

Center for Entrepreneurship & Technology  
University of California, Berkeley

Technical Brief

Exhaust Data: Credit Score 2.0

Engineering Leadership Professional Program

Number: 20xx.n.v.1.0  
Revision Date: April 21, 2015

Author-Reviewer Names  
(Aleksey Yanovich, Jon Gibbons and Francisco Juarez)

This paper was created in an open classroom environment as part of the Engineering Leadership Professional Program (ELPP) developed and led by Prof. Ikhtlaq Sidhu at UC Berkeley. There should be no proprietary information contained in this paper. No information contained in this paper is intended to affect or influence public relations with any firm affiliated with any of the authors. The views represented are those of the authors alone and do not reflect those of the University of California Berkeley.

Introduction: Exhaust Data ..... 3

Exhaust data as part of big data eco-system ..... 3

Exhaust data Collection ..... 4

    Data Brokers ..... 4

    Data Broker Products ..... 5

    Company Profiles ..... 6

        I. Experian ..... 6

        II. Acxiom ..... 7

        III. Spokeo ..... 9

    Regulation..... 10

    Data Collection Summary ..... 11

Emerging Players ..... 11

    Kred, Klout, PeerIndex..... 11

    TrustCloud..... 12

Emerging Opportunity - Credit Score 2.0 ..... 14

Credit Score 2.0 Market & Future ..... 15

References ..... 16

About UC Berkeley Center for Entrepreneurship & Technology. 17

## INTRODUCTION: EXHAUST DATA

The brief will focus on how big data has revolutionized the collection and analysis of exhaust data, how personal data has become a commodity sold by businesses for the purposes of targeted advertising, marketing, and customer qualification, and the business opportunity to create a new transparent credit score for the sharing economy, Credit Score 2.0, based on exhaust data.

Exhaust data is the electronic trail of information you leave as you go about your day using mobile devices, credit cards, reward cards, surveys and online activities. Data in the form of cookies, selections or preferences leave a trail preferences that are valuable. Exhaust data is not restricted to electronic data but could include data from rewards cards, paper surveys, mortgages or loan applications and credit reports.

Data is growing at an incredible rate and exhaust data is growing proportionally. In 2014 Intel<sup>1</sup> predicted that by 2017, mobile traffic will have grown 13 times in just five years, and by 2017 there will be 3 times more connected devices than people on earth. All digital data created 4 zettabytes in 2013<sup>2</sup> (that is 2 to the 70<sup>th</sup> power or 10<sup>21</sup> bytes). Most of the data is unstructured but all this electronic data leaves a digital fingerprint.

## EXHAUST DATA AS PART OF BIG DATA ECO-SYSTEM

The exhaust data eco-system consists of data collectors, data storage and processors, data analysis and data sales. Data collectors gather our exhaust data with or without our knowledge or permission. The data is stored and processed using software such as Apache Hadoop which is an open-source software framework written in Java for distributed storage and distributed processing of very large data sets<sup>3</sup>. Data analysis includes sorting, analyzing and packaging the data for specific purposes or end customers. Data sales is just that, sales to end customers inside or outside the company that collects, stores and analyses the information.

Data can be purchased from data brokers. Data can be used internally to help target customers or improve some business process, or data can be sold to an outside customer. For example, CVS Caremark and Walgreens sold data about its pharmacy customers, such as patient names, contact information, doctor's names and

prescription details, to data brokers who then sold this information to pharmaceutical vendors such as Eli Lilly and Co., Merck, AstraZeneca, Bayer and other drug manufactures<sup>4</sup>.

## EXHAUST DATA COLLECTION

Every company who has a service or a product collects data from its customers. This is not new, but what has changed is the quantity and quality of customer specific data being generated and the emergence of customer data being a product that can be bought and sold.

Data collectors include traditional companies such as credit rating agencies like Experian, Equifax and TransUnion. Marketing agencies like Acxiom and Epsilon Business information and news agencies such as Bloomberg and Dun & Bradstreet. Retailers like Walmart, Target, and Walgreens. New players include Google, Facebook, and Amazon use offline data and profile information to target ads and market products. Telecommunication companies such as AT&T, Verizon, Comcast and BT use mobile data that adds GPS location to the exhaust data being generated and stored. Agencies such as the NSA and Department of Homeland Security collect similar information for security reasons.

