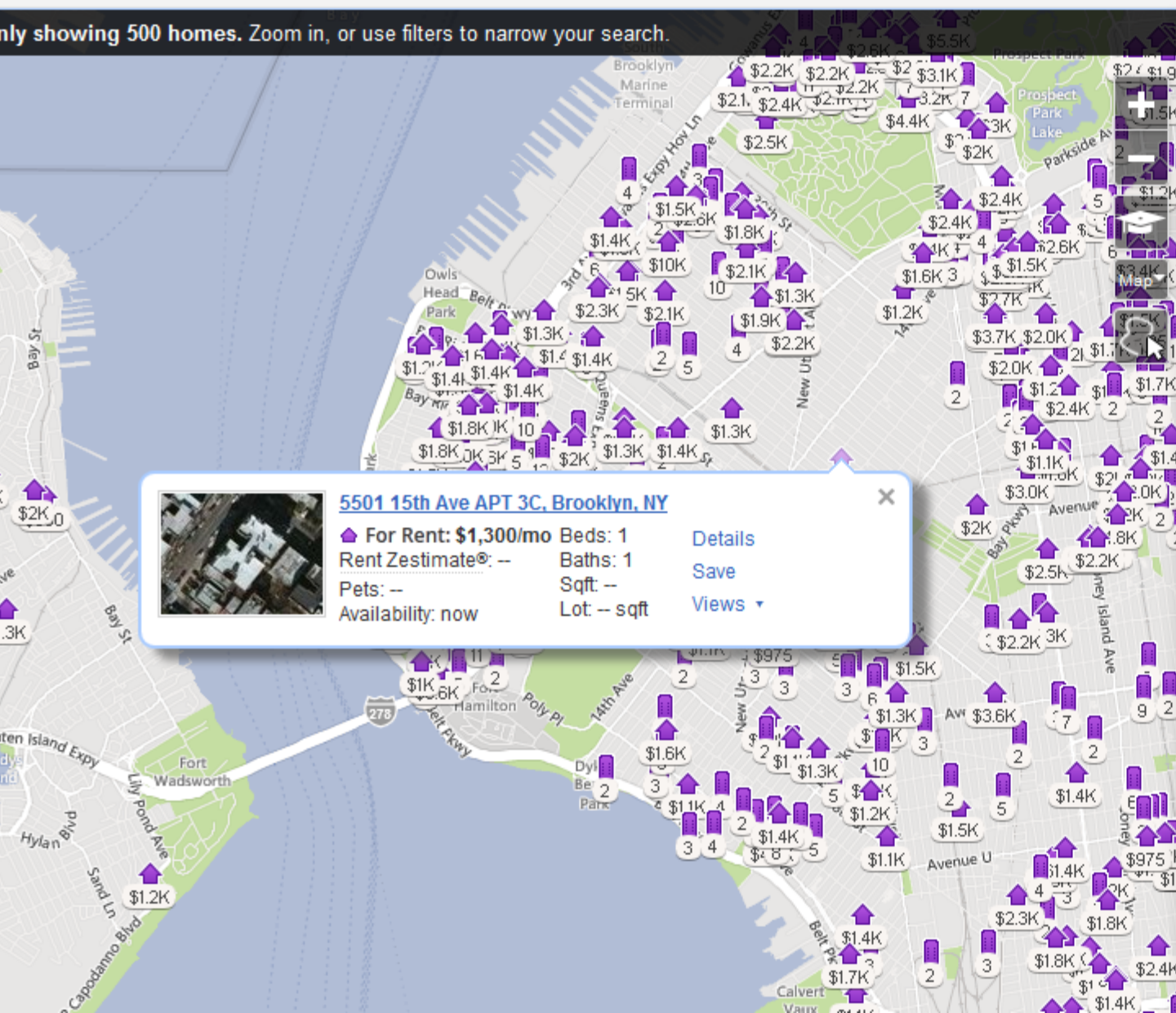
Any Price ▾

0+ Beds ▾

more ▾

Save Search

Only showing 500 homes. Zoom in, or use filters to narrow your search.

