

# PRODUCT BULLETIN

Inventory, filtering, workflow, review, classification and documentation of personal data in unstructured data using collaborative cognitive learning.

Aigine

## INTRODUCTION

### A unique solution to a new need

Aigine is a unique solution to an unaddressed problem – managing and documenting personal data in unstructured data sources.

Aigine is packaged to deliver customer value with minimal service effort.

GDPR did not end the 25<sup>th</sup> of May 2018, instead that's when it started. Aigine will therefore continuously offer solutions and features that solve new customer needs, also in the governance phase. These solutions use the metadata layer and the cognitive abilities that are built up in Aigine, which gives the customer a continuously increased value of the solution.

In California, the California Consumer Privacy Act was passed on 29th of June 2018. It implements the same principles as the European GDPR and affects companies holding data of US citizens.

Other countries are in the process of preparing similar bills in the upcoming year, making privacy a global prioritized issue – especially since non-compliance can mean consequences in several jurisdictions.

### Buying the solution

Aigine is delivered through IBM distributors and resellers. Aigine is provided as a subscription with 12 months commitment period. Contact your normal IBM reseller for a quotation.

### GDPR – a new data protection regulation

The purpose of the GDPR regulation is to protect the privacy and personal data of every citizen of an EU country. GDPR will influence and have direct effect on all EU organizations and all legal entities that in any way handle personal data concerning an EU citizen, wherever they are in the world.

When GDPR entered into effect the same rules apply to all personal data (PD) saved in any searchable source. Personal data can be processed, used and managed provided that the organization:

1. Knows where the PDs are located
2. Has secured access to the PDs through permissions and similar
3. Knows what the legal ground is for processing the PDs

Each document and e-mail that an organization stores must therefore be reviewed to determine whether they contain personal data according to the new definition in GDPR. If they do, a legal analysis must determine whether there is a legal ground for storing the personal data. This legal analysis must also be documented, serving as the register of records.

Auditing, analysis and documentation will take at least ten man per terabyte.

That's such an extensive workload that most organizations have chosen not to address the problem at all and no tools have existed on the market to reduce this challenge.

Deterrent fines have been introduced to ensure there is no incentive to violate, or ignore, GDPR as an EU-wide regulation. Any shortcomings in each organization, may lead to a fine of up to 20 million Euro or 4% of the annual turnover at group level, per occasion. The amount of the penalty payment is therefore dissuasive for any organization. The public sector may have slightly reformulated penalty payments according to the government's proposal (for instance Prop. 2017/18:105) but will still be so high that there is no possibility not to obey and establish compliance towards the regulation.

In addition to these fines, there is also a right for any victim to claim damages even for non-material damages, i.e. offence.

## THE SOLUTION AIGINE

Aigine provides a solution for managing and document data in identified data sources.

- Aigine filters all data sources and sorts out documents that do not contain PD. Therefore, the staff do not have to read through documents that do not need to be reviewed.
- Aigine presents the documents with the suspected PDs marked. This makes the review much faster.
- Aigine helps the staff find and propose the legal ground with a contextual knowledge database and artificial intelligence. This means that the analysis goes faster, but also that the analysis can be done by staff who do not have legal training.
- The legal ground for processing PD in the document is directly specified in the Aigine user interface. This saves time since there is no need to document the decision separately.
- Aigine saves all meta-data, creating a ground for register of records, and making subject access rights, automatic information management and monitoring possible.
- Aigine changes artificial intelligence. By creating the code required for the training of artificial intelligence locally, it is ensured that no information ever leaves the customer. In this way, the actual work is used to make the artificial intelligence better. These sharpened cognitive abilities are then shared with everyone that uses the solution; through improved filtering, marking and suggestions for legal ground.

## CONTENT

Introduction.....	1
A unique solution to a new need .....	1
Buying the solution .....	1
GDPR - a new data protection regulation .....	1
The solution Aigine.....	2
1    Background.....	6
1.1    GDPR and CCPA.....	6
1.1.1    A new definition of personal data.....	6
1.1.2    A new definition of data source .....	6
1.1.3    Processing of personal data .....	6
1.1.4    Privacy records.....	6
1.1.5    Lawfulness of processing .....	7
1.1.6    Possible consequences .....	7
1.1.7    What to do with the unstructured data.....	8
1.2    Creating long term business values .....	9
1.3    SAMBRUK and SALAR project.....	9
1.3.1    Background of the pilot project.....	9
1.4    A solution to the problem, not merely show the scope of the problem.....	10
1.4.1    No solution on the market .....	10
1.4.2    Does not support daily operations .....	10
1.4.3    Structured and Unstructured data .....	10
1.4.4    No obvious system supplier .....	10
1.4.5    Takes too much time and requires legal training.....	10
1.5    Requirements for a solution .....	11
1.5.1    Unstructured data .....	11
1.5.2    Structured data (not covered by the project's proposed solution) .....	11
1.5.3    Identifying data sources .....	11
1.5.4    Inventory workflow.....	11
1.5.5    Metadata.....	11
1.5.6    Indication of legal basis.....	11
1.5.7    Link to external sources for support in decision making .....	11
1.5.8    Reduced need for legal expertise .....	11
1.5.9    Non-previously identified personal data.....	12
1.5.10    Artificial intelligence (local).....	12
1.5.11    Artificial intelligence (cloud).....	12
1.5.12    Generated Content .....	12
1.5.13    No data may leave the customer .....	12
1.5.14    A way to solve the problem .....	12
2    Functional description .....	13
2.1    Installation and Configuration.....	13
2.1.1    How the installation is done .....	13
2.1.2    Connect data sources .....	14
2.2    Filter .....	14
2.2.1    Cognitive abilities downloaded from the cloud .....	15
2.2.2    Why regular expressions are not enough.....	15
2.2.3    Filters out of documents containing suspected personal data .....	16
2.2.4    Filters out documents containing suspected personal data .....	16
2.2.5    Timesaver .....	16

2.3	Review .....	16
2.3.1	Workflow .....	17
2.3.2	GUI .....	18
2.4	Decide legal basis .....	18
2.4.1	Contextual knowledge database.....	19
2.4.2	Time savings and the ability to distribute work.....	19
2.5	Document .....	19
2.5.1	Integrated smart digital form for direct documentation .....	20
2.5.2	Legal basis suggestions from AI .....	20
2.5.3	Time savings and the ability to distribute work.....	20
2.6	Collaborative cognitive learning.....	20
2.6.1	Annotation coding done locally .....	21
2.6.2	Cognitive learning in the cloud .....	21
2.6.3	Cognitive abilities improve.....	22
2.6.4	Everyone gets benefits from the improved skills.....	22
2.7	Timesaver .....	22
2.7.1	Example .....	22
2.7.2	Make your own simulation .....	23
2.8	The silver bullet approach.....	23
2.8.1	Data Discovery .....	23
2.8.2	Data Classification .....	24
2.8.3	Records of Processing.....	24
2.8.4	Information management.....	24
2.8.5	Monitoring.....	24
2.8.6	Compliance training .....	24
2.8.7	Subject access request.....	24
2.8.8	Anonymization and pseudonymization .....	24
2.8.9	Data protection impact assessment.....	24
2.8.10	Rights Management, Breach.....	24
2.8.11	Secure transfers .....	24
2.8.12	Reporting of personal data incidents .....	24
3	Technical description.....	25
3.1	License restriction .....	25
3.2	Included components.....	25
3.2.1	Aigine Unstructured Data Inventory Engine.....	25
3.2.2	StoredIQ.....	25
3.2.3	Watson Ai and Watson NLU .....	25
3.2.4	Elinar AI Miner.....	26
3.2.5	Cloud AI .....	26
3.3	System design .....	26
3.3.1	Encryption and certificate management.....	27
3.4	System.....	27
3.4.1	Downloading installation packages .....	27
3.4.2	IP addresses .....	27
3.4.3	Space that needs to be allocated for installation packages .....	28
3.4.4	Requirements for ports to enable communication to the cloud .....	28
3.4.5	Other system requirements: .....	28
3.5	Browser support.....	28
3.6	File format and protocol support .....	29
3.6.1	File formats that can be scanned by Aigine .....	29



3.6.2	Cloud services and local systems which can be scanned by Aigine .....	34
3.6.3	TIFF and other image formats.....	35
3.7	Updates.....	35
3.7.1	Distribution.....	35
3.7.2	Downloadable package .....	35
3.8	Server Governance.....	35
3.9	Installation in outsourced environments.....	35
4	Demonstration of the solution.....	36
4.1	End-User cloud demo environment .....	36
5	Documentation and marketing materials.....	37
5.1	Marketing material.....	37
5.1.1	This Product bulletin.....	37
5.1.2	Process design .....	37
5.1.3	Website .....	37
5.1.4	Blog.....	37
5.1.5	PowerPoint presentations .....	37
5.1.6	Flyers .....	37
5.1.7	Video .....	37
5.2	Technical documentation .....	37
5.2.1	Debugging scheme to exclude non Product Fault .....	37
5.2.2	Installation and maintenance manual.....	38
5.2.3	Technical guidance and checklist .....	38
5.2.4	Configuration guidance.....	38
5.2.5	Test schedule.....	38
5.3	User documentation.....	38
5.4	Release notes .....	38
5.4.1	Release Notes .....	38
5.4.2	Road map .....	38
6	Support.....	39
6.1	End-users.....	39
6.2	Technical support .....	39
7	Upcoming Products from Aigine.....	40
7.1	Monitoring .....	40
7.2	Subject Access Rights.....	40
7.3	Structured data .....	41
7.4	Deletion and masking.....	41
7.5	Pseudonymization .....	41
7.6	Other areas besides GDPR.....	41
8	PRICES.....	42
8.1	Pricing mechanisms .....	42
8.2	Quotation .....	42
9	Order.....	43

# 1 BACKGROUND

## 1.1 GDPR and CCPA

The 25<sup>th</sup> of May 2018 [GDPR](#), General Data Protection Regulation, entered into effect. The regulation is directly applicable in all EU member states.

The purpose of GDPR is to protect the privacy and personal data of every citizen of an EU country. GDPR will influence and have direct effect on all EU organizations and all legal entities that in any way handle personal data concerning an EU citizen, wherever they are in the world.

GDPR can therefore be considered as a global legislation.

In California, the California Consumer Privacy Act was passed on 29th of June 2018. It implements the same principles as the European GDPR and affects companies holding data of US citizens.

Other countries are in the process of preparing similar bills in the upcoming year, making privacy a global prioritized issue – especially since non-compliance can mean consequences in several jurisdictions.

### 1.1.1 A new definition of personal data

GDPR changes the definition of personal data. Where previous legislation defined personal data as "any kind of information directly or indirectly attributable to a natural person who is in life is considered as personal information." GDPR extends this definition by including information that is not in themselves sufficient to identify an individual, but can do so in combination with other pieces of information.

"Personal data means any information relating to an identified or identifiable natural person. It is crucial that the task, individually or in *combination with other data*, can be tied to a living person."

Therefore, what is considered to be personal data under the GDPR becomes a matter of what other information the organization holds, since it is the theoretical combination that determines whether or not it is a personal data.

### 1.1.2 A new definition of data source

Where previous legislation had its strongest regulation on what you call register, which means database, GDPR is neutral. Therefore, the GDPR applies to any source of information that can be searched. This includes, of course, databases, but also unstructured data such as files and emails, but even analog filing cabinets where the information is sorted in such a way that search is possible.

### 1.1.3 Processing of personal data

GDPR does not contain any prohibition for the processing of personal data. Instead, it sets specific requirements on how and when personal data may be handled. In short, it requires:

1. To know where the PDs are located
2. To have secured access to the PDs through permissions and similar
3. To know, and document, what the legal ground is for processing the PDs

If these main rules are followed, there are no obstacles to process personal data.

### 1.1.4 Privacy records

Under GDPR, it's prohibited to 'stumble' over a personal data, and subsequently investigate whether or not one was entitled to process it. For this reason, the GDPR requires that a register of all personal data processing is created. These records shall be in writing, be available in

electronic format and kept up to date. On request, the Privacy records shall be made available to the national authority. What is to be included in the privacy record is described in article 30 of GDPR.