## DATA BROKERS

In 2012 the FCC defined a Data Broker as “companies that collect information, including personal information about consumers, from a wide variety of sources for the purpose of reselling such information to their customers for various purposes, including verifying an individual’s identity, differentiating records, marketing products, and preventing financial fraud.”<sup>6</sup> The number of companies and size of the industry is not clear but estimates indicate that this is multi-billion dollar industry consisting of hundreds of companies<sup>7</sup>.

As previously noted the industry includes credit reporting agencies, digital marketing companies, retailers, and various online media companies. Data brokers collect traditional offline data and combine it with online data to create detailed customer profiles. Customer profiles are used to predict if an individual is a credit risk or is likely to be a customer for a high end SUV.

In a 2013 US Senate, Staff Report for Chairman Rockefeller<sup>5</sup> it states the following:

*Acxiom alone has “multi-sourced insight into approximately 700 million consumers worldwide,”<sup>54</sup> and Datalogix asserts its data “includes almost every U.S. household.”<sup>55</sup> Some of the companies maintain thousands of data points on individual consumers, with one providing the Committee a list of approximately 75,000 individual data elements that are in its system.<sup>56</sup> Data*

*collected by these companies includes detailed and personal information including data on consumers' health and financial status.*

The 2013 US Senate, Staff Report for Chairman Rockefeller<sup>5</sup> concluded the following:

- Data brokers collect a huge volume of detailed information on hundreds of millions of consumers.
- Data brokers sell products that identify financially vulnerable consumers.
- Data broker products provide information about consumer offline behavior to tailor online outreach by marketers.
- Data brokers operate behind a veil of secrecy

## DATA BROKER PRODUCTS

Data brokers are in the business of collecting and analyzing consumer data to provide their customers insight into groups or individual consumer's characteristics and predicted behaviors. Customer profiles consist of "actual" or "modeled" elements. Actual data could include date of birth, address, marital status, Etc... Modeled elements could be something like "an interest in a sport based on a magazine subscription"<sup>8</sup>. Customer data is provided in the form of a list of individuals who may be interested in a project or service, or specific customer information to augment information that a company may already have on an individual. The consumer databases are provided with analytic tools, and multi-platform tracking tools to enable data broker customers to advertise and market to their potential customers.

Data broker products are known to characterize a consumer's economic status. According to the Rockefeller Senate Report. Consumers who are financially stable may be characterized as "Power Couples", "American Royalty, or "Just Sailing Along"<sup>5</sup>. The Rockefeller Report also identified Data Broker profiles that targeted financially vulnerable populations as shown in the Figure I below<sup>5</sup>.

Sample List of Targeting Products Identifying Financially Vulnerable Populations			
"Burdened by Debt: Singles"	"Struggling Elders: Singles"	"Meager Metro Means"	"Very Elderly"
"Mid-Life Strugglers: Families"	"Retiring on Empty: Singles"	"Relying on Aid: Retired Singles"	"Rolling the Dice"
"Resilient Renters"	"Tough Start: Young Single Parents"	"Rough Retirement: Small Town and Rural Seniors"	"Fragile Families"
"Very Spartan"	"Living on Loans: Young Urban Single Parents"	"Financial Challenges"	"Small Town Shallow Pockets"
"X-tra Needy"	"Credit Crunched: City Families"	"Credit Reliant"	"Ethnic Second-City Strugglers"
"Zero Mobility"		"Rocky Road"	"Rural and Barely Making It"
"Hard Times"			
"Enduring Hardships"			
"Humble Beginnings"			

**I. Vulnerable Population Profiles**

**COMPANY PROFILES**

It is worth taking a quick look at three companies identified as Data Brokers. Experian and Acxiom are well established and Spokeo is an emerging player.

**I. EXPERIAN**

Experian is a \$4.8B corporation headquartered in Dublin Ireland. Most consumers would identify Experian as a credit reporting agency, but Experian is really an information services company and operates in four segments: Credit Services, Consumer Services, Marketing Services, and Decision Analytics. Experian’s business customers include financial services, automotive, insurance, media, telecoms, healthcare, education, and others.

