Who will help you Survive the Age of Data Regulation?
- ## China's new cybersecurity law took effect June 1, 2017
- ## The EU will enforce their Regulations (GDPR) May of 2018
- ## EU and US Privacy Shield Regulations now being Released

BigDataRevealed is built to assist in todays World of Data Regulations as well as Tomorrows

The premier application for identifying Person or Private Information, Sequestering/Securing/Encrypting solutions. BigDataRevealed's flagship product and it's Intelligent Catalog / Metadata is the central core of all capabilities **Plus Discovery of Indirect Identifiers both Intra-File and Cross Files!**

There is overlap, confusion, heavy fines and even criminal penalties identified for those who violate these regulations. There is a world full of advancing regulations, but not an abundance of solutions. You will need a very nimble and informed IT department to keep up with just one country's set of regulations, let alone all three and the others that will follow. Sophisticated software will be required to protect against excessive fines, or from the embarrassment of poor data governance being publicly disclosed. Is your company capable of swiftly constructing such sophisticated software? Remember Google was just fined $2.7 billion dollars by the EU for violating their laws. I fully believe the EU will be very willing to enforce their General Data Protection Regulations (GDPR), and the cost of non-compliance may be severe.

You may be a very informed company and have begun to look for available software packages to ease the burden on your IT staff. If you have started that search, you represent a small percentage of companies. And I will also assume you have been very disappointed with the quality of software you have discovered in the market place. Most available products offer a partial solution, a solution designed for a single data platform, or a solution requiring massive support from expensive technical staff.

What is needed, and needed Now, is an application that performs the vast majority of tasks and extensibility necessary for regulatory compliance. You may ask what are those tasks? Each country mandates that you fully inform clients, customers, business partners etc. as to your intended use of the data you collect from them. This would include, but is not limited to uses such as marketing research, re-selling their data to others, targeting promotional material to them, and many other uses. I am sure your legal staff is capable of developing material that will meet these regulations. We therefore will concentrate on regulations dictating how a company must protect data from misuse or theft. Below is our understanding of the major data collection, storage, and processing requirements needed to comply with regulations;

- Personally, Identifiable information (PII) must be protected, either by removal, Data Swapping, Data Reduction or simply encryption. Data Swapping and Data Reduction are difficult techniques used mostly to deliver research findings. Encryption is our preferred method to protect PII.

- **The Right to be Forgotten. The EU requires the removal of all PII data from your databases if an individual makes such a request. Finding all this data amongst the disparate data stored these days is much easier thought about than delivering upon,**

Looks like a small list that should be easy to accomplish, correct? Let's look at it more closely and then decide how easy it will be.

- **What is PII data really? It is any Direct Identifier, or single data column, that uniquely identifies an individual; such as name, address, Tax-id, bank account etc. PII also includes any combination of Indirect Identifiers, which are single data rows that would uniquely identify an individual when combined; such as Date of Birth, Gender and Mailing Code.**

- **The unfortunate reality of PII is that no definitive list has been published to give you confidence that all forms of PII contained in your various data stores have been removed or encrypted. Your software must be flexible enough to identify additional fields or combination of fields, at any time, as new PII are understood.**

- **Another troubling aspect of <u>Indirect Identifiers</u> is that they might not be in a single file to be considered a violation. If you have a person's Gender in one file, their Mailing Code in another file and their Date of Birth in a third file it would be considered a violation of GDPR. This is described as <u>Cross File Indirect Identifiers</u>. As of this time we know of no vendor that even claims to address these violations, except for BigDataRevealed.**

- **Please keep in mind these regulations are not limited to your main processing systems, they include data contained in PDF files, Emails, Word Documents and any other form of unstructured data maintained by your company.**

What will your IT infrastructure look like in order to find all the PII data hidden in your multiple data stores?

- **Finding a product that will discover PII data across all your various platforms would be ideal. You don't want to struggle using, maintaining and integrating software for each of your legacy systems, such as IBM Mainframe, Teradata, and AS400; and then another set of products for SQL Server, and another for PDFs and emails, and even a set for IoT.**

- **How will you preserve your production systems speed, reliability and integrity while complex pattern searches and cross file data comparisons are being executed.**

- **We believe the sensible approach would be to use a single platform that accepts any type of digitized data from all of your various data stores. A platform with enormous storage capacity to house all the data being analyzed and all the intermediate tables that are necessary to fully complete Indirect Identifier discovery. A platform that is extensible and capable of providing incredible processing power. We believe that choice should be Apache Hadoop.**

- **You may need an ETL tool to extract data from your various platforms and deliver it to Hadoop, but many fine quality, mature ETL tools are available at reasonable prices and of course the No cost Apache products like Sqoop.**

Understanding the basic nature of current and future data regulations and of the many IT ecosystem hurdles you need to negotiate, BigDataRevealed would like to offer our single software application that address each of the issues raised in this article. We are fully imbedded in the Hadoop ecosystem and take full advantage of its unique attributes. We are capable of discovering, removing or encrypting Direct Identifiers, Indirect Identifiers, and Cross File Indirect Identifiers. We process most any form of structured or unstructured data. We can assist in keeping you safe.

Infractions of the EU GDPR regulations will carry fines up to 4% of last year's total receipts for each individual violation. A potentially crippling amount. See BigDataRevealed in action.
https://vimeo.com/224269768   info@bigdatarevealed.com 847-791-7838