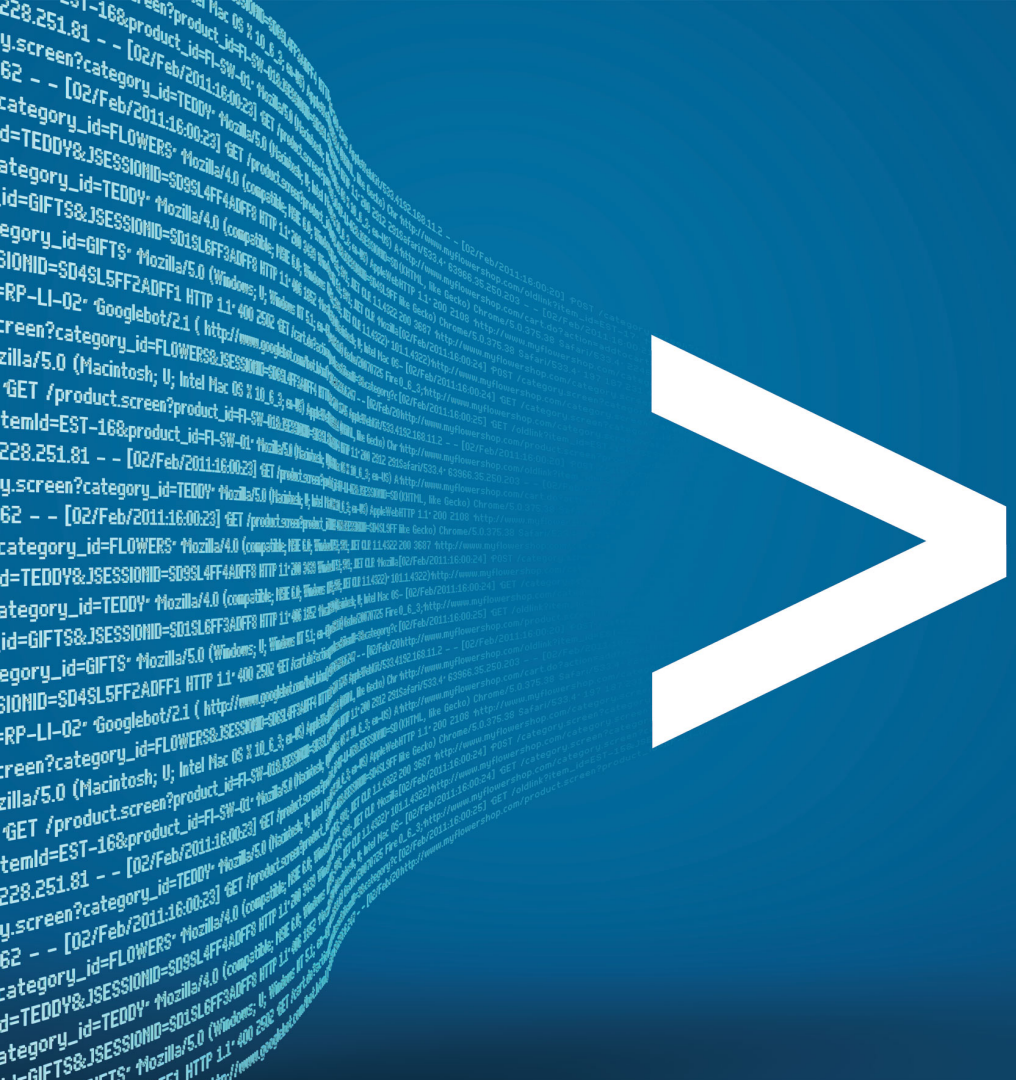
Copyright © 2014 Splunk Inc.