As part of the Marketing Services business Experian has a suite of products that provide detailed consumer profiles, analytics, and tools that allows Experian customers to track their customer across multiple platforms<sup>8</sup>. See Figure II



## II. Experian Marketing Services Products<sup>14</sup>

In a 2013 US Senate, Staff Report for Chairman Rockefeller<sup>5</sup> it states the following:

*Experian manages a database open to catalog sellers as well as brick and mortar and e-commerce retailers. Participants provide customer transactional records, which may include consumer's name, address, gender, e-mail address, phone number, channel of purchase (e.g. online or in-store), dollar amount, payment method, transaction date, and transaction product category.<sup>85</sup> Experian summarizes the information to describe buying behaviors at the household level within general product categories – such as “Kitchen and Tabletop,” “Books,” or “Vitamins/Health Products.”<sup>86</sup>*

## II. ACXIOM

Acxiom is \$1.1B Corporation headquartered in Little Rock Arkansas. The company operates as an enterprise data, analytics, and software-as-a-service company in the United States, Europe, South America, and the Asia-Pacific. The company operates in three segments: Marketing and Data Services, IT Infrastructure Management, and

Other Services<sup>12</sup>. Similar to Experian Acxiom provides consumer profiling, analytics, and online platform tracking tools that allow its customers to target market to consumers. Its products include:

- PersoniX: Consumer Segmentation
- AbiliTec: Customer matching & identify resolution
- Audience Operating System (AOS): Device platform tracking

Acxiom maintains databases for approximately 700 million individual's world wide<sup>5</sup>. Its consumer profiles consist of around 1,500 data points<sup>10</sup>. The PersoniX classification systems assigns consumers to one of 70 socioeconomic clusters<sup>10</sup>. An example is given in Figure III<sup>13</sup>. Acxiom also maintains the consumer database for 47 of the top fortune 500 companies<sup>10</sup>. Individuals can see some of the data that is being held by Acxiom at [www.Aboutthedata.com](http://www.Aboutthedata.com). After entering information to identify yourself, e.g. social security number, individuals can view and correct information on file.

Like Experian and Eqifax Acxiom admitted to the Senate Commerce Committee panel in 2013 that it passed on sensitive information of individuals in the "X-tra needy" category. The information was un-specified but included income, credit denials, and if individuals had been victims of crimes<sup>10</sup>.

## GROUP 6X

### MIXED SINGLES

GROUP	6X	MIXED SINGLES	HOUSEHOLDS: 3,151,700	(2.55% OF U.S.)
CLUSTERS	61	Urban Scramble		
	69	Pennywise Mortgagees		
	70	Resilient Renters		



#### ABOUT MIXED SINGLES

The three clusters of Group 6X, "Mixed Singles," struggle with some of the lowest incomes and often little more than minimal savings or investments; all occupy the lowest rungs of the socio-economic ladder. It should be noted, however, that all three have high concentrations of students when compared to national averages. All depend on low echelon, white-collar and blue-collar jobs, often part-time. Members of this life stage group have no children. This group shows strong urbanicity, two clusters extremely urban and one moderately dense.

#### MIXED SINGLES, A DAY IN THE LIFE

*Name of protagonist: "Malcolm"*

*Wakes up...* warms up some Entenmann's cake for breakfast and washes it down with skim milk.

*Spends the day...* looking for a new job on the Internet, and listening to Urban radio, missing his first class at Phillips Community College. He hates the instructor and figures that he was going to drop that class anyway.

*Talks about weekend plans to...* hang out with his buddies, checking out the latest dance club and maybe go to the opening weekend of the latest movie.

*Has a meeting with...* his financial aid counselor to find out if he could still receive financial aid if he drops the course.

*Spends the evening...* working the late shift in the grocery store. It's a boring job, but the two M.O.D. (Managers on Duty) are okay to work for. He's a checker for part of the evening, then a stocker. He gets off work at 2:00 a.m.

*Goes to bed at...* 3:00 a.m., watching...*Family Guy.*

#### WHEN THEY GREW UP...

- Oprah Winfrey debuts
- *Back to the Future* is released
- *Miami Vice*
- Bobby Brown releases *My Prerogative*

### III. Personix Profile

## III. SPOKEO

Spokeo is a privately held company based in Pasadena CA. Their product is a web based search platform that enables individuals and business to search people online. Enter a name into their search engine and it will identify potential matches with information like date of birth, previous homes, and family members to narrow the search. Pay a small fee and the site will provide aggregated online and offline information. See Figure IV.

Results Are Ready!

latest report as of 03/26/2015

**Ikhlaq Sidhu**  
Age 40's, Danville, CA

**Full Results May Include**

- ✓ Full Address
- ✓ Email Address
- ✓ Family Members
- ✓ Month/Year of Birth
- ✓ Phone Number
- ✓ Location History

**Get Instant Information on Ikhlaq Sidhu Now!**

#### IV. Spokeo Search Hit

Spokeo has had issues with the FTC. In 2012 It paid a \$800,000 in civil penalties after it was accused of selling personal profiles to third parties in violation of the Fair Credit Report Act (FCRA)<sup>10</sup>.