The free,
Post your

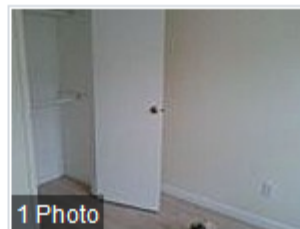
Rentals

3,001 results. 4 un

Verified Source

Newest

Rent (low



1 Photo

Apartm

\$1,100

1 day on Z

[E 14th St, Brooklyn, NY](#)

6 Photos

Apartm

Pets: cats



7 Photos

Apartm

Pets: No

Show 8 m



Zillow

盈利模式



- Zillow主要以广告盈利；
- 它与雅虎家居合作，运营着全美最大的房地产和房屋租赁网络；
- “该公司于2005年创立，在2011年约有6600万美金年收入。在2013年公司资产估价为23亿美金。



Zillow

声誉

2013年：

11月17日：Zillow总裁Spencer Rascoff被提名为EY国家级企业家；

10月10日：Zillow的APP，Zillow Mortgage Marketplace，获得贷款技术奖；

6月13日：Zillow在Webby Awards上获得民声奖和房地产类最佳奖；

5月4日：Zillow获得Webby Award最佳房地产网站奖。



+ Zillow与R



- WebTechnologies task view下有zillow包
 - 根据地址和邮政编码给出房屋的估价
 - 查询给定房产的信息
 - 搜索同类型的房产

<http://www.omegahat.org/Zillow/>



汽车



CARFAX

公司背景介绍



- 该公司每年为数以百万计的消费者提供车辆历史信息。
- 该公司网站每月访问量达百万。
- 在美国境内，数千汽车经销商注册使用该网站。



CARFAX

数据来源

- 美国机动车管理机构
- 加拿大省级机动车管理机构
- 机动车拍卖会
- 机动车碰撞修理机构
- 服务维修机构
- 保险公司
- 汽车回收机构
- 汽车租赁公司
- 州检查站





CARFAX

数据处理方式

■ 将他们的数据统一格式并形成报告。报告中包含如下信息：

1. 牌照信息；
2. 损坏记录；
3. 事故记录；
4. 仪表盘度数；
5. 使用历史；
6. 事故指标，如气囊被使用的次数；
7. 历任所有者情况等



CARFAX

为客户提供的服务



- 以一定价位出售报告；
- 通过提供二手车信息获利；
- 同样提供不同品牌型号的汽车调查。



CARFAX

盈利模式



- 主要通过销售报告盈利。
- 同时提供CARFAX Hot ListingsTM 和 Safety & Reliability RatingsTM 等产品。



CARFAX

声誉



- 具有北美最全的汽车历史数据库。
- 最受消费者认可的五家汽车信息网站之一。



SHOW ME THE CARFAX®



公共服务领域



Consumer Reports ConsumerReports®

公司背景介绍



- Consumer Reports成立于1936年，是一个专业、独立的非营利组织。
- 它致力于为消费者提供一个公平、公正、安全的媒介。
- 它现有超过500名员工。



Over **5,000** Products Rated

Join today & get 24/7 online & mobile access to:

Expert Ratings • Buying Advice • Shop Online • User Reviews • Much, much more!

Subscribe Now!

Find Ratings

Cars ▶

Appliances ▶

Electronics ▶

Home & Garden ▶

Babies & Kids ▶

Money ▶

Shopping ▶

Health ▶

GENERATORS THAT KEEP YOUR HOUSE RUNNING

Hurricane season starts on June 1.
Are you prepared for an emergency?

Get Ratings and watch our buying guide video



Recent Tests



Air conditioners



Freezers



Gas grills



Tablets



Tires

Policy & Action



**SUBSCRIBE
ONLINE**

Join today & get 24/7 online access to:

- ▶ Expert Ratings
- ▶ Buying Advice
- ▶ Mobile (now included in your subscription)

CLICK HERE TO
SUBSCRIBE

Latest From Consumer Reports



WASHING MACHINE REVIEWS
Memorial Day is deal time for

Videos



+ Consumer Reports

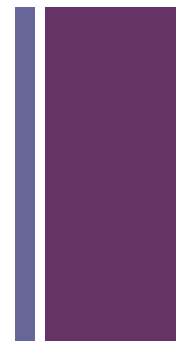
盈利模式

- Consumer Reports是一家非营利机构。
- 它不接受广告费用，没有股东。
- 其收入仅来源于会员的服务使用费。



+ Consumer Reports

声誉



- 2008年度职业记者协会（Society of Professional Journalists）颁发的公共服务领域Sigma Delta Chi奖。
- 2007年度国家媒体俱乐部（National Press Club）荣誉提名。
- 2006年度国际数据艺术与科学学术协会（International Academy of Digital Arts and Sciences）颁发的人民之声奖。
- 2005年度美国皮肤学协会（American Academy of Dermatology）颁发的金三角奖。



GetRaised

公司背景介绍

GETRAISED Promote Yourself

- 该公司于2010年创立，目前拥有不超过10名成员。
- 旨在通过将大量复杂数据转化为一个简明而可用性强的用户界面，来帮助人们提高薪金。

Find out if you're underpaid and get the right raise, for free.