A brief history of data

Damien Dallimore
Worldwide Developer Evangelist @ Splunk

splunk>

Who Am I

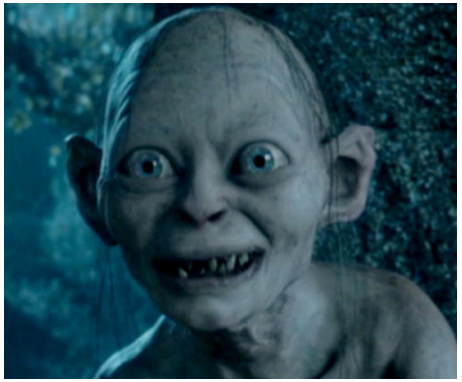
Worldwide Developer Evangelist @ Splunk

I code
I talk about coding

Community, Collaboration, Open

splunk> listen to your data™

From Aotearoa (New Zealand)



splunk> listen to your data™

Where did the “BIG DATA” come from

splunk> listen to your data™

Let’s go on a journey

splunk> listen to your data™

20,000 BC Arithmetic

We start to count things

splunk> listen to your data™

15,000 BC Cave Painting

We start to record data visually

splunk> listen to your data™

3,500 BC Written Language

We start to record and “transmit” knowledge

splunk> listen to your data™

2,500 BC Sumerian Calendar

We start to organize and track time

splunk> listen to your data™

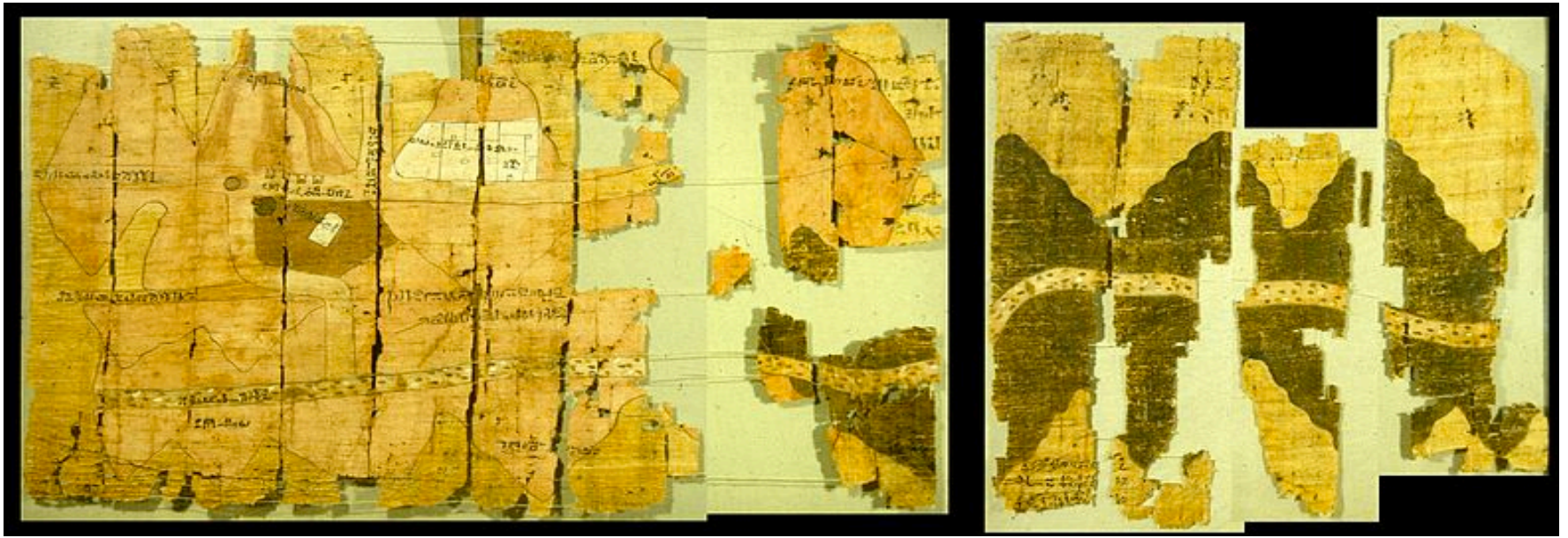
1,250 BC Library at Thebes

We start to store data in mass

splunk> listen to your data™

1,150 BC Egyptian Maps

The origin of Google Maps



splunk> listen to your data™

1000 BC Era Math, Computation, Logic

We start to develop understanding through numbers

500 BC Pythagoras
300 BC Euclid

And how numbers can be used to compute data

250 BC Archimedes
100 BC Antikythera Mechanism

We start to classify objects and use logic to derive insights

350 BC Aristotle

splunk> listen to your data™

0 - 1600

78 Pliny : “all” the world’s knowledge captured
105 Paper : bulk recording of data
340 Codices : making data browseable with sections and indexes
1350 Nicole Oresme : turning data into picture
1453 Guttenberg : mass distribution of data

