

# Making Big Data Useful Data

*Speaker Anna Bongenhielm*





**Click 'engage'  
to rate sessions  
and ask questions**

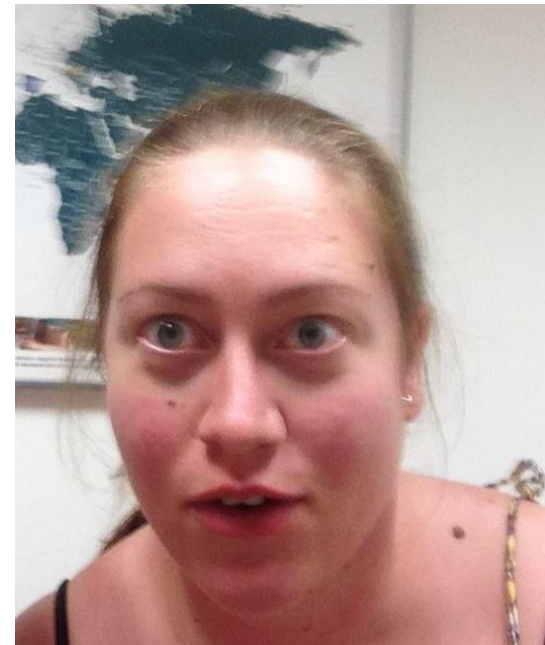


Kung

# Anna Bongenhiehl

## Producer Data Analytics @ King

- Little over 4 years with King
- Mainly worked with reporting and downstream data chain
- Personal favorite King game right now – Candy Crush Soda Saga



# We make great games

- We have developed more than **200 fun titles** and offer games in over **200 countries and regions** around the world.
- **340 million** average monthly unique users (Q2 2015).



- Studios in Stockholm, London, Barcelona, Bucharest, Malmo, Berlin, Singapore and Seattle.
- Offices in San Francisco, Malta, Tokyo, Seoul and Shanghai.



# The evolution of King



## Online skill

- Founded in **2003**
- Originally games were only available through our site and portals including AOL and Yahoo!



## Social

- Launched first game on Facebook in Q2 **2011**



## Mobile

- Launched first game on mobile H2 **2012**

# Presentation overview

1. The Daily Pipeline
2. Brief Reporting History
3. Dealing with Data Size in the Reporting Layer
4. Governing the User End of a Big Data Environment
5. Getting the Stakeholders the Right Thing

Credit goes to Jonathan Palmer and Jakob Berglund who have contributed several of the slides in this presentation

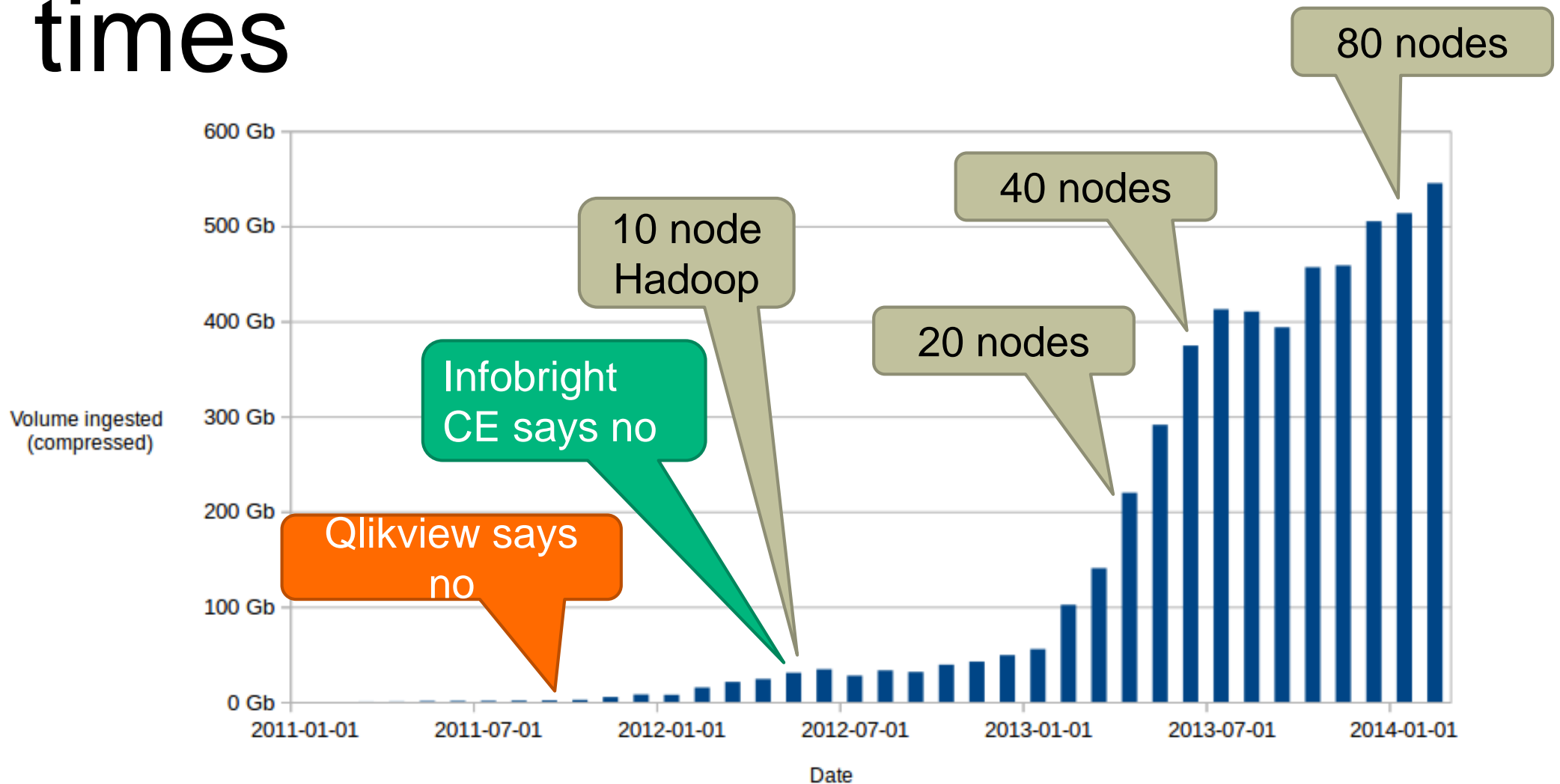


# The Daily Pipeline





# Fast changing times



Yes, our data volume has grown *significantly* over time...



