

Are you ready for Data Privacy?



Most organizations are aware of the stiff penalties associated with data privacy regulations, passed to force organizations to deal with the barrage of data theft occurring around the globe. But, according to Gartner, more than 50% of organizations will not be able to meet their data privacy mandates (GDPR, PrivacyShield, IDP, Pipedata, etc.)



Data privacy regulations are not limited to protecting your data stores from intrusion within your four walls. Regulators assume you will eventually be compromised, either from outside the organization or from within. The new regulations demand that Personal Data maintained in your environment, or kept by others for your use, be obscured, removed or encrypted so that it will have little value to individuals when it is illegally obtained. Personal data falls into two categories.

Direct Identifiers, single fields that directly identify an individual, and Indirect Identifiers, fields that when combined can identify an individual.

Indirect Identifiers may appear across multiple files and constitute a violation when dispersed anywhere in your data stores.



The least understood, yet most demanding requirement, of these regulations, first requiring implementation within the European Union, is that an individual may exercise their right to be forgotten, requiring erasure or permanent obfuscation of personally identifiable information stored about a consumer.



What this means to you – in simple terms

Using GDPR (European Union) as an example

If you have any Data Subjects (Consumers, Employees, Suppliers, Partners, Citizens, Patients etc.) sharing their identity with you from Europe or while traveling in Europe, you have to take the new GDPR regulations seriously, which requires you to be able to identify all the information you have about a Data Subject, whether identified by key data items (national insurance number, tax payer id, name, address, credit card numbers, etc.) or indirect information that allows a Data Subject to be identified when multiple fields are grouped together (address, professional affiliations, HIPAA and other medical records or any other information that allows the identification of an individual).

GDPR calls for an individual to have the ‘Right to be Forgotten’ , the Right to Erasure of data that has been entrusted to you, whether stored within your four walls, in the cloud, or stored with partners who participated in your digital ecosphere (PayPal, Priceline, Saavis, Amazon, Google, or any other firm which houses information on your behalf or on the behalf of your platform partners storing Data Subjects information in their environment). If your consulting partner has told you that they have you covered and they will build a database to house all your Data Subject information so that they can be forgotten, you will be unfortunately surprised when the regulators impose stiff fines due to your non-compliance. Organizations have until May 2018, or less than a year to implement a program that allows Data Subjects to request to be forgotten.

What must you do to comply

1

Identify where in your portfolio of systems within your four walls, within cloud environments you use, within your backups and within partner environments are information that uniquely identifies a consumer, either directly or indirectly.

2

Develop a plan to secure personally identifiable information sufficiently documented to satisfy regulators and execute the plan, documenting milestones achieved, again to satisfy regulators

3

Take the necessary steps to protect personally identifiable information from being stolen or altered inappropriately through a well devised program

4

Have a facility that details where information about a consumer is located so that should a consumer request to be forgotten, you can fulfill that right within 48 hours

What must you do to comply (continued)

To meet your privacy obligations as imposed by GDPR, PrivacyShield, IDP, and a host of other privacy regulations surfacing around the globe, some key components which your consulting partner forgot to tell you about are also required:

5

If a system change was implemented which stores credit card information in an address field because it was too hard to change the system, you must be able to identify and delete this information.

6

If you are stripping information from Facebook, Twitter, Instagram, or other social media sites, you must be able to identify and delete this information when requested by an individual exercising their 'Right to be Forgotten'.

7

And if you use Amazon as a sales platform, PayPal as a billing or ShipStation as a distribution platform and they store information on your behalf, you must be able to delete or obfuscate information that personally identifies a consumer.

What are the global privacy regulations and why are they coming onto the scene?

<u>WHY</u>	<u># of consumers effected</u>	<u>What</u>	<u>Where</u>	<u>What are the consequences</u>
Equifax:	143 million	GDPR	EU	up to 4% of revenue per occurrence
Yahoo	500 million	PrivacyShield	USA	July of 2016, the European Commission deemed the EU-U.S. Privacy Shield Framework adequate to enable data transfers under EU law
Sony	77 million			
Target	70 million	Cybersecurity Law	China	June 2017 China's new Cybersecurity Law is one of the most important pieces of privacy and cybersecurity legislation of 2017
Ashley Madison	37 Million			
eBay	145 Million	PIPEDA	Canada	Personal Information Protection and Electronic Documents Act (PIPEDA), the federal private-sector privacy law.
LinkedIn	164 Million			
Anthem Health	78 Million			

All of these regulations are coming onto the scene to force companies to take the protection of personally identifiable information (PII) entrusted to them in order to participate in the digital economy.