splunk> listen to your data™

1600 - 1900

1640 Napier : before logs there were logarithms
1662 Graunt : father of statistics
1796 Watt : recording data with a machine
1801 Jacquard : programming !!
1830 Babbage : the first mechanical and programmable computer
1844 Morse : data encodings
1850 Reuter : first “WAN” (the CSMA/CD was a bit messy)
1876 Dewey : data classification

splunk> listen to your data™

1900 - 2000

1930's Fisher : modern statistics
1936 Turing : the universal computer
1950's Programming Languages : Fortran et al
1962 Tomlinson : first standard for Geo Data
1963 ASCII : a standard for representing letters and numbers
1969 ARPANET and other Protocols
1970's RDBMS : ETL , BI , Data Warehouses , I am your father
1970's/ 80's Personal Computing : the foundations for alot of today's data
1982 TCP/IP standardized and the Internet came to be
1989 The Web and HTML blink tags
1991 Unicode : all languages captured

splunk> listen to your data™

To infinity and beyond

Web 2.0

Google

Social Networking

IOT

splunk> listen to your data™

Data Today ?

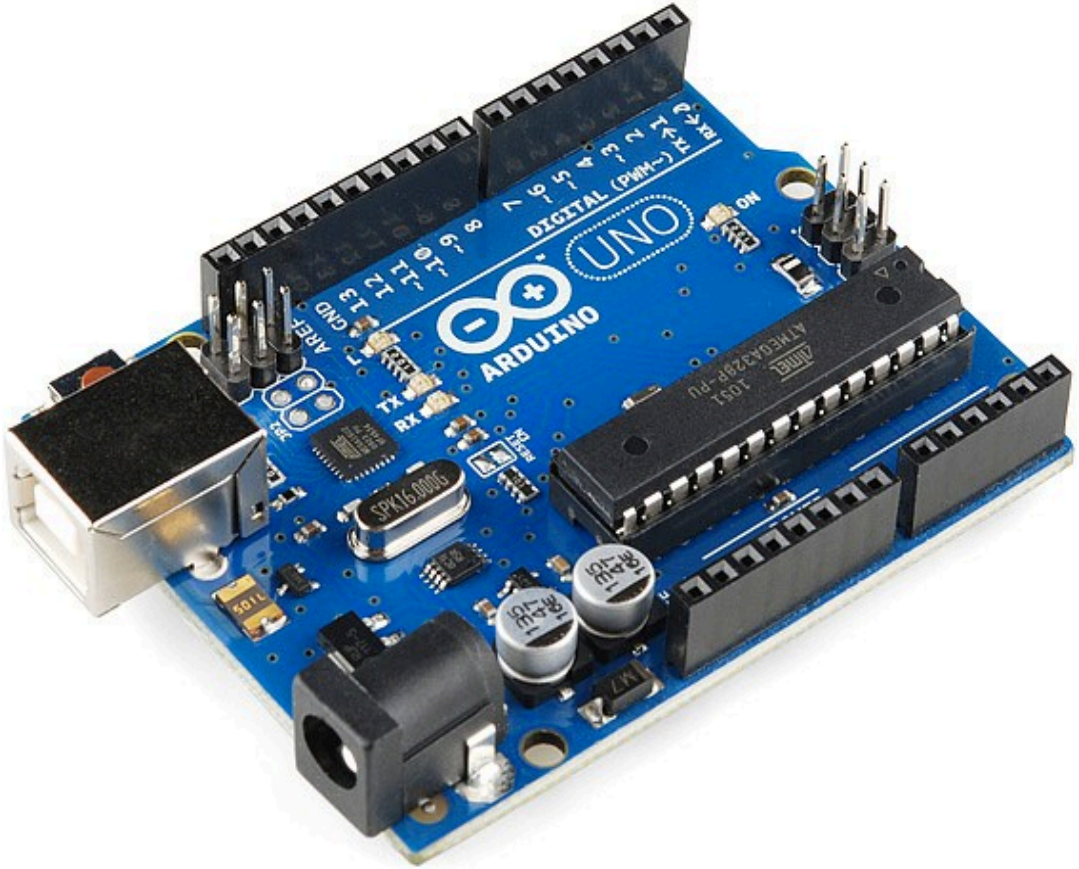
splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™



splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™

splunk> listen to your data™

Data Stats Candy

Every day 2.5 quintillion bytes of data (1 followed by 18 zeros) are created

A full 90 percent of all the data in the world has been generated over the last two years

2.7 Zetabytes of data exist in the digital universe today.

Facebook stores, accesses, and analyzes 30+ Petabytes of user generated data

Akamai analyzes 75 million events per day to better target advertisements.

