



Who Really Owns Your Data?

WWW.LONE-STAR.COM



CIO

Most read:

Home > Business Analytics > Big Data

FEATURE

Who Legally Owns Your Big Data?

Commingling internal and external data can produce problems if your contract isn't clear about who owns what -- and for how long.

InfoWorld

Most Popular:

[Twitter](#) | [LinkedIn](#) | [Facebook](#) | [8+](#) | [RSS](#) | [Email](#) | [Search](#)

TIBCO
Jaspersoft

Cloud BI, Starting <\$1/hr: 5-Minute Demo Video

[EXPAND](#)

[awsmarketplace](#)

Home > Big Data

Big Data mining: Who owns your social network data?

An attractive application of Hadoop and other Big Data technologies is to analyze users' social activities, sometimes without their express knowledge

MORE LIKE THIS

Self-service BI catches on

A conversation with IBM's Mr. Big Data

UBM
Tech
[informationweek.com/healthcare](#)

InformationWeek
Healthcare
SEPTEMBER 2014

Who Owns Patient Data?



Who owns the data in the Internet of Things?

Data are at the heart of the Internet of Things (IoT). Without connected devices being able to capture, transfer, analyse, report and act on data, the benefits of the IoT cannot be achieved.

January 2014

Data are hugely valuable. But who owns them? The answer is: no one – there is no property in a piece of data itself. The owner of a smart thermostat does not, for example, own the data but how he uses it. The only thing that is 'ownable' is an aggregation or collection of such data, provided there has been a relevant investment in carrying out that aggregation or collection (the individual user is very unlikely to have made that investment). It is that investment and who carries it out that are the focus of this article.



Related articles

[There's spam in my fridge! ▶](#)

[Connecting the Internet of Things ▶](#)

Several different ways to think about this...

WWW.LONE-STAR.COM



So...Big Data Brings This Topic Up...but...

- It's not new
- Big Data is not the main reason Lone Star cares
- ***All of our businesses involve blending information from several sources***
- Different markets and technology communities use different terms to discuss the problem of owning, sharing and controlling data
- And..... “Big Data” is making more people aware

Objectives of this presentation

- Make some general observations independent of technology/markets
 - Look at how SIPs & SLURPs promote real data ownership



Why people ask... “Who really owns the data?”

- Can corporations, governments, and others claim to “own” and tightly manage their data
- What about data blended from many sources?
- Within locked boxes of proprietary software?
- Within closed standards?
- What about customers?
- Privacy?
- Regulation?

What does “owning” data mean???



What is “Ownership?”

- Own – to have legal title or right to something
- In particular, Owners have the right to

1. Possess

2. Use

3. Convey

Important – We'll Come Back to These

- But these ideas are complicated
 - “Possession” alone is not ownership
 - Ownership can be shared, even distributed



Some Complications

- Ownership of different things can have unique shades of meaning (mineral rights vs. surface rights, spectrum ownership, ownership of restricted technology...)
- Shared ownership; common & often awkward
 - Example – Land owner with surface rights can have a say about how mineral rights are exercised – which varies depending on national laws
- Owning data can be very complicated
 - In particular shared ownership – more on that later
- International data ownership is even more complicated
 - Legal systems, such as the EU and US have somewhat different principles of data ownership





Limits to Unfettered Ownership Privacy & Proprietary Concerns

- Privacy; HIPPA and other privacy laws limit the extent to which information can be shared
 - Transfer of data by an owner must not violate privacy laws and regulations
- Governmental limits on release (or governments can require release)
- Proprietary limits on public statements; most B2B purchase agreements state neither side can make a public announcement about a transaction without prior approval
- Internet of Things data flows and exchanges between dozens of vendors and millions of individuals.
- All Essentially the same problem for “ownership” – Ownership is “shared” limiting the options and authority of the data aggregator



- If data resides in a “locked box”
 - Fees to hold it
 - Fees to manipulate it
 - Fees to format for export
 - Workers using the data must pay to acquire certifications for the locked box
 - Recipient/Client must also pay fees
 - Box vendor may even mine “my” data
- Who really owns this data?
 - Roughly like your stock broker who makes money regardless of how your investments perform
 - Or like a toll booth... which usually implies **ownership by whoever owns the toll gate**