What is tracking really?

# Tracking measures what players do

For example...

- ...when players start the game
- ...when players start a level
- ...who players interact with
- ...what players buy
- ...and a lot of other stuff



We currently use ~415 event types  
(415 measurement points)

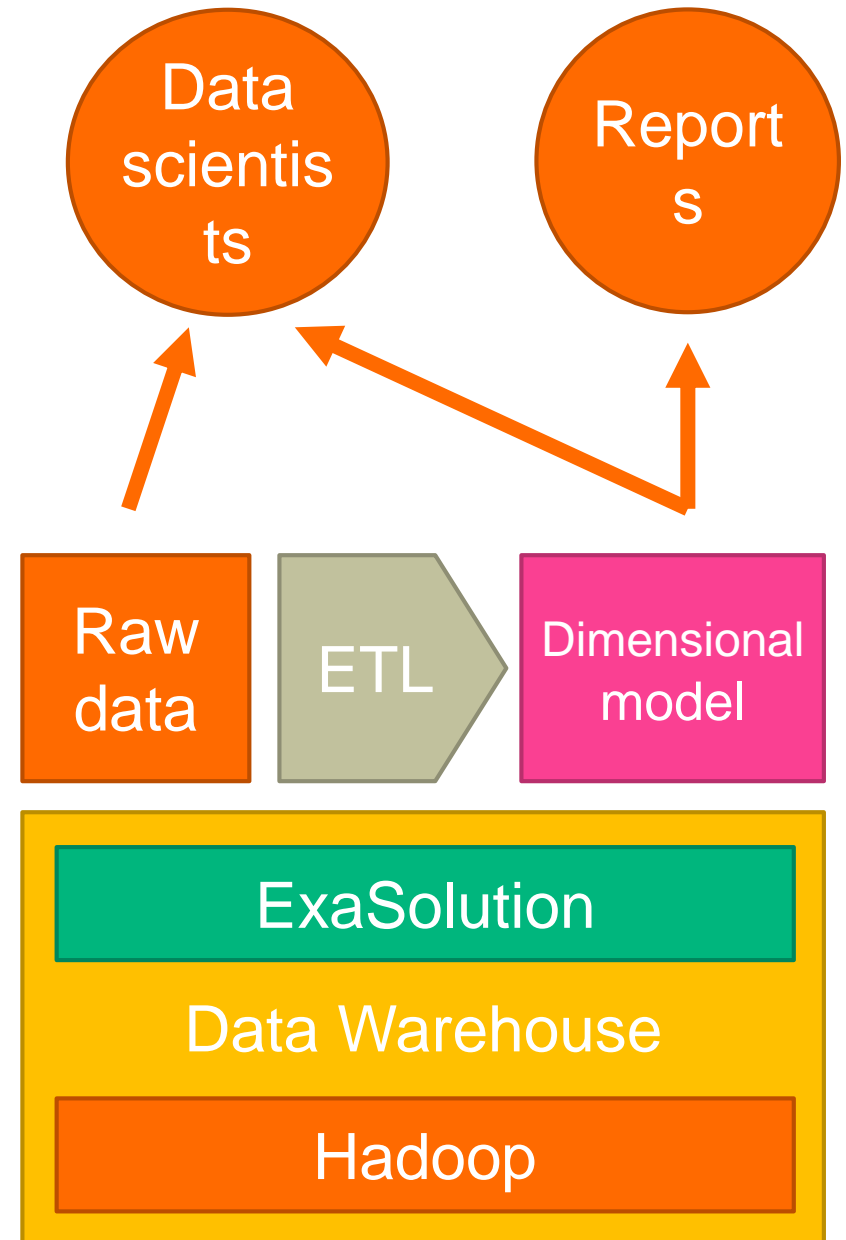
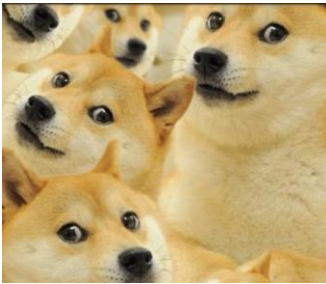
More than a 100 million daily players...  
...create over than **15 billion** events  
daily!

```
AppAppleStoreTransactionBegin 20130923T112512.053+0200
AppAppleStoreValidationStart 20130923T112512.055+0200
AppAppleStoreValidationDone 20130923T112513.013+0200
AppAppleStoreTransaction 20130923T112513.014+0200
AppAppleStoreTransactionDone 20130923T112513.016+0200
AppItemTransaction2 20130923T112516.654+0200
AppItemTransaction2 20130923T112516.656+0200
SagaAppGameStart2 20130923T112516.657+0200
SagaAppGameEnd2 20130923T112516.657+0200 240017
AppChangedCoreUser2 20130923T112705.928+0200
AppSignInUser 20130923T112705.945+0200 250017
MessageProcessed3 20130923T112722.120+0200
ToroHandleLink 20130923T112722.148+0200 250017
MessageProcessed3 20130923T112722.159+0200
ToroHandleLink 20130923T112722.183+0200 250017
MessageProcessed3 20130923T112722.196+0200
ToroHandleLink 20130923T112722.220+0200 250017
MessageProcessed3 20130923T112722.230+0200
ToroHandleLink 20130923T112722.251+0200 250017
MessageProcessed3 20130923T112722.260+0200
ToroHandleLink 20130923T112722.281+0200 250017
AppSignInUser 20130923T112752.685+0200 250017
AppClientVersion 20130923T112752.837+0200
AppStart5 20130923T112753.981+0200 250017
AppSignInUser 20130923T112754.605+0200 250017
AppItemTransaction2 20130923T112836.232+0200
SagaAppGameStart2 20130923T112836.233+0200
SagaAppGameEnd2 20130923T112836.235+0200 250017
```

What does our pipeline look like?