In order to address GDPR you must...

- 
- Have a means of Identifying Personally Identifiable Information (PII) on File at a moment's notice

- 
- Provide a means to protect PII data (that directly or indirectly identifies an individual) from cyber-threats

- 
- Identify cross-file indirect identification of individuals who have entrusted their identity to you and provide a means to protect that data from cyber-threats and wandering eyes

- 
- Be able to meet the demands of a citizen exercising their GDPR right to be forgotten from all points of identification, whether on your premises, on a backup environment or in a partner's environment storing PII data on your behalf (e.g., PayPal, Shipstation, Amazon, eBay, DoubleClick, Eventbrite, Hotels.com, etc.)



Addressed through our Intelligent Catalog



Addressed through our Sequester & Encrypt processes which employ the intelligent catalog



Employ the functionally rich metadata created with the Intelligent Catalog

How we approached the "Right to erasure ('right to be forgotten')" and Indirect Identifiers

- 
- We use Hadoop as a staging area for interrogating data for privacy concerns. Hadoop removes the challenges that hide privacy concerns due to the complexity of data

- 
- We use our intelligent catalog mechanisms codify rules that define patterns of data that represent privacy concerns and schedule the interrogation of files and streams of data

- 
- We use pattern detection to identify potential privacy concerns and use a process we call sequester, encrypt and secure, which encrypts the exposed data so that privacy information is not in harms way and sequester it into a highly protected environment

- 
- We allow the creation of false positive lists so you do not have to revisit potential issues as new data enters your environment

We store the metadata necessary for you to respond to regulators and consumers

• The privacy regulations being passed around the world (GDPR, PrivacyShield, IDP, Cloud-A) call for an ability to prove to regulators that you can administer the right to be forgotten. That means you can identify where a customer's data exists at a moment's notice.

• We create a functionally rich metadata layer as part of the intelligent catalog which we use for what would otherwise be an administrative nightmare.

• We also provide you with the ability to enrich our metadata by augmenting it with metadata from other sources and tools, such as your technical data integration (ETL) tools, your compliance engines, your regulatory reporting engines and any other source you feel aids in your ability to respond to consumers and regulators in meeting the obligation to execute the right to erasure now endowed upon consumers.

How to integrate us into your information fabric

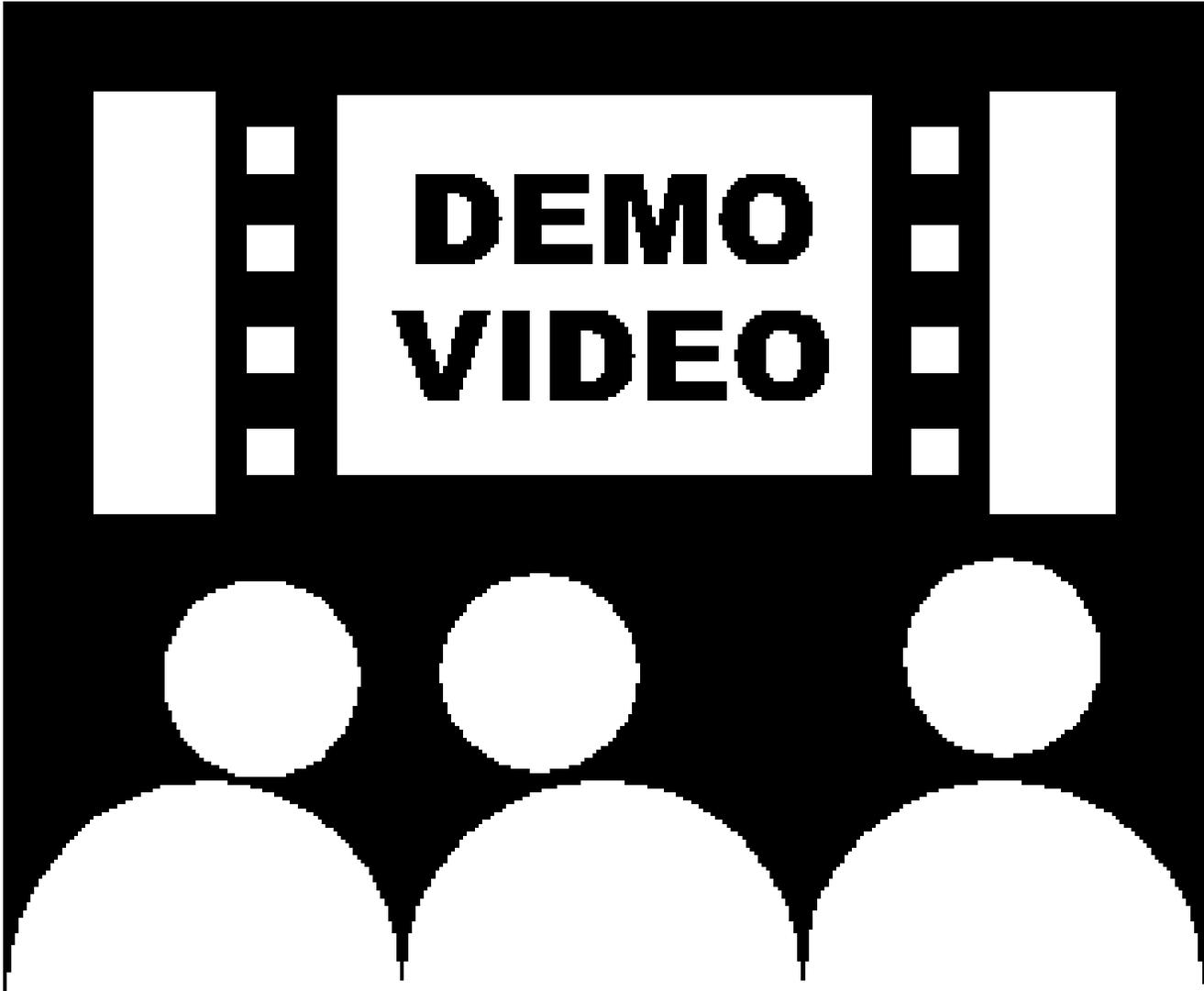
BigDataRevealed's Intelligent Catalog/Metadata and The Simple to use Callable API's integrate into most any existing application and workflows