- (a) the name and contact details of the controller and, where applicable, the joint controller, the controller's representative and the data protection officer
- (b) the purposes of the processing;
- (c) a description of the categories of data subjects and of the categories of personal data;
- (d) the categories of recipients to whom the personal data have been or will be disclosed including recipients in third countries or international organizations;
- (e) where applicable, transfers of personal data to a third country or an international organization, including the identification of that third country or international organization and, in the case of transfers referred to in the second subparagraph of Article 49(1), the documentation of suitable safeguards;
- (f) where possible, the envisaged time limits for erasure of the different categories of data;
- (g) where possible, a general description of the technical and organizational security measures referred to in Article 32(1).

### 1.1.5 Lawfulness of processing

The purpose of the processing is of great importance, since processing only can be carried out where the purpose can be attributed to one of the legal grounds set out in article 6 of the GDPR. These purposes, and their respective legal grounds, shall be shown in the Privacy records.

#### 1.1.5.1 Consent

The data subject has given consent to the processing of his or her personal data for one or more specific purposes;

#### 1.1.5.2 Agreement

Processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;

#### 1.1.5.3 Legal obligation

Processing is necessary for compliance with a legal obligation to which the controller is subject;

#### 1.1.5.4 Vital interest

Processing is necessary in order to protect the vital interests of the data subject or of another natural person;

#### 1.1.5.5 Public interest

Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

#### 1.1.5.6 Legitimate interest

Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

### 1.1.6 Possible consequences

GDPR is a law of rights designed to protect the privacy and personal data of EU citizens. The legislator has been aware that compliance with previous legislation has been far from good. The GDPR therefore clearly stipulates that failure to comply should lead to 'dissuasive' consequences.



#### 1.1.6.1 Fines

Private companies and organizations risk fines of up to EUR 20 million or 4 Percent of the annual turnover at group level per failure. As this is a fine, the fine can be issued by the authority, without any legal process. It is up to the fined party to commence a legal process in order to evade payment liability.

#### 1.1.6.2 Damages

In GDPR there is a right for the damaged party to receive damages, described in article 82. GDPR explicitly gives the victim the right to compensation for both direct and indirect damage. Indirect damage also includes mental suffering. Non-compliance towards GDPR can therefore lead to very large claims for damages from injured parties.

In addition to this, shareholders can also file damage claims towards the management of a company, that has not taken sufficient actions to comply towards the GDPR.

#### 1.1.6.3 Bad Will

GDPR is, as mentioned, a law to safeguard privacy. In time, therefore, it is likely that loss of Good Will, or extensive Bad Will, will constitute the most significant effects of the legislation. Serious organizations take privacy seriously, whether the PDs belongs to customers or citizens.

### 1.1.7 What to do with the unstructured data

As has already been stated, unstructured data has been excluded from previous legislation. As this change, each organization must secure its unstructured data with respect to the new regulatory framework in the GDPR. An organization must therefore assume that PD can be present in each document and in each data source.

The option to simply delete old files, without knowing what they contain, does not exist. Partly because it may be bad for business, but above all, because legal obligations make it directly illegal to delete some data.

Structured data, where each database has a clear purpose, only needs to be described at a level of the input tables. For unstructured data, each document can theoretically be its own data source, since the legal ground and the need for protection is different for each document, even if they happen to be in the same file storage area, or in the same directory.

As a part of continuous compliance with GDPR under a governance phase, it is therefore appropriate to determine and communicate how personal data may be stored in different file storage areas and directories. With such a structure, the privacy records only need to be on an area and directory level, provided that these file storage areas are continuously monitored in order to prevent incorrect information from being saved on them.

#### 1.1.7.1 Read everything

In order to know if the unstructured data contains PDs, all documents, emails and files must be reviewed.

#### 1.1.7.2 Determine if there is personal data

The review must determine if the document contains personal data in relation to the GDPR's new definition; data in combination with other data may constitute a piece of information allowing identification of an individual.

As far from all the documents de facto contain personal data, this means that many documents that does not actually contain personal data has to be reviewed.

### 1.1.7.3 Find legal basis

If the document contains personal data, a qualified legal assessment is required to determine if a legal basis exists to process the data. Such an assessment must be made in relation to both GDPR and other laws and regulations affecting the business. Over time, this legal assessment must also include the prejudiciary rulings from other European countries and the EU court.

### 1.1.7.4 Document

In the event that the legal assessment fails to identify a legal basis to process the data, the PD in document must either be masked, or the whole document must be deleted. This task must be ensured, since the processing of personal data, when the lack of legal basis has been identified, constitutes a serious incident and therefore implies a direct qualification for the higher fines in GDPR.

When a legal basis is found, it must be documented, often together with information on how long the PD will be saved. If the legal basis is a contract or consent, a reference should be made to the location of the consent or agreement.

This documentation constitutes a Records of process activities, which is a requirement under the GDPR. Before a storage structure and policy is implemented in the organization, any document containing personal data risks constituting a separate 'register' in respect of Legal basis, purpose and saving/masking time.

## 1.2 Creating long term business values

The number of regulations is increasing, with GDPR and other legislation, and the impact of non-compliance are becoming more and more severe for businesses, spanning from fines and damages, to liquidation.

Wise organizations do not only calculate risk costs of non-compliance. Instead, modern technology as Aigine, makes it possible to create business value and competitive advantages.

For instance, GDPR gives companies a unique possibility to upgrade their information governance framework, differ efforts (and thereby costs) for different sources, and increase the possibilities for as-yet-unknown uses of data. Among these possibilities are automated information management, substantially cutting costs and generating revenue over time. Aigine also makes it possible to both save and use more data, by making both anonymization and pseudonymization possible in large amounts of data.

What lays in front of us is difficult to know, but one thing is sure – the future will be data driven, and data is the new gold. What would it mean for a company's future competitive advantages if they could save and use all their data, whilst the competition needs to delete theirs.

## 1.3 SAMBRUK and SALAR project

### 1.3.1 Background of the pilot project

The association Sambruk, which is a member organization with about 100 Swedish municipalities, early became aware of the large workload GDPR would create for its members. Especially the task of making an inventory of processed and stored personal data occurrences in many forms and in large numbers of information systems. Mapping this would be a very large and complex task, which, for all organizations, means hundreds to tens of thousands of hours of work, to review and assess the legal basis for the processing of personal data.

That is why the association, together with 10 member municipalities and a regional association and SALAR, has performed a pilot project with the goal of producing requirements for a solution which can lead to significant increase in efficiency in the extensive mapping and documentation

work, which is part of preparations to secure organizations' compliance with the new GDPR requirements.

## 1.4 A solution to the problem, not merely show the scope of the problem

### 1.4.1 No solution on the market

One of the initial activities in the project was to identify potential, existing and available technical solutions. The different system solutions on the market had a focus to present the extent of the problem by carrying out searches for personal data using regular expressions. However, no systems could resolve the actual work task since they lack support for delegation, analysis, legal assessment, and documentation. The usage of regular expressions also creates problems, as personal data under GDPR is not a fixed term.

### 1.4.2 Does not support daily operations

This means that these tools fit single inventories linked to projects with the aim of mapping the scope of occurrences of (suspected) personal data rather than a continuous and iterative approach that directly supports daily operations.

### 1.4.3 Structured and Unstructured data

Another observation made early in the project was that unstructured data was the biggest and most pressing challenge when it comes to the screening of data sources. 70-80% of the data in an organization is unstructured.

Many organizations that started the initial GDPR work believe that personal data handled in business systems, classified as structured data, can be searched within the systems. The legal basis for the processing of this personal data is mostly included in the definition of the support processes that form the basis of the system. In addition, the existing systems have a pronounced and existing vendor who can provide services to change the processes and functions of the systems in a necessary way.

Therefore, the scope of the project focused on unstructured data.

### 1.4.4 No obvious system supplier

Unstructured data does not have a pronounced and existing system vendor or supplier that addresses the problem or offers a solution. Nor can OS or hard drive vendors be obliged to take responsibility for the content of common storage areas or file management system. The problem has therefore not been addressed in the past and has, for that reason, not been prioritized.

### 1.4.5 Takes too much time and requires legal training

It is clear that extensive resources are required in the work to review, carry out legal assessment and document legal decisions. Estimating the exact time required is difficult. The project has tried to make a broadly visible time estimation on these activities by producing a basis for calculation. The equation that is developed requires a number of assumptions to visualize the required resource needs.

*Assumption 1.* An average document is 2 MB.

*Assumption 2.* It takes on average 2 minutes to read through a document

*Assumption 3.* 10% of the data files contains some form of personal data

*Assumption 4.* The average time for legal decision-making on legal basis is 3 minutes per document

*Assumption 5.* The average time for documentation of decisions is 2 minutes per document

This results in 2 604 Working days or 11.84 Year of work per terabyte, most of it having to be performed by legal professionals.

Assuming that an average Swedish municipality with 35 000 inhabitants stores between 15 and 20 terabytes of data on average, it results in a much larger workload.

## 1.5 Requirements for a solution

### 1.5.1 Unstructured data

The system should be able to handle various forms of unstructured data even those stored outside the local environment such as OneDrive and Office 365. Also e-mail systems are considered as unstructured data and also needs to be managed.

### 1.5.2 Structured data (not covered by the project's proposed solution)

Data stored in databases, so-called structured data, must be managed at the data table level. Within databases, also unstructured data exists in the form of free text fields and attached files. This form of data also needs to be managed within the system at some point.

### 1.5.3 Identifying data sources

It must be possible to connect the data discovery system to the respective identified data source with the appropriate permissions to read files and data stored. There must also be a function to enter the email address of the person responsible for each data source. This is to be able to distribute search results for further review.

### 1.5.4 Inventory workflow

Each responsible for a data source will be alerted by e-mail. This e-mail contains a link to the gross list to be audited. Document content must be presented in a web GUI, so other software is not needed to review the content. This GUI also highlights occurrences of personal data through color markings that visualize different types of suspected personal data that the system has identified. This is to make the review of documents more efficient.

### 1.5.5 Metadata

Metadata is stored data about data. Each review must result in a number of metadata inputs. The metadata layer that is created contains information such as rejection of personal data, when a document that the system suspects contains personal information does not actually do so. Such documents are exempted from future analyses. The principle is that a human review always trumps the machine's assessment. Other types of metadata stored are the type of personal data identified and also the identified legal basis. Also personal identifiers are stored in the metadata layer, making it possible to perform Subject Access Rights and pseudonymization directly by crawling the metadata.

### 1.5.6 Indication of legal basis

The system gathers information about legal bases; agreement, consent, powers of public authority, legal obligation, protection of fundamental interest and legitimate interest.

### 1.5.7 Link to external sources for support in decision making

In examining and stating the legal basis, the system shall be able to show contextual support text for decisions in the form of existing checklists and support texts. These should be able to be downloaded from internet.

### 1.5.8 Reduced need for legal expertise

The decision support text facilitates work on the basis of a legal basis so that the need for legal training is reduced, which further creates the opportunity to make the work more efficient.

### 1.5.9 Non-previously identified personal data

The person who reviews the data will most likely identify definitions of personal data that have not been identified by the system in connection with data discovery. This new definition is intended to be used in future analyses.

### 1.5.10 Artificial intelligence (local)

Through cognitive learning and annotation coding, the system analyzes how the definition is grammatically constructed and composed. This annotated description of the new definition of personal data is sent to the centrally located AI engine.

### 1.5.11 Artificial intelligence (cloud)

The centrally located AI engine stores new and existing definitions of personal data in a database of constantly updated common definitions that are used as a search algorithm by all installations of the system.

