



# Oracle Endeca Information Discovery

Chris Lynskey  
Senior Director  
OEID & OBI Product Management

May 2014

Oracle Endeca Information Discovery | Chicago Facilities

Food Inspections Building Violations

Search Box

Selected Refinements

- Your search
- clean
- Inspection Type
- Canvass

Reset

Available Refinements

- General
- AKA Name
- Address
- City
- DBA Name
- Facility Type

Search for specific value...

- Restaurant (3694)
- School (687)
- Grocery Store (598)
- Bakery (72)
- Daycare (2 - 6 Years) (48)
- Daycare Combo 1586 (38)
- Catering (30)

Show More Select All

Inspection Date

- Year
- Year-Month
- Year-Month-Day

Between 4/7/2013 - 8/24/2013

2215 results

Map

Food Inspections (Top 250)

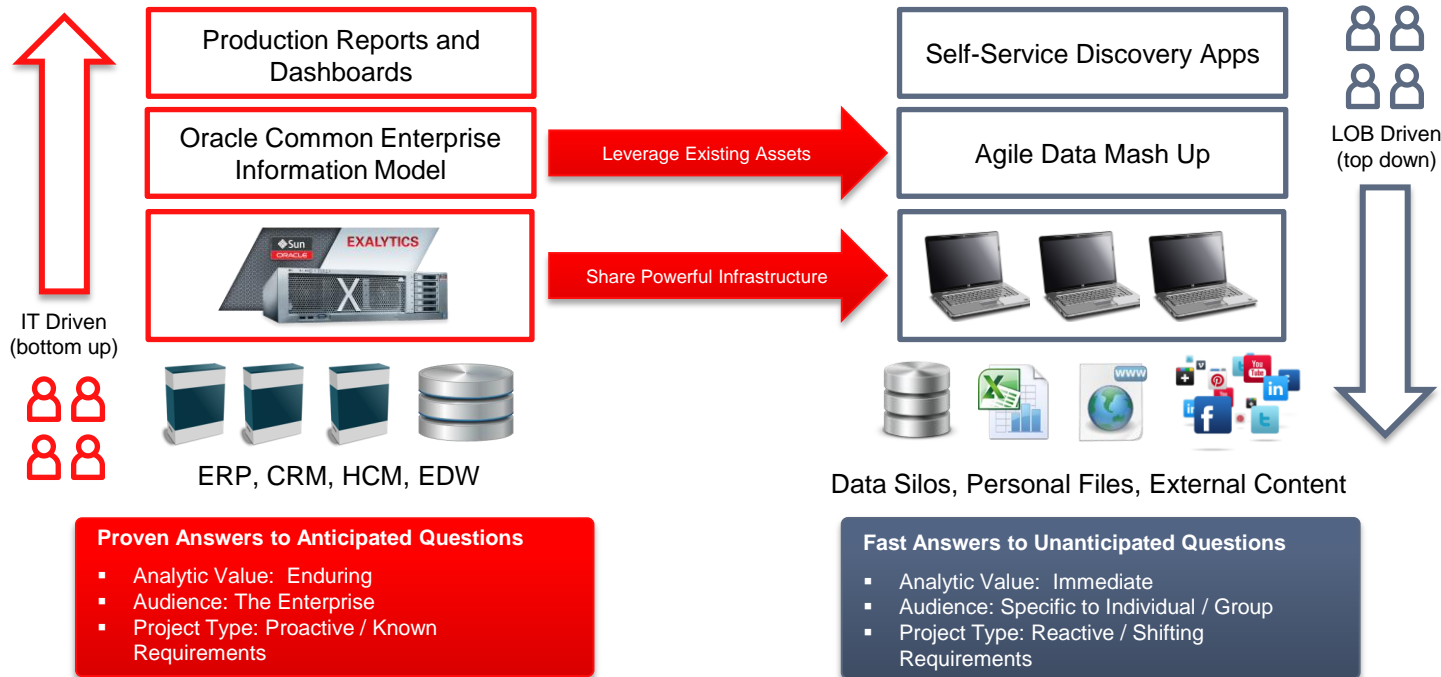
Inspection Details

Inspection ID	AKA Name	Address	City
Data Set	Data Source / Type	Last Loaded	
Building_Violatio...	Building_Violatio... Excel	10/10/2013 11:17 AM (UTC)	
Food_Inspections	Food_Inspections Excel	10/10/2013 11:04 AM (UTC)	