Decoding the human genome originally took 10 years to process; now it can be achieved in one week.

Data production will be 44 times greater in 2020 than it was in 2009

splunk> listen to your data™

Data Characteristics

VOLUME
VELOCITY
VARIETY
VERACITY

splunk> listen to your data™

Data Tomorrow ?

splunk> listen to your data™

splunk> listen to your data™

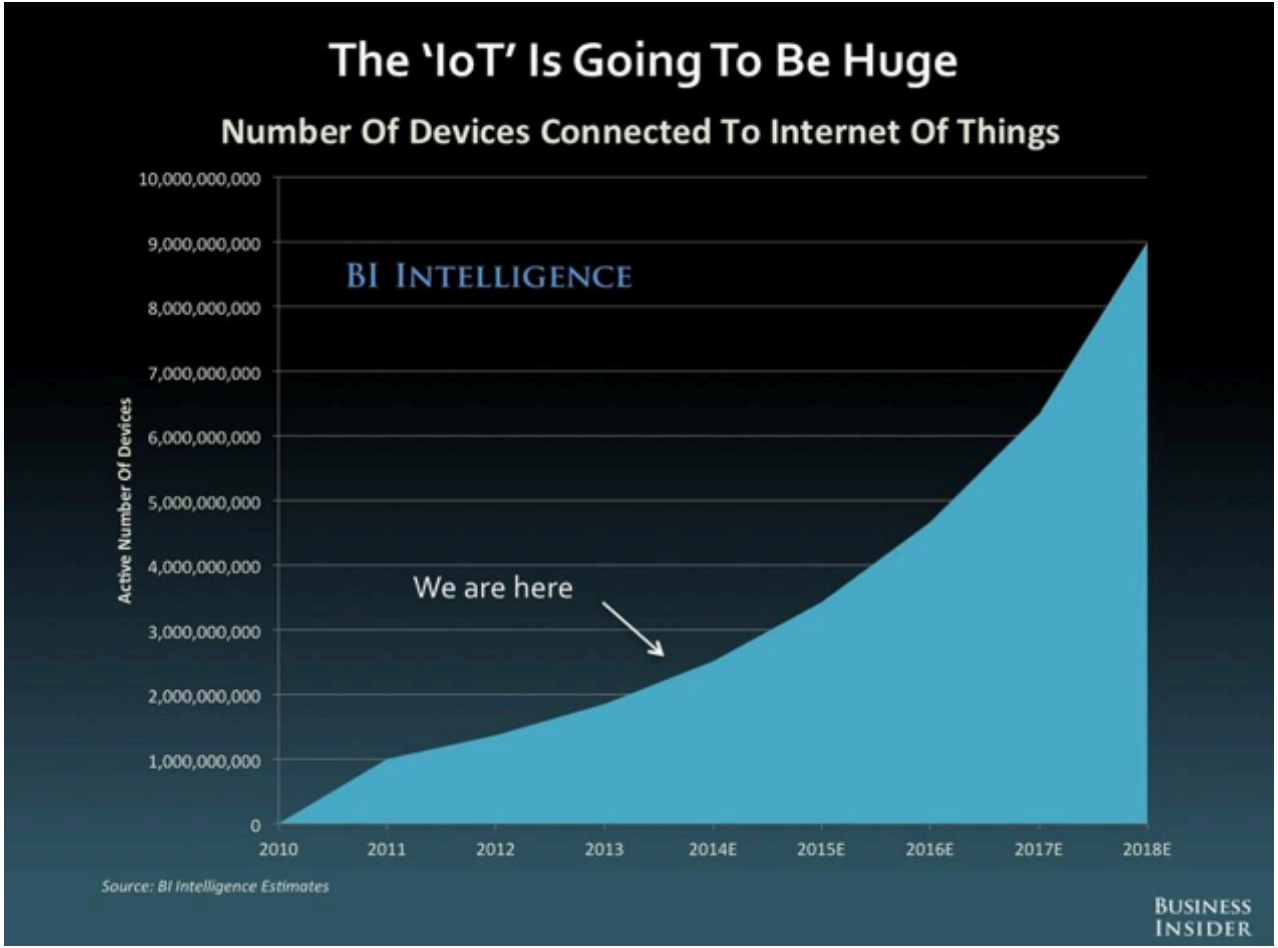
DATA

UNDERSTANDING

INSIGHTS

ACTIONS

splunk> listen to your data™



splunk> listen to your data™

How are we going wrangle this data ?

splunk> listen to your data™