### 1.5.12 Generated Content

The system is based on synergy, because all who use the system will also add new and sharper definitions, which means that the search results, and the system, will improve over time. This applies to:

- Constantly updated common definitions
- Definitions of different personal data
- Metadata about existing and new patterns
- Annotations-Cognitive Learning (AI)

### 1.5.13 No data may leave the customer

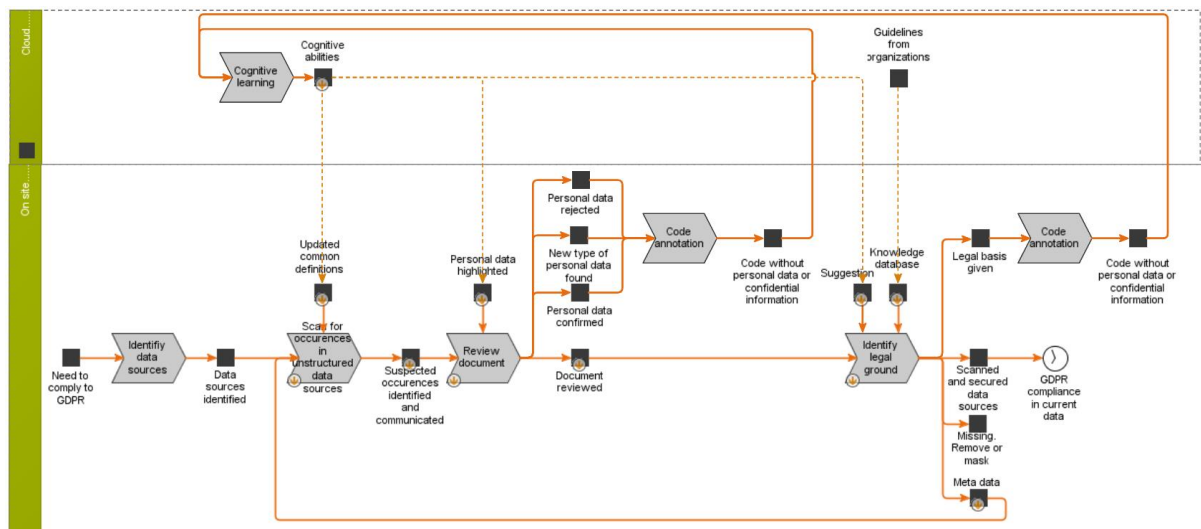
A very important part of the requirement is that the solution is legally impeccable, also based on GDPR. That is to say that no personal data or sensitive information may in any circumstances leave the municipality.

All reviews, listing of search filtering, results of review or meta data of legal basis are implemented locally and stays in the local installation.

### 1.5.14 A way to solve the problem

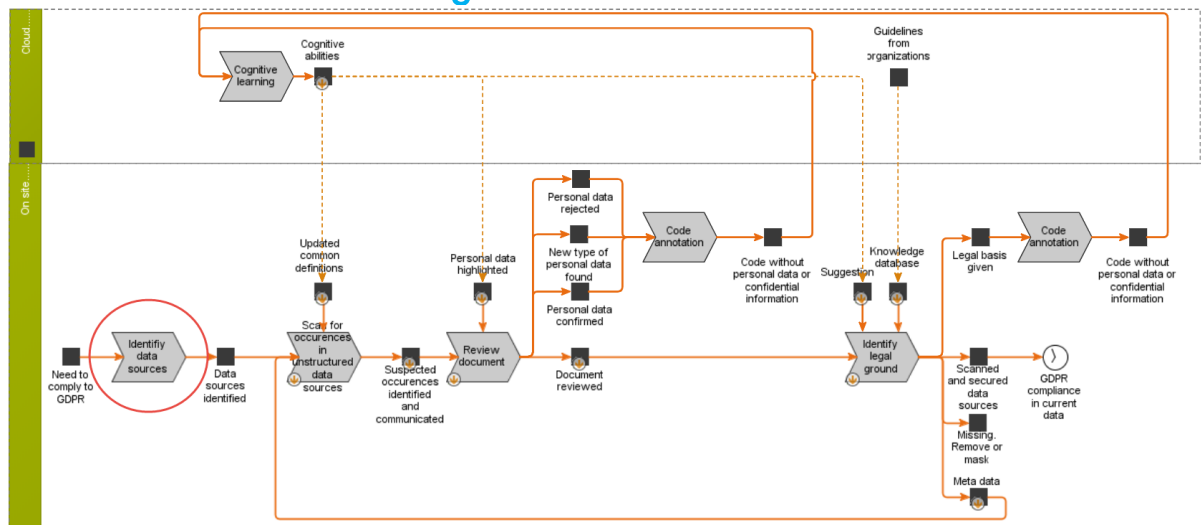
These requirements for the system to be developed create the conditions for a tool that facilitates work with mapping for GDPR, but also supports daily work and activities.

## 2 FUNCTIONAL DESCRIPTION



[http://aigine.se/Aigine\\_GDPR\\_Process/index.html](http://aigine.se/Aigine_GDPR_Process/index.html)

### 2.1 Installation and Configuration



An absolute goal in the development of Aigine has been to minimize the services needed to get the solution implemented. GDPR in general, and above all in relation to unstructured data, represents a sufficiently large challenge.

The solution is therefore created to minimize costs and time for things that do not add value to the GDPR work. We have therefore eliminated, as far as possible, such features and functionality that requires a service operation and customer interactions.

This provides business opportunities for retailers to add value to the solution, whenever customer needs arise.

#### 2.1.1 How the installation is done

Aigine is delivered by downloading images of four Linux virtual servers. Linux has been selected in order not to require customers to subscribe to additional operating system licenses, or to use existing space in license pools.



Installation is done by VMware and all software delivered with Aigine are pre-installed and configured on these machines.

Installation therefore takes, excluding preparation of the server environment and downloading the software packages, about 1-2 Hours. Installation is done through a detailed instruction and an associated test schedule and can be carried out by technicians with rudimentary knowledge of VMware.

## 2.1.2 Connect data sources

In order for Aigine to be able to make discovery in data sources, access to these sources is required. GDPR requires the customer to identify its data sources, as well as assigning a designated responsible for each one of them. This is included in the establishment of the Register list. If the customer has not performed this work, resellers may offer to carry out this inventory.

Aigine needs access to these sources via Superuser Account. If there is no such account, resellers may offer to create these.

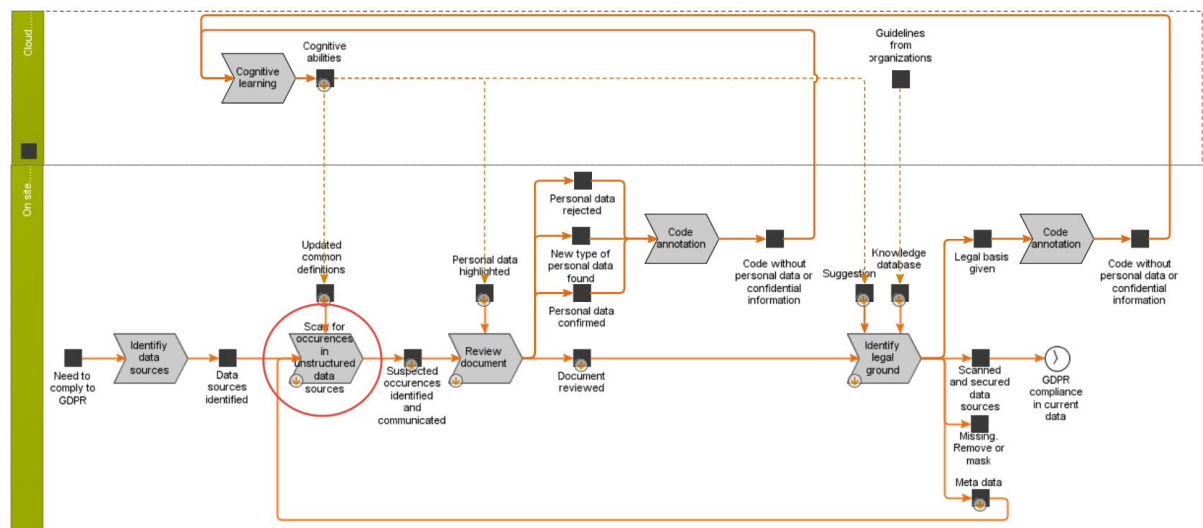
Different permissions may be required depending on different source systems connected to Aigine.

Additional information on permissions and system specs can be found at the following link: [Http://www.redbooks.ibm.com/redpapers/pdfs/redp5316.pdf](http://www.redbooks.ibm.com/redpapers/pdfs/redp5316.pdf)

Aigine can only handle normal access, not enhanced security such as three-party login and the like. In the event that such security solutions make access impossible, your reseller may offer to establish functioning access.

The e-mail address of each data source owner is required. These addresses are used by Aigine to delegate auditing tasks by e-mail to managers.

## 2.2 Filter



Aigine uses StoredIQ From IBM for inventory and discovery of connected data sources. On the first run, a full-text index is created. The time for this discovery session is estimated to be approximately 4 hours per TB.

The purpose of filtering is to ensure that staff do not have to review documents that do not contain personal information.

### 2.2.1 Cognitive abilities downloaded from the cloud

For the scanning, Aigine downloads the latest definitions, e.g. algorithms, from the cloud part of Aigine. This will increase the performance of the search at each scan.

### 2.2.2 Why regular expressions are not enough

The new definition of personal data in GDPR makes it contextual. It depends on the context where the information is present, together with other information in all other data in the organization's possession, and elsewhere.

With previous regulations, a specific piece of information that could identify an individual was described as personal data. Most of the time, that made it sufficient to search for a social security number, and that is fairly easy to do with regular expressions. A regular expression, regex is a sequence of characters that define a search pattern.

With GDPR, personal data is any piece of information that, together with any other information in the organization's possession, or publicly available, that can be used to identify a person. This makes the sentence "Lola Andersson, social security number 770812-0232, her husband and mother-in-law" a sentence that actually contains personal data from three different persons. Lolo, her husband that can be found in public records, and her husbands' mother, that also can be found in public records. Finding social security number and even name is possible but writing search strings for "husband" and "mother-in-law" is more complicated, especially since it's the context where the writing is found that determines if this is personal data or not.

It becomes even more complicated when you have writings of the kind "who won the national French championship in free-diving 1992". This phrasing is personal data, since it theoretically can be used to identify a person. At the same time, none of the words are personal data.

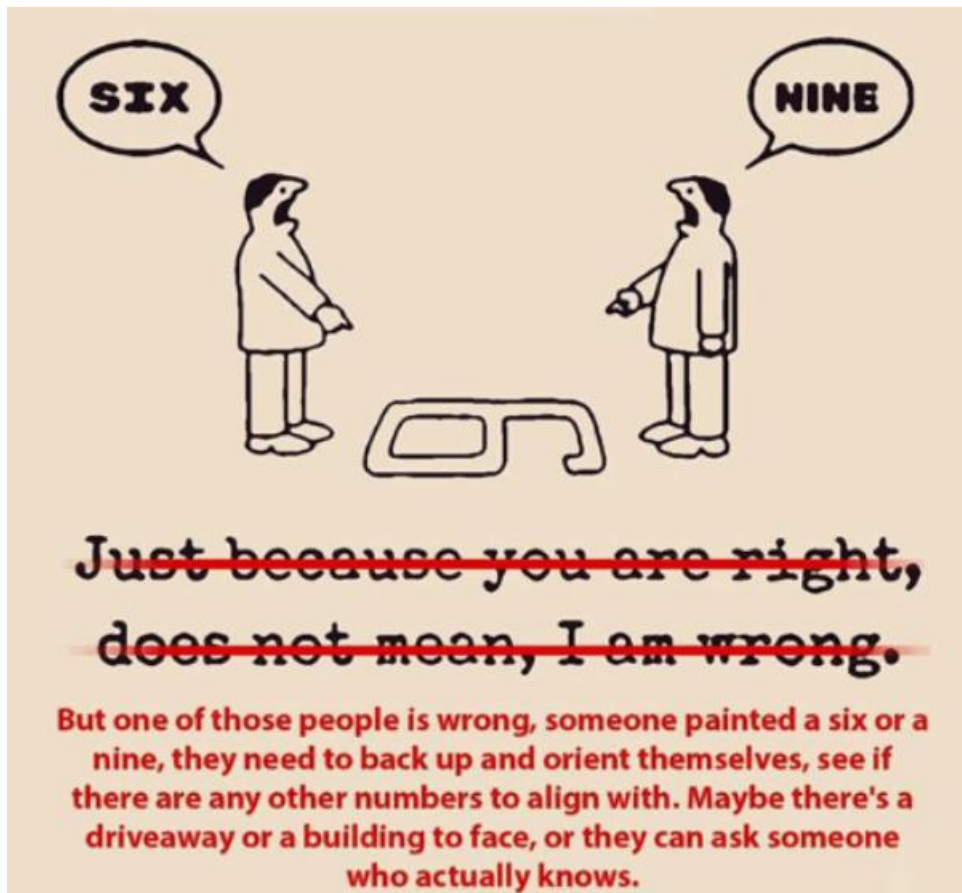
#### 2.2.2.1 Deep learning algorithms

Aigine gathers knowledge from the context where specific words becomes personal data. This knowledge is used in a neural network where deep learning is performed. Deep learning is a way to achieve pattern recognition, and the process is fueled by a large volume of high quality data. The result of this deep learning is refined algorithms, that are deployed in the local installations StoredIQ search.