## REGULATION

Data collection technology has raced ahead of regulation. In the U.S. there is a mixture of regulations at the Federal and State level. For example the Fair Credit Reporting Act, FCRA regulates the reporting financial transactions. The Video Privacy Protection Act of 1988 regulates the reporting of video rentals as a result of the Robert Bork Supreme Court nomination hearings<sup>15</sup>. However the collection and usage of vast quantify of offline and online data is not regulated. This includes data captured by browsers, social media posts, location data provide to applications, retail purchases, home purchases, Etc...

There are several question with respect to the ownership of data. For example do you give up your privacy rights when you agree to a multipage click through agreement or loyalty program? Do you own your image if it is posted by another individual on Social Media? Can an individual get a report of what data is being held by an organization and how it is being used?

The U.S. Federal and State regulators are just beginning to debate the regulation of Exhaust Data. The EU has been ahead of the U.S. with respect to regulation. The EU has passed regulations on the “Right to be forgotten” and “The Right to Know”. The “Right to be forgotten” is targeted at Google to remove information from search results if it is requested by an individual. It does not regulate the data sources. The “Right to Know” regulations allow individuals to know what data is being held by companies. What is not clear is how effective regulations will be given changing technology and that data is being held across international boundaries.

## DATA COLLECTION SUMMARY

The Data Collection and Broker business is a multi-billion dollar industry that is collecting vast quantiles of consumer data. A few key observations:

- Online and Offline data is being sold & merged to target market to consumers
- Data brokers create detailed “profiles” to target consumers
- Consumer have limited visibility of their online “profile”
  - No transparency
  - May be incorrect
- Regulation has not caught up with the technology

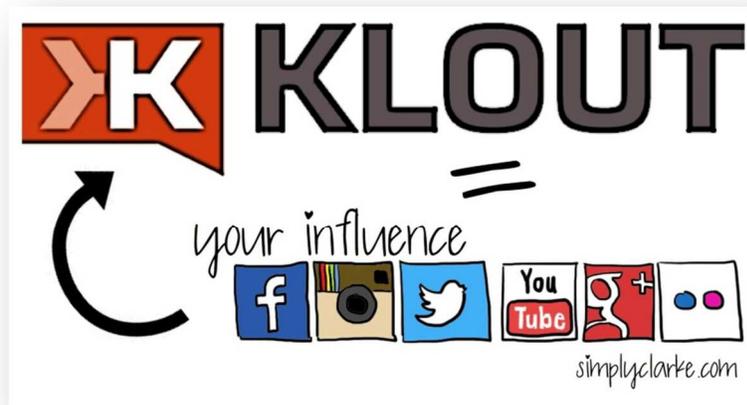
This may sound harmless if the only downside it miss targeted advertising or paying a different price for a product, but an incorrect profile data or analysis could be used to set insurance rates, deny access to a service, or impact employment. The industry and regulation needs to provide individual control, transparency, security, and access to accuracy.

## EMERGING PLAYERS

Now after we looked at existing companies in the market analyzing personal data exhaust to build personal profile, let's talk about emerging players, what they do and how this may enable new business opportunity and potentially build new ecosystem and bring new players around such well-known and established terms as credit score and credit report. We will look at several companies who are using data exhaust to build personal profile.

## KRED, KLOUT, PEERINDEX

There is a group of companies Kred (<http://kred.com>), Klout (<http://www.klout.com>), PeerIndex (<http://www.peereindex.com>, recently acquired by BrandWatch) which are measuring personal influence score based on activity of the person in social media. Once influence metrics is built - those companies help to connect "influencers" with other companies who are interested to advertise their new product/service or increase awareness on specific topic in social media.



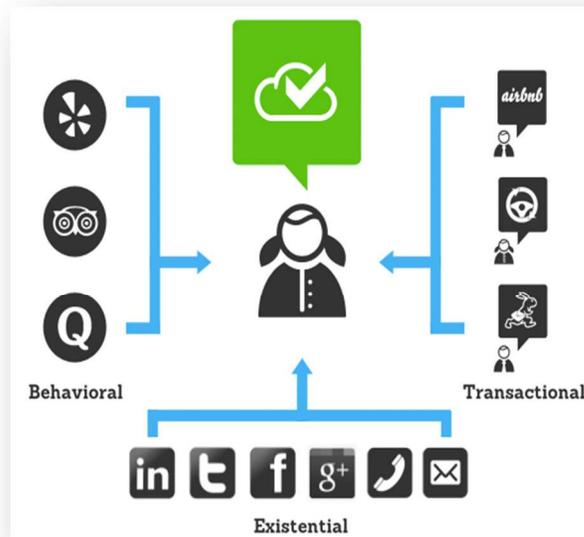
While this does not sound like personal profile or worthiness metrics – this is an example of an emerging personal metric in the world of IoT and “sharing economy”.

