EXONAR

GDPR – Why It's About More Than Regulation

And why you should put Data Management at the heart of your compliance plans

White Paper

GDPR has been described by some as being the most significant regulatory framework to hit companies since the Sarbanes-Oxley Act. With a stated objective to "give citizens back control of their personal data, and to simplify the regulatory environment for business", it will impact every single European individual who has shared their personal data with an organisation, and every single organisation that holds information on any European individual.

This whitepaper discusses why any attempt to be compliant must start with the data an organisation holds, and gives practical guidance to help organisations prepare for their GDPR responsibilities.

Looking Beyond The Regulation

Why Data Should be Your Primary Concern

The requirements that GDPR places on companies are wide-ranging and will impact everything from the people employed by the organisation, through to policies and processes (see Fig.1). However, we believe that one of the key elements of GDPR is being overshadowed by a preference for organisations to only frame the GDPR conversation in legal terms. For the purposes of this white paper we will investigate and demonstrate how data management is central to the success of organisations to plan, and successfully implement, their GDPR strategy.

We'll begin by setting out the importance of rights. As with current data protection legislation, GDPR is clear that individuals have a series of rights when it comes to how their data is collected, stored, used and disposed of by organisations. When considering GDPR in the context of your business, it is fundamental to be able to fulfil each of these rights. In Fig.1 we've summarised these rights and their impacts.

Key Elements of GDPR	People (Roles)	Operational Processes	Policy and Governance	Legal and Compliance	Data Discovery
Mandatory roles such as Data Protection officer	✓			✓	
Mandatory PIA (privacy impact assessments)	✓		✓	✓	✓
The right to be informed		✓	✓		✓
The right of access		✓			✓
The right to rectification	✓				✓
The right to be forgotten		✓			✓
The right to restrict processing		✓			✓
The right to data portability		✓			✓
The right to object		✓			✓
Rights in relation to automated decision making and profiling		✓	✓		✓
Data breach notification	✓		✓	✓	✓
Accountability	✓		✓	✓	
Redress- fines and class actions	✓		✓	✓	✓
Internal governance, assurance and compliance testing	✓		✓		✓

Figure 1

What should be clear from this illustration is not only do business leaders have a lot to consider in making sure their organisation is ready to fulfil their GDPR obligations from a people, process and policy perspective, but that if they don't understand where their data is, they won't be able to comply.

The financial penalties of non-compliance have been frequently reported – 4% of annual turnover is a headline grabbing number – however, the risk is far greater than one fine. With GDPR allowing individuals to take class actions against organisations that mistreat their data, any organisation that is subject to a data leak / hacking incident can expect to receive individual lawsuits which will not only increase the financial loss, but also consume vast amounts of time in settling individual litigation.

With this understanding in place, data management becomes the primary activity for any organisation getting ready for GDPR.

Data Management Begins with Discovery

Any data management process requires you to know precisely what data you have. At a macro level this seems simple – most organisations are arranged by function; ie. HR, finance, sales, marketing; and most would expect that each team has associated data servers and assets. What GDPR forces business leaders to consider, however, is where every single piece of personal data is across their IT estate – including the Cloud. Taken in this context, the question of the data that an organisation holds on individuals becomes a far more complex one to answer, and one that is going to require time, resource and budget.

A thorough approach to data discovery, properly implemented, will lead you to data that you did not know about. We have uncovered instances where organisations have multiple gigabytes of data 'hiding' throughout their network, including company sensitive information, personally identifiable data and duplicated information; all of which could be misappropriated or mistakenly shared.

No Excuses for Not Knowing

Should your organisation suffer a breach, "not knowing" that you have unseen data or inconsistencies in the treatment of data is not a permissible excuse in the eyes of regulatory bodies such as the Information Commissioner's Office (ICO) in the UK. This means that not only do organisations have to set aside adequate time and money to undertake discovery, they need to be prepared to make time to assess, understand and decide what to do about unexpected data. As an example, what would you do if you found that your marketing organisation had three databases which contained a mixture of duplicated and unique data? How would you consolidate and organise the data and how long would it take to go through that process?