Average raise earned using GetRaised: \$6,726

GET STARTED →



1. TELL US ABOUT YOURSELF

Tell us what you do and how much you make—we'll tell you if you're underpaid.



2. ASK FOR MORE

Create a custom Raise Request that helps you start the conversation with your boss.



3. REAL ADVICE

Get our expert guidance through the raise process.

FEATURED IN



Forbes
Woman



The
New York
Times

Mashable



WHAT GETRAISED USERS ARE SAYING

When I saw the letter that was compiled I was

SUPPORTED BY

ORIGINATE

CONNECT WITH GETRAISED



GetRaised

数据来源



- 来自美国劳动统计局（U.S. Bureau of Labor Statistics）的开放数据；
- 用户输入的信息；
- 各类网上的招聘信息。



GetRaised

数据处理方式



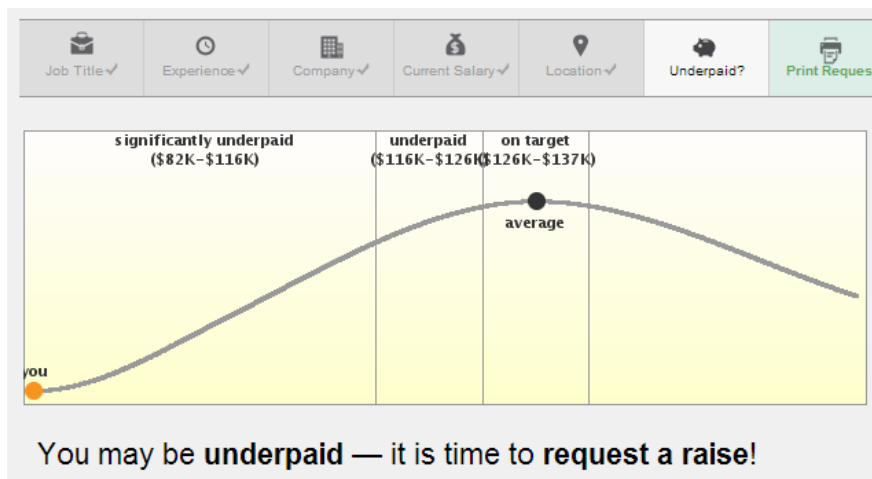
- 该公司通过其收集的信息，开发了一个薪金的搜索引擎。
- 它根据相关领域专家（如人力资源和研究机构）的分析，为用户提出了很多加薪申请。



GetRaised

为客户提供的服务

- 给用户提供的信息，告知他们应得的薪金，是否应提出加薪申请等。
- 大量女性利用由该公司提供的信息，提出了加薪申请并成功得到了加薪。





GetRaised

盈利模式

- 作为一个非营利组织，该公司提供的服务是免费的。
- 它接受了两个机构和四个个人的捐款。





GetRaised

声誉



- 通信策略专家Dave Clarke提到，“使用GetRaised的81%的女性真正得到了加薪。平均获得的加薪金额是6473美元。”