## TRUSTCLOUD

Another company building personal profile is TrustCloud. TrustCloud is a credit platform for Peer Marketplaces that measures virtuous behavior online to build a portable and contextual trust score you can use anywhere.



TrustCloud collects your publicly-available data from networks like Facebook, Google and LinkedIn, then analyzes it and assigns your trustworthiness a ranking between 1 and 1000 and builds 3 metrics out of it.

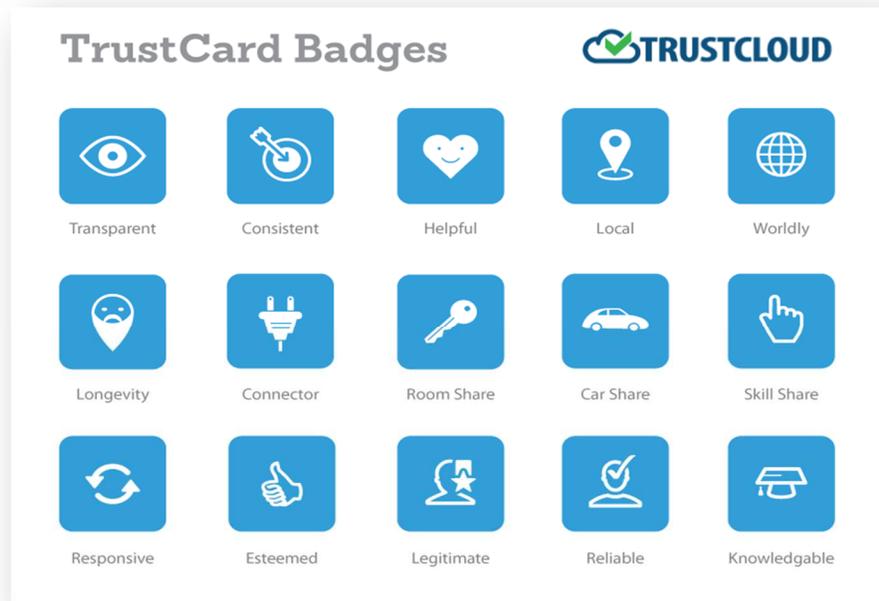


A TrustScore which measures trustworthiness on a scale of 1 to 1000, with 1000 being the most trustworthy. The score represents a composite of universal trustworthiness “virtues” like consistency, generosity, and transparency. In essence this is very similar to regular credit score except the source of information, but represents mainly your worthiness in online/sharing world.

TrustCard is an online widget that shows your profile picture, name, TrustScore, data sources, and the TrustCard Badges you’ve earned. The card can be embedded almost anywhere to let others know that you’re trustworthy: across participating sharing site user profiles, personal identity sites, blogs — even your email signature.



A TrustCard Badge represent trustworthiness in different contexts. Some badges highlight behaviors like consistency, generosity and transparency. Other Badges represent specific areas such as room-sharing, car-sharing, or task-sharing.



Overall as you can see what TrustCloud is providing is "online" version of credit score and credit report built based on your personal data exhaust. The TrustScore along with TrustCard/TrustBadges are relevant in context of new "sharing" economy and online services which most of us are using in everyday life – renting the apartment through such services as AirBnB and HomeAway, sharing your car for a ride or selling something on eBay.

## EMERGING OPPORTUNITY - CREDIT SCORE 2.0

Now let's talk about market opportunity which we see in this sector. We think that there is a need in new kind of credit score - "Credit Score 2.0". Credit Score 2.0 should be less sensitive and confidential compared to classic credit report to be used widely online. It should be more relevant in context of social networks, online marketplaces and sharing services which people use - the fact that you have no debt and high salary does not mean you're careful apartment renter or that you accept returns once you sell something on eBay.