# Our fancy data pipeline!

## (much wow!)



wow   such amaze   much boxes   very statistics   so mystery



# Brief Reporting History



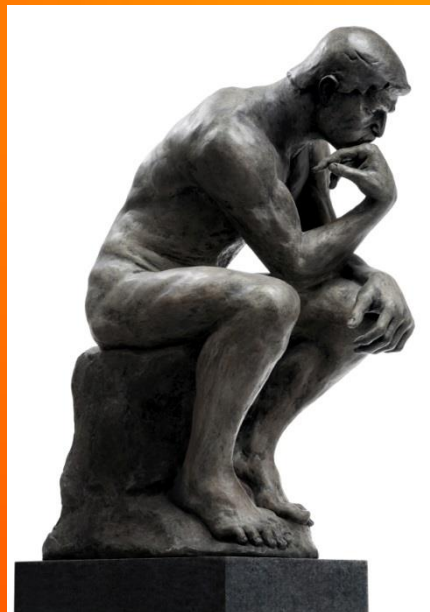
# In the beginning:

- Most data processed directly in Qlikview
- Few developers, all located in the same office
- No need for a DWH
- Information and dimensionality on a user level basis



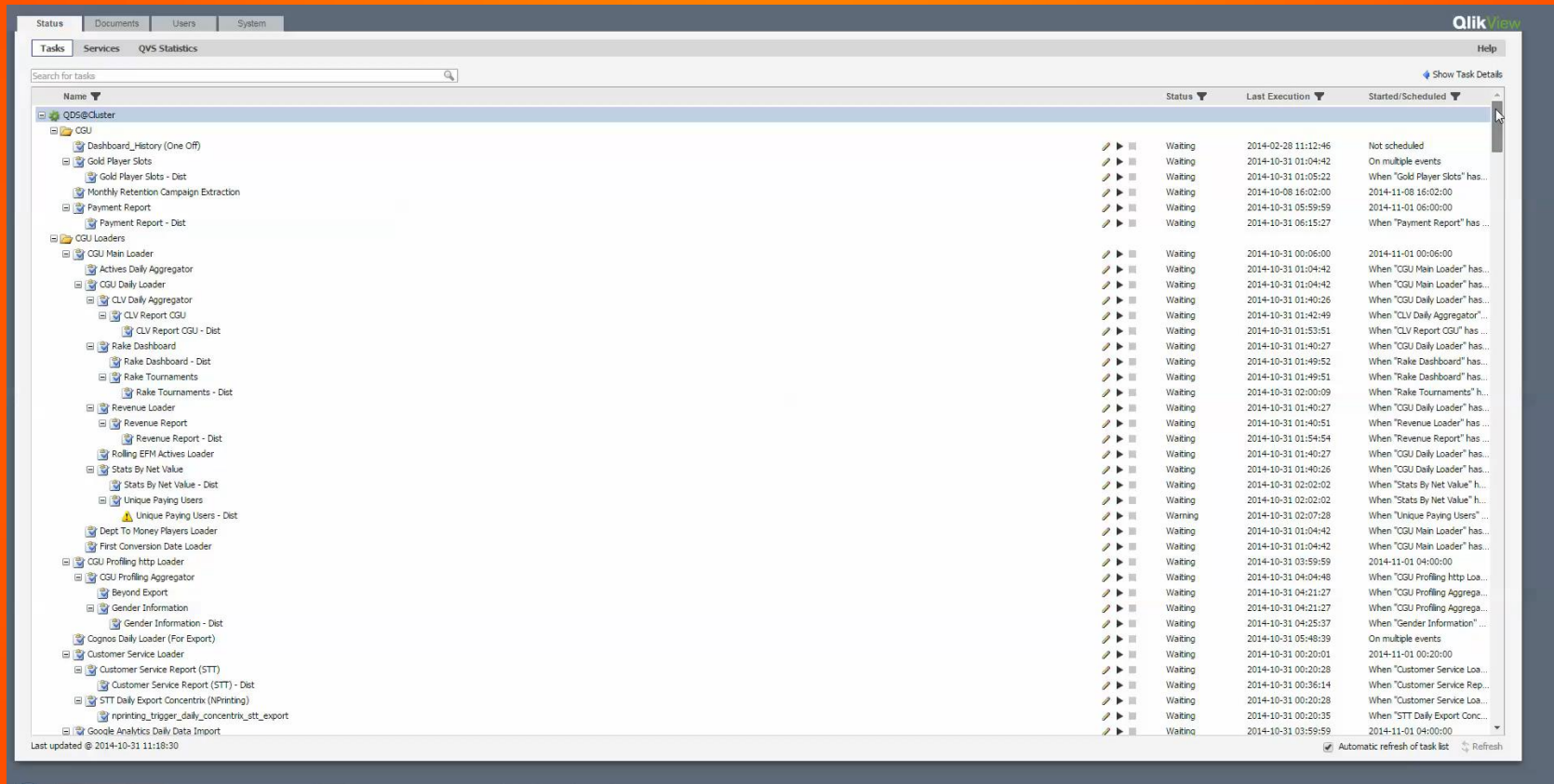
# The problems we face with size:

- Adding more games
- Adding new types of data
- Knowing what to look for and when to group the data



$$1+2(3+2) + \dots$$

# Our QMC:



The screenshot displays the QlikView QMC interface. The top navigation bar includes 'Status', 'Documents', 'Users', and 'System'. Below this, the 'Tasks' tab is active, showing a search bar and a 'Show Task Details' link. The main area is a table of tasks, organized into a tree view on the left. The table columns are 'Name', 'Status', 'Last Execution', and 'Started/Scheduled'. The tasks are categorized under 'QDS@Cluster' and include various loaders, aggregators, and reports. The status of each task is indicated by a green checkmark (Waiting) or a red X (Error). The last execution time and the scheduled start time are also provided for each task.