2 Data Sets (Last loaded 10/10/2013)

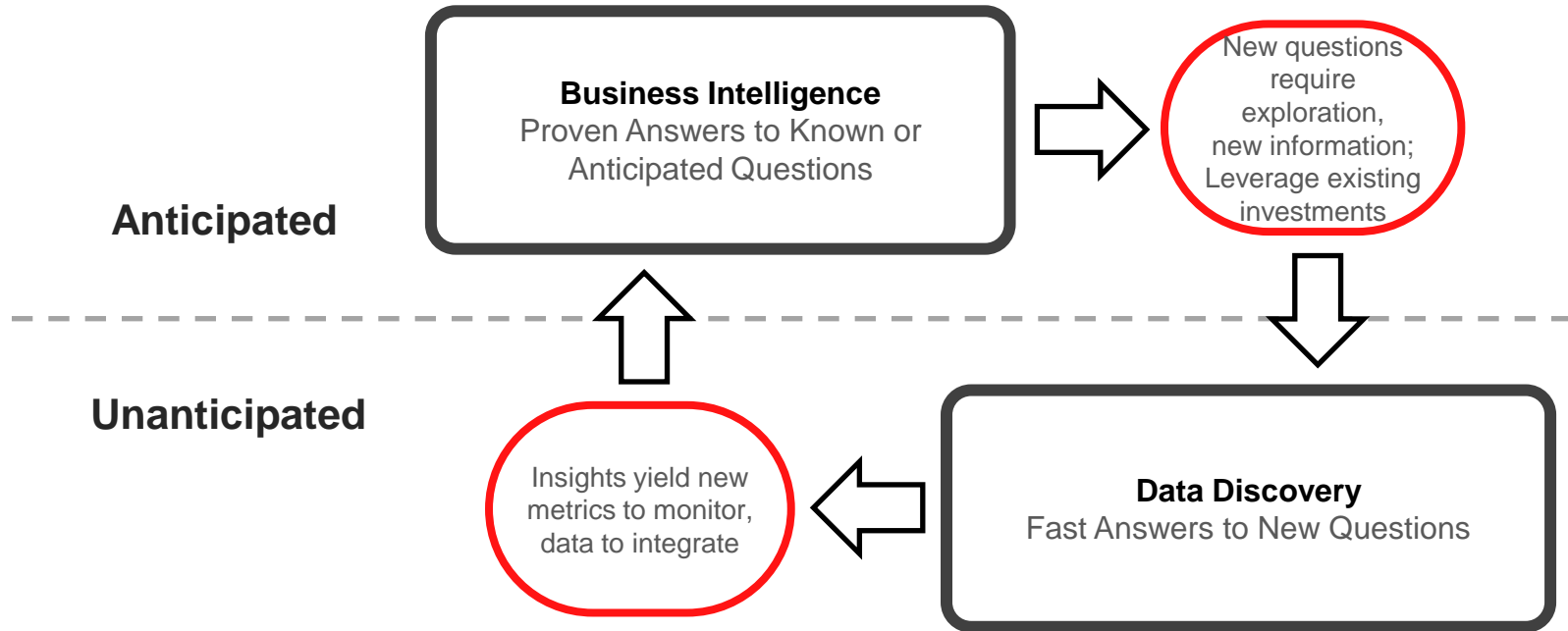
# Self-Service is Critical to Efficient IT Strategy

## Business Users Can Help IT Bridge the 'Last Mile'



# Business Intelligence and Data Discovery

Complimentary Concepts – Technology is More Powerful Together



# Unstructured Data Unlocks New Insights

Combining Structured Data and Unstructured Dialogues Creates New Perspectives

YOY Sales: -45% ↓

- Marketing Spend: +21%
- # Platinum Customers: +7%



".. zoom feature must be broken, seems stuck on wide angle setting.."

".. changes in compensation plan is not incenting sales team to target SMB customers.."

".. our new model surpasses the D150 in quality, functionality and ease of use.."

"..posted some more landscapes from my D150. Love it, picture quality is awesome.."

