Getting personal with your customers and GDPR

A practical approach to a secure, governed 360 degree customer view

Darren Brunt – Presales Director UK&I, Talend
Colm Moynihan – Partner Presales Manager EMEA, Cloudera
Jean-Michel Franco – Sr. Director for Product Marketing, Talend
Disclaimer

GDPR is a complex and detailed regulation. Cloudera can only help address a portion of it.

This presentation is intended to help organizations understand how Cloudera software and services can be used to help comply with certain aspects of EU General Data Protection Regulation (GDPR) requirements.

Applicability of any of these capabilities depends on each organization’s own requirements specific to their business. Every organization should determine its own needs with regard to GDPR and then evaluate solutions for suitability to those needs. The information contained in this presentation is not intended to be and should not be construed to be legal advice. Customers and prospective customers must not rely on the information herein and they should obtain legal advice from their own legal counsel or other professional legal services provider.
Data is the new oil…

“Personalization can deliver 5 to 8* ROI on marketing spend and lift sales by 10% +”

…but beware of the oil spill

BREAKING NEWS

Data Leaks, Privacy Violations and Data Flaws

Equifax breach exposed data for 143M consumers

Last year’s privacy fines would be 79 times higher under GDPR
General Data Protection Regulation - GDPR

• Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and replacing the Data Protection Directive (95/46/EC). Implementation date: 25 May 2018

• Scope – organisations which have EU establishments where personal data is processed in the context of the activities of such establishment

• Extended to non-EU established organisations that process personal data about EU residents in connection with: i) the offering of good or services, or ii) monitoring their behaviour within the EU.

• Fines up to 4% of total global revenues or €20 million whatever is larger

• Includes obligations for both Data Controller and Data Processor
GDPR is a data management challenge

- Right to be Forgotten
- Data Breaches
- Data Portability
- Access Requests
- Accuracy
- GDPR Data Hub
- Consent

EU citizens
Employees
Poll 1: Where are you in your GDPR journey? (one response)

a) Not started
b) Conducting assessment
c) Building the data map and data hub
d) Operationalising compliance (consent, access requests, right to be forgotten…)

Load and transform GDPR data into Cloudera

Portability

Data discovery, integration & transformation

Structured
Billing and Ordering
Supply Chain and CRM
Enterprise

Semi-Structured
Apps and Clickstreams
Log Files
Machine and Sensor Data

Unstructured
Social
Other Sources
Accelerate data ingestion, reconciliation, cleansing and stewardship with Talend Data Fabric.

Native on Hadoop, unified and open-source.
Consent management – individual plus children

MDM & consent management

- Single person view
- A complete, accurate & consistent view of person data
- J.Q. Jones
- J. Quincy Jones
- J. Jones
- Lucy Jones

Children

Customer
Employees

Rights & consent management

Web form

My Details
- Personal details
  - Name: Colin MacPhail
  - Contact details
    - Address: 41 Parkmore Drive

Mobile app

Pros
- View details
- Manage permissions
- Consent

Consent

EU citizens

Update
What’s the role of MDM in the GDPR data hub?

A registry that links to detailed data within and beyond the data lake

A system of record to match all personal data within the 360° view

A rule based engine that propagates updates beyond the data lake

An audit trail that delivers record level lineage
Data Subject 360 and all their data

All EU citizen data
Files, PDF, Word, email
Social, voice, video
Structured data, Oracle
Third party data

Access

360 View & consent management

Single person view
A complete, accurate & consistent view of person data

J.Q. Jones
J. Quincy Jones
J. Jones
Lucy Jones
Children

Customer
Employees

Linked
Seamless Business user, Data Subject and API access

- Business view
- Linked
- Data Subject view
- Developer view

All EU citizen data
- Files, PDF, Word, email
- Social, voice, video
- Structured data, Oracle
- Third party data

Data Subject 360

Consent management

cloudera
Access request to all data – cloud & on-premise
Give the DPO & team self service access to all data

Access

Governance & lineage
Enterprise search
Data analysis

CORE SERVICES
DATA SCIENCE
ANALYTIC DATABASE
OPERATIONAL DATABASE
DATA ENGINEERING
EXTENSIBLE SERVICES

SECURITY
GOVERNANCE
WORKLOAD MANAGEMENT
INGEST & REPlication

DATA CATALOG

STORAGE SERVICES
Amazon S3
Microsoft ADLS
HDFS
KUDU

Data prep
Data stewardship
Business glossary

Web form

DPO
Req. access

EU citizens
Email

Employees

Data stewards

Delegated
360 Enables data portability – download all

Portability

All EU citizen data
Files, PDF, Word, Email
Social, voice, video
Structured data, Oracle
Third party data

Data Subject 360
Consent management

‘Download all’ request
Send all GDPR data

Cloudera

EU citizens
Allow EU citizen full edit rights on their data
Reduce data breaches
Use next generation cyber security powered by Cloudera

- Apache Spot - cyber security done at petabyte scale
- Real-time alerting of breaches
- Built using Cloudera

- Spot expedites threat detection, investigation, and remediation via machine learning and consolidates all enterprise security data into a comprehensive IT telemetry hub based on open data models.