Limits on Unfettered Ownership Proprietary Locked Box

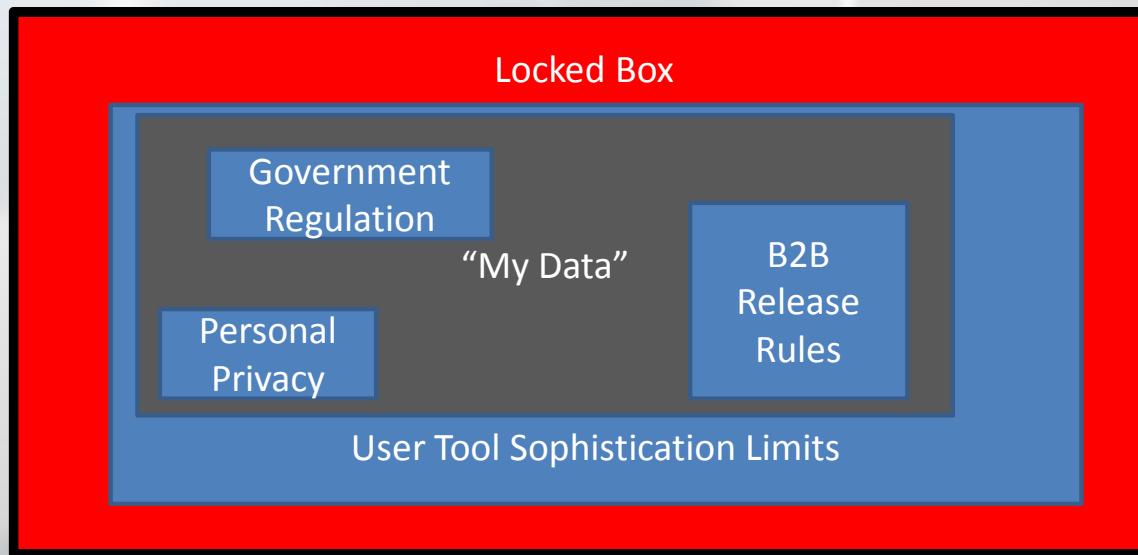


The Dulwich Toll Both, South London

The Dulwich Estate charges a toll to pass along this road, the main artery through its land, in part to control commuter congestion – tolls slow commerce and reduce traffic... perhaps not what we want for our data

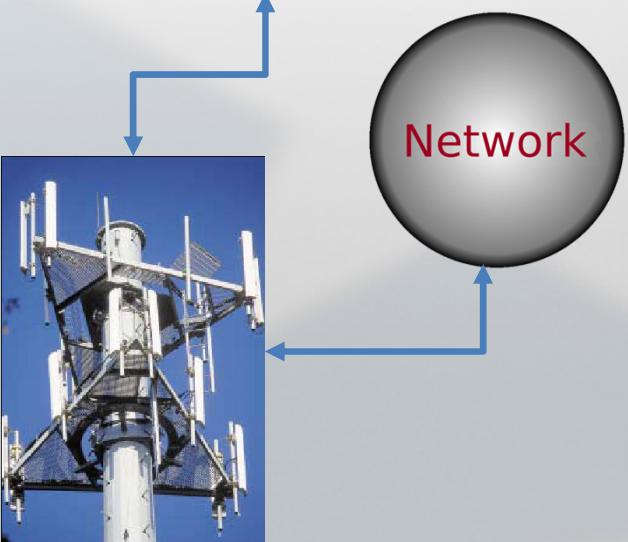


Data Co-Ownership?



If “My Data” can be held, used, transferred, assigned, or shared only with the cooperation of several other parties

.... Then... this construct is NOT ownership – maybe not even “co-ownership”



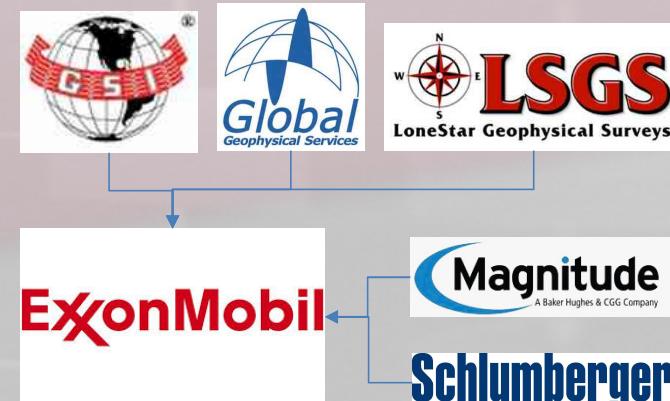
Why does an iPhone (or Samsung) work with an Ericsson base station, connected to an Alcatel/Lucent network core?