Sales History



Customer Reviews



Employee Surveys



Competitor FUD



Public Opinions



# Oracle Endeca Information Discovery

Agile and Confident Data Discovery that Drives Innovation across the Enterprise

Personal Excel Files



Oracle BI Server Models



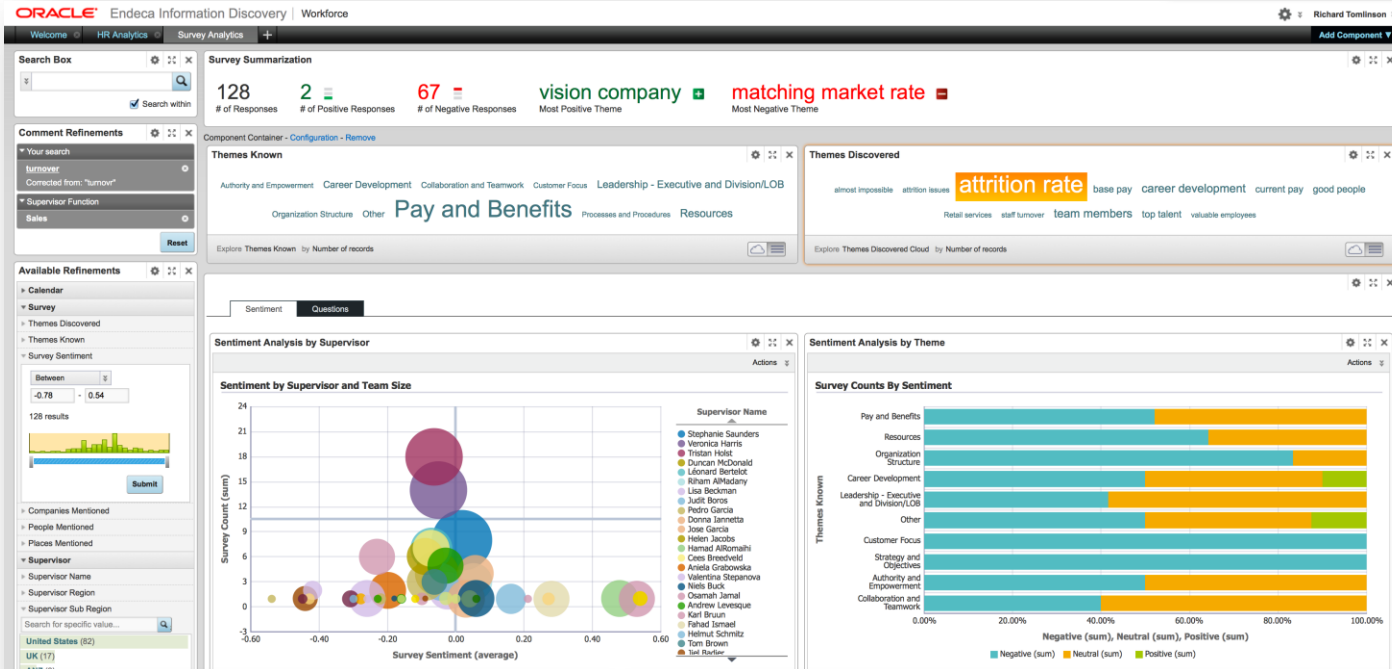
IT Provisioned Enterprise Data



Cloud and Social



Documents and Web Pages



## Agility

Quickly combine diverse content. Create a dynamic relationship between business and IT

## Confidence

Explore trusted enterprise data blended with personal and external sources using powerful managed infrastructure

## Innovation

Expose patterns in unstructured data and text heavy content, including themes and sentiment, for broader perspectives and deeper insights

ORACLE

# DEMO

**ORACLE** Endeca Information Discovery | VoC

Welcome Admin Admin | + Add Page

Sales Social

**Breadcrumbs**

- Search
- canon\_powershot
- Corrected from: "canon powershot"
- Delete All

**Search Box**

Search Within

**Guided Navigation**

- Social Media Discovered
- Themes
  - Search for specific value...
  - lens unit (1161)
  - lens unit image (896)
  - digital (803)
  - quality case (606)
  - digital (380)
  - Show More
  - Select All
- Tagged Topics
- Interaction Sentiment
  - Between: -1.72 - 1.83
  - 20135 results
  - Submit
- People Discovered
- Places Discovered
- Companies Discovered
- Products Discovered
- Sentiment Flag
- Product
- Social Media
- Calendar

**Social Summary**

- Total Mentions: 20,299
- Positive Mentions: 12,614
- Negative Mentions: 5,322
- Average Sentiment: 0.13
- Engaged Users: 11,216

Configuration - Remove

**High Influence Authors by Sentiment and Followers**

# Followers (000's) vs Sentiment (average)

Legend: Klout Bucket (1-5), Influence (Low to High), Sentiment (Very Low to Good)

Y-AXIS: Followers (Log) (sum) | COLOR: Klout Bucket | DETAIL: Author Name | BUBBLE SIZE: Interaction ID (sum)

**Themes Cloud**

20x image canon powershot case classic straight compact digital digital battery digital case digital zoom image quality internet tablet lens unit lens unit image lens unit viewfinder lithium-ion battery low light memory card point-and-shoot quality case registered trademark slim case stereo sound video quality written permission

**Search Results**

Sorted by: Search Relevance

- Interaction Title: @Blackcurrantpix @rivergunjewelry It's a Canon PowerShot G12. Not as fantastic as the cameras you would use ;)
- Date: 4/1/12
- Author Name: RiverGumJewelry
- Interaction Content: @Blackcurrantpix @rivergunjewelry It's a Canon PowerShot G12. Not as fantastic as the cameras you would use ;)
- Brand: Canon
- Product Make: Canon PowerShot
- Date: 6/3/12
- Author Name: Yasher
- Interaction Content: 366 of 2012 - Day 156. Our challenge this week is to turn an every day snap shot into a portrait. Super!Webb! Taken with my Canon PowerShot G12. 84, 1/50 sec., ISO-160, no flash, camera hand held, shot outdoors. Edited with Pictometry and Photoshop. Picasa's Weekly Challenge is posted here: 8 Steps to Help You Turn Your Everyday Snapshots Into Portraits Photography Tutorial .
- Brand: Canon
- Product Make: Canon PowerShot
- Interaction Title: Extra Powerful 4.5X Telephoto Lens for The Canon PowerShot G10 G11 G12 Digital Camera Includes protective 67MM U... http://t.co/HR3Nagvp
- Date: 7/31/12
- Author Name: andreasaxwood
- Interaction Content: Extra Powerful 4.5X Telephoto Lens for The Canon PowerShot G10 G11 G12 Digital Camera Includes protective 67MM U... http://t.co/HR3Nagvp
- Brand: Canon