Name	Status	Last Execution	Started/Scheduled
QDS@Cluster			
CGU			
Dashboard_History (One Off)	Waiting	2014-02-28 11:12:46	Not scheduled
Gold Player Slots	Waiting	2014-10-31 01:04:42	On multiple events
Gold Player Slots - Dist	Waiting	2014-10-31 01:05:22	When "Gold Player Slots" has...
Monthly Retention Campaign Extraction	Waiting	2014-10-08 16:02:00	2014-11-08 16:02:00
Payment Report	Waiting	2014-10-31 05:59:59	2014-11-01 06:00:00
Payment Report - Dist	Waiting	2014-10-31 06:15:27	When "Payment Report" has ...
CGU Loaders			
CGU Main Loader	Waiting	2014-10-31 00:06:00	2014-11-01 00:06:00
Actives Daily Aggregator	Waiting	2014-10-31 01:04:42	When "CGU Main Loader" has...
CGU Daily Loader	Waiting	2014-10-31 01:04:42	When "CGU Main Loader" has...
CLV Daily Aggregator	Waiting	2014-10-31 01:40:26	When "CGU Daily Loader" has...
CLV Report CGU	Waiting	2014-10-31 01:42:49	When "CLV Daily Aggregator"...
CLV Report CGU - Dist	Waiting	2014-10-31 01:53:51	When "CLV Report CGU" has ...
Rake Dashboard	Waiting	2014-10-31 01:40:27	When "CGU Daily Loader" has...
Rake Dashboard - Dist	Waiting	2014-10-31 01:49:52	When "Rake Dashboard" has...
Rake Tournaments	Waiting	2014-10-31 01:49:51	When "Rake Tournaments" h...
Rake Tournaments - Dist	Waiting	2014-10-31 02:00:09	When "Rake Tournaments" h...
Revenue Loader	Waiting	2014-10-31 01:40:27	When "CGU Daily Loader" has...
Revenue Report	Waiting	2014-10-31 01:40:51	When "Revenue Loader" has...
Revenue Report - Dist	Waiting	2014-10-31 01:54:54	When "Revenue Report" has...
Rolling EFM Actives Loader	Waiting	2014-10-31 01:40:27	When "CGU Daily Loader" has...
Stats By Net Value	Waiting	2014-10-31 01:40:26	When "Stats By Net Value" h...
Stats By Net Value - Dist	Waiting	2014-10-31 02:02:02	When "Stats By Net Value" h...
Unique Paying Users	Waiting	2014-10-31 02:02:02	When "Unique Paying Users" ...
Unique Paying Users - Dist	Warning	2014-10-31 02:07:28	When "Unique Paying Users" ...
Dept To Money Players Loader	Waiting	2014-10-31 01:04:42	When "CGU Main Loader" has...
First Conversion Date Loader	Waiting	2014-10-31 01:04:42	When "CGU Main Loader" has...
CGU Profiling Http Loader	Waiting	2014-10-31 03:59:59	2014-11-01 04:00:00
CGU Profiling Aggregator	Waiting	2014-10-31 04:04:48	When "CGU Profiling http Loa...
Beyond Export	Waiting	2014-10-31 04:21:27	When "CGU Profiling Aggrega...
Gender Information	Waiting	2014-10-31 04:21:27	When "Gender Information" ...
Gender Information - Dist	Waiting	2014-10-31 04:25:37	When "Gender Information" ...
Cognos Daily Loader (For Export)	Waiting	2014-10-31 05:48:39	On multiple events
Customer Service Loader	Waiting	2014-10-31 00:20:01	2014-11-01 00:20:00
Customer Service Report (STT)	Waiting	2014-10-31 00:20:28	When "Customer Service Loa...
Customer Service Report (STT) - Dist	Waiting	2014-10-31 00:36:14	When "Customer Service Rep...
STT Daily Export Concentrix (NPrinting)	Waiting	2014-10-31 00:20:28	When "Customer Service Loa...
nprinting_trigger_daily_concentrix_stt_export	Waiting	2014-10-31 00:20:35	When "STT Daily Export Conc...
Google Analytics Daily Data Import	Waiting	2014-10-31 03:59:59	2014-11-01 04:00:00