Why can the USAF launch a missile designed by Hughes, built by Raytheon, from an aircraft built by Fairchild and rebuilt by Boeing?

How can several firms collect seismic data over a period of years, which is in turn used by another firm for exploration, and then be blended with well logs to guide development and production?

Answer – data ownership management is not unique – focus on critical interfaces it's how complex systems owners/operators avoid being hijacked

Is this Unique???





Complex Systems.... What Interfaces Need Managing?

- Complex Systems (and simulations) require interface management across the entire system
- Two Perspectives OSI and DHE help illustrate where particular attention is needed
 - OSI defines communications systems (just one kind of complex system)
 - Deaton-Hubka-Eder scale (DHE scale) is generic – applies to any product vs. Open System Interconnection(OSI) Layers
- The corners of this matrix help show high risk areas

	DHE 1 Components Parts	DHE 2 Sub-Assy Mechanism	DHE 3 Apparatus Functional Device	DHE 4 System of 3's	DHE 5 System of Systems
OSI 1 Physical Layer	x	x	x		
OSI 2 Data Link	x	x	x		
OSI 3 Network			x	x	
OSI 4 Transport				x	x
OSI 5 Session				x	x
OSI 6 Presentation				x	x
OSI 7 Application					x

Opportunities for Mischief

- **Can't Connect**
- **Can't Operate**



Really Owning Data Step 1 Open Interfaces



Telecom network operators are vigilant to maintain open interface standards for “air interfaces” and for “core network” connections

MIL-STDs (like 1760) keep the interfaces between platforms and payloads “open”

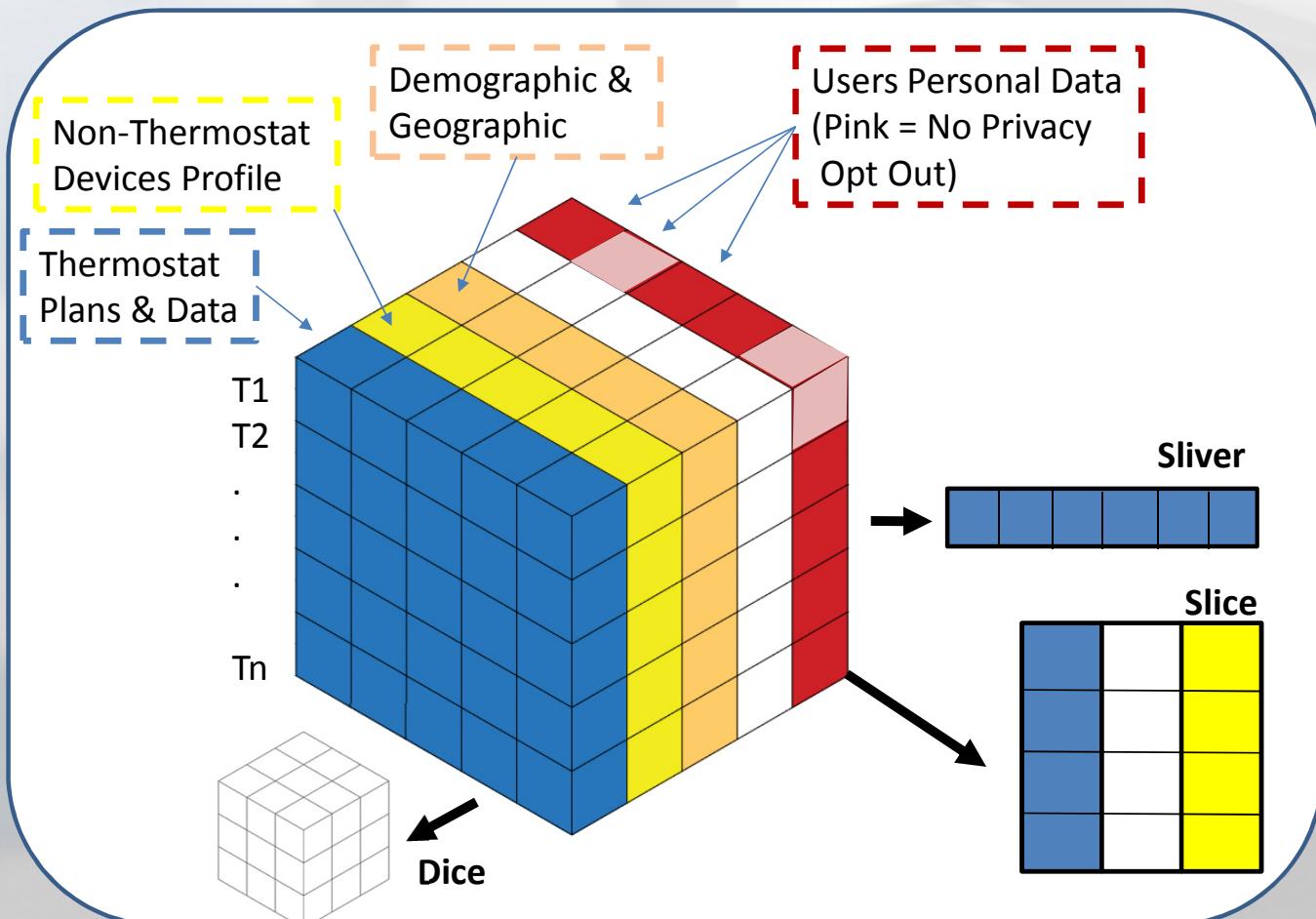
Standards like LAS and SEG Y promotes blending data from multiple surveys, well logs...