#### 2.2.2.2 Learning semantics

Words are all about context, and one word can change meaning depending on the context where it appears. The english word "set" actually has 464 different definitions in the Oxford English Dictionary. What we actually mean when using the word "set" becomes clear in the context we use it. There is a big difference in "Set your alarm" and "He won the set". By using deep learning, Aigine learns the context where specific words are present. It can also happen that the same word is used in different ways depending on branch of industry. However, this also means that the context where the specific word is present will be different between these branches, Aigine will learn the different meaning of the words depending of branch.

The only time Aigine will struggle with learning a specific meaning of a word is if the context is not present in the document being scanned. This is a situation where two identical documents can have two different meanings. This is a strictly hypothetical situation, but its important to mention that also a human would have great difficulties in such situations, since it requires them to search for context elsewhere.



### 2.2.3 Filters out of documents containing suspected personal data

Aigine uses the downloaded definitions to determine whether the document contains suspected personal data. Only suspicious documents are included for review and assessment. When the scan is made iterative documents already audited are excluded by the use of metadata.

#### 2.2.3.1 Language independent

Since Aigine uses deep learning to understand the contextual meaning of words, Aigine will also learn different languages. This make the solution language independent, but of course, in order to learn context from new languages, annotation code from those languages are needed.

As with semantics, the same words are present in different languages, with a completely different meaning. This has given reason to some less successful branding mistakes, and the problem is solved in Aigine by securing waste amounts of high quality data in different languages.

This is needed for many organizations, as documents, files and emails can exist in many different languages within the same organization.

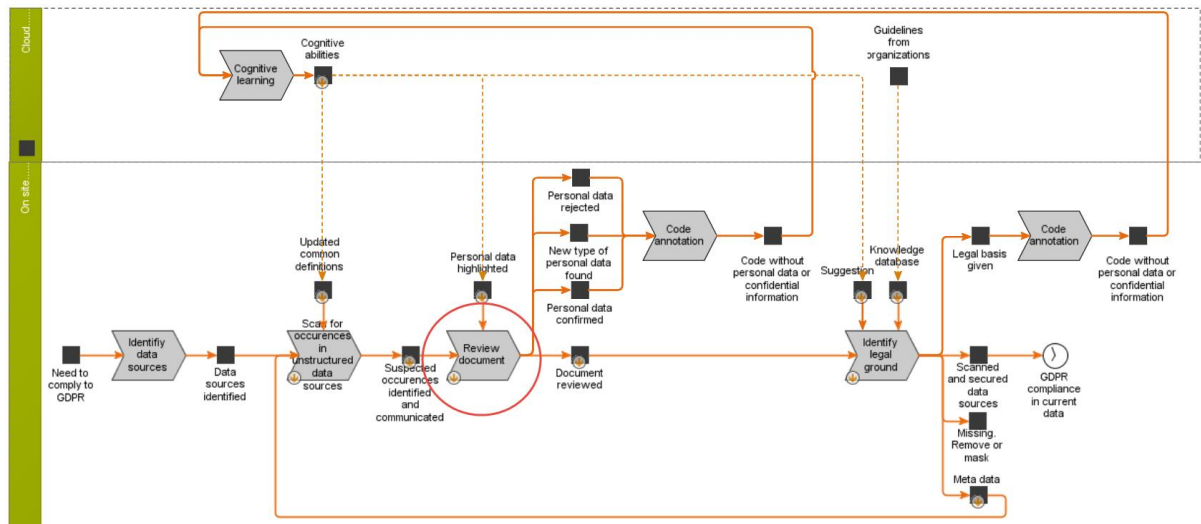
### 2.2.4 Filters out documents containing suspected personal data

Aigine uses the downloaded definitions to determine whether the document contains suspected personal data. Only suspicious documents are included for review and assessment. When the scan is made iterative documents already audited are excluded by the use of metadata.

### 2.2.5 Timesaver

Aigines filtering of documents that do not contain personal data will reduce the number of documents that need to be audited by 85-95%.

## 2.3 Review



Aigine has been created to offer a silver bullet solution to the challenge of unstructured data. For this reason, we have taken functionality from BPM and built these into the solution. Aigine therefore has a digital process that automates some work, increases the efficiency of the manual work and secures that review tasks and documentation are actually done.

### 2.3.1 Workflow

Aigine's workflow engine divides the hit list by data source and sends a unique list to each data source responsible by e-mail. A data source can be defined down at the folder level.

The list contains the documents with suspected personal data, which have not yet been reviewed by a human.

Each document in the list can be clicked to open the Web-based graphical interface.

#### 2.3.1.1 Delegate to anyone

The work flow engine also contains a function for the data source owner to delegate a specific document, folder or folders to any other person in the organization. This makes it possible to route responsibility to where it actually belongs and can be handled. The assigned person gets an email with a direct link to the delegated documents.

#### 2.3.1.2 Personal folders and e-mails

Firstly, it's important to make a distinction between personal and private folders and emails. Under GDPR, it's strongly recommended that the organisations servers does not contain private information at all.

Personal folders and emails are emails sent to or from a specific person within the organization, and files saved in locations where only a specific person has access, for instance a "My documents" folder.

It is clear in all jurisdictions that any information produced by an employee in the line of work is the sole property of the employer. However, it is not clear to what extent, and for what purposes, an employer has the right to read such personal emails and documents in personal folders.

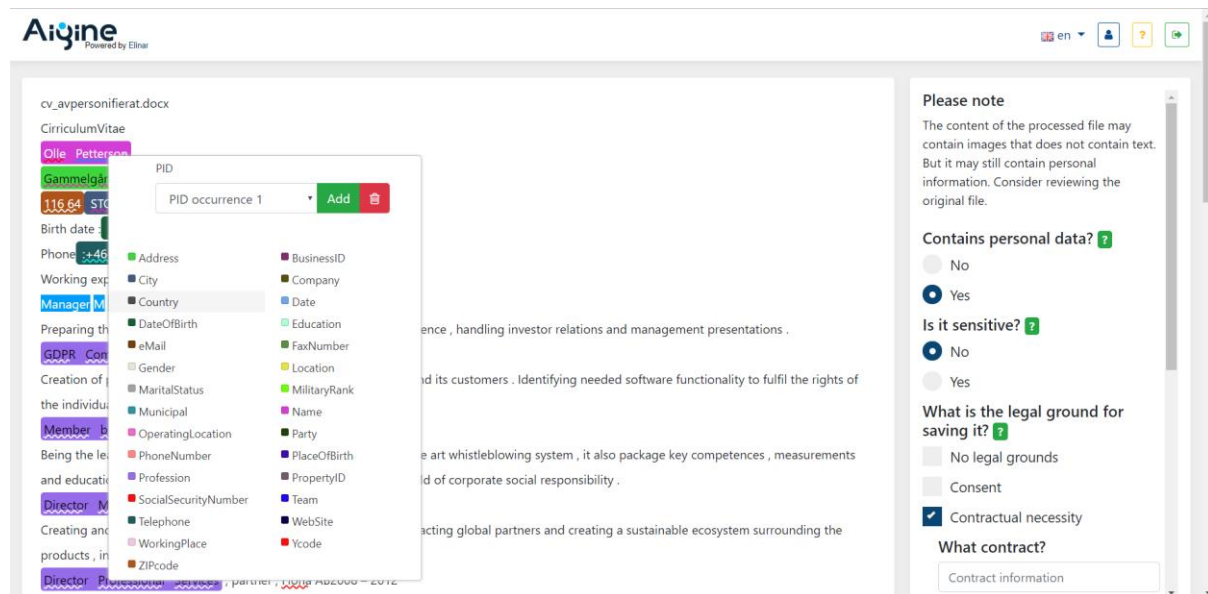
Aigine is not only based on legal insights, it's a solution that also takes process and competences into account in order to solve the challenge with personal data.

Our recommendation is therefore that personal file and email folders are delegated to the individual person that has been using them. This person is the most competent to decide whether

an email or document should be deleted or masked, or if it exists a legal ground to continue processing it, and whether it should be moved to a new location.

Such an approach also minimizes any legal questions about the right to read personal emails and folders.

## 2.3.2 GUI



*Aigine contains a web-based graphical user interface to facilitate the review of the document.*

### 2.3.2.1 Personal data highlighted

The suspected personal data is displayed highlighted, which speeds up the review.

### 2.3.2.2 Built-in support for confirming or rejecting personal information

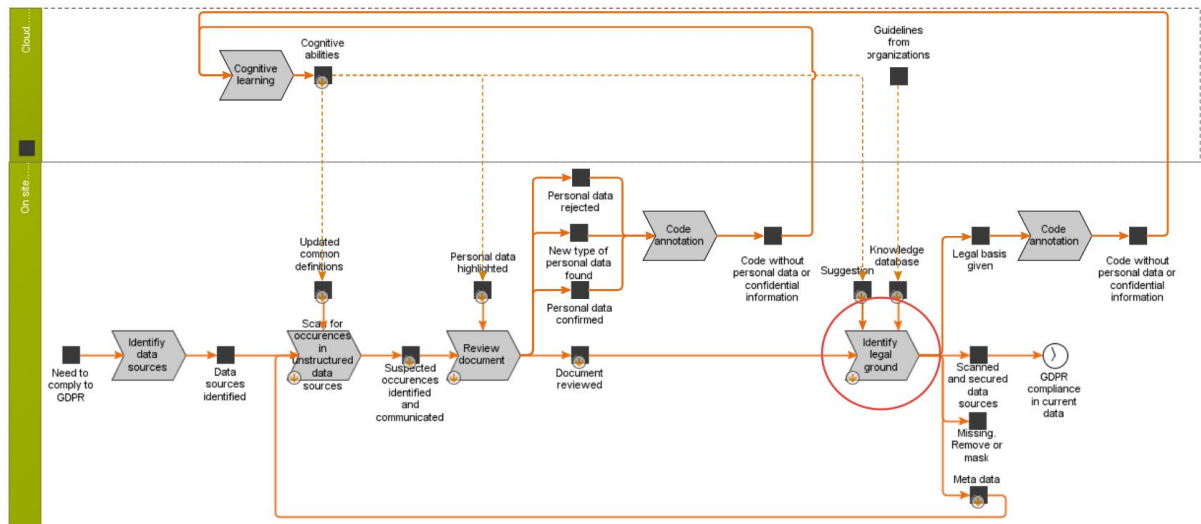
The tool provides intuitive features to mark, reject, or confirm personal information. You also specify what kind of personal data it is. This information can later be used for subject access rights and anonymization/pseudonymization.

### 2.3.2.3 Timesaver

By displaying the suspected personal data highlighted directly in the document, the review time is accelerated by 80-95%.

## 2.4 Decide legal basis





The requirements on the records of process makes it mandatory to know, and document, the legal basis for processing the personal data.

### 2.4.1 Contextual knowledge database

Aigine contains a contextual knowledge database that shows relevant information for the different legal decisions that needs to be taken in relation to the personal data in the document.

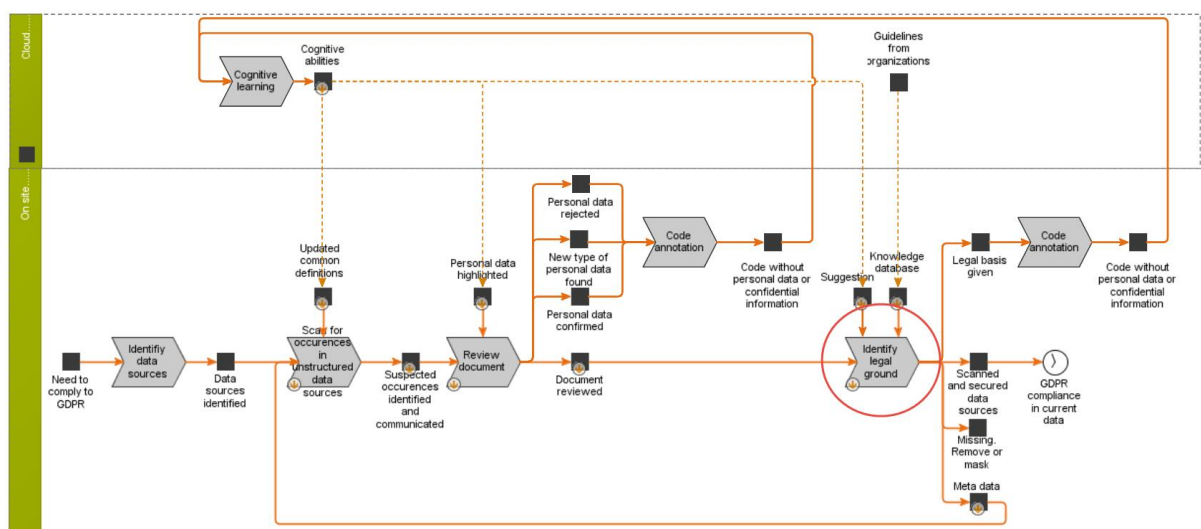
By default, the guidelines from the customers national privacy authority are showed.