Last updated @ 2014-10-31 11:18:30

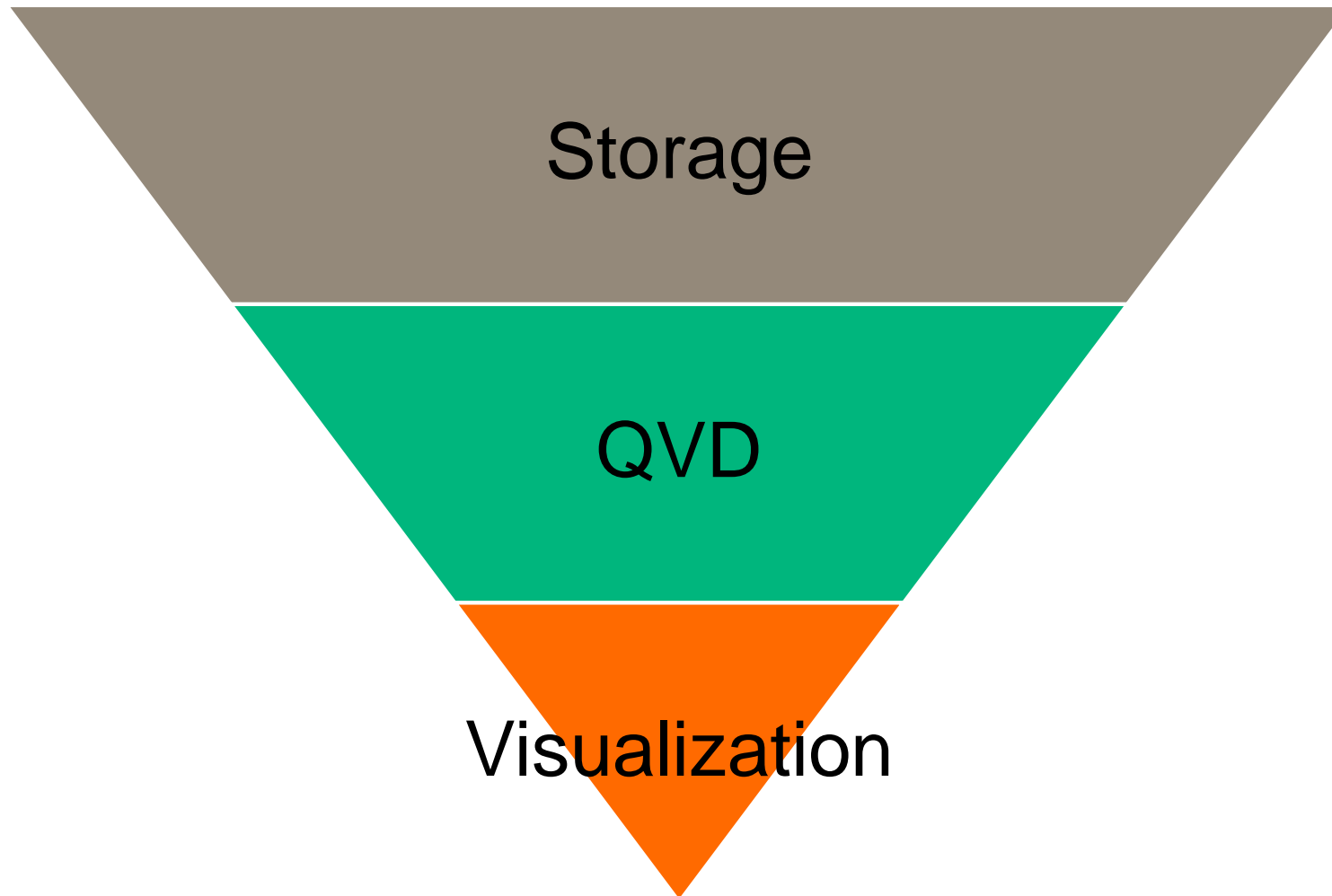
Automatic refresh of task list Refresh

# Dealing with Data Size in the Reporting Layer

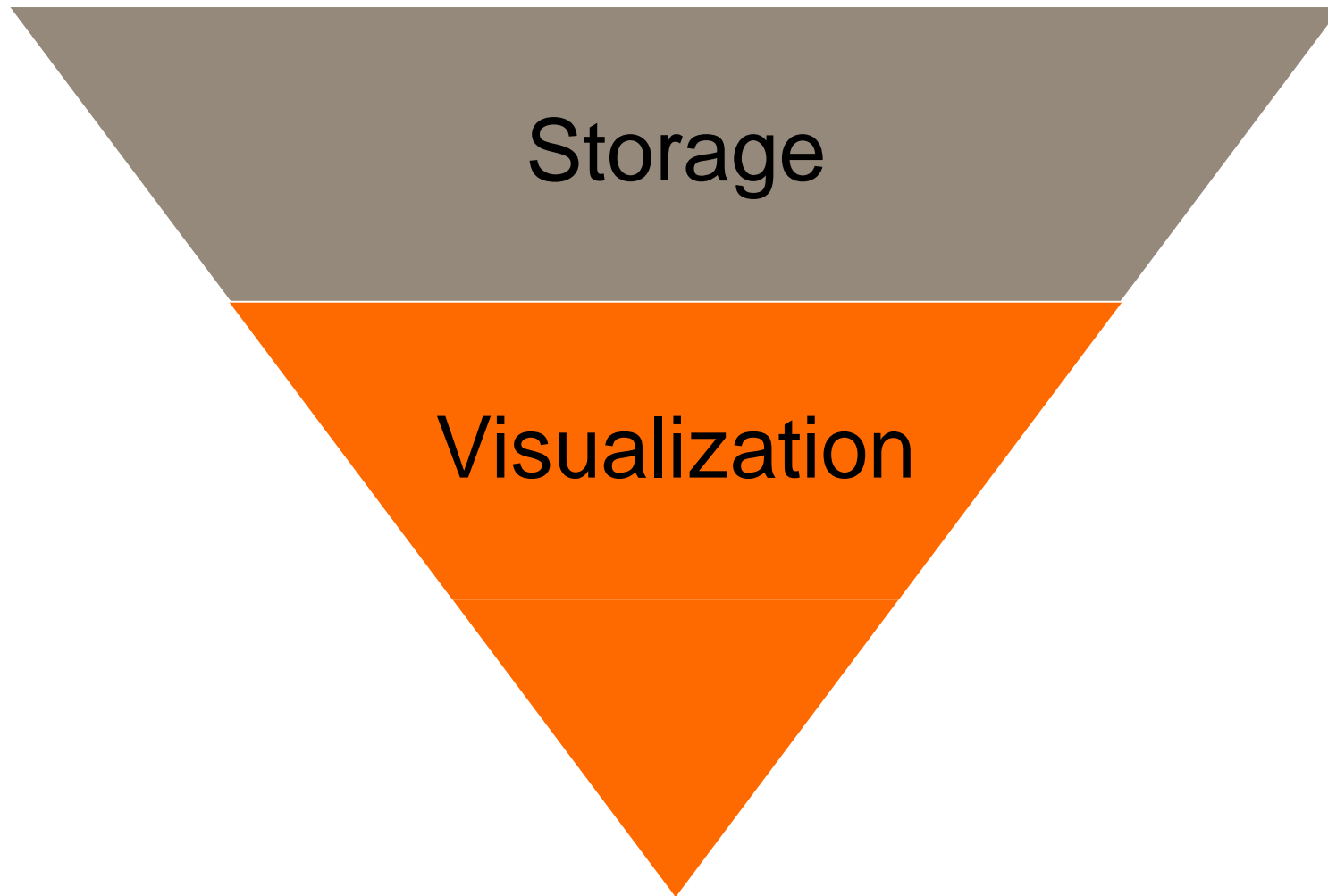




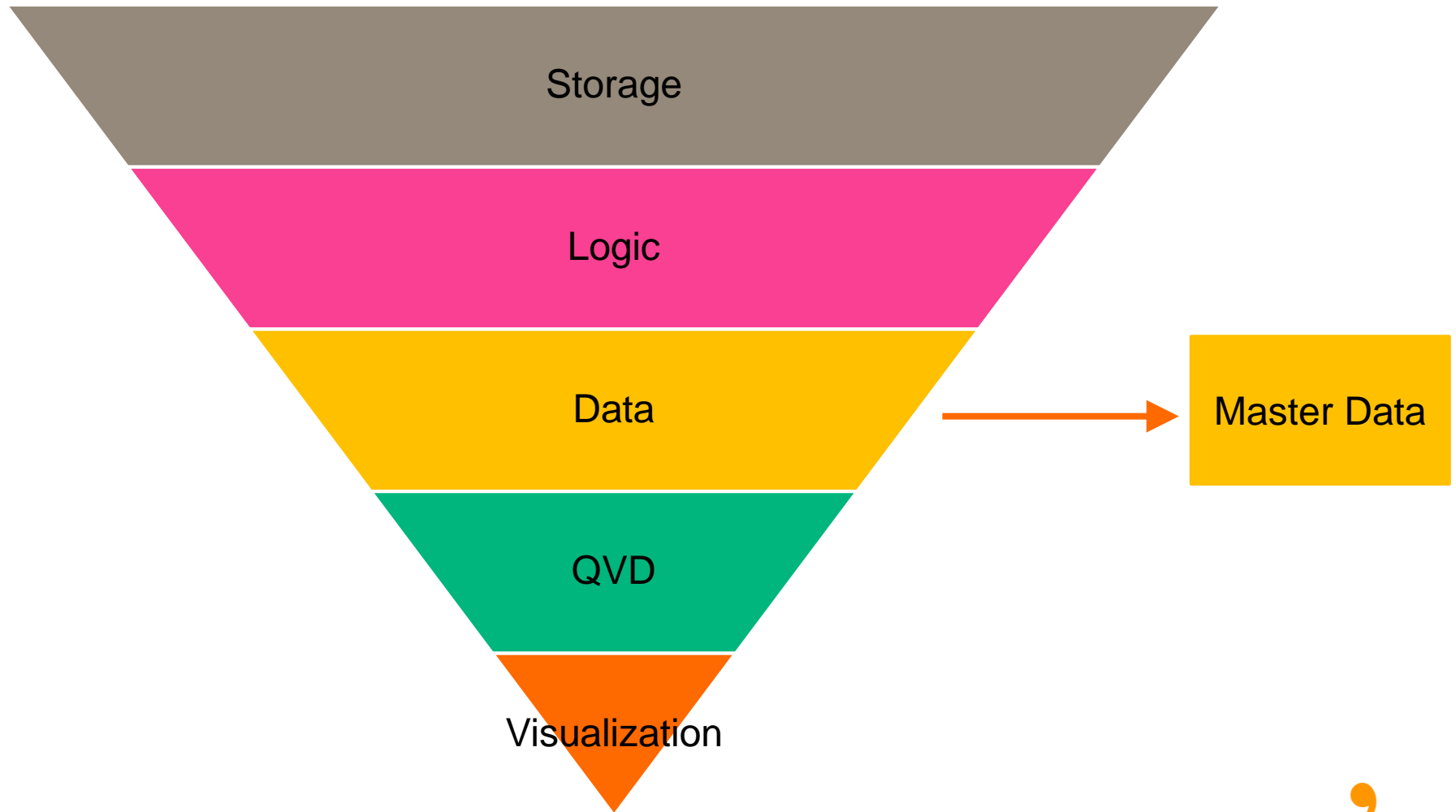
# Original architecture



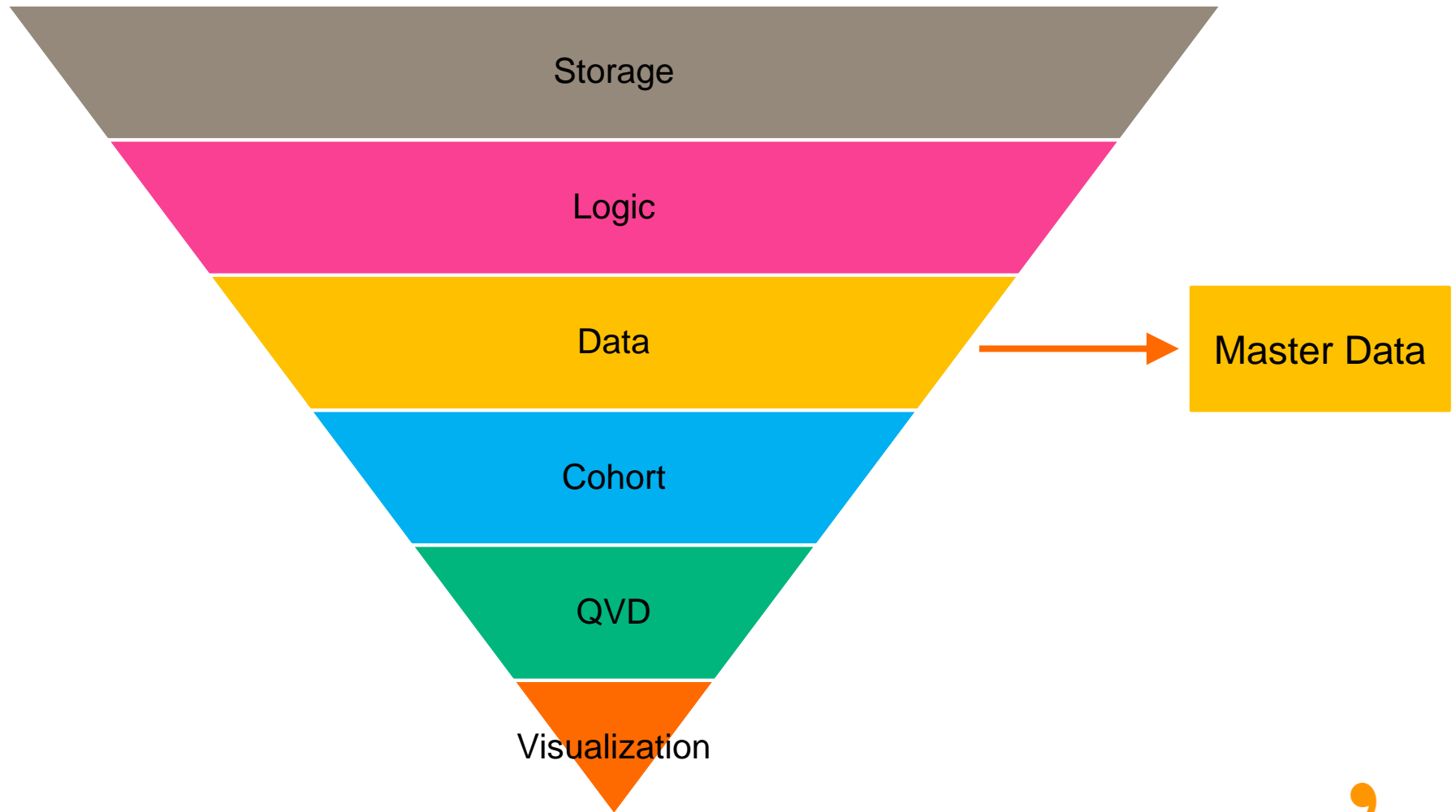
# Occasional architecture



# Architecture with a logic layer



# Architecture with a cohort layer



# User experience

- Performance
- Dependability
- Trust



# Governing the User End of a Big Data Environment



# Our Framework: What issues does it address?

- Governance
- Development overhead
- Scalability
- User engagement



# How do we do this?

- Global KPI expression dictionary
- Common code/tool sharing
- Holistic monitoring
- Governed folder structure
- Design consistency





# Tools – QV Helper


Introduction Charts QVD Sniffer

QV HELPER


Last reload: 2015-10-04 22:41:28 | Data from: 2012-06-01 | Up to: 2015-10-03

Q Search Script

Q Search Task



CURRENT SELECTIONS

Task Type  ReloadTask

Q Search filters

Task Name

Task Category

Task Type **ReloadTask**

DB source

Schema

Table

QVD Generator

QVD Reader

Folder

File Name

FWed?