BigDataRevealed is a company whose mission is to deliver tools to improve the usability of the big data environment.

BigDataRevealed's flagship product, the intelligent catalog is the central core of all capabilities offered, including out of the box analytic capabilities extended to the big data environment. BigDataRevealed is driven by a core team relentless in devising capabilities not offered through the core capabilities available from the traditional big data vendors. One of these capabilities is the Secure/Sequester and Encrypt facilities, which extends the intelligent catalog through processes devised to ensure the identification and capture of potential PII issues whether introduced to big data through data feeds, real time data streams or other means.

All of BigDataRevealed's Intelligent Catalog / Metadata, Discovery, Secure/Sequester/Encryption are callable API Modules (JAR files), fully documented with the call parameters and where and how to obtain the results of jobs your streams have called, or jobs BigDataRevealed has already ran and stored the results.

<https://vimeo.com/234330294>



**DEMO
VIDEO**

A short
demonstration of
the GDPR Gotcha's
delivered with
ease

Next steps

- Remember all those recent studies showing a shortage of over 2 million Data Scientists and Data management individuals that are needed to meet EU GDPR and other Compliance Regulations? With an Application like BigDataRevealed to perform the great majority of tasks, we believe companies can utilize current resources and be prepared to pass an audit in a matter of months. Without such an application your clock may have already run out for the May 2018 implementation of GDPR.
- BigDataRevealed offers Free Trial use of the technology and has resellers with trained resources ready to work on your companies regulatory compliance now. BigDataRevealed and it's resellers and services Partners can step in, conduct 3-5 day audit assessments, share this with you and get started asap on delivering results in an Agile methodology. Your C levels can and will see results in days and gain the comfort, confidence that regulatory compliance is achievable.

BigDataRevealed is reachable at info@bigdatarevealed.com or call (847) 440-4439

Appendices

The Article 17 EU GDPR regulations

A Technology overview for your technical team

"Right to erasure ('right to be forgotten')"



(1) The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:

- (a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;
- (b) the data subject withdraws consent on which the processing is based according to point (a) of [Article 6\(1\)](#), or point (a) of [Article 9\(2\)](#), and where there is no other legal ground for the processing;
- (c) the data subject objects to the processing pursuant to [Article 21\(1\)](#) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to [Article 21\(2\)](#);
- (d) the personal data have been unlawfully processed;
- (e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;
- (f) the personal data have been collected in relation to the offer of information society services referred to in [Article 8\(1\)](#).



"Right to erasure ('right to be forgotten')" Cont.