Aigine can also be configured to display information from other sources. Such configuration, and also the production of relevant scripts, can be offered as a consulting service by resellers.

### 2.4.2 Time savings and the ability to distribute work

Access to a contextual knowledge database directly in the user interface streamlines the assessment. Those who do the assessment do not need to search for context and information from scattered sources. The contextual knowledge database also serves as a real-time competency enhancement, which means that the task can be performed by staff with lower legal training. This means that an efficiency increase of 80-90 % can be achieved, while legal practitioners do not become bottlenecks in the process.

## 2.5 Document





### 2.5.1 Integrated smart digital form for direct documentation

The graphical user interface of Aigine contains a smart, digital, form that is directly populated in accordance with requirements from the records of process. By using modern information gathering techniques, we can ensure that the documentation actually takes place, while facilitating the work performed by the user. Since each document, potentially, is a separate data source, and therefore has to be a separate item in the register, it is unclear how this is to be resolved in some other way.

### 2.5.2 Legal basis suggestions from AI

The cognitive abilities of Aigine also suggests legal basis for the process of personal data in the individual document. However, the final decision is always taken by a human being. These suggestions also speed up the legal assessment process by pointing the examiner in the right direction. This means, along with the contextual database, that the competency requirements of the person performing the task are reduced.

### 2.5.3 Time savings and the ability to distribute work

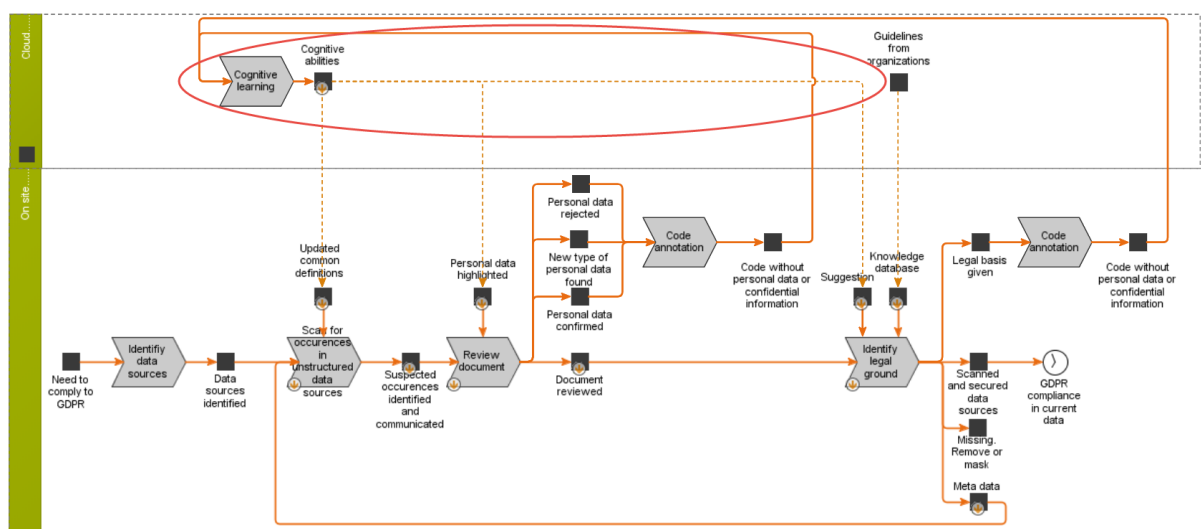
The cognitive abilities that suggests a legal basis streamlines the legal assessment process. The actual efficiency gain depends on many factors, where the most important is the reviewer's reliance on the AIs Ability. The accuracy of the proposals will increase with time, as the cognitive abilities are sharpened. An increase in efficiency of 50-95% is possible. Over time, the AI with its cognitive abilities, are expected to surpass humans in accuracy, making it possible for customers to totally automate this activity.

## 2.6 Collaborative cognitive learning

The functionality offered by Aigine is intended to reduce the required time to achieve a specific result. Therefore, to ensure increased performance over time, Aigine contains cognitive learning and artificial intelligence.

A prerequisite for cognitive learning is to assure sufficient data, both in quality and quantity, available for patterns recognition with deep learning. At the same time, Aigine must deliver customer value immediately after installation, with minimal customer and service effort. Therefore, a time-consuming traditional AI training project can't be a part of the delivery.

Instead, Aigine uses real human work from the actual process to train and improve its cognitive abilities.



### 2.6.1 Annotation coding done locally

In order to ensure compliance to both GDPR and other laws regulating secrecy, it is a prerequisite that personal data or prospective confidential information never leaves the customer. The architecture of Aigine is therefore designed in such a way that customer data is never sent from the client. Therefore, annotation code is created locally. The annotation code resembles machine code and cannot be recreated in any way to understandable information.

Annotation code is created every time a document is reviewed.

Annotation code is created every time a legal basis is given.

Note that this is not an encryption solution. Prepared data is exported in anonymized form since we are training on the structure of privacy data, not actual data. This means we do not need actual values like Peter for names. All possibly interesting tokens are replaced with place holders that are consistent within the document like FirstName001, FirstName002 and so on...ending with Integer001,Integer002).

Encryption and certificate management are managed in another chapter.

#### 2.6.1.1 Example of annotation code before Aigines anonymization

Annotation code traditionally used by AI contains code with personal data in plain text:

```
4|I Knudsen|N ,|U Elise|F <Nl>|X 5|I Sletteveien|T 10|In the 1617|Z Fredrikstad|N <Nl>|X 6
|In the 30073347897|'s <Nl>|X 7|U 410% 2032% 20911|O <Nl>|X 9|I
EliseKnudsen@teleworm.us|E <Nl>|X 28|I Lima|N Ellen|F <Nl>|X 29|I Kvalamarka|T 30|In
the 5514|Z HAUGESUND|T <Nl>|x 30|In the 11094847615|'s <Nl>|X 31% 20979% 2031%
20963|O <Nl>|X 111|I Sara|F The Fog|N <Nl>|X 112|In the 14077548891|'s <Nl>|X 116|I
Even|C P|C <Nl>|X 117|In the 05026548367|'s <Nl>|X 121|I Philip|F Mountain|N <Nl>|X 122
|In the 22117547918|'s <Nl>|X 126|I Leonard|F Hult|N <Nl>|X 127|In the 22117548388|'s
<Nl>|X 131|I Victoria|F Høydal|N <Nl>|X 132|In the 14049847875|'s <Nl>|x
```

#### 2.6.1.2 Example of annotation code after Aigines anonymization

To ensure that no personal data or confidential information ever leaves the customer, code generation takes place locally. This makes the code incomprehensible in relation to the underlying Data. Below is the above code washed with Aigines new technology.

```
Integer1|In LastName001|N|U FirstName001|F <Nl>|X Integer2|In Location001|T Integer3|
In ZipCode001|Z LastName002|N <Nl>|X Integer4|In SocSecNum001|'s <Nl>|X 7|U
PhoNum001|O <Nl>|X Integer5|In Email001|E <Nl>|X INTEGER6|In LastName003|N
FirstName002|F <Nl>|X Integer7|In Location002|T Integer8|In ZipCode002|Z Location003|T
<Nl>|X Integer8|In SocSecNum002|'s <Nl>|X PhoNum002|O <Nl>|X Integer10|In
FirstName003|F LastName004|N <Nl>|X Integer11|In SocSecNum003|'s <Nl>|X Integer12|C
<Nl>|X Integer13|In SocSecNum004|'s <Nl>|X Integer14|In FirstName004|F LastName006|N
<Nl>|X Integer15|In SocSecNum005|'s <Nl>|X Integer16|In FirstName005|F LastName007|N
<Nl>|X Integer17|In SocSecNum006|'s <Nl>|X Integer18|In FirstName006|F LastName008|N
<Nl>|X Integer19|In SocSecNum007|'s <Nl>|x
```

### 2.6.2 Cognitive learning in the cloud

Cognitive learning takes place in the cloud, by using the anonymized annotation code from each review and legal assessment. AI in the cloud is installed on servers controlled by Aigine AB.

Actual AI training is done using IBM Power HPC appliances. Each Appliance has 20-36 Power8/9 cores and 4 high end NVidia GPUs connected directly to Power CPUs using Nvidia proprietary NVLink bus for ultra-high speed communication. These appliances, with Volta V100 GPUs, have ~500 TeraFLOPS of deep learning capacity each.

### 2.6.3 Cognitive abilities improve

Through this solution, we can ensure that the AI gets Sufficient data to quickly improve its cognitive abilities. This makes Aigine continuously improve its ability to understand what constitutes personal data, how to highlight personal data, and, based on the structure and composition of the document, the legal basis that may exist for processing it.

### 2.6.4 Everyone gets benefits from the improved skills

Aigines sharpened cognitive abilities are shared with each customer and are downloaded from the cloud to the customer's local installation. This means that the accuracy of the filtering is constantly increasing, which can mean more or less suspected documents found, depending on the content of the documents.

A continuously improved understanding of the context also increases the ability to highlight relevant personal data in each document, which further streamlines the human review process.

Aigines suggestions for legal basis will also be improved with higher relevance and accuracy. In the foreseeable future, people will still take the legal decision, but Aigines improved suggestions will speed up this decision-making further.

Everyone teaches. Everyone benefits from enhanced abilities.

## 2.7 Timesaver

Aigine has combined experience from Business Process Management, organizational development and legal assessment work in order to reduce the time it takes to perform data discovery, review and document the content of unstructured data sources. The reduction in needed time is so great that it is both practical and theoretically possible for organizations to approach the challenge.

Aigines knowledge database and suggestions for legal support also mean that the legal training requirements of the resources that carry out the work are reduced. This also helps to make the work effort both theoretically and practically possible as the work can be allocated to more resources in the customer's organization.

It is always difficult to assess the real time savings as it depends on a large number of factors. At the same time, performance, and thus efficiency, will increase with enhanced cognitive abilities.

### 2.7.1 Example

#### Work effort without support

Amount of unstructured data	1	Terabytes
Average document size	2	MB/Document
Average time for review	2	Minutes/Document
Percentage that contains personal data	10	Percent
Average time for legal decision-making on legal basis	3	Minutes/Document
Average time to perform documentation	2	Minutes/Document
<b>Total work</b>	<b>11.84</b>	<b>Man years</b>

#### Efficiency with Aigine

Efficiency filtering	90%	Percent
Streamlining review	90%	Percent
Efficiency increase decision making	90%	Percent
Efficiency increase documentation	90%	Percent
<i>Total efficiency</i>	97%	<i>Percent</i>
<b>Remaining work</b>	<b>0.33</b>	<b>Man years</b>

The calculation refers to 1 terabyte of data. The assumption has been made that each document is 2MB large on average. To read through each document to determine whether or not it contains personal data is assumed to take 2 minutes per document.

Assumption is made that 10% of the documents actually contains personal data, which is in line with the inventories made during the development of the solution.

Time to make the legal assessment is assumed to 3 minutes. This includes information collection from sources such as the privacy authority. The documentation of each document is assumed to take 2 minutes. In this documentation type of personal data, purpose, legal basis, deletion time, hash linking for reference etc. are included.

Aigine streamlines the work by presenting only documents that actually contains suspected personal data by filtering. This cut work time by 90%.

Aigine streamlines the review of the document by 90% by displaying suspected personal data highlighted.

Aigine streamlines the decision-making process to find a legal basis with 90% by cognitive abilities making suggestions, and providing contextual help directly from relevant sources.

Aigine streamlines the documentation by 90% by providing documentation directly in the graphical user interface. The information is gathered in a smart, digital, form where relevant fields are displayed.