QVD Date

Hive/Exasol Table Dependencies

QVD Dependencies

Map Dependencies

QVD Orphans

Show Script (select 1 task)

From Map

Map used by

01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\COUNTRY_AND_CONTINENT_2011-10-20.XLS	clv (x-platform) report; weekly forecasts report; x-platform user activity report
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\COUNTRYMAP.QVD	clv (x-platform) report; mobile_cross_promotion_report; weekly forecasts report; x-platform user activity report
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\EP_LVL_TO_SEQ_LEVEL.XLS	hive maps loader
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\EXTRAGAMEENDREASON.QVD	hive maps loader
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\FLAVOURMAP.QVD	clv (x-platform) report; mobile_cross_promotion_report; weekly forecasts report; x-platform user activity report
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\FORECAST_TYPE.XLS	weekly forecasts report
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\GAMEENDMAP.QVD	hive maps loader
01\PRIVATE_DATA\DATA STORAGE\MAPPING TABLES\SPINNER_TEMPLATE_MAP.XLS	cgu daily loader

# Tools – Master Portal

## MASTER PORTAL

**CURRENT SELECTIONS**

Metric Installs

Dimension Install Country, Install Date, Install Type

Audience Marketing

**Category**

- Crm
- Game Analytics
- KPI
- Customer Service
- External Data
- Internal
- Miscellaneous
- Player Experience Quali

**Subject**

- Churn
- Customer Value
- Dashboards
- Real Time Bidding
- Ab Tests
- App Annie Market Insig
- Browser
- Campaigns
- Expert Weekly Forecast

**Type**

- Cross Platform
- Mobile
- Qlikview Usage Insights
- Qv Helper
- Qvpr
- Reports
- Web

**Tags**

- %dpu
- 3rd party
- ab testing
- analysis
- android
- annie
- app annie
- attributed install

**Audience**

- Marketing**
- CRM - Customer Relationship
- Finance
- Game Teams
- All
- Argonauts
- BI - Business Intelligence
- CRM - Customer Relationship

**Business Owner**

- Jessica Hayes
- media
- Ryan Topping
- Alexander Von Roth
- Anders Johansson
- BPMS
- Brian Burford
- Carlos Caicedo (Creator...

Report Na...

Dimension

Metric

**Dimension**

- ☒ Install Country
- ☒ Install Date
- ☐ Install SubType1
- ☐ Install SubType2
- ☐ Install SubType3
- ☐ Install SubType4
- ☒ Install Type
- ☐ Install.date
- ☐ Install.weekday
- ☐ Install.year
- ☐ Install.yearmonth
- ☐ Install.yearquarter
- ☐ Install.yearweek
- ☐ install\_day
- ☐ install\_month
- ☐ install\_month\_muw
- ☐ install\_subtype5
- ☐ install\_year
- ☐ install\_yearweek
- ☐ Is Juego
- ☐ Is Payer

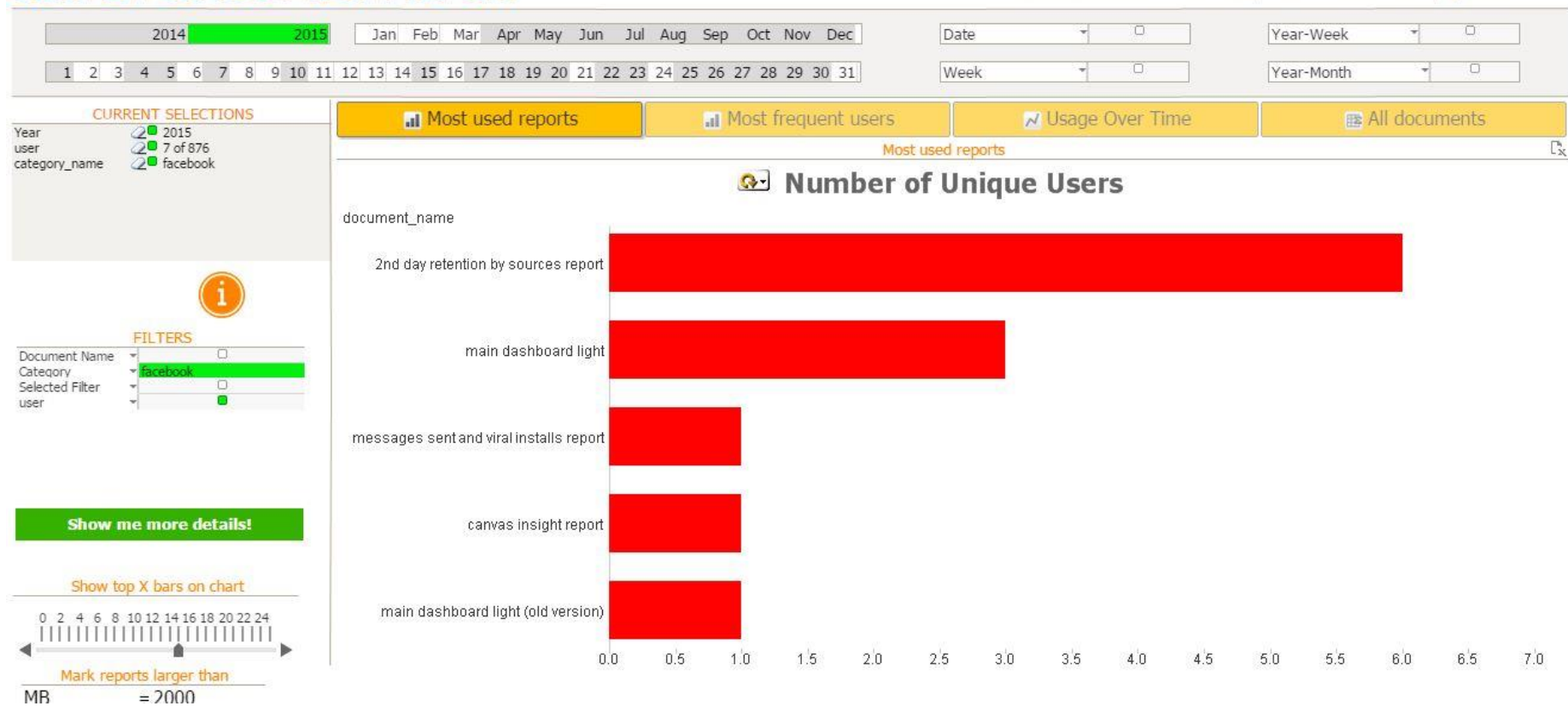
	Category	Subject	Type	Report Name	Audience	Click to open
	Crm	Real Time Bidding	Mobile	Mobile RTB Analysis Report	marketing	<a href="#">Mobile RTB Analysis Report</a>
	Game Analytics	Churn	Cross Platform	Cross Platform Rule Segmentation Report	Marketing and game teams.	<a href="#">Cross Platform Rule Segmentation Report</a>
	KPI	Customer Value	Mobile	Mobile Customer Value Report	Finance; Marketing	<a href="#">Mobile Customer Value Report</a>
	KPI	Dashboards	Mobile	Mobile Extended Dashboard Report	Marketing	<a href="#">Mobile Extended Dashboard Report</a>