Page 1 of 2233 | 1-10 of 22323 | 10 per page

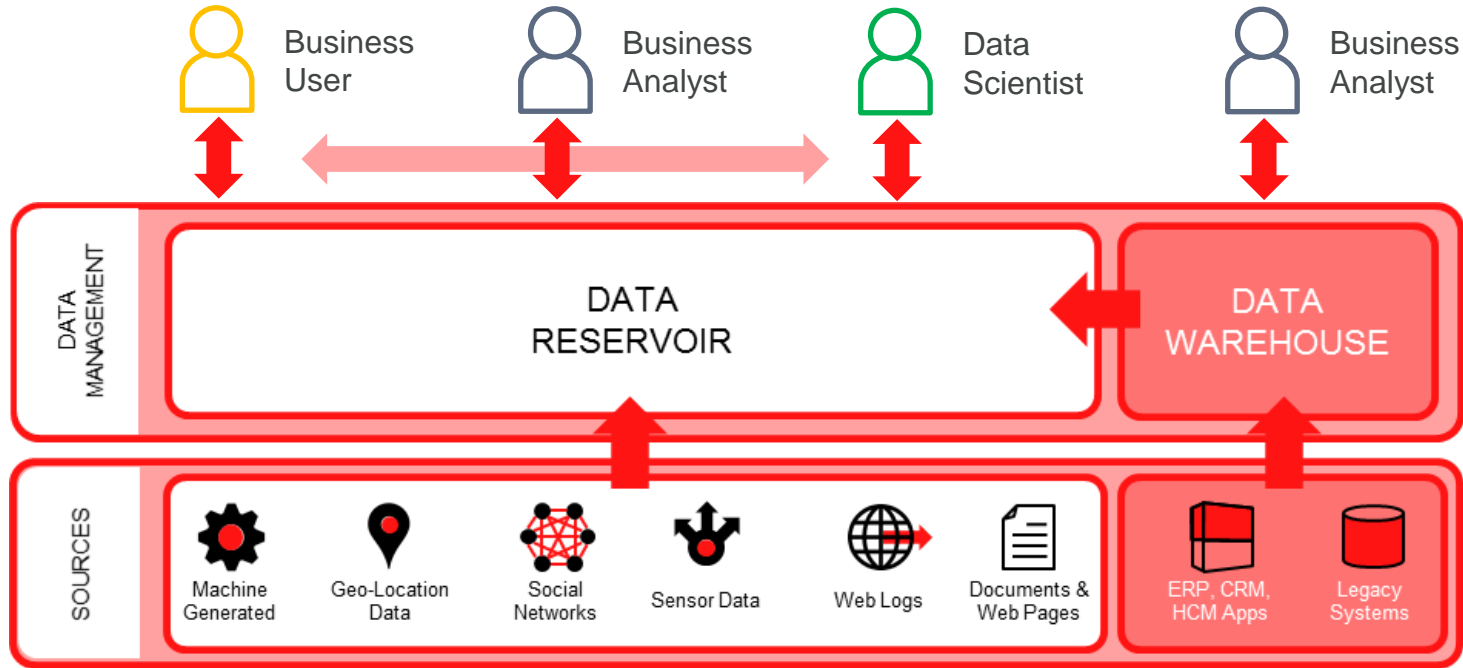
ORACLE®

# Big Data + Discovery

New Strategies for Changing Landscape



# The Next Generation of Data Analysis



**Old World**  
Business Analysts make independent decisions based on highly curated data warehouse

**New World**  
All data sources, structured and unstructured, flow into Data Reservoir, where teams with varied skills collaborate



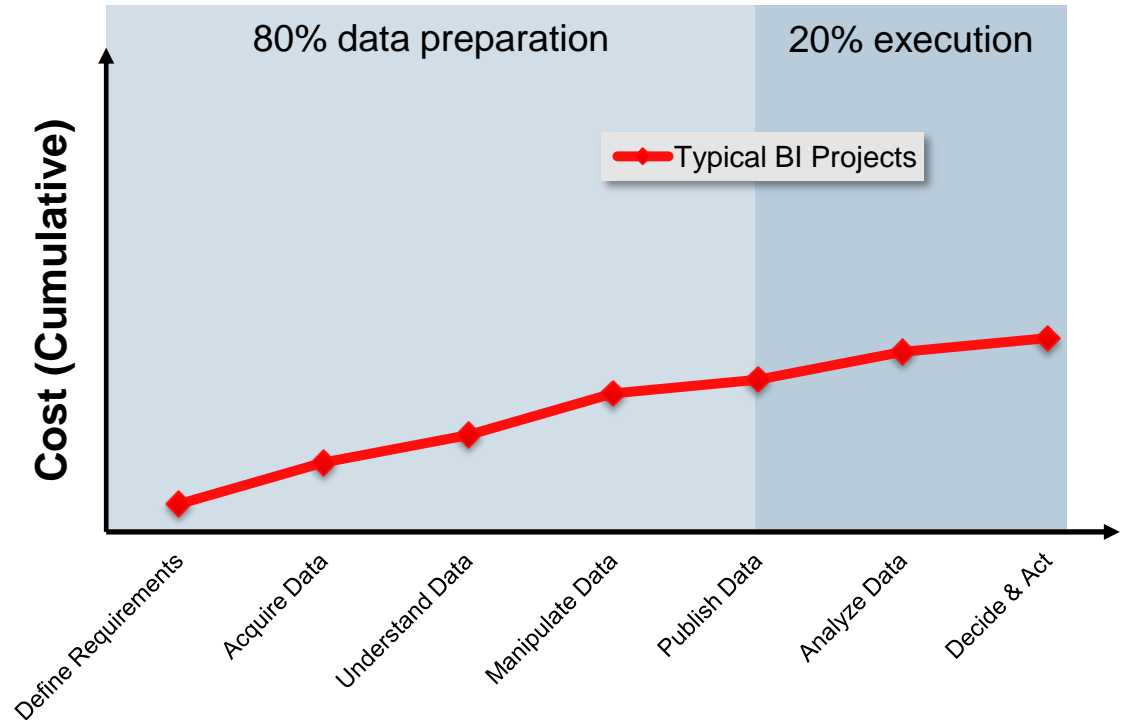
# Analytics Projects Place Most Effort Upfront