Palantir

公司背景介绍



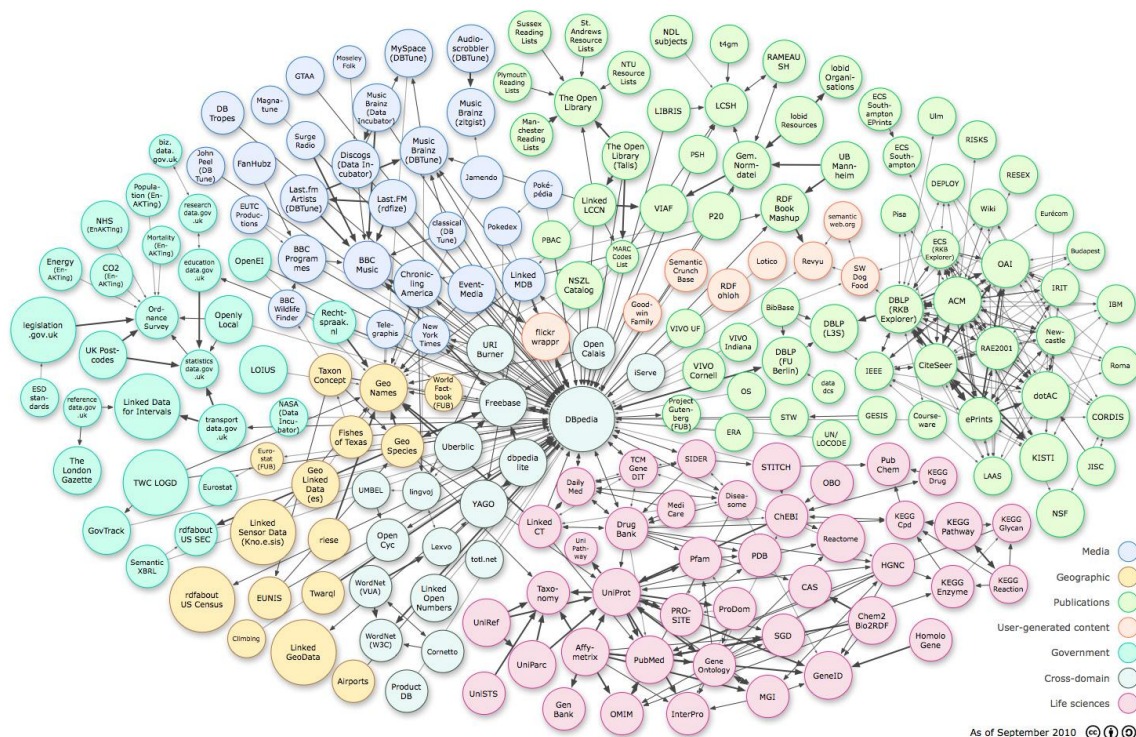
- 于2004年创立，Palantir目前有超过500名员工。
- 它致力于开发从大量分散数据中提取信息的软件。
- 它所提供的产品和服务，有助于反恐、遏制犯罪和诈骗等。



Palantir

数据来源

■ 政府的开放数据



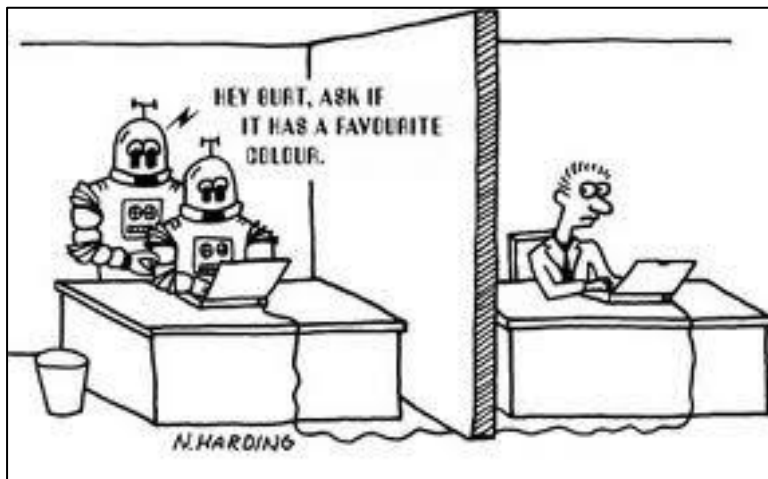
As of September 2010 © ⓘ ⓘ



Palantir

数据处理方式

- 用户输入各类数据到Palantir的软件中；
- 它将生成用户可直接理解的报告和分析；
- 虽然不能将该软件称为人工智能，但是有人将其称为“人工理论”。





- [illegible]

They're in your data. We'll show you where.





Palantir

盈利模式

- 股东的投资；
- 软件的销售。





Palantir

声誉



- 据报道，在2013年底，该公司的价值达到了90亿美金。
- 2009年度VAST2009 可视化分析奖。
- 2010年10月，在摩根大通银行技术革新座谈会上获得创新奖。



英国的开放数据



Mastodon C

公司背景介绍

MASTODON C



BIG DATA DONE BETTER

- Mastodon C是一家大数据分析公司。
- 该公司是第一家加入英国开放数据协会(Open Data Institute, ODI)的此类公司。
- 公司总裁Fran Bennett为搜索引擎公司工作多年，帮助他们从数据中获利。



Mastodon C

数据来源

- 他们的数据来源是来自英国国家卫生局（National Health Services, NHS）的开放数据。





Mastodon C

数据处理方式



- 使用云计算分析数据；
- 使用Hadoop和Cassandra技术来整合实时探测器数据，网络服务数据和来自客户的表格数据；
- 他们的一个重要特征是低碳环保地进行数据分析。



Mastodon C

为客户提供的服务



- 从杂乱无章的开放数据或客户的私有数据中提取有用的信息；
- 将数据整合转换成一个更清晰的格式。



Mastodon C

盈利模式



- “Mastodon C刚刚完成一个由政府资助的分析项目，分析了在英国境内一些医生无故开昂贵药物的情况。”



Mastodon C

声誉

- 2012年度在伦敦绿色黑客马拉松活动中获得Best Hack和PeopleFund.it奖。



Peoplefund.it
START IT. FUND IT. BE PART OF IT.