# Tools – Usage Insight

## QLIKVIEW USAGE INSIGHT REPORT

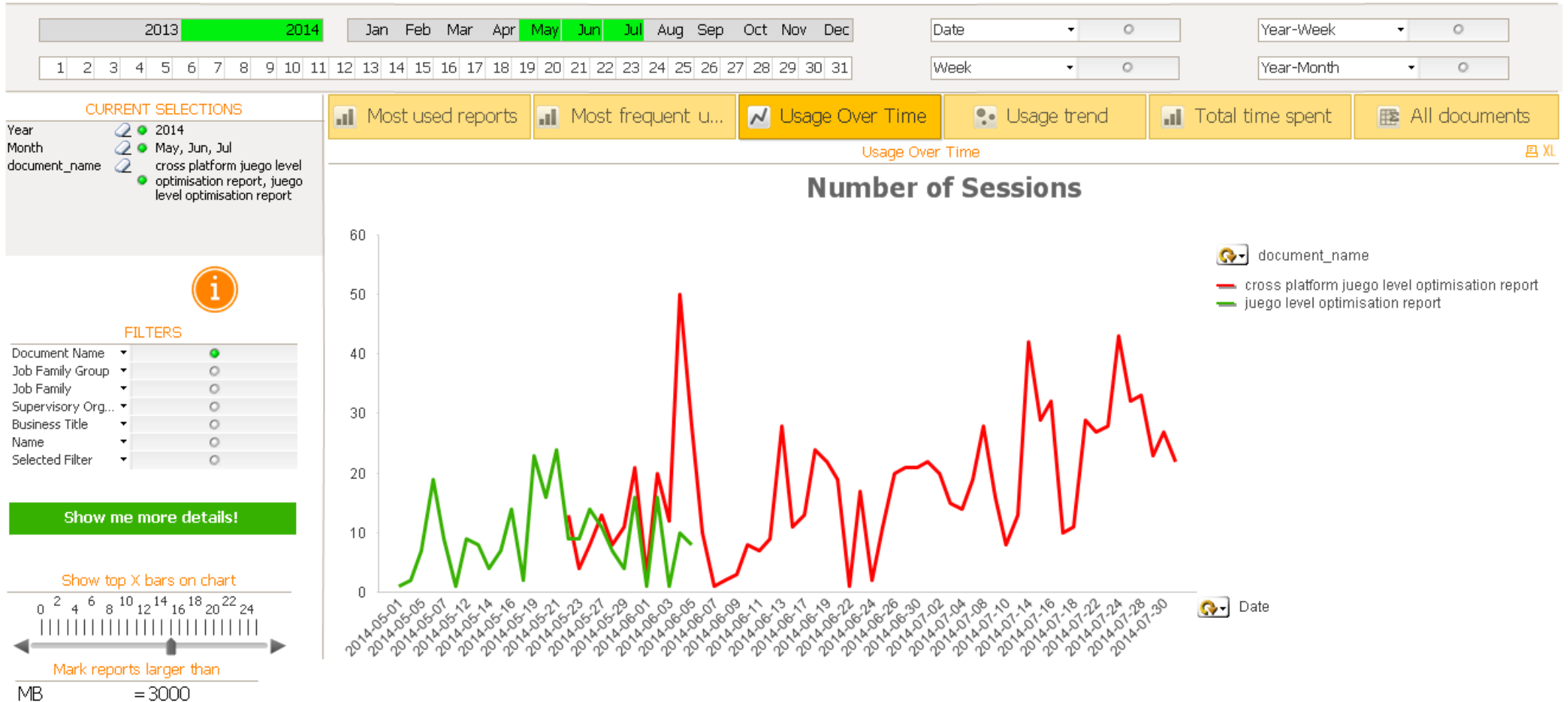
Last reload: 2015-10-05 02:07:26 | Data from: 2014-02-17 | Up to: 2015-10-04



# Tools – Usage Insight

## QLIKVIEW USAGE INSIGHT REPORT

Last reload: 2014-10-21 02:03:02 | Data from: 2013-07-01 | Up to: 2014-10-20



# Getting the Stakeholders the Right Thing



# Trade off between specialization and generalization

- Don't get locked in too much on how it should be
- Don't be afraid to try things, and iterate improvement
- Make sure that it serves the right purpose, not all stakeholders need all information – Smart Data
- Don't let future use limit what you develop and use today



# Questions?

[Anna.Bongenhielm@king.com](mailto:Anna.Bongenhielm@king.com)



*Please*

**Remember to  
rate this session**

*Thank you!*