- Partner ISV – SYNPR Securonix built on Cloudera
- Log management, next gen – SIEM
- User Entity Behavior Analytics - UEBA
- Insider threat management
- Real-time reporting and notification
- Fraud detection in one system
- Lineage back to source with Cloudera Navigator
360 Enables right to be forgotten

Delete or mask at source

EU citizens

End 2 End Metadata Management
Maps All Data from Hub to Source
All GDPR Tagged with GDPR Tags
Source system, Table, Key stored
Deliver accurate data lineage – always!

Talend Metadata Manager & operational processes

Allows to harvest, govern, and share metadata across data platforms (Big Data, Applications modelling tools, ETL, BI, etc.)

Provides:

- **Business Glossary**
  - Shared definition and collaboration around data definition and policies
- **Enterprise Data Governance & Auditing**
  - Data Flow Lineage, across systems and tools
- **Automatic harvesting with change management**
  - Through model versioning, impact analysis, change notifications and reporting

Leverage metadata insight

Develop operational processes

Assured execution & consistency
Use search to ensure everything was deleted

Cloudera Search

Full-text search

All data searched
Emails, PDF, Structure data, messages

Delete
Poll 2 : What are your main challenges for GDPR compliance?  
(3 responses max)

a) Consent  
b) Data breaches  
c) Right to be forgotten  
d) Data portability  
e) Data Subject Access Request
GDPR Demo

Challenges
(source Deloitte GDPR Survey)

- **10%** Believe they current Consent are accurate
- **11%** Have a plan to automate Data Subject Access Request
- **64%** Are yet to prepare estimations on how many Requests for Erasure they’ll receive

GDPR Data Hub Solution

- Master the Data subject 360° view with consent
- Open your data through a Data Subject Portal
- Track, trace and erase data subjects footprint across the data lake
Demo Part 1: Master the Data subject 360° view with consent

Where can I find the opt-in data?

How can I share opt-ins across all customer facing systems?

IT architect

Metadata Manager

Talend MDM

Know where your data is and deliver it across systems
Demo Part 2: Deliver a Data Subject Portal

How can I review and port my personal data?

Data Subject

MDM Data Platform

Leverage governed data as a single, consistent source of personal information
Demo Part 3: Manage the right to be forgotten

IT architect

How can I establish the right to be forgotten?

Talend MDM

Managed governance

Enterprise Data Hub

Managed, cascading governance with audit trail down to the finest grain
Right to Be Forgotten

Trigger Workflow to delete all data on Hub and then back to source

- Fs
  - Delete
    - /user/customer_data/

- Shell
  - impala_delete_customer.sh

- Java program
  - sap_delete.jar

- Email
  - To: sfdc-admin@abcompany.com
  - About: Please delete Customer Data ID = 544338899, Name = Colm Moynihan
  - 35 Parkmore Drive
The GDPR Data Hub: Powered by Cloudera & Talend

DATA MAPPING & TRACEABILITY
ACCOUNTABILITY & STEWARDSHIP
DATA ACCESSIBILITY

DPO  Data Stewards  Employees  EU citizens

METADATA MANAGEMENT
DISCOVERY & PROFILING
STEWARDSHIP & CURATION
SEARCH
DATA ANALYSIS

DATA SECURITY & BREACH NOTIFICATION
RIGHTS & CONSENT MANAGEMENT
ANONYMIZATION & PSEUDONYMIZATION

RIGHTS OF ACCESS & PORTABILITY
RIGHTS TO BE FORGOTTEN & RECTIFICATION

SaaS
ERP
Documents
File System
SQL Store
Streams and IOT
Bus. Intelligence

The GDPR Data Hub [Diagram]
What’s next in your GDPR journey?

Visit our Solution Pages


Self-assess your readiness: http://talend.gdprevaluation.com/

Create your GDPR Data Hub & Map
Set accountabilities & orchestrate data governance
Address Consent and Data Subject Access Requests

Email Cloudera if you have any current GDPR Challenges gdpr@cloudera.com
Email Talend if you have any current GDPR Challenges webinar@talend.com
Thank you