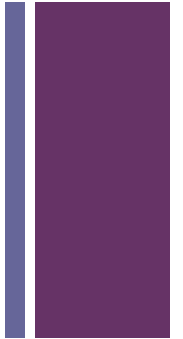


Locatable

公司背景介绍



Locatable



- 另一个加入了开放数据协会的公司。
- 起初他们希望建立一个网站帮助人们决定住在哪里，现在他们的服务转变为帮助人们更好管理自己的房屋。
- 公司创立者Vasanth Subramanian和David Prime曾经都是物理系学生。



Locatable

数据来源



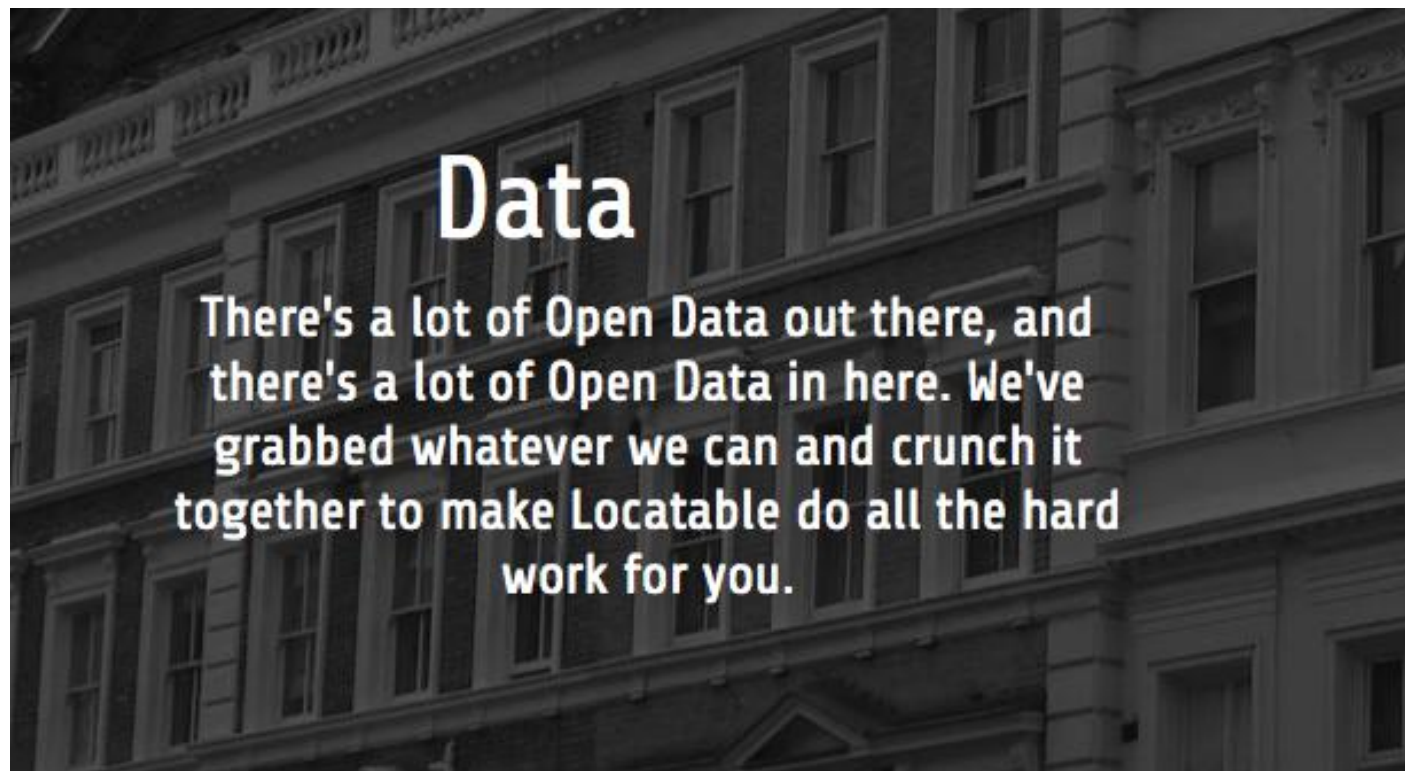
Layer	Data	Source	License	Coverage
Places	National Public Transport Data Repository	data.gov.uk	UK Open Government License	England, Scotland, Wales
	OpenStreetMap road network (all roads)	© OpenStreetMap contributors	Open Database License	Greater London
	OpenStreetMap road network (major roads)	© OpenStreetMap contributors	Open Database License	UK
Amenities	OpenStreetMap point-of-interest data	© OpenStreetMap contributors	Open Database License	UK
Schools	Independent school inspections and outcomes	Ofsted	UK Open Government License	England
	Maintained school inspections and outcomes	Ofsted	UK Open Government License	England
	School performance data	Department for Education	UK Open Government License	England
House Prices	Price Paid Data	Data produced by Land Registry © Crown copyright 2013.	UK Open Government License	England
	Postcode Data (Code-Point Open)	Contains Ordnance Survey data © Crown Copyright and database right 2013	OS Open Data License	England
Property	Zoopla API	Powered by Zoopla	Zoopla Terms of Use	UK
Electricity Tracker	NEED report - Local Authority consumption tables 2011	Department of Energy & Climate Change	UK Open Government License	UK
	Annual domestic energy bills - QEP 2.2.3	Department of Energy & Climate Change	UK Open Government License	UK