### 2.7.2 Make your own simulation

To make your own simulations, with custom values, spreadsheets for calculation can be found on the link below:

<https://1drv.ms/x/s!At8KNeGh0QotjKZfj-tPnUabaz0pCw>

## 2.8 The silver bullet approach

The development of Aigine has been preceded by several years of process analysis to better understand what processes need to be adopted in an organization to comply with the regulations and individual rights in GDPR. We have chosen to focus on the practical deliverables needed to be created, and Aigine does not provide a platform for policy development. We handle what is produced by the customers business processes, not the design of the processes, nor why they have been developed. Some features described below will be offered as add-ons to the inventory engine.

### 2.8.1 Data Discovery

Aigine performs state of the art data discovery, using collaborative cognitive deep learning to create continuously improved cognitive abilities and by that increase data discovery performance.

### 2.8.2 Data Classification

Aigine contains a meta data layer for data classification, and uses dynamic forms, a contextual knowledge database and collaborative cognitive deep learning to create continuously improved cognitive abilities to suggest classification.

### 2.8.3 Records of Processing

Aigine saves all meta data, creating a record of processing even down on document/file level.

### 2.8.4 Information management

By using cognitive abilities together with meta data, Aigine can use these resources to perform automated information management, making it possible to handle anonymization, pseudonymization and deletion of personal data and files with add-ons.

### 2.8.5 Monitoring

By using cognitive abilities and meta-data, Aigine makes it possible to monitor all data sources and ensures compliance towards both GDPR and internal routines with add-ons.

### 2.8.6 Compliance training

By having workflow connected to monitoring, Aigine enforces continuous compliance training, not only by information personnel when they are making mistakes, but also providing an explanation why it was a mistake – together with guidance to correct the mistake. These features will be available in upcoming add-ons, and will combine meta data, cognitive abilities and your internal policies for rules management.

### 2.8.7 Subject access request

By combining the meta data layer and cross-references from the datasets of PD, Aigine can deliver results for subject access rights in seconds.

### 2.8.8 Anonymization and pseudonymization

By holding data regarding legal ground and placement of PD in the data sources, Aigine can offer one click anonymization or pseudonymization.

### 2.8.9 Data protection impact assessment

Aigine does not contain methods to perform data protection impact assessment.

### 2.8.10 Rights Management, Breach

Aigine does not contain functions for identity nor access management.

### 2.8.11 Secure transfers

Aigine does not contain functions to ensure secure transfers in network.

### 2.8.12 Reporting of personal data incidents

Aigine does not contain functionality to report personal data incidents.

## 3 TECHNICAL DESCRIPTION

### 3.1 License restriction

Aigine consists of several highly advanced software products, packaged to solve the challenge of personal data in unstructured data and offered at a manageable cost.

The right to use all included software is therefore strictly limited to the functions described in this document and the EULA that the customer receives and accepts. Customers who wish to use additional abilities in any of the included software must therefore acquire a separate license for these functions.

### 3.2 Included components

#### 3.2.1 Aigine Unstructured Data Inventory Engine

Locally installed software that uses abilities from StoredIQ Watson NLU and Elinar AI Miner together with custom functions and GUI.

##### 3.2.1.1 User interface

The software also includes a GUI for system configuration as well as GUI for end users.

This GUI, in turn, consists of several parts, including the selection of suspected personal data, simple confirmation of data, knowledge database and data collection forms.

##### 3.2.1.2 Work flow

The software includes a work flow function that delegates responsibility by e-mail to persons being identified as responsible.

##### 3.2.1.3 Meta data layer

The software includes a meta data layer and reporting by email.

Basic metadata and workflow-related metadata are saved for each document in a database (Postgresql).

#### 3.2.2 StoredIQ

Runs as a service and performs search for personal data in connected data sources. Increases the ability to make informed decisions based on large amounts of information. Creates patterns and manages all unstructured data in the same system.

#### 3.2.3 Watson Ai and Watson NLU

Artificial Intelligence (AI) and that annotation technique Natural Language Understanding NLU, also known as Machine Intelligence, MI, is intelligence seen in learned machines. Artificial intelligence is very useful in iterative processes, where the technology independently creates information needed in the next iterative sequence. In some cases, tasks with large amounts of data where iterative processes with elements of recursion (iterative calculations with result feedback and requirements for experience learning) be so complex and voluminous that a human being is unable to come to a corresponding result within a reasonable time. Artificial intelligence in these cases means a direct and extensive time saving. In addition, high quality gains are being made in that a task can be moved from a human being to a computer that, according to a learning process, imitates and refines the human work of the learning process, and does as it should and cannot be careless.

##### 3.2.3.1 Cognitive learning

Cognitive IT and artificial intelligence AI are concepts in Cognitive Computing. The expression was coined by IBM, during the work with the Watson computer. Cognitive IT and AI are used interchangeably to describe the same things. Depending on whether different types of algorithms,



platforms, machine learning or neural deep learning (Deep Learning) network is meant. AI is usually used as a name for technical solutions that can perform tasks where usually human intelligence is required.

Cognitive IT is defined by machine intelligence, which can be described as algorithmic functionality used to improve people's performance, automate complex tasks and create so-called cognitive agents that can simulate human thinking and action.

Three main areas of cognitive IT:

- Automation of repeatable tasks, with the intention of increasing efficiency, quality and accuracy.
- Create insights and discover, identify hidden patterns and relationships.
- To create enhanced solutions for personalization (Hyperpersonalization).
- All applications for advanced classification and predictive analysis, and present suggested solutions.

#### 3.2.4 Elinar AI Miner

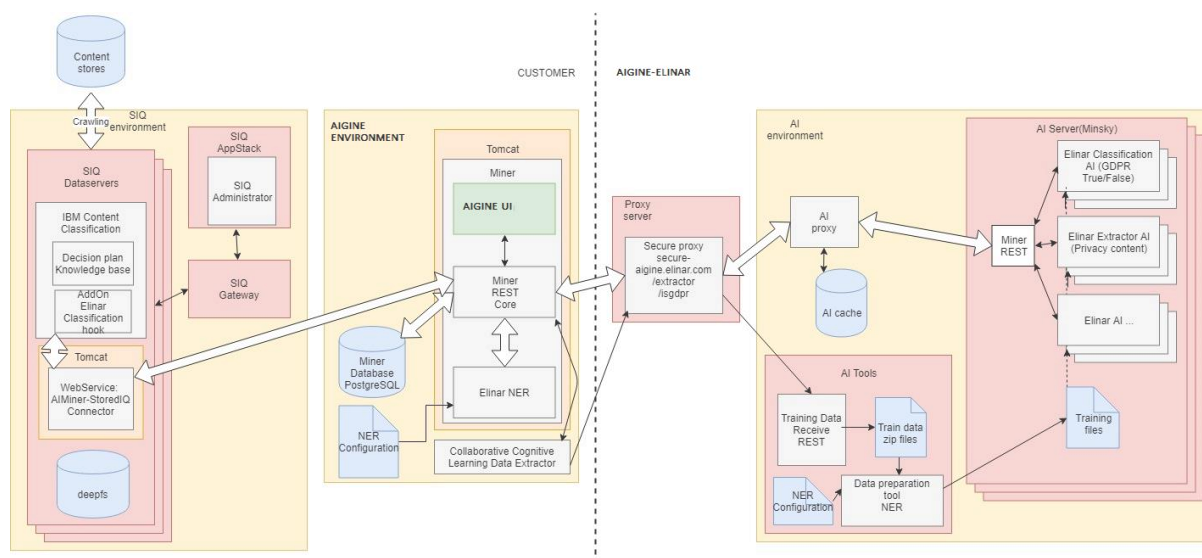
System that provides annotation code from the various local systems to the central cloud solution's artificial intelligence (AI). AI Miner uses a technology called Multilayer Artificial Neural Network (ANN) to create machine understanding and converts the local data into person-independent annotation code in processes that form the basis of learning processes in the cloud by Natural Language Understanding NLU.

#### 3.2.5 Cloud AI

The cloud instance of AI receives the annotation code from local installations and uses it to perform deep learning for both personal data and legal basis. The deep learning produces algorithms that are sent to customer installation and are used in StoredIQ to find personal data and in the GUI to highlight personal data. Cognitive abilities to suggest legal basis are also used in the GUI.

Actual AI training is done centrally using IBM Power HPC appliances. Each Appliance has 20-36 Power8/9 cores and 4 high end NVidia GPUs connected directly to Power CPUs using Nvidia proprietary NVLink bus for ultra-high speed communication. These appliances (With latest Volta V100 GPUs) have ~500 TeraFLOPS of deep learning capacity each.

### 3.3 System design



Aigine combines IBM StoredIQ, IBM Watson NLU annotation module and Elinar AI-Miner to a comprehensive solution for filtering, reading, decision-making on legal basis and documentation of decisions taken.

The system is installed and stores all locally and communicates only anonymized annotation code with the cloud-based AI.

### 3.3.1 Encryption and certificate management

Aigine Uses HTTPS (SSL) Certificates issued by [GlobalSign](#). This solution uses so-called Wild Card Certificate.

Communication between Aigine UDIE and AI-Miner uses a method with pre-installed root certificates, which means that both components of the communication chain have been part of the certificate structure that enables communication.

## 3.4 System

Aigine is packaged in Linux based VMware to facilitate and streamline the installation.

### 3.4.1 Downloading installation packages

The installation package is available on an SFTP server (This download requires an SFTP Client).

Ip Address	95.175.106.41
Port	1222
Protocol	Sftp
Username	Aigine
Password	<i>Delivered to end customer when ordering</i>

### 3.4.2 IP addresses

4 Ip Addresses need to be allocated

- Aigine Unstructured Data Inventory Engine
- IBM StoredIQ Appstack
- IBM StoredIQ Gateway
- IBM StoredIQ Data server

### 3.4.3 Space that needs to be allocated for installation packages

Name	Disk size (Thin Commission), GB	Disk size (Thick Commission), GB	OS version
Aigine Unstructured Data Inventory Engine			
Ai-PM-Medium-1.0.1	5.0	80	Ubuntu Server 14.04
IBM StoredIQ			
IBM SIQ Appstack	3.0	35	<i>Centos 6.5 (Final)</i>
IBM SIQ Gateway	3.1	215	<i>Centos 6.5 (Final)</i>
IBM SIQ Data server	9.7	2100	<i>Centos 6.5 (Final)</i>
Total	20.8	2430	

#### Other software included in the download package is:

- Content Classification 8.8
- Embedded WebSphere Application Server 8.5.5

### 3.4.4 Requirements for ports to enable communication to the cloud

From	To	Port(s)	Comments
Aigine Unstructured Data Inventory Engine			
	Ai-PM-Medium-1.0.1	8080	User interface access for AI
IBM StoredIQ			
IBM SIQ Appstack	IBM SIQ Gateway	8765, 5432	
IBM SIQ Gateway	IBM SIQ Appstack	8765, 5432	
IBM SIQ Data server	IBM SIQ Gateway	11103	
IBM SIQ Admin Console	IBM SIQ Appstack, IBM SIQ Data server	80, 443, 22	Used by administrator for administration of IBM StoredIQ
IBM SIQ Admin Console	IBM SIQ Gateway	22	

### 3.4.5 Other system requirements:

Name	Memory allocation	CPU Power allocation
Aigine UDI Engine	6GB	4 VCPU
IBM SIQ Data server	16GB	4 VCPU
IBM SIQ Gateway Server	8GB	2 VCPU
IBM SIQ Application Stack	4GB	1 VCPU

#### Server Capacity for StoredIQ

Each Data server handles up to 30 TB of stored data (this can vary based on the number of volumes, the number of items per volume, and the type of object).

Each gateway handles up to 50 Data servers.

Only one application server is needed – Regardless of the number of other servers.

#### VMware Supported Versions:

- VMware Esxi 5.0
- VMware Esxi 6.0
- VMware VSphere 5.0
- VMware Vsphere 6.0

## 3.5 Browser support

We recommend the latest version of any of the following browsers:

- Mozilla Firefox
- Google Chrome
- Apple Safari
- Microsoft Edge
- Internet Explorer

All of the above browsers are supported by Aigine. Performance may vary between browsers. Best performance is achieved with Chrome and Firefox Quantum.