Example – Is this my data?

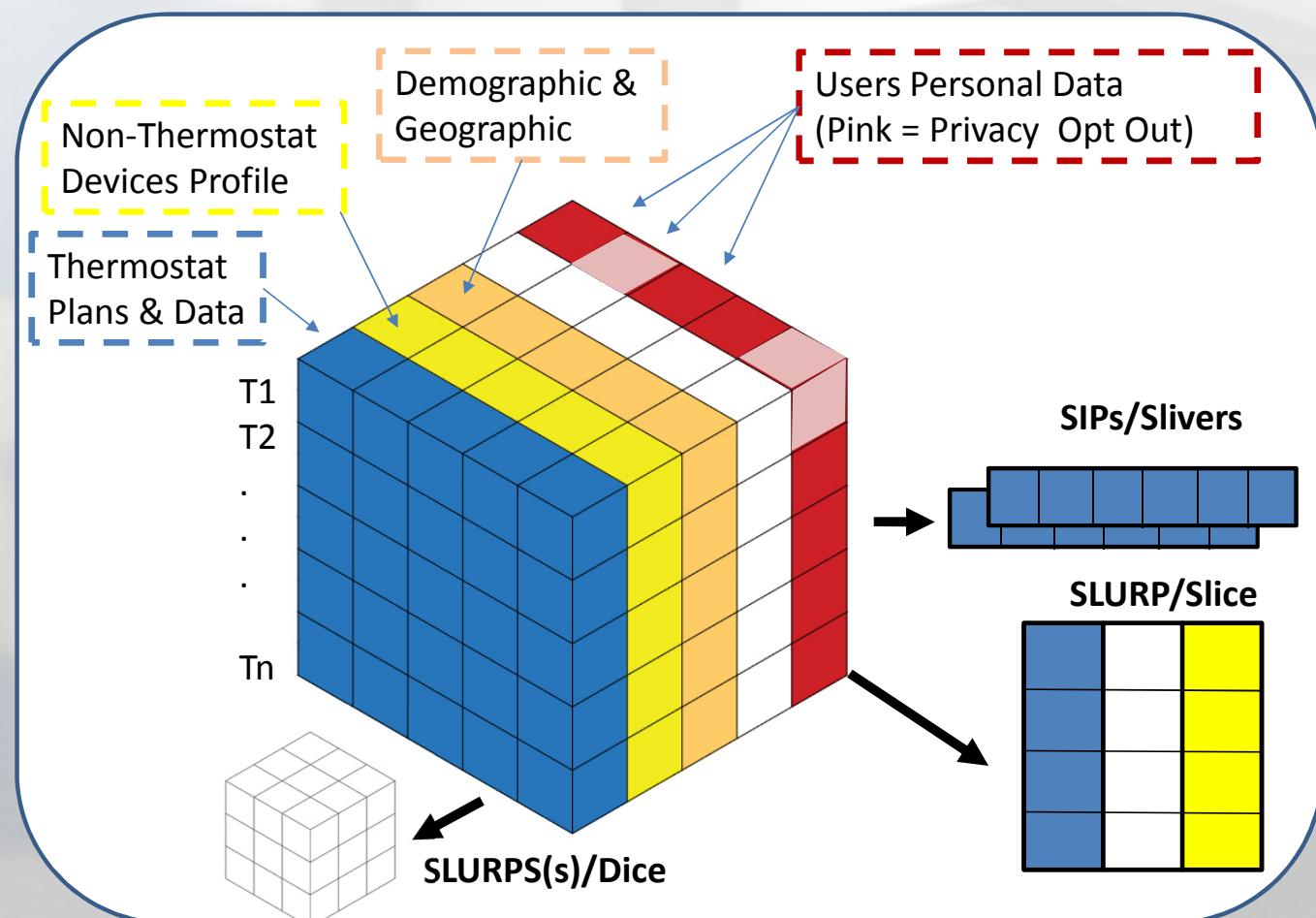
- Imagine an Internet of Things (IoT) data warehouse
- In this case, smart home thermostat information
 - Programmed plans
 - Heating & Cooling data
 - With demographics and geographic
 - Information from other systems, like lighting
- Our query engine can create **slivers**, **slices** and **dice(s)**
 - But...In our thought experiment, the ever present legal department has some warnings
 - Can we really just convey the slivers, slices and dice?
 - Is this my data?