Locatable

数据处理方式

- 把不同来源的数据整合到一个简单的数据仪表板中。





Locatable

为客户提供的服务



- 给客户提供一个简单的数据表，来帮助他们管理并优化他们的房屋；
- 提高昂贵开销（如贷款、保险和物业等）的透明度；
- 帮助他们的客户找到能将费用降低的方式。



Locatable

盈利模式

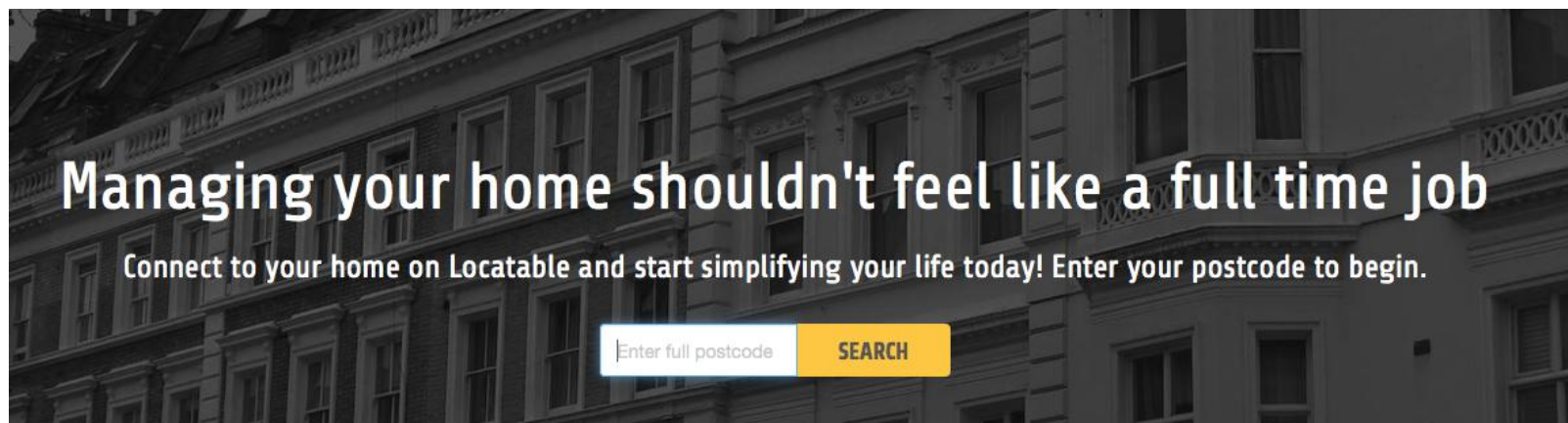
- 为房地产经纪提供商业模式；
- 网站提供的搜索服务与收费：
 1. 消费者登录网站并确定自己希望居住的位置；
 2. 搜索结果会告知消费者那些最符合其要求的位置；
 3. 网站将消费者推荐到其他机构并收取中介费用。



Locatable

现状

- 目前只针对伦敦市内的房屋提供服务，但该公司有意向将业务范围推广到全英。





OpenCorporates opencorporates

公司背景介绍

- OpenCorporates是Chrion Ltd的门户网站, 由Chris Taggart和Rob McKinnon建立。
- Rob也创立了TheyWorkForYou.nz和WhosLobbying.com这两个网站。
- Chris创立了OpenlyLocal.com, OpenCharities, 以及英国政府网站的地方开放数据板块、税收透明度公开板等。