Release the Developers

splunk> listen to your data™

New approaches to data platforms are needed

Open and easily extensible to cope with a variety of data sources and use cases, API oriented

Traditional ETL / Data Warehouses = **schema at write time**
To cope with data today = **schema at read time**

Elasticity

A Data language , the new SQL

Platforms that support an **ecosystem of developers** , content creators , data knowledge sharing

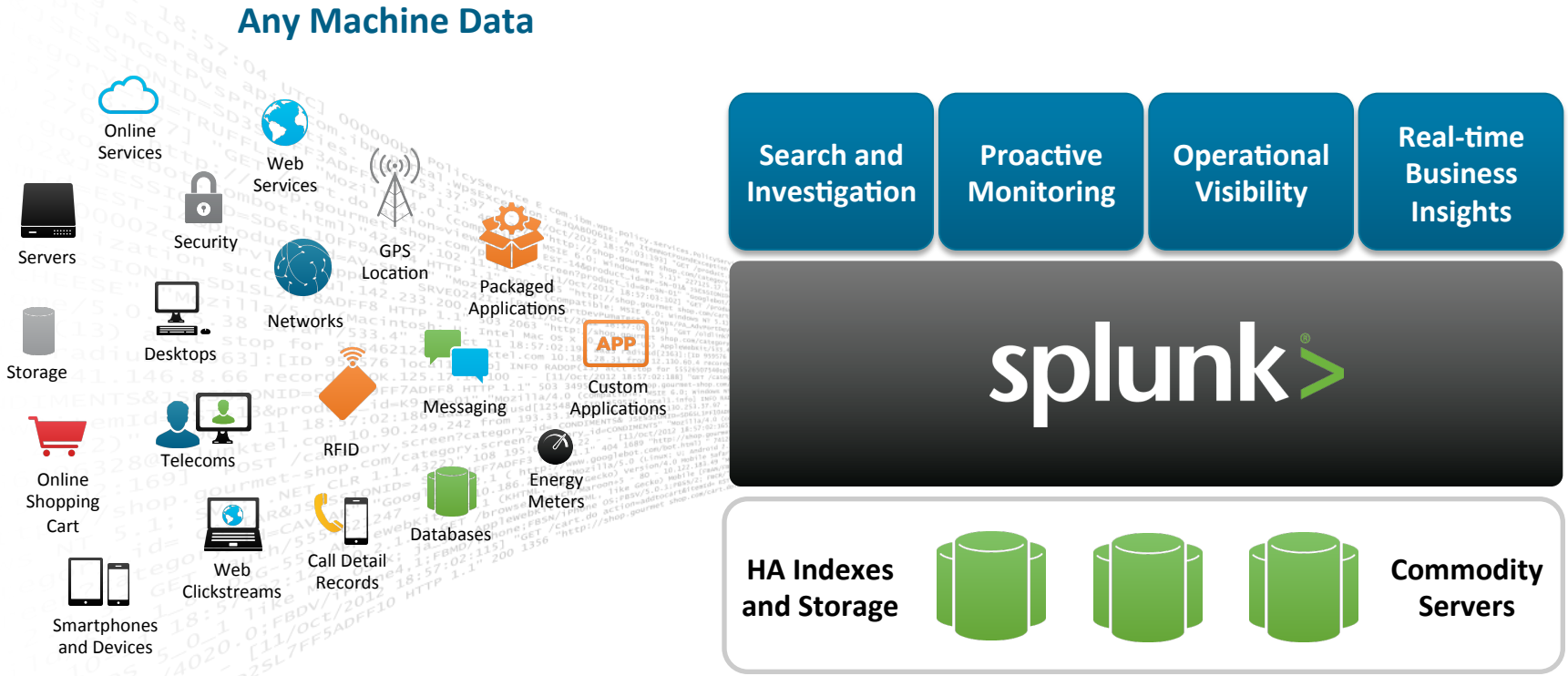
splunk> listen to your data™

Make machine data accessible, usable and valuable to everyone.