Example – Set My Data Free!!

- Let's sell data, to demonstrate our ownership and ability to take independent action
 - Three buyers
- A/C use profiles for the Gas Trader – two SIPs
 - One SIP is the average duty cycle for each unit of time.
 - One SIP is the ratio of units operating at 100% duty cycle
- The gas company trading desk buys a slice of data
 - We delete the electric company clients in accordance with the rules
- We sell the mobile phone company a Die/SLURP
- Why do we need SIPMath 2.0™?**





Probability ManagementSM

Standard Specification
for
Stochastic Information Packets (SIPs)
and
Stochastic Library Units with Relationships
Preserved (SLURPs)
Version 2.0

APPROVED		
Title	Signature, Date	
Chair, Standards Committee	 1 October 2014	
CHANGE HISTORY		
Revision	Date	Change Highlights
Version 2.0	2014-10-01	Initial Published Standard

Copyright © ProbabilityManagement.org, 2014. All rights Reserved
Version 2.0: 2014-10-01
Page 1 of 23

Why Does SIPMath™ Set My Data Free?

- The Slivers, Slices and Dice can be created many ways, SQL, Oracle, Hadoop...
 - They provide part of what we need to exercise ownership rights
- SIPMath 2.0 provides
 - Royalty Free Open Standard
 - Barrier free interface; XLSX, CSV, and XML
 - Provenance; easy to sign/certify and specify the rules which were obeyed
 - Provenance which travels with the data product
 - Privacy on both sides of the interface
- Goal = Minimum Barriers, Maximum Protections



Really Owning Data - Four Steps

- **Step 1** is Standards based Open Interfaces
- **Step 2** is producing data products which comply with contracts, regulations and laws
- **Step 3** is production of the products in a manner which segregates intellectual property
- **Step 4** is embedding documentation of the compliance within the data product



Summary

- SIPMath allows us truly own our data
 1. Possess
 2. Use
 3. Convey
- SIPMath 2.0 makes it easy to
 - Use Standards-based Open Interfaces
 - Producing data products which comply with contracts, regulations and laws
 - Segregate intellectual property while using the open interface
 - Embed compliance documentation within the data product



Sources & Credits

Photos and images are the product of Lone Star, Probability Management, or were obtained from open sources for license free reuse, or are screen shots used under Fair Use doctrine as commentary. Logos and trademarks are the property of their respective owners. Probability Management, SIP, SLURP and SIPMATH are trademarked by Probability Management.

The Deaton-Hubka-Ede (DHE) scale is from unpublished work of a colleague Jennifer Deaton. She based her work on extending the earlier systems complexity studies of Hubka and Ede. Jennifer worked with Dr. Jerry Krasner of the EMF company. The DHE scale and research behind it will be published in the future.

The SIPMath 2.0 specification can be downloaded free from the ProbabilityManagement.org website.

Other standards are available from their respective sponsoring organizations.

Most legal definitions were taken from Black's Law Dictionary, 10th edition