OpenCorporates

数据来源



- 他们的数据主要来自政府和其他机构的资源，包括：
 1. 直接从政府网站和API中下载；
 2. 收集对公共开放的数据集；
 3. 依据《信息自由法案》获取所需数据。



OpenCorporates

数据处理方式



- 公司创始人之一**Chris**解释到：“我们依据《信息自由法案》收集公司注册信息以及政府网站发布的信，做整理之后使其可被公众使用。”

+ OpenCorporates

为客户提供的服务



Get data access to over
60 million companies

Open data

- All the data on the world's largest open database of companies
- Available as either share-alike attribution open data or commercially
- All data includes sources, allowing checking of data

Quality data

- Data is from primary public sources
- Fully provenanced, providing context and confidence
- 'Many eyes' on OpenCorporates adds quality control for free
- Versioned API

Unique data

- Open, transparent and highly granular model
- Many unique datasets and sources
- Enrich your existing data, or build new services



OpenCorporates

盈利模式



- 尽管OpenCorporates免费发布他们所收集的所有开放数据，但是他们提供的商业模型则需要额外付费。

+ OpenCorporates

现状

We have information on
67,193,547 companies

☒ search companies ☐ search officers

Filter by jurisdiction

1,297 [Abu Dhabi \(UAE\)](#)
144,280 [Alaska \(US\)](#)
40,157 [Albania](#)
877,977 [Arizona \(US\)](#)
46,509 [Aruba](#)



DueDil

公司背景介绍



- 创立于2011年的创业公司DueDil，其名字源于英文“Due Diligence”，意为“双方调查”。
- 他们的目的是为有意于向中小型公司注资的投资者们提供所需信息以增强其投资信心。



DueDil

数据来源



- “DueDil基本就是依赖于开放数据而建立的，这些数据包括私人公司的房产以及英国政府控股企业的在册房产。”
- 此外，他们也使用一些有专利性质的数据。



DueDil

数据处理方式

- 他们收集了20年内的经济数据并在他们的网站上公布。

Financials			
Downloadable accounts			
Download the Mar 1999—Mar 2013 accounts for Cooper Steatham Limited in CSV format.			
Download			
Select filing periods			
Key financials	December 2009 ✓	31 December 2010 ✓	31 December 2011
Employees	416441 ↑ +2.55%	406088 ↑ +1.07%	401791 ↑ +4.5
Turnover	62,826,000,000 GBP ↑ +0.44%	64,539,000,000 GBP ↑ +0.97%	63,916,000,000 GBP ↑ +4.9
Cost of Sales	50,737,000,000 GBP ↑ +2.46%	59,278,000,000 GBP ↑ +1.30%	58,519,000,000 GBP ↑ +4.7
Gross Profit	12,089,000,000 GBP ↑ +22.28%	5,261,000,000 GBP ↑ +2.52%	5,397,000,000 GBP ↑ +6.6
Operating Profit	2,188,000,000 GBP ↑ +45.09%	3,985,000,000 GBP ↑ +4.71%	4,182,000,000 GBP ↑ +9.7
Pre-Tax Profit	1,866,000,000 GBP ↑ +48.89%	5,836,000,000 GBP ↑ +5.07%	4,038,000,000 GBP ↑ +14



DueDil

为客户提供的服务



- 他们提供关于中小型公司的数据，以此为这些公司带来新的近千亿元的投资。

Learn how DueDil can work for you

Explore some of the ways that you can use DueDil to benefit you and your business.

Get started 

Introduction to DueDil



Start tutorial 

Find new customers



Start tutorial 

Conduct a credit check



Start tutorial 

Scope out the competition



Start tutorial 

Conduct a compliance check



Start tutorial 



DueDil

盈利模式

- 他们利用开放数据来填补中小型公司与投资者间信息的空白。
- 除了这些信息之外，他们提供的分析和硬件上优势也为他们带来了收益。
- 此外，他们也为那些希望能吸引到潜在的商业合作伙伴并稳定自身根基的中小型企业提供了展示的平台。



DueDil

声誉



- 该公司在一些针对初创公司颁发的奖项中受到提名，例如：
 1. 在2012年获得数据创新保卫者奖两项提名；
 2. 在2012年获得国家商务奖橙色创新奖最终提名。



CarbonCulture

公司背景介绍



Join in, team up, make a difference.