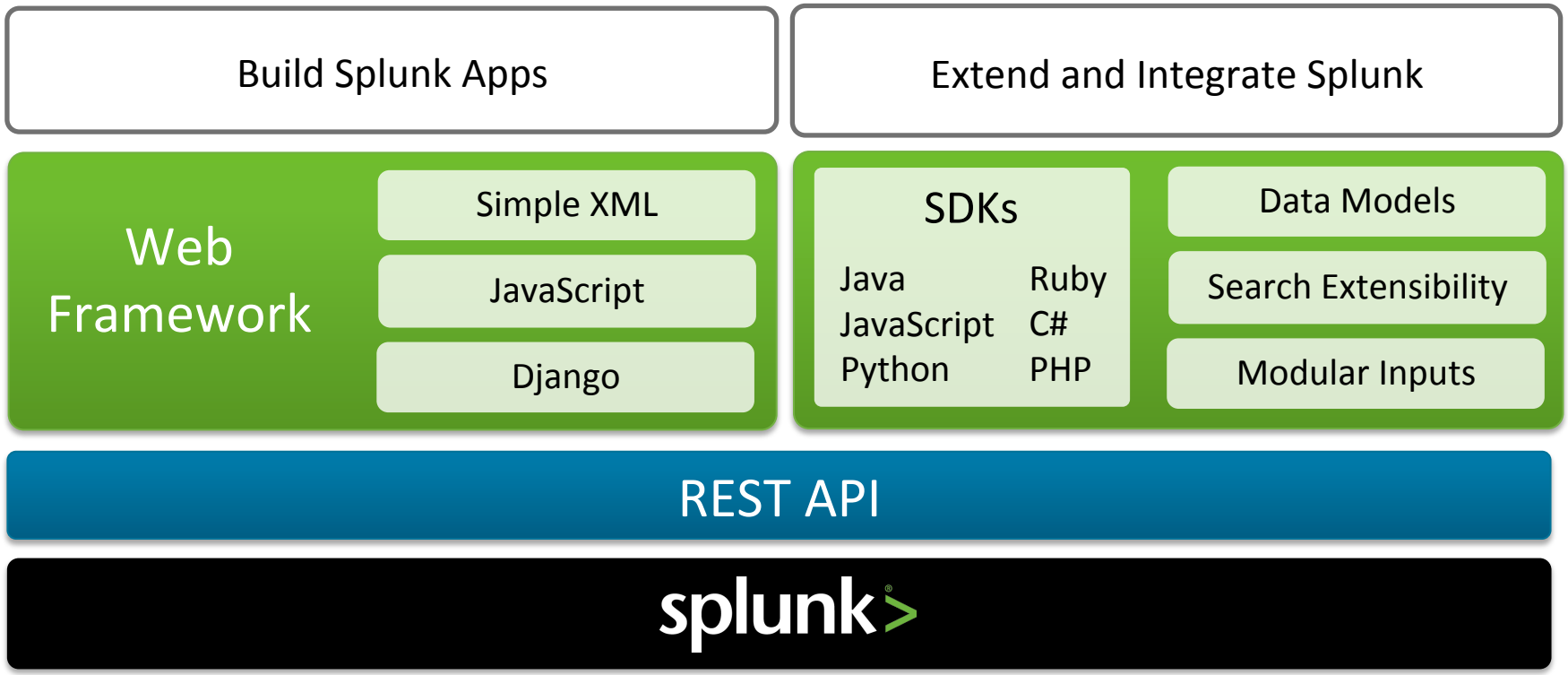
Spelunking

splunk> listen to your data™

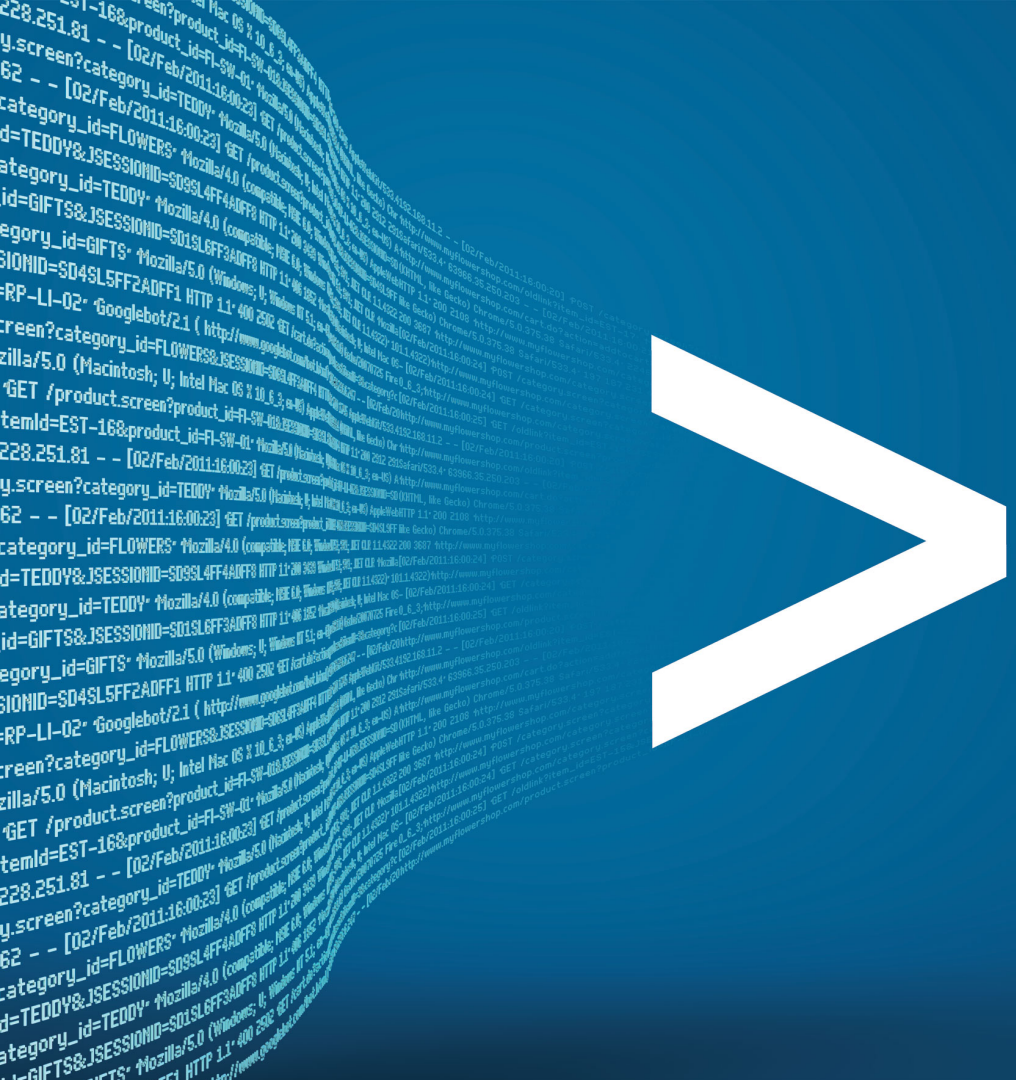
Platform for Machine Data



Powerful Platform for Enterprise Developers



splunk> listen to your data™



Enough talking
DEMO TIME !!

splunk>

The Developer Opportunity in Data

We talk alot about the how , what , where and who ….. but what about the **WHY**

It's fun to make cool things
Get a job , build a business , make money !
Promote yourself , Promote your company
Get involved in community projects
Do Good
Think of new data sources and tap into them
Democratize data
Discover new things & drive society forward

splunk> listen to your data™

Contact Me

ddallimore@splunk.com

@damiendallimore

http://dev.splunk.com

http://blogs.splunk.com/dev

splunk> listen to your data™


Thankyou !
splunk>