### 3.6 File format and protocol support

IBM StoredIQ is used for scanning and inventory of data sources,. This software has a very large support for different file types and protocols. Additionally, the number of supported file types and protocols is continuously extended, and new versions of StoredIQ, with support for new file types, will be included in the updates of the solution.

Additional information on Supported file formats in the solution can be found at the following link: <http://www.redbooks.ibm.com/redpapers/pdfs/redp5316.pdf>

#### 3.6.1 File formats that can be scanned by Aigine

Format	Extension	Category	Version
Adobe Acrobat	Pdf	Graphic	2.1 3.0-7.0 Japanese
Adobe FrameMaker Graphics	Fmv	Graphic	Vector/raster Through 5.0
Adobe FrameMaker Interchange Format	Mif	Word Processing	3.0-6.0
Adobe Illustrator		Graphic	Through 7.0 9.0
Adobe Photoshop	Psd	Graphic	4.0
Ami Draw	SDW	Graphic	All
Ansi	Txt	Text and Markup	7-and 8-bit
Ascii	Txt	Text and Markup	7-and 8-bit
Autocad	Dwg	Cad	2.5-2.6 9.0-14.0 2002 2004 2005
Autoshade Rendering	Rnd	Graphic	2.0
Binary Group 3 Fax		Graphic	All
Bitmap	BMP, RLE, ICO, CUR, DIB, WARP	Graphic	All
CALS Raster	GP4	Graphic	Type I, II
Comma-separated Values	Csv	Spreadsheet	
Computer Graphics Metafile	Cgm	Graphic	Ansi Cals NIST 3.0
Corel Clipart	CMX	Graphic	43226
Corel Draw	Cdr	Graphic	3. x-8. x
Corel Draw (CDR with Tiff header)		Graphic	2. x-9. x
Corel Presentations	Shw	Presentation	Through 12.0 X3
Corel WordPerfect Windows	Wpd	Word Processing	Through 12.0 X3
DataEase		Database	4. X
Dbase Database		Database	Through 5.0
DBXL		Database	1.3
DEC WPS PLUS	Dx	Word Processing	Through 4.0

DEC WPS PLUS	Wpl	Word Processing	Through 4.1
DisplayWrite (2 and 3)	Ip	Word Processing	All
DisplayWrite (4 and 5)		Word Processing	Through 2.0
Dos Command Executable	Com	System	
Dynamic Link Library Files	Dll	System	
Ebcdic		Text and Markup	All
ENABLE		Word Processing	3.0 4.0 4.5
ENABLE		Database	3.0 4.0 4.5
ENABLE Spreadsheet	Ssf	Spreadsheet	3.0 4.0 4.5
Encapsulated Post-Script (raster)	Eps	Graphic	Tiff Header
Executable Files	Exe	System	
First Choice		Database	Through 3.0
First Choice		Word Processing	Through 3.0
First Choice		Spreadsheet	Through 3.0
FoxBase		Database	2.1
Framework		Database	3.0
Framework		Word Processing	3.0
Framework		Spreadsheet	3.0
GEM Bit Image	Img	Graphic	All
Graphics Interchange Format	Gif	Graphic	All
Graphics Environment Manager	GEM VDI	Graphic	Bitmap and Vector
Gzip	Gz	Archive	All
Haansoft Hangul	HWP	Word Processing	1997 2002
Harvard Graphics (DOS)		Graphic	2. x 3. x
Harvard Graphics (Windows)		Graphic	All
Hewlett-Packard Graphics Language	HPGL	Graphic	2
Html	.htm	Text and Markup	Through 3.0
IBM FFT		Text and Markup	All
IBM Graphics Data Format	Gdf	Graphic	1.0
IBM Picture Interchange Format	Pif	Graphic	1.0
IBM Revisable Shape Text		Text and Markup	All
IBM Writing Assistant		Word Processing	1.01
Initial Graphics Exchange Spec	IGES	Graphic	5.1
Java Class Files	CLASS	System	
JPEG (not in TIFF format)	JFIF	Graphic	All
Jpeg	Jpeg	Graphic	All
JustSystems Ichitaro	Jtd	Word Processing	5.0 6.0 8.0-13.0 2004
JustSystems Write		Word Processing	Through 3.0
Kodak Flash Pix	Fpx	Graphic	All
Kodak Photo CD	Pcd	Graphic	1.0

Legacy		Word Processing	Through 1.1
Legato Email Extender	Emx	Email	
Lotus 1-2-3	WK4	Spreadsheet	Through 5.0
Lotus 1-2-3 (OS/2)		Spreadsheet	Through 2.0
Lotus 1-2-3 Charts	123	Spreadsheet	Through 5.0
Lotus 1-2-3 for Smartsuite		Spreadsheet	1997-Millennium 9.6
Lotus AMI Pro	Sam	Word Processing	Through 3.1
Lotus Freelance Graphics	PRZ	Presentation	Through Millennium
Lotus Freelance Graphics (OS/2)	Pre	Presentation	Through 2.0
Lotus Manuscript		Word Processing	2.0
Lotus Notes	Nsf	Email	
Lotus Pic	Pic	Graphic	All
Lotus Snapshot		Graphic	All
Lotus Symphony		Spreadsheet	1.0 1.1 2.0
Lotus Word Pro	Lwp	Word Processing	1996-9.6
LZA Self Extracting Compress		Archive	All
Lzh Compress		Archive	All
Macintosh PICT1/2	PICT1/PICT1	Graphic	Bitmap Only
Macpaint	PNTG	Graphic	Na
MacWrite II		Word Processing	1.1
Macromedia Flash	Swf	Presentation	Text Only
MASS-11		Word Processing	Through 8.0
Micrografx Designer	Drw	Graphic	Through 3.1
Micrografx Designer	Dsf	Graphic	Win95, 6.0
Micrografx Draw	Drw	Graphic	Through 4.0
MPEG-1 Audio Layer 3	Mp3	Multimedia	ID3 metadata Only These files can be harvested, but there is no data in them that can be used in tags.
MS Access	Mdb	Database	Through 2.0
MS Binder		Archive	7.0-1997
MS Excel	Xls	Spreadsheet	2.2-2007
MS Excel Charts		Spreadsheet	2. X-7.0
MS Excel (Macintosh)	Xls	Spreadsheet	3.0-4.0 1998 2001 2004
MS Excel XML	Xlsx	Spreadsheet	
Ms Multiplan		Spreadsheet	4.0
MS Outlook Express	Eml	Email	1997-2003
MS Outlook Form Template	OFT	Email	1997-2003
MS Outlook Message	Msg	Email	All
MS Outlook Offline Folder	Cheese	Email	1997-2003
MS Outlook Personal Folder	Pst	Email	1997-2007
MS PowerPoint (Macintosh)	Ppt	Presentation	4.0-2004
MS PowerPoint (Windows)	Ppt	Presentation	3.0-2007
MS PowerPoint XML	Pptx	Presentation	
MS Project	Mpp	Database	1998-2003
MS Windows XML	Docx	Word Processing	
MS Word (Macintosh)	Doc	Word Processing	3.0-4.0 1998 2001



MS Word (PC)	Doc	Word Processing	Through 6.0
MS Word (Windows)	Doc	Word Processing	Through 2007
Ms Wordpad		Word Processing	All
MS Works	S30/S40	Spreadsheet	Through 2.0
MS Works	Wps	Word Processing	Through 4.0
MS Works (Macintosh)		Word Processing	Through 2.0
MS Works Database (Macintosh)		Database	Through 2.0
MS Works Database (PC)		Database	Through 2.0
MS Works Database (Windows)		Database	Through 4.0
Ms Write		Word Processing	Through 3.0
Mosaic Twin		Spreadsheet	2.5
MultiMate 4.0		Word Processing	Through 4.0
Navy Dif		Word Processing	All
Nota Bene		Word Processing	3.0
Novell Perfect Works		Word Processing	2.0
Novell Perfect Works		Spreadsheet	2.0
Novell Perfect Works (Draw)		Graphic	2.0
Novell WordPerfect		Word Processing	Through 6.1
Novell WordPerfect (Macintosh)		Word Processing	1.02-3.0
Office Writer		Word Processing	4.0-6.0
Openoffice Calc	SXC/ODS	Spreadsheet	1.1 2.0
Openoffice Draw		Graphic	1.1 2.0
Openoffice Impress	SXI/SXP/ODP	Presentation	1.1 2.0
Openoffice Writer	SXW/ODT	Word Processing	1.1 2.0
OS/2 PMMetafile Graphics	Met	Graphic	3.0
Paint Shop Pro 6	Psp	Graphic	5.0-6.0
Paradox Database (PC)		Database	Through 4.0
Paradox (Windows)		Database	Through 1.0
PCFile Letter		Word Processing	Through 5.0
PC-File + Letter		Word Processing	Through 3.0
Pc Paintbrush	PCX, DCX	Graphic	All
Pfs: Professional Plan		Spreadsheet	1.0
Pfs: Write		Word Processing	A, B, C
Portable Bitmap Utilities	PbmLlll	Graphic	All
Portable Greymap	Pgm	Graphic	Na
Portable Network Graphics	Png	Graphic	1.0
Portable Pixmap Utilities	Ppm	Graphic	Na
Postscript File	Ps	Graphic	Level li
Professional Write		Word Processing	Through 2.1
Professional Write Plus		Word Processing	1.0
Progressive JPEG		Graphic	Na
Q & A (database)		Database	Through 2.0

Q & A (DOS)		Word Processing	2.0
Q & A (Windows)		Word Processing	2.0
Q & A Write		Word Processing	3.0
Quattro Pro (DOS)		Spreadsheet	Through 5.0
Quattro Pro (Windows)		Spreadsheet	Through 12.0 X3
R:BASE 5000		Database	Through 3.1
R:BASE Staff		Database	1.0
R:BASE System V		Database	1.0
Rar	Rar	Archive	
Reflex Database		Database	2.0
Rich Text Format	Rich text	Text and Markup	All
SAMNA Word IV		Word Processing	
Smart Ware li		Database	1.02
Smart Ware li		Word Processing	1.02
Smart Ware li		Spreadsheet	1.02
Sprint		Word Processing	1.0
Staroffice Calc	SXC/ODS	Spreadsheet	5.2 6. x 7. x 8.0
Staroffice Draw		Graphic	5.2 6. x 7. x 8.0
Staroffice Impress	SXI/SXP/ODP	Presentation	5.2 6. x 7. x 8.0
Staroffice Writer	SXW/ODT	Word Processing	5.2 6. x 7. x 8.0
Sun Raster Image	RS	Graphic	Na
Supercalc Spreadsheet		Spreadsheet	4.0
Text Mail (MIME)	Various	Email	
Total Word		Word Processing	1.2
Truevision Image	Tiff	Graphic	Through 6
Truevision Targa	Tga	Graphic	2
Unicode Text	Txt	Text and Markup	All
Unix TAR (tape Archive)	Takes	Archive	Na
Unix Compressed	Z	Archive	Na
UUEncoding	Uue	Archive	Na
Vcard		Word Processing	2.1
VisioPreview)		Graphic	4
Visio 2003		Graphic	5 2000 2002
Volkswriter		Word Processing	Through 1.0
Vp Planner 3d		Spreadsheet	1.0
WANG PC		Word Processing	Through 2.6

Wbmp		Graphic	Na
Windows Enhanced Metafile	Emf	Graphic	Na
Windows Metafile	Wmf	Graphic	Na
Winzip	Zip	Archive	
Wml		Text and Markup	5.2
WordMARC Word Processor		Word Processing	Through Composer
WordPerfect Graphics	WPG, WPG2	Graphic	Through 2.0, 7. and 10
Wordstar		Word Processing	Through 7.0
Wordstar 2000		Word Processing	Through 3.0
X Bitmap	Xbm	Graphic	X10
X Dump	Xwd	Graphic	X10
X Pixmap	Xpm	Graphic	X10
XMLGeneric)	Xml	Text and Markup	
XyWrite	XY4	Word Processing	Through III Plus
Yahoo! IM Archive		Archive	
Zip	Zip	Archive	PKWARE-2.04 g

### 3.6.2 Cloud services and local systems which can be scanned by Aigine

Platform/Protocol
Box
CIFS/SMB or SMB2
CMIS 1.0
Connections
Emc Centera
Emc Documentum
HDFS (HADOOP)
Hitachi HCAP
IBM Content Manager
IBM Domino
IBM Filenet
Jive
Microsoft Exchange
Microsoft SharePoint
Newsgator
Nfs
OneDrive For Business
OpenText Livelink/Content Server
Salesforce Chatter
Symantec Discovery Accelerator
Symantec Enterprise Vault

### 3.6.3 TIFF and other image formats

To manage images, Aigine uses the built-in OCR functionality of StoredIQ. StoredIQ is a software designed to handle large amounts of data, the built-in OCR function is designed for performance and not quality.