- 成立于2009年的CarbonCulture是一家数据初创公司。
- “公司的创始人兼主管Luke Nicholson是一个在设计交流与可持续发展创新方面有一定经验的社会企业家。”
- 现在，“该公司队伍由4名全职员工和几名兼职成员组成。此外，该公司也与其他机构、商业伙伴和非政府组织保持合作关系。”



CarbonCulture

数据来源

- 他们的数据是从不同组织的建筑管理系统和自动读表系统中获取的。
- 正如Luke Nicholson所说：“我们可以整合任何系统，包括没有自动水电煤等读表系统的建筑大楼。我们把从多个政府部门、地方管理局和大学获取到的开放数据发表出来，并且与一些公司客户合作帮助他们做同样的事情。”



CarbonCulture

数据处理方式



- 正如ODI对该公司的描述：“他们面临着一个全球性挑战——使用数字化技术和优秀设计，使大范围向可持续发展加速转变成为可能——并受此机遇激励。”
- 它们使用高科技测量技术来监测工作环境中的碳使用状况。



CarbonCulture

为客户提供的服务



- 它利用数据帮助客户更好地做出能源和可持续发展的决策，并帮助他们节省资金。
- 他们同样强调设计和使用界面，让同一机构的人们能够更好地沟通，使员工们对可持续发展有一个共同的理解。



CarbonCulture

盈利模式

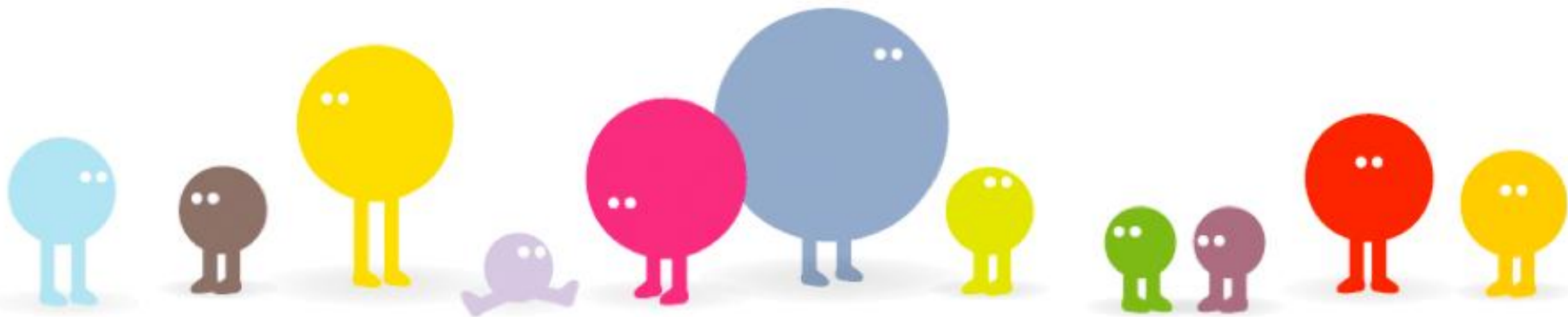


- 监测并报告客户的低碳和节能表现，在互联网和客户的局域网上将其公开。
- 为客户开发APP，帮助他们合理安排能源使用。

+ CarbonCulture

声誉

- 该公司达到的比预期更高的能源节省量——减少了40%的员工的占用率和10%的天然气消耗。
- 现有7家政府机构在使用CarbonCulture提供的服务。





上海大数据局

Big Data Bureau



- 针对新一代信息技术的重要新兴领域——大数据，上海市经济和信息化委员会（上海市经信委）已经在研究成立大数据局，成立后将推进上海政府层面的数据公开和信息共享，通过大数据挖掘来更好发挥政府管理职能。
- 值得关注的是，近日，上海市公共信用信息服务平台正式面向社会开通运行。这是全国覆盖信息主体最多、提供数据单位最多的公共征信系统。



台湾开放数据联盟

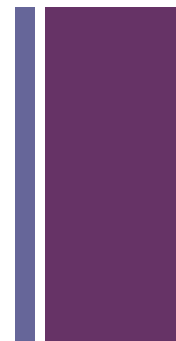
Taiwan's Open Data Alliance





台湾开放数据联盟

Taiwan's Open Data Alliance



- 英国开放数据协会（ODI）和台湾开放数据联盟（ODA）于2013年12月11日签署了合作协议。
- 他们将在诸多领域进行合作：
 1. 交流技术和经验；
 2. 开展合作项目；
 3. 开放相互支持的合作系统来驱动开放数据领域商业的发展；
 4. 共同开发开放数据技术。



Q & A

Email: contact@supstat.com

幻灯片会发布在

<http://blog.supstat.com/cn>