SmartData Fabric® Use Case
EU General Data Protection Regulation (GDPR)
With and Without Data Security*

August 2019

*GDPR includes severe penalties for personal data breaches/theft

GDPR compliance is at the intersection of core data management

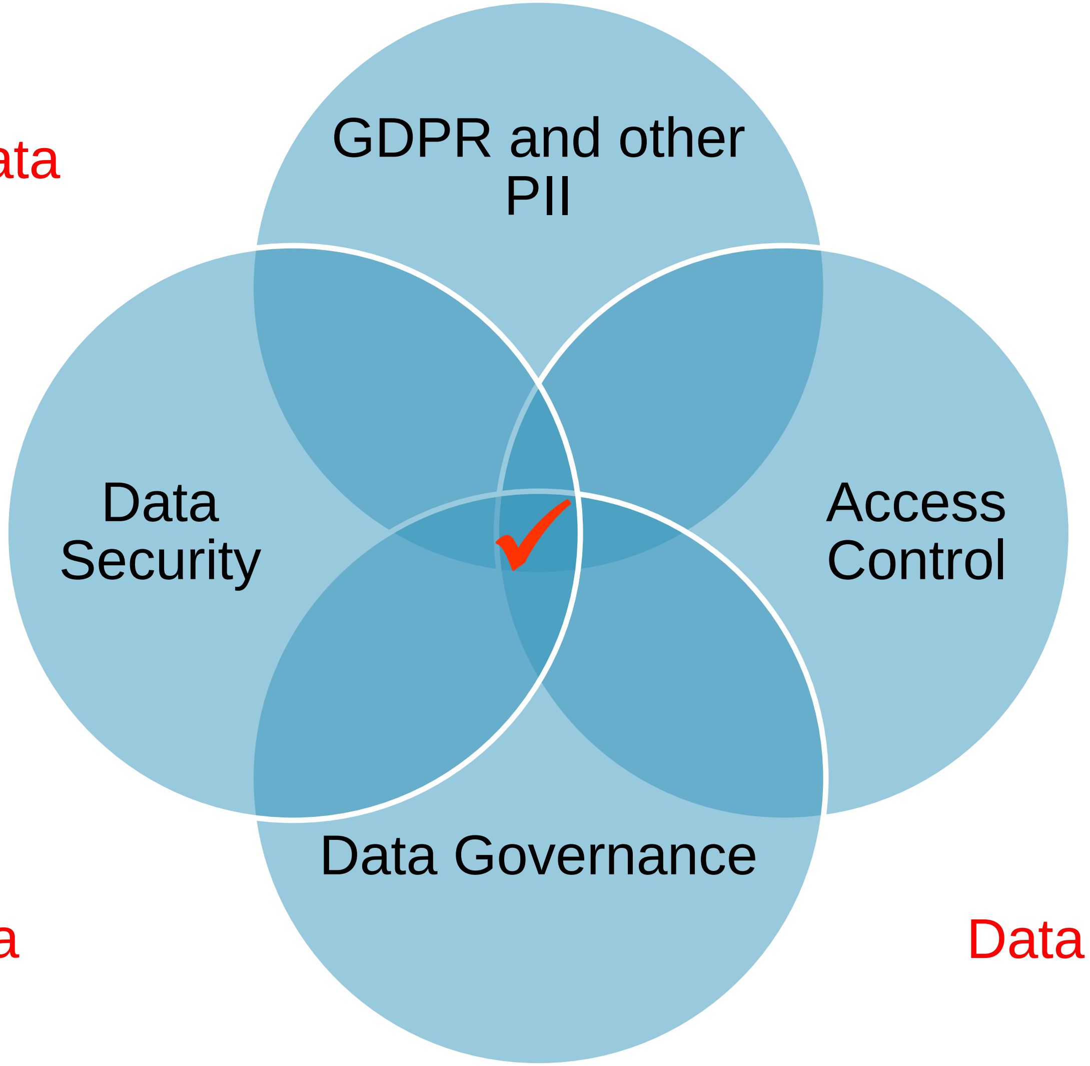
Structured Data

Unstructured Data

Reference Data

Master Data

Graph Data



Data Discovery

Data Classification

Data Masking, Tokenization and/or Encryption

Data Integration

Data Retention and Disposal

There will be
RELATIVELY FEW CUSTOMERS
who will want to see their personal data and even
FEWER CUSTOMERS
who will want their personal data “erased”