Data Type	HR	Finance	Marketing	Sales
Name	✓	✓	✓	✓
Job Title	✓	✓	✓	✓
Telephone Number	✓	✓	✓	✓
Email Address	✓	✓	✓	✓
Home Address	✓	✓		
Business Address	✓	✓	✓	✓
Salary	✓	✓		
Personal Bank Details	✓	✓	✓	
Business Bank Details		✓		
Purchases		✓	✓	✓
Payments	✓	✓		✓

Figure 2

Starting with What You Know

Most organisations have distinct functional areas with distinct processes and tools for holding data on individuals (see Fig.2). Once this initial dataset is understood, it becomes important to identify what is personal data, and what is not. This is further broken down into data that could be used to identify an individual, and information that would be classified as sensitive.

With GDPR, these definitions of data have been broadened to reflect the ways in which many organisations now retrieve and store information, including location data that is automatically harvested by online organisations. The below, taken from a study by White & Case provides an excellent representation of one way in which this change will manifest itself:

Personal Data or Sensitive Personal Data?

Personal Information: Personal data enables you to be identified. The definition is broad ranging and includes (but is not limited to) your name, a visual image of you, your address, place of work, date of birth or even your computer's IP address. The EU defines personal data in this document.

Sensitive Information: Sensitive data is defined as that revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life. In some countries, details about criminal activity and/or criminal offences is included within this definition.

	The Directive	The GDPR	Impact
Personal data *This definition is critical because EU data protection law only applies to personal data. Information that does not fall within the definition of "personal data" is not subject to EU data protection law.*	***Art.2(a)*** *"Personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.*	***Art.4(1)*** *"Personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person.*	*The definition of personal data is, for the most part, unchanged under the GDPR.* For some organisations, the explicit inclusion of location data, online identifiers and genetic data within the definition of "personal data" may result in additional compliance obligations (e.g., for online advertising businesses, many types of cookies become personal data under the GDPR, because those cookies constitute "online identifiers").

Source: http://www.whitecase.com/publications/article/chapter-5-key-definitions-unlocking-eu-general-data-protection-regulation

The New Rules of Ownership
Data Processors in the Picture

Where GDPR brings about significant change from previous legislation is that it places a number of direct obligations on data processors as well as data controllers. This means that individuals can enforce their rights directly against data processors. Whilst data processors may have a wide range of business models from on-premise to Cloud-based, the same obligations and provisions will apply to them however they are processing individuals' data. To avoid non-compliance, Data Processors will have to put in place a robust and accurate data audit trail.

Processor or Controller?

Processor: "A natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller."

Controller: "A natural or legal person, public authority, agency, or other body which, alone or jointly with others, determines the purposes and means of processing of personal data."

Creating Your Audit Trail
Why the Clipboard Approach No Longer Works

In its official guidance to preparing for GDPR, the ICO states "You should document what personal data you hold, where it came from and who you share it with. You may need to organise an information audit."

If you have inaccurate personal data and have shared this with another organisation, you will have to tell the other organisation about the inaccuracy so it can correct its own records. You won't be able to do this unless you know what personal data you hold, where it came from and who you share it with. You should document this. Doing this will also help you to comply with the GDPR's accountability principle, which requires organisations to be able to show how they comply with the data protection principles, for example by having effective policies and procedures in place.

Within this sentence is a clear message for any organisation that holds data on individuals: you need to clearly account for **every piece of data** that you hold and **how you handle it**. Reflecting the fact that the world is networked, the ICO provides good examples of the amount of detail it expects organisations to go into when creating their asset register and audit trail.

As with registers for fixed assets, creating data or process audit trails has traditionally been a manual task for many organisations. External consultants can help with making this process less of a time burden in terms of documenting results and presenting evidence, but using manual discovery and interviews is a flawed approach, not least because data is being generated and manipulated faster than any human could be reasonably expected to keep up with. The single biggest reason why the 'clipboard approach' no longer works is because it's not able to uncover data that you do not know is there.

Unknown Data
Your Biggest Risk

Understanding your 'known data' is just the first part of the process and it is easy for organisations to forecast the potential impact of inaccuracies or poor governance on the data that they already have. The risk is tangible because the data is visible. The greater danger is for organisations to stop at simply recording and securing what they have, as the biggest threat to their business could be from the data that they cannot see.

