DocYard

Anonymize with DocYard

Productsheet
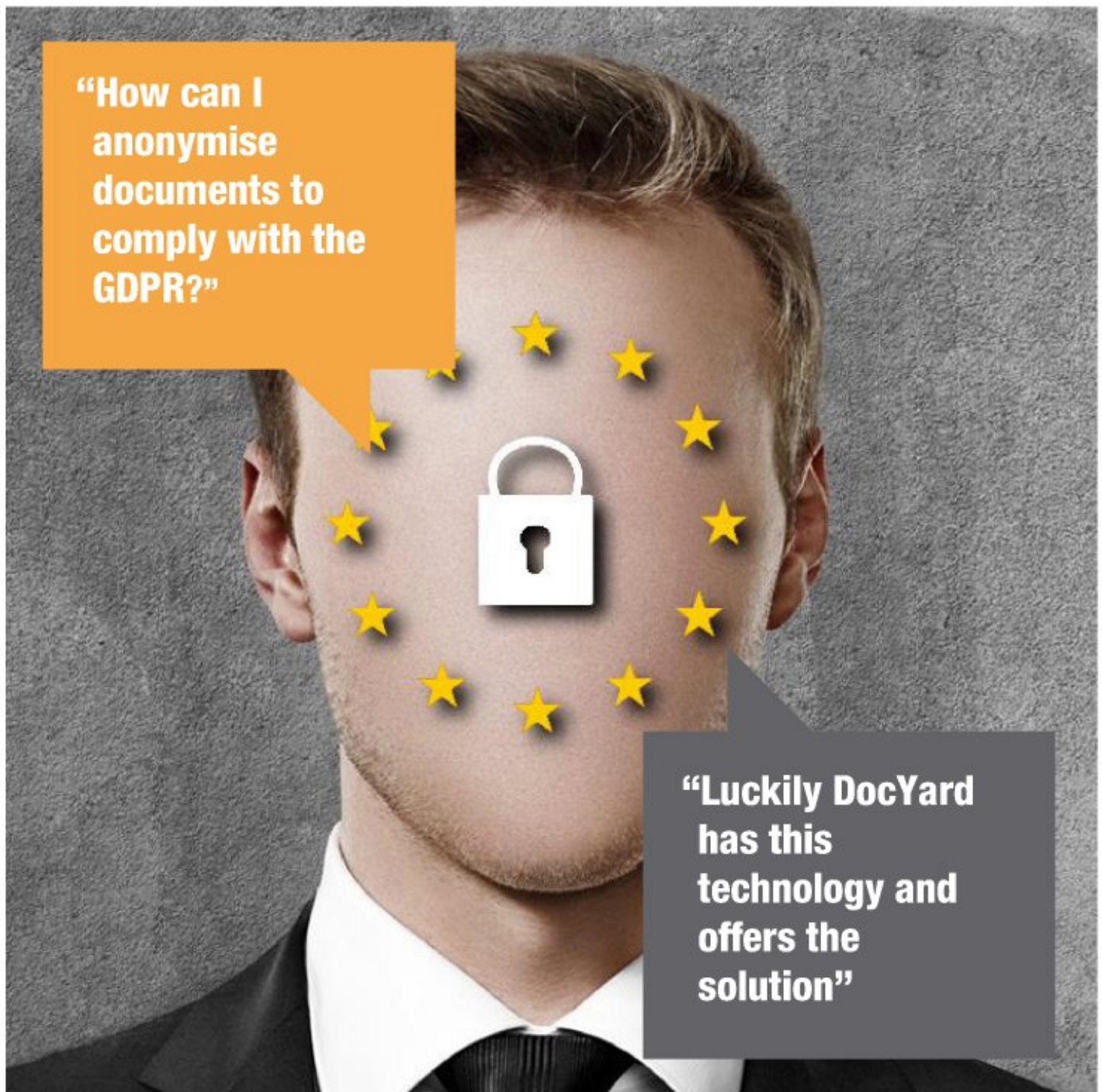
GDPR implemented? What's next?