Relatively long time regulated to perform + Limited scale =
Ad hoc ID resolution/MDM
(vs. a large-scale pre-built MDM solution)

WHAMTECH®

Keys to successful GDPR compliance and data security

Process Step	Technical Comment
• Discover where all data resides	• Device, data source and data discovery
• Leverage a standard data model	• Standard data view of entities, attributes and relationships
• Profile and classify all data	• Metadata gathering and semantic mapping to a Standard Data View
• Secure all personal data	• Decision on data masking, tokenization and/or encryption
• Develop data transforms	• Data cleansing, transformation, standardization and security
• Build and maintain indexes – content, link (and master data)	• Initial batch and then, incremental/near real-time through CDC/polling
• Enforce access control	• AD/LDAP, SSO, multidomain (virtual directory), RBAC, RLS/ABAC, etc.
• Match data within and across data sources	• Deterministic and probabilistic
• Match and merge key entities	• Identity/master data management (batch or ad hoc/on demand)
• Logically combine source data	• Data integration
• Connect and present all personal data to a customer, and allow all or some to be "erased" by customer	• Interactive 360 single person view, e.g., graph visualization and dashboard, and file download option
• If customer wants complete or near-complete "erasure", provide a unique "tracking number"	• Log "tracking number", status and date-time
• Perform customer "erasure"	• Soft or hard delete, tokenize, deny access/use, etc.
• If provided, allow customer to enter "tracking number" to see status	• Report of "erasure" progress

WHAMTECH®

Some of the information needed in addition to personal data

- Purposes of the processing for which the personal data are intended as well as the legal basis for the processing
- Recipients or categories of recipients of the personal data, if any
- Where applicable, the fact that the controller intends to transfer personal data to a third country or international organization, with reference to the appropriate suitable safeguards and the means by which to obtain a copy of the personal data transferred or where it has been made available
- Period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period
- Existence of automated decision-making, including profiling, and meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject

INDEXES are key to obtaining a complete understanding of all data, enabling processes and driving value

Including being able to discover, access, secure and "erase" GDPR-defined personal data

Source: Part from #WhamTech SmartData Fabric Power of Indexes

GDPR is an extension of SmartData Fabric®

Not all capabilities required, but many related to indexes are

Use Case: EU General Data Protection Regulation (GDPR)

- EU GDPR is being adopted worldwide, including a few US states
- US companies with an EU entity or employees must comply
- Severe fines of up to 4% of gross worldwide revenue
- Personal data definition is very broad and includes cookies, email and IP addresses
- Includes private, public and work data
- Marketing is opt-in, not opt-out

	THE CHALLENGES	THE SOLUTIONS	THE BENEFITS
BUSINESS SUMMARY	• Difficult to know where private data is located • Difficult to determine if data can be readily used to identify individuals • Broad definition of personal data • Impacts marketing • Severe fines	• Discover and profile data • Ensure any and all data found • Mitigate risk data • Enable individual opt-in to marketing and opt-out to data • Generate compliance reports • RBAC-based data security • Establish auditable processes	• Addresses a massive regulatory problem • Minimally intrusive solution • Rapid and incremental deployment • Could be a short-term fix that tends to longer-term solution • Automatically connect people to their personal data • Explore additional risks • Open up new channels of customer interaction • Could be seen as competitive advantage
TECHNICAL SUMMARY	• Multiple disparate systems • Missing single person views of all related data • Poor data quality – avoids detection • Poor access control • Poor data anonymization • No way to anonymize or delete opt-out data	• Use indexes to discover data • Dynamic data masking, tokenization and encryption • Single person views of data • Individual opt-in and opt-out, and write back to systems • Index or data anonymization, or delete data for opt-outs • Graph database/visualization • Could be Cloud-based	• Advanced data-centric systems now available • Potential Cloud management of on-premise systems access

The End