Many organisations have gigabytes of unknown or 'hidden' data. This may take the form of redundant information on decommissioned servers, duplicate data held on file shares, information emailed within the business or to 3rd parties, data that has been used or held outside of agreed compliance processes, and data that has expired, but has not been deleted. See Fig.3

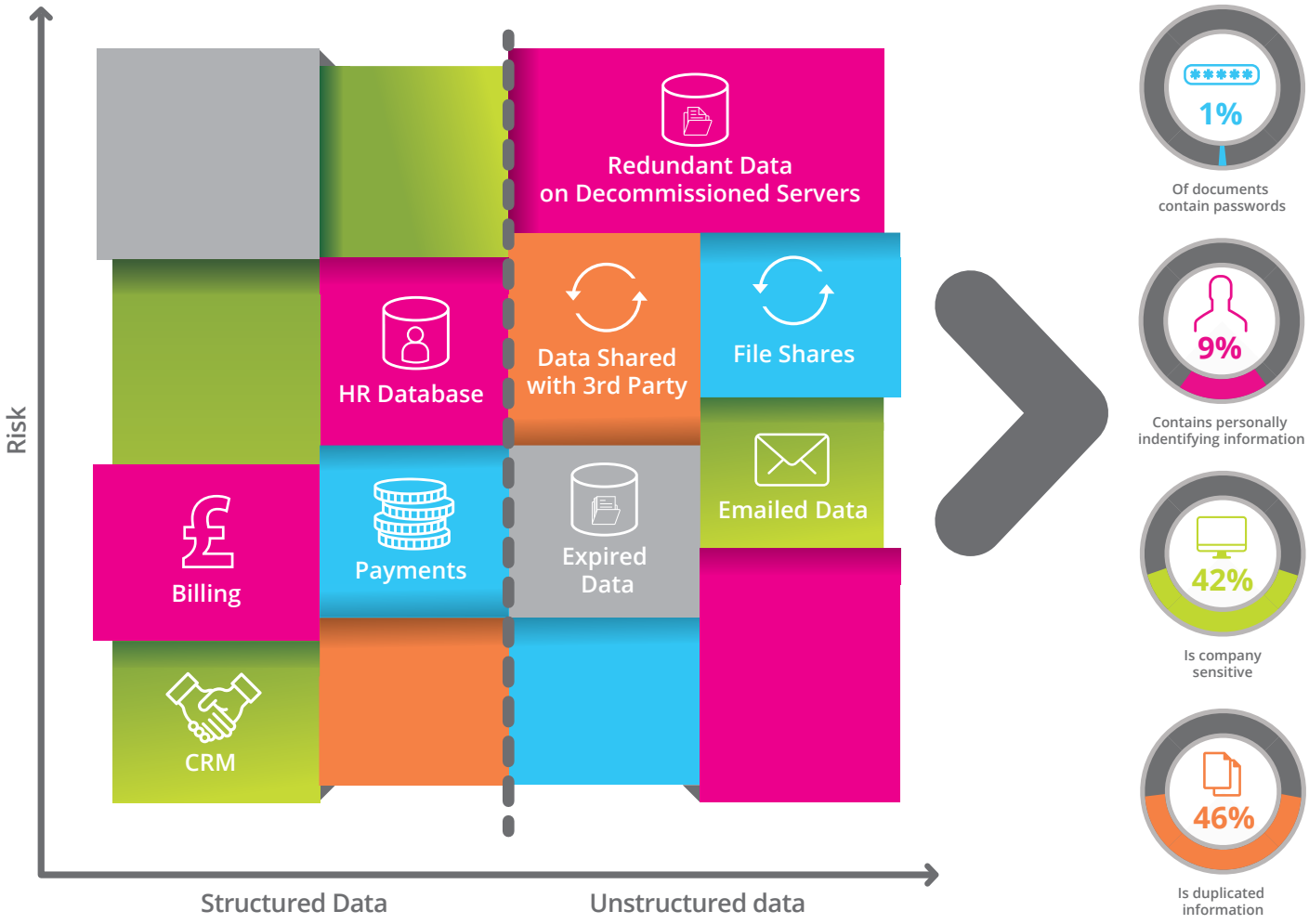


Figure 3

Because this data is not sat within the 'known' systems outlined in Fig.3, many organisations simply don't factor it into their data management plans. This is a common mistake as data that is not in use, or not known about, could act as an open door to malicious attack or accidental damage.

In the companies that engage our services, it is not unusual to find data hidden across their estate which contains passwords, company sensitive information, duplicated data, and personal and sensitive data relating to individuals.