80% of effort focused on **data preparation**

- Try to predict all possible business questions upfront, modeling, locating and acquiring data, ETL, etc
- Need to design a model to answer predicted questions
- Then need to hunt down data sources and manipulate feeds to fit into model designed

20% of effort actually focused on **execution**

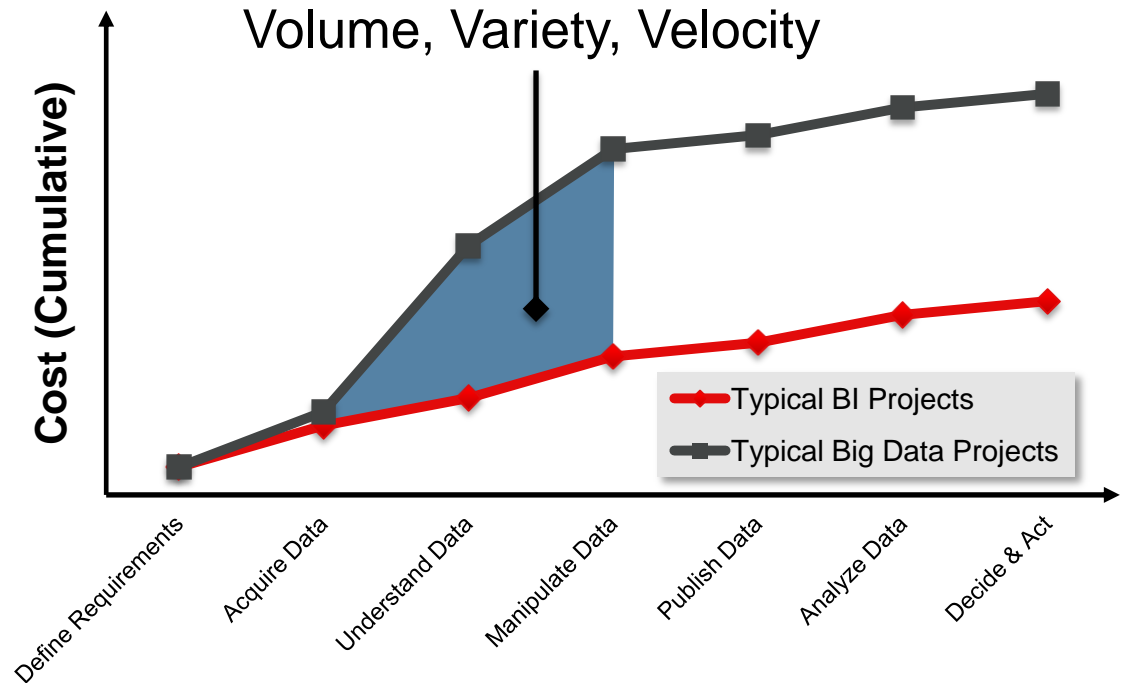
- Analyze data, discover insights, decide and act



# Big Data Complexity Compounds Upfront Cost

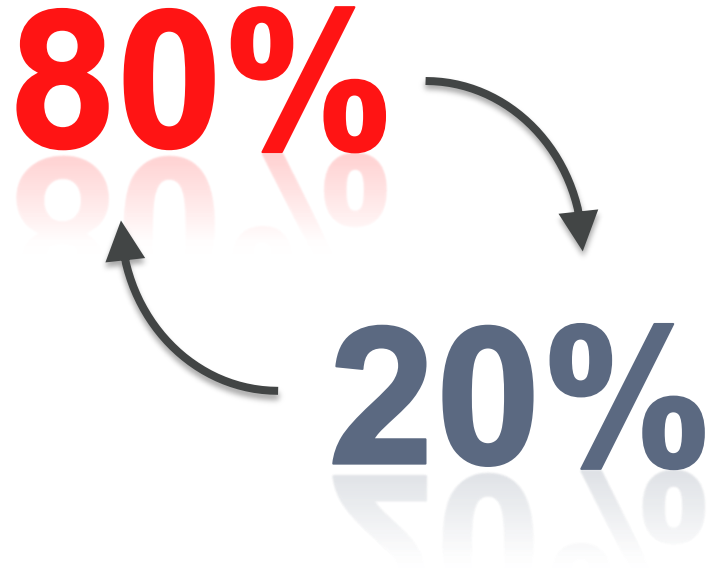
Increased volume, variety and velocity of data only adds complexity, effort and cost