Therefore, a customer who holds, and intends to handle, larger amounts of TIFF or other image formats in its unstructured data, should ask for a separate OCR project by the reseller.

In this project, the text content is digitized and saved to a storage area before Aigine is used for filtering and manage the content.

## 3.7 Updates

Aigine continually updates the solution with bugfixes, security enhancements, performance optimizations, improved interfaces, and possible additional functionality.

When a new update of any of the included components is available, a qualification is made whether the updates of the included components contains relevant changes that should be incorporated into the solution. New versions will therefore be released two to four times a year.

### 3.7.1 Distribution

The availability of a new version of Aigine is communicated via e-mail containing information on the content of the new version.

E-mail is sent both to the customer's contact person and to the contact person at the reseller, so resellers are given an opportunity to address the question of updating with the customer.

### 3.7.2 Downloadable package

Updates are distributed as downloadable packages from Aigine, with included softwares already configured and installed, and includes custom upgrade instructions and relevant test schedules to ensure functionality after update.

The package is installed locally at the customer.

## 3.8 Server Governance

Aigine is supplied as images on four servers installed on VMware with Linux as the operating system. Aigine does not undertake the management of these servers, including continuous installations of patches etc.

## 3.9 Installation in outsourced environments

In the event that the end customer has outsourced its IT-environment, Aigine can still be installed in the outsourced operating environment provided that only licensed customer data is connected. We are assuming that the legal conditions between the customer and the reseller, with regard to the creation of a personal data-processing agreement, have been clarified. Aigine assumes no responsibility for this being the case.

## 4 DEMONSTRATION OF THE SOLUTION

Aigine is a comprehensive solution including a large number of competent software. Aigine also adds business value to the customer on first use, through the filtration process using Aigine's cognitive abilities, partly through the feedback of only those documents containing suspected personal data.

For this reason, there is no "try it" or Proof-of-concept including installation and test runs in customer environment with customer data.

For demonstration of user interfaces, abilities and functionality, therefore, we are referring to a Web-based cloud installation of Aigine.

In this demo environment, customers can upload individual documents to see how cognitive abilities highlights personal data, and how personal data is confirmed in the user interface. It is also possible to see how the knowledge base interacts with the user and how legal ground is proposed by AI and documented by the user.

Aigine does not offer the signing of a personal data processor agreement with the customer for these test runs. Instead, uploaded data is deleted at latest when the web session is terminated.

### 4.1 End-User cloud demo environment

Demo environment can be accessed at the following URL:

<https://demo-aigine.elinar.com>

Instructions and credentials will be given by your reseller or distributor.

## 5 DOCUMENTATION AND MARKETING MATERIALS

Aigine works continuously with the development and improvement of documentation. Aigine addresses a global market for its value proposition, which assumes that customer management must be done by resellers who own both customer and relationship. A precondition for this is that the dossiers are of such a level that this is possible.

### 5.1 Marketing material

#### 5.1.1 This Product bulletin

This product bulletin is available in its latest version for download here:

[http://aigine.se/files/PRODUCT\\_BULLETIN\\_AIGINE\\_UDIE\\_EC\\_ENGLISH.pdf](http://aigine.se/files/PRODUCT_BULLETIN_AIGINE_UDIE_EC_ENGLISH.pdf)

#### 5.1.2 Process design

Aigine's functionality, architecture and abilities are described in a published process design that is accessed here. Note that the process design consists of different layers, and is clickable into objects. The design also contains different layers.

[http://aigine.se/Aigine\\_GDPR\\_Process/index.html](http://aigine.se/Aigine_GDPR_Process/index.html)

#### 5.1.3 Website

Our website is continuously expanded, including Customer Testimonials. It is reached here

<http://aigine.se/en/>

#### 5.1.4 Blog

A blog is present on our homepage. It contains news about activities, new customers, new products and functions and general info about GDPR and other legislations.

<http://aigine.se/blog/>

#### 5.1.5 PowerPoint presentations

Current presentations can be found on our website.

#### 5.1.6 Flyers

Flyers can be found on our website.

#### 5.1.7 Video

Produced video material is presented in Aigine's YouTube Channel.

[https://www.youtube.com/channel/UCFkDMg9k0EcaAz2wmJPQ\\_Lg](https://www.youtube.com/channel/UCFkDMg9k0EcaAz2wmJPQ_Lg)

### 5.2 Technical documentation

#### 5.2.1 Debugging scheme to exclude non Product Fault.

Aigine only takes responsibility for pure product faults. Reseller, or customer, must therefore ensure that any fault is related to the product. This is done through a troubleshooting schedule found here:

[http://aigine.se/files/TROUBLESHOOTING\\_SCHEME\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/TROUBLESHOOTING_SCHEME_AIGINE_UDIE_ENGLISH.pdf)

Please note that Aigine reserves the right to charge resellers or customers for the time worked for troubleshooting errors related to Non-Product-Fault.



### 5.2.2 Installation and maintenance manual

Aigine provides an installation and maintenance manual. This manual includes step-by-step instructions for package installation. The maintenance describes recommended services for server environment management tasks.

[http://aigine.se/files/DEPLOYMENT\\_GUIDE\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/DEPLOYMENT_GUIDE_AIGINE_UDIE_ENGLISH.pdf)

### 5.2.3 Technical guidance and checklist

The technical guidance also includes a checklist to ensure that all operations are completed.

[http://aigine.se/files/DEPLOYMENT\\_GUIDE\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/DEPLOYMENT_GUIDE_AIGINE_UDIE_ENGLISH.pdf)

### 5.2.4 Configuration guidance

For configuring installation packages and mapping of data sources, there is a step-by-step guidance.

[http://aigine.se/files/DEPLOYMENT\\_GUIDE\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/DEPLOYMENT_GUIDE_AIGINE_UDIE_ENGLISH.pdf)

### 5.2.5 Test schedule

To ensure the operation of the solution after installation, there is a test schedule and associated action points for correcting any shortcomings.

[http://aigine.se/files/TEST\\_SCHEME\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/TEST_SCHEME_AIGINE_UDIE_ENGLISH.pdf)

## 5.3 User documentation

Aigine offers no end customer training. Responsibility in the system is distributed through self-instructing e-mail messages, and the user interface consists solely of one view, which itself is self-explanatory.

User training is therefore performed by self-training by reading user documentation.

[http://aigine.se/files/USER\\_GUIDE\\_AIGINE\\_UDIE\\_ENGLISH.pdf](http://aigine.se/files/USER_GUIDE_AIGINE_UDIE_ENGLISH.pdf)

## 5.4 Release notes

### 5.4.1 Release Notes

Release notes are communicated with new updates and includes bug fixes, dependencies, new feature, changed features, compatibility and known errors. With each new version specific installation instructions are also provided.

### 5.4.2 Road map

An updated road map for future development and functionality can be obtained through contact with Aigine at [info@aigine.se](mailto:info@aigine.se)

## 6 SUPPORT

### 6.1 End-users

No support is offered to end users. Communication takes place through self-instructing e-mail messages, and the user interface consists of only one view.

If the need for end-user support occurs, this should be resolved internally by the customer. The Data Protection Officer is the natural owner of the solution, and also the person who can continually respond to questions about both what is meant by personal data and legal basis. It may therefore be assumed that end-users have continuous contact with the DPO.

Of course, resellers are free to provide support to end users.

### 6.2 Technical support

Aigine provides technical support for faults in the product to retailers. Such support is conditioned by a performed troubleshooting according to the debugging scheme to exclude Non-Product-Fault.

The technical support is contacted by sending an email to:

[aigine-service@elinar.com](mailto:aigine-service@elinar.com)

The support is manned:

Mon-Fri

08:00-16:00 (Business Hours EET, Local Time Finland)

## 7 UPCOMING PRODUCTS FROM AIGINE

Aigine is designed to deliver increased value over time.

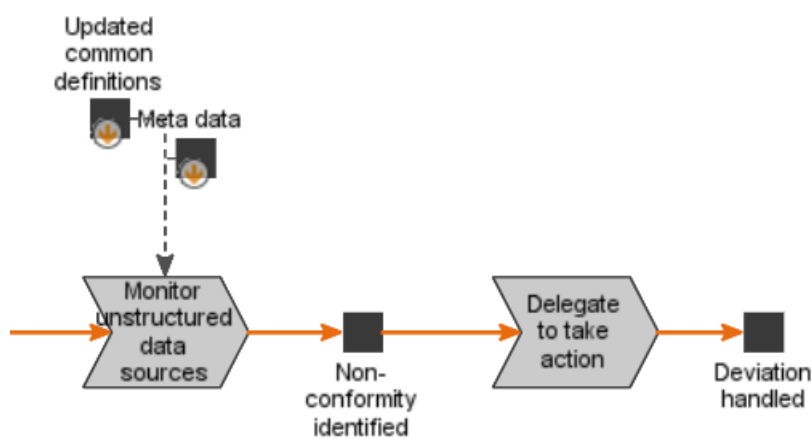
The strategy is, however, that the meta data layer created at the customer, together with continuously improved cognitive abilities, will continue to deliver value and efficiency to the customer.

A large number of additional offerings are already in the road map and will be available in the closest year.

Please note that additional products may have as a condition that the customer has performed the inventory of unstructured data and therefore has updated metadata.

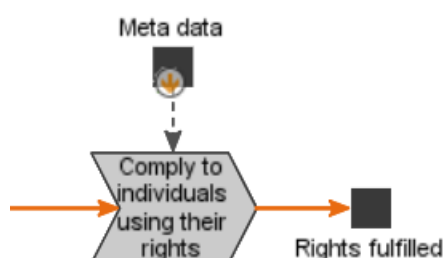
### 7.1 Monitoring

After 25/5 2018, and after the inventory of existing unstructured data is completed, the customer will need to continuous control their data. Customer must ensure that any data source designated as free of personal data are actually free. This provides a continuous need for monitoring, where Aigine uses metadata and cognitive abilities to monitor that routines and internal policies really are followed. Aigine Unstructured Data Monitor Engine also includes a digital process for easier deviation management.



### 7.2 Subject Access Rights

With GDPR personal data subject are giving a series of rights. These can be used by individuals use against the controller of personal data. One of these rights is the right to register records. Without an updated meta data layer, an organization must perform a full search of all its data to ascertain whether the individual who asks the question is in a document or e-mail message. Such search will take very large resources. With Aigines meta data layer, the search will only take a few seconds, as the search can be performed directly in the metadata. This also simplifies retracted consents, terminated contracts and data portability.



### 7.3 Structured data

Aigine's abilities and architecture can also be applied to structured data sources such as databases, both for inventory and for monitoring.

### 7.4 Deletion and masking

As Aigine also collects data for when documents, or specific personal data in a document, are to be masked or deleted, this administrative and resource-intensive process can be fully automated.



### 7.5 Pseudonymization

Pseudonymization is a data management and de-identification procedure by which personally identifiable information fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. The keys for this pseudonyms are stored in a secure location.

Pseudonymization decreases costs, limits risk exposure and makes it possible to save and process all data in the organization's possession.

The metadata layer in Aigine contains information on the location of each personal data within the data sources. This meta data will be used to offer customers real-time pseudonymization of all their data.

### 7.6 Other areas besides GDPR

Aigine's architecture and design does not limit the solution to only GDPR. By adding other metadata, the solution can also be used to manage e-archives, open data, insider information, trade secrets, legal assessments, etc.

## 8 PRICES

### 8.1 Pricing mechanisms

Aigine's pricing mechanism is based on the assumption that it's difficult to fully understand what data sources need to be handled, and the amount of data they contain, neither initially nor over time.

The pricing mechanism is therefore based on the type of organization, i.e. the nature of the activities they carry out, and the number of employees.

The assumption is that the more employees one organization has, the greater the production of unstructured data is.

Aigine is delivered on 12 months term against a monthly fee. Invoicing is done yearly in advance.

### 8.2 Quotation

Contact your normal IBM vendor for quotation.

## **9 ORDER**

Orders are made to your normal IBM vendor.