In a simple way Credit Score 2.0" is similar to current seller/buyer rating on eBay. The problem with the eBay rating is you cannot easily benefit from it within other online services and it is limited to reflect your trustworthiness as seller/buyer of goods on the internet, but tells nothing about your apartment rental history or car sharing experience. The Credit Score 2.0 should fill exactly those gaps to be widely accepted and useful - it should be portable/sharable between online services and should be universal. In this respect data exhaust is the perfect data source for it - the

data exhaust we leave behind has a lot of details and dimensions allowing the ability to make such metrics very powerful and relevant from many perspectives. At the same time most "exhaust" data is public today, so there should be less concerns using it vs. sources of classic credit reports which many people consider confidential.

How to monetize Credit Score 2.0? We think there are many opportunities, but in the simplest form - online services such as AirBnB or HomeAway could pay to Credit Score 2.0 providers (for example, to TrustCloud) to link user profiles with TrustCard, so that renters could see landlord's TrustScore and TrustBadges and could use it while selecting among potential places to rent. Same is true from landlord's perspective - if there are multiple candidates for your apartment, it is useful to have some score helping to decide which candidate is better future tenant.

## CREDIT SCORE 2.0 MARKET & FUTURE

The current market is new and has few established players. We expect more companies to appear in this area to bring newer/better technology and to address specific scenarios and concerns. Emerging players will substitute/complement current established services like Experian and Equifax in many services where "classic" credit report is not relevant. Once monetization is clear there will likely be many mergers and acquisitions done by current established Big Data players.

## REFERENCES

1. Intel infographic. 2014.  
“<http://www.intel.com/content/www/us/en/communications/internet-minute-infographic.html>”. Intel Corp. Communications.
2. 2013 US Digital Future in Focus Whitepaper. February 14, 2013.  
<http://www.comscore.com/Insights/Presentations-and-Whitepapers/2013/2013-US-Digital-Future-in-Focus>.
3. Marudubshinki. 2015. “Apache Hadoop”. Wikipedia.
4. Ridgeway, James. 2011. “Is Your Drugstore Selling Your Private Information to Big Pharam?”. Mother Jones.
5. Office of Oversight and Investigations, Majority Staff. 2013. “A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer data for Marketing Purposes. United States Senate Committee on Commerce, Science, and Transportation.
6. Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change*, at 68 (Mar. 2012) (hereafter “FTC Privacy Report”). These companies may also be referred to as “information resellers.” See Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, GAO-13-663 (Sept. 2013) (hereafter “GAO Information Resellers Report”).
7. GAO Information Resellers Report, *supra* n.1, at 5 (noting: “Several privacy related organizations and websites maintain lists of data brokers—for example, Privacy Rights Clearinghouse lists more than 250 on its website—but none of these lists claim to be comprehensive. The Direct Marketing Association, which represents companies and nonprofits that use and support data-driven marketing, maintains a proprietary membership list, which it says numbers about 2,500 organizations (although that includes retailers and others that typically would not be considered information resellers)”).
8. Acxiom Narrative Response to Senate Commerce Committee (Feb. 15, 2013).
9. Experian Web Site.
10. Digital Exhaust: What Everyone Should Know About Big Data. Digitization, and Digitally Driven Innovation. Dale Neef.
11. TrustCloud (<http://www.trustcloud.com>)
12. Yahoo Finance (<http://finance.yahoo.com/q/pr?s=ACXM+Profile>)
13. Acxiom: Personix\_Binder.pdf  
([http://reference.mapinfo.com/software/anysite\\_segmentation/english/2\\_0\\_1/Personix\\_Binder.pdf](http://reference.mapinfo.com/software/anysite_segmentation/english/2_0_1/Personix_Binder.pdf))
14. Experian: ([http://www.experian.com/marketing-services/consumer-insights.html?cat1=customer-acquisition&cat2=market-research&intcmp=hptn\\_cust\\_acq\\_simm](http://www.experian.com/marketing-services/consumer-insights.html?cat1=customer-acquisition&cat2=market-research&intcmp=hptn_cust_acq_simm))
15. Epic.org; <https://epic.org/privacy/vppa/>

## ABOUT UC BERKELEY CENTER FOR ENTREPRENEURSHIP & TECHNOLOGY

The Center for Entrepreneurship & Technology (CET) seeks to foster entrepreneurship within the University and to bring Berkeley's research capability to industry collaborations. To these ends, the CET hosts multi-disciplinary research projects in collaboration with industry stakeholders; as well as provides mentoring and support to new Berkeley ventures.



National rankings consistently place UC Berkeley's undergraduate and graduate programs among the world's best. Berkeley is home to top scholars in every discipline, accomplished writers and musicians, star athletes, and stellar scientists—all drawn to this public university by its rich opportunities for groundbreaking research, innovative thinking and creativity, and service to society.