- Significantly more time needed to **understand** and **manipulate** (wrangle) new data sets before they are ready to be published
- Requires highly **specialized resources** (i.e. Data Scientists vs. ETL Developers)
- Tools they use require **expert skills** (Pig, Python, Hive, R, etc)



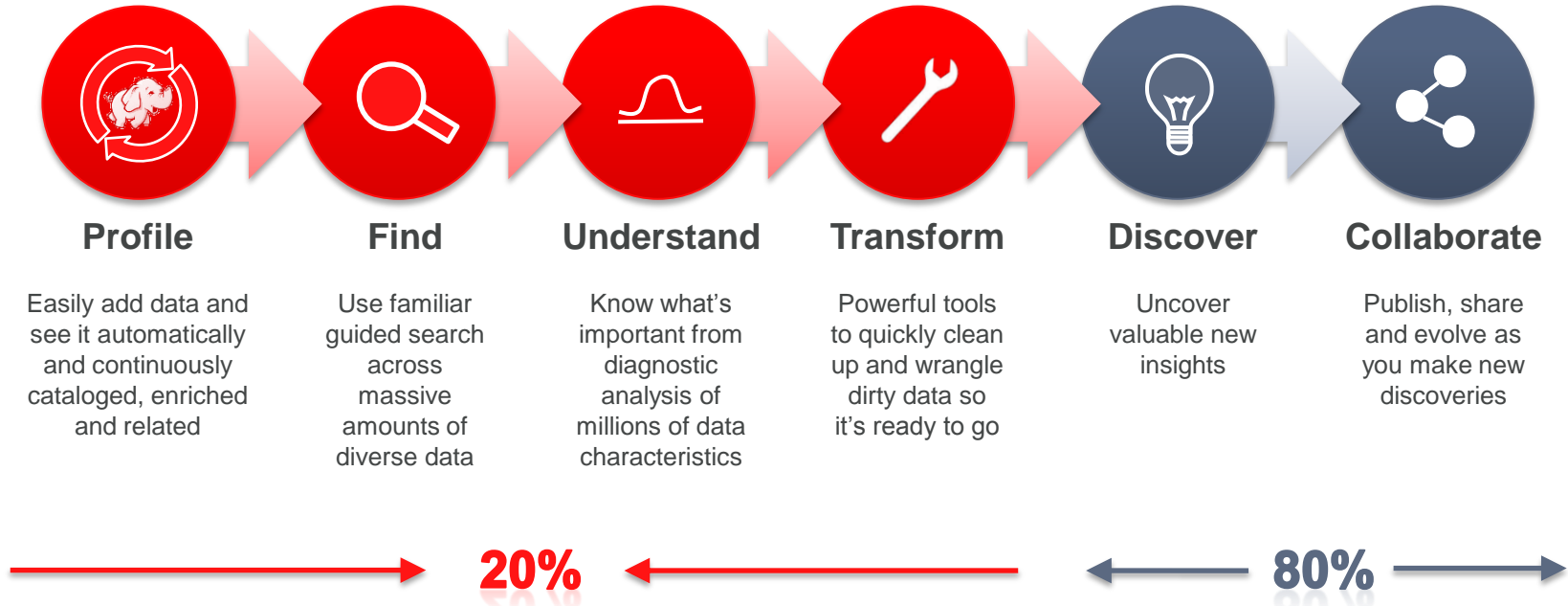
# The Challenge for Big Data Projects

How do we make data preparation 20% of the effort so you can focus 80% of your time on executing and improving your business?



# Tackling Big Data Variety

Changing the game for agile business innovation on big data



# Key Requirements – Profile



## Profile

- Automatic attribute profiling and enrichments
  - Primary key detection
  - Related data sets
  - Data type detection
  - Geo-location
  - Salient term extraction
  - Sentiment analysis
  - Language detection
  - Named entity recognition

Bleed Lines **Recharge Battery** Replace Alternator  
**Replace ABS Control Module** Replace Clock Spring Replace Control Module  
Replace Filter and Fluid Replace Front Caliper **Replace Fuse/ Reset Circuit Breaker**  
Replace Harness Replace Ignition Module  
**Replace Mass Air Flow Sensor**

HELLO 

buenos días 

BONJOUR 

*guten Tag* 



# Key Requirements – Find



## Find

- Search and navigate across data reservoir
- Surface relevant data sets
- View data set summaries
- Jump to team projects