The reasons for data being hidden are wide ranging and can include:

- Deliberate avoidance of corporate data management policy in order to save time
- Accidental duplication due to human or system error
- Company acquisition where data assets are not accurately catalogued
- Accidental breaches of data management policy due to lack of process automation
- Lack of understanding of data management policies
- Inability for data management policies to be automatically enforced

Hidden data presents numerous risks to an organisation, and every organisation will have some stored somewhere on their network. The fact that unknown data exists makes it impossible for an organisation to quantify its risk of exposure without specialist tools to turn that unknown data into a known data asset.

The Solution to Discovering Hidden Data

Search and Machine Learning

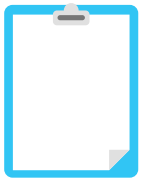
The rapidity with which we can capture, create, manipulate, store and share data outstretches our ability to manually identify data, let alone make informed decisions on how to manage it. In the time it would take one person to interview your database administrator (DBA) on the approved processes for 'known' data, an employee could accidentally load an entire database to the Cloud, or a malicious coder could breach your network.

The identification of 'unknown' or hidden data requires a method and rapidity of discovery that only a machine can provide, which is why we have dedicated our business to leveraging AI, Big Data and Machine Learning to solve the problems inherent with data growing beyond the capacity with which an organisation can rationally deal with it. It is not commercially viable for an organisation to slow the speed of data creation, and it's not possible for every employee to remember every data compliance process or piece of legislation. This is where Machine Learning comes in.

At Exonar, we've compressed the amount of time, removed the hidden cost, and taken away the potential for human error in data collection and management by leveraging Machine Learning and Big Data to undertake data discovery and classification. Our unique platform allows organisations to audit, identify and classify hidden, high-risk data held within a network – whether it's Cloud, on-premise or hybrid.

This is not about using technology for its own sake, but to speed up the process by which businesses can make smart decisions, especially in light of the timelines involved in GDPR. This gives organisations the scalability to process millions of documents across multiple platforms, use drag and drop functionality that makes classification intuitive, and search like you're using a search engine – not something that requires a database admin.

Approaches to Data Discovery

Method	Approach	Output
Clipboard Method 	• Interview all staff • Document known data • Assume all processes adhered to • Resource intensive	• Documents all current and known data and processes at a single point in time
Digital Search 	• Searches all digital storage • Matches simple known patterns and words	• Identifies many incidents of possible unknown data BUT also many false positives
Intelligent Classification 	• Understands context in which data is provided • Considers the structure of a document to inform identification • Uses surrounding words and phrases to pinpoint the right data	• Identifies all hidden data • Updates in near real-time • Continues to identify and manage data

The GDPR Data Dashboard

Fig.4 is an example of how Exonar's GDPR dashboard provides a high-level view of an organisation's information in relation to EU GDPR law. It shows an overall picture of all the data held by an organisation, which is subject to GDPR, where it is held and its characteristics. This approach takes organisations beyond spreadsheets and interviews, and into the realm of making well informed decisions, rapidly.

Figure 4

Conclusion – Your Data Will Determine Your Compliance

It could be easy to feel paralysed by such significant legislation, or to believe that everything related to GDPR compliance can be dealt with by your legal team. It would also be a mistake. We want to put organisations in control of their data so that they can be **compliant by design**, allowing them to apply the right duty of care to their customer's data, and avoid becoming the subject of a breach or action taken by a whistle blower.

With the right foresight, some helpful tools and some expert knowledge, getting your data right in the context of GDPR, and therefore improving your compliance position overall, is achievable. Note, it's more cost effective and simple than you might think. We've summarised our top 5 actions...

1. Review any contracts, customer agreements or contact strategies that you are likely to implement in the next 6 months and/or will be in place when GDPR legislation comes into force – this will determine how fast you need to act
2. Map out the data that you know you have
3. Review our list (Fig.3) to help you establish where data may be hidden in your organisation
4. Move quickly to secure budget for a thorough data discovery exercise
5. Ask yourself whether your organisation has the time, tools or resource to undertake this kind of discovery alone – or whether you'd like a partner who can confidently guide you.

Exonar will help you to reduce the risk associated with GDPR by getting right to the heart of the matter: finding and managing your data.

EXONAR

14 West Mills, Newbury
Berkshire, RG14 5HG

Tel: 01635 888581

Email: tellmemore@exonar.com

www.exonar.com

For a demo, **click here**.
For a free trial, **get in touch**.