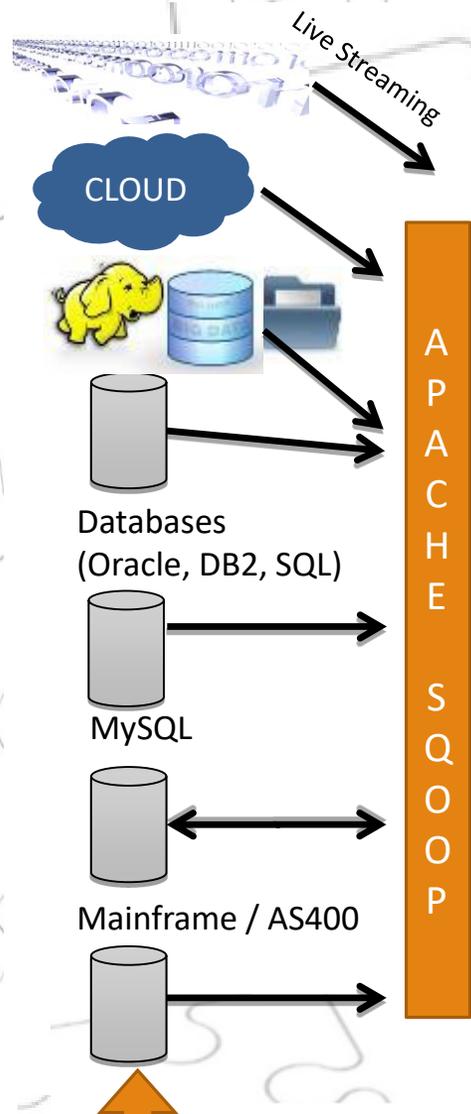
(2) Where the controller has made the personal data public and is obliged pursuant to paragraph 1 to erase the personal data, the controller, taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures, to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data.

(3) Paragraphs 1 and 2 shall not apply to the extent that processing is necessary:

- (a) for exercising the right of freedom of expression and information;
- (b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
- (c) for reasons of public interest in the area of public health in accordance with points (h) and (i) of [Article 9\(2\)](#) as well as [Article 9\(3\)](#);
- (d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with [Article 89\(1\)](#) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing; or
- (e) for the establishment, exercise or defence of legal claims.

BigDataRevealed Architecture For GDPR / Legacy - for All Regulatory Compliances, Powered by Apache™ Hadoop®

Slay the GDPR Dragon for Both Hadoop & Legacy Systems with BigDataRevealed The Intelligent & Quickest Path to jump start your GDPR & Regulatory Compliance

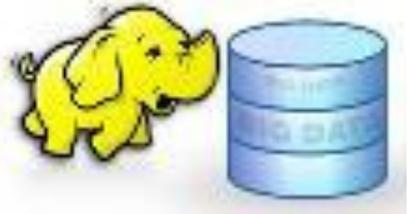


A P A C H E S Q O O P

OR

B D R - S p a r k S t r e a m i n g

Apache™ Hadoop® Staging ODS



Keep your existing legacy systems functioning without disruption or degradation and begin meeting the demands of GDPR and other Regulatory Compliances, using Apache Hadoop as your central Data / Operational File Store.

With BigDataRevealed : For Existing Hadoop Data Lakes or Staged Legacy Data to create Catalog / Metadata for legacy Compliance usage

- Discover Personally Identifiable Information (PII) by searching **every column in every row**. We don't use a randomizing algorithm that searches only a fraction of your columns looking for undiscovered PII data.
- Encrypt PII wherever it is found
- Process Streaming data
- Allow Data Scientists to drill down and view suspected PII data in any column
- Sequester the Original file in a Hadoop managed Encrypted Zone for reference
- Remove files and historical versions where PII was discovered
- Provide Workflow Management screens for task assignment and completion control
- Provide an Intelligent Catalog/Metadata for:
 - o collaborative efforts and file/columnar naming
- Discovery for Peoples rights to be forgotten
- Provide an Intelligent Catalog/Metadata for:
 - **Indirect file matching to determine Indirect Identifiers**
- The use of Hadoop Encrypted Zones for additional sequestering of sensitive Data

B D R - A p a c h e H a d o o p E n g i n e s

P r o c e s s H a d o o p D a t a L a k e

P r o c e s s S t a g e d L e g a c y D a t a

Use the BDR-Intelligent Catalog/Metadata to revert back to your Legacy Data for remediation for GDPR and other Regulatory Compliances



BDR-Intelligent-Catalog/Metadata