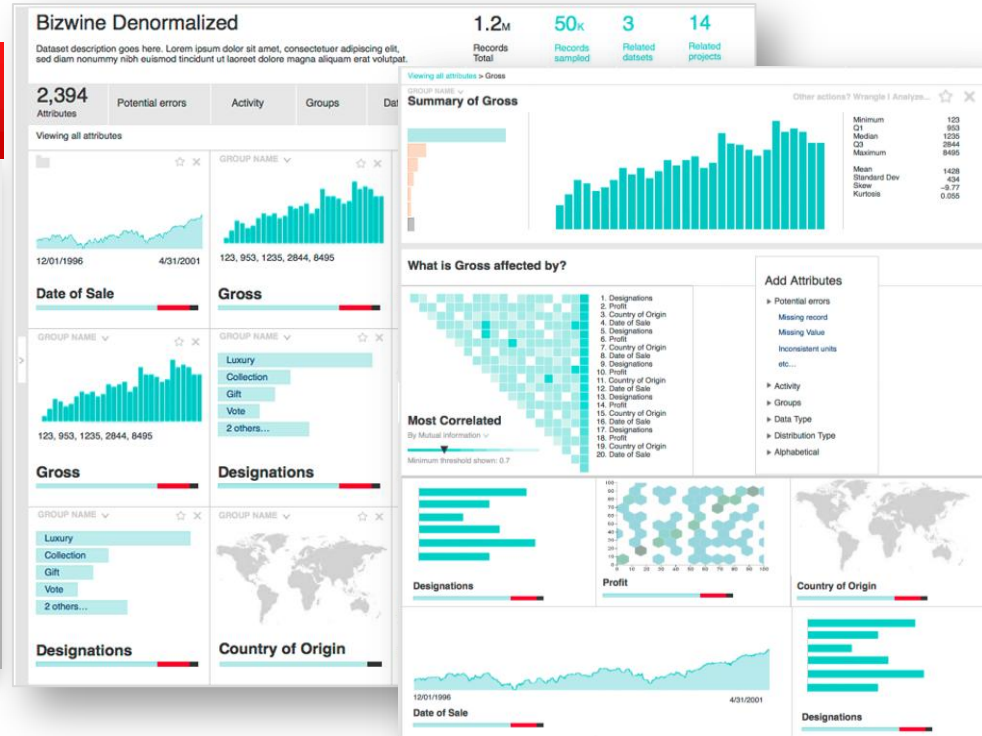
The screenshot displays the Oracle Data Catalog interface. At the top, it shows 'ORACLE' branding and a search bar with the text '128 52'. Below this, there are two main sections: 'All DATA SETS' and 'All PROJECTS'. The 'All DATA SETS' section is further divided into 'Most used DATA SETS' and 'Recently added DATA SETS'. Each data set entry includes a 'Data Set Name', 'Attributes', 'Records', 'Data Source', and 'Tags'. The 'All PROJECTS' section is divided into 'My PROJECTS' and 'My DATA SETS'. Each project entry includes a 'Project Name', 'Attributes', 'Records', 'Created by', and 'Tags'. The interface is clean and modern, with a dark header and light content area.

# Key Requirements – Understand



## Understand

- Get a feel for the data
- Spot dirty data
- Find important attributes
- Uncover relationships
- Validate assumptions





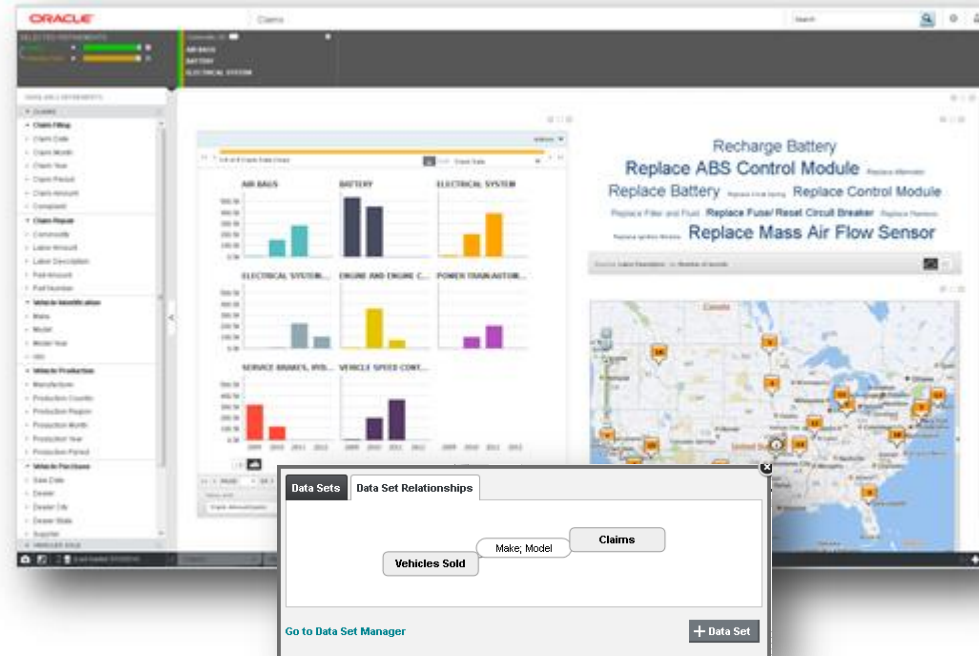


# Key Requirements – Discover



## Discover

- Advanced data blending
- Interactive visualizations
- Summarize data in a variety of ways
- Powerful search and data driven navigation



# Key Requirements – Collaborate



## Collaborate

- Publish data to Hadoop for use in other tools
- Share data transform scripts
- Save and share bookmarks
- Take snapshots with annotation
- Save offline or email images

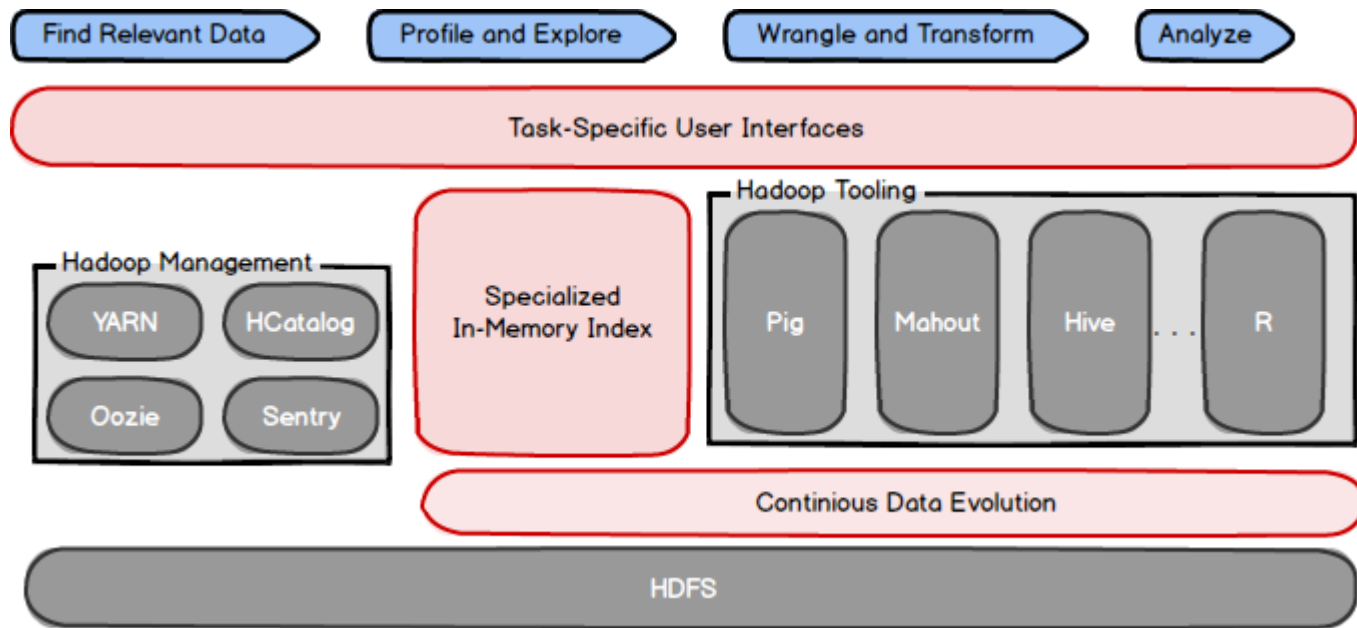
The screenshot displays a data management interface with several overlapping windows:

- PERSONAL BOOKMARKS:** A window with a search filter and a bookmark titled "Big Three".
- Save Snapshot:** A dialog box for saving a snapshot, showing "Showing 1-1 of 4" items. Fields include "Name" (Screen-Chart\_2014-03-10 15:42:21) and "Description". Buttons for "Cancel" and "Save" are visible.
- Big Three:** A detailed view of the bookmark, showing fields for "Description" (US Manufacturing), "Created at" (03/10/2014), "Data Source" (Claims | Vehicles\_Sold), and "Refinements". It also lists "Claims" for "GENERAL MOTORS CORP." and "DAIMLERCHRYSLER CORP".
- Export (undefined records):** A dialog box for exporting data to a CSV file. It includes fields for "Name" (BatteryClaims\_2014-03-10.csv), "CSV Format" (Standard or Custom separator), "Fields separated by" (Pipe), "Fields enclosed by" (Single Quote), and "Value Formatting" (Raw data or Formatted data). Buttons for "Cancel" and "Export" are visible.

# Leveraging Hadoop Ecosystem

Embedded with Hadoop for Fast Analysis on All the Data

- **Full-featured platform** that integrates with the Hadoop ecosystem to address the entire analysis life cycle
- **Analysis where the data lives** means not having to move data to analyze it
- **Auto enrichment** available to all Hadoop tools
- **Enables easy interaction** with other Hadoop ecosystem tools



**ORACLE®**