The vortex caused by the implementation of the GDPR legislation has a little bit subsided. But even if you have organized everything well, it remains an agenda item for all organizations. Essentially, the GDPR determines how to deal with identifiable personal data. Privacy by Design now requires software and / or services for applications which have to focus on collecting the minimum amount of personal data for a particular purpose.

Many government agencies and other large organizations still have to anonymize their current documents with personal information (author name, teacher, etc.) in order to publish these files through their portal while protecting personal information. For example archives, books, and theses. But which information do you have to anonymise according to GDPR/DSGVO and how can you do this correctly and efficiently? For completely anonymized data, the GDPR is no longer applicable. Anonymization can be used as a technique when an organization wants to keep the data for publication or statistical purposes but reducing it to individuals is no longer necessary or lawful. However, full anonymization is not easy. For each situation, it has to be checked which data has to be exchanged so that a natural person cannot be identified.

What does the legislator say about anonymisation? The General Data Protection Regulation (GDPR), which has been in effect since 25 May 2018, applies the principles of data protection to all information concerning an identified or identifiable natural person. (Article 26):

The principles of data protection should apply to any information concerning an identified or identifiable natural person.

Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person.

To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly.

To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.

The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.

This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.

Anonymisation means: exchanging personal data with data that can no longer be used to identify a natural person, taking into account "all means that can reasonably be used", both by a controller and by third parties. An important factor here is that the processing must be irreversible. (Wikipedia, 2017).

Pseudonymization means: Replacing personal data in a defined procedure by encrypted data (pseudonym) using an algorithm. The algorithm can always designate the same pseudonym for a person so that information about the person, even from different sources, can be combined. (Wikipedia, 2017)

" The leading factor: processing has to be irreversible "

New DocYard Module

A new module has been developed for the DocYard document processing platform that allows customers to efficiently anonymise person-related data in bulk. Irreversibility is guaranteed. The anonymized personal data can no longer be traced or reproduced. Some examples: building permits, environmental permits, etc., which the public administration would like to make available online to citizens.

The modular structure of the DocYard software allows to perform anonymization either as part of the digitization process or as a separate processing of an existing digital archive. Due to the hardware-independent licensing, the production capacity can also be easily scaled up to the desired level (by additional module managers).
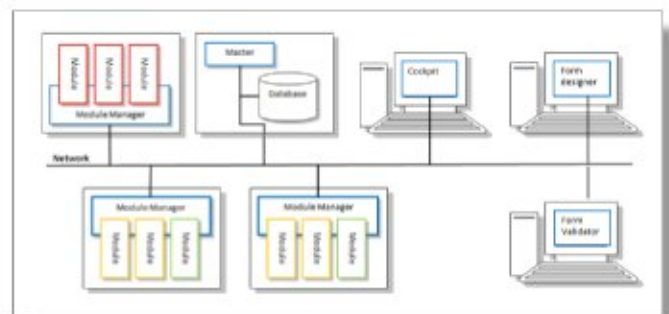
Picture 1: DocYard architecture

DocYard Workflow

To anonymize an existing digital archive, a workflow can be configured and used in DocYard including the following process steps:

1. Analysis of existing content and determination of which personal data in which documents (types) must be anonymized. However, it is also possible to submit the entire archive. In this case, the document type is selected based an automatic classification.

2. Establish one or more form designs (i.e. rules) to automatically recognize personal information. In this way you can save the process for later usage which saves time and the change of failure.

3. Import and configure the workflow in DocYard. Then start job and processing data.

4. The documents (all formats are possible, such as Tiff, JPEG or PDF format) are read and the software determines the search criteria to be used for each document.

5. Finally the software determines for each document which search criteria apply.

6. The results of this search are displayed to the user in the validation module. The user can confirm the fields, cancel them or add more fields manually.

7. The anonymization module performs two actions on the confirmed fields: graphically, the selected text is replaced by a black (or other colour) bar, and the associated OCR text in the document is removed.

8. Finally, the document is converted into a full-text searchable PDF/A file using Foxit MRC compression technology.

9. Optionally, it is possible to write all anonymized fields per document to an audit file. This file can also be used if the fields have been pseudonymized and the original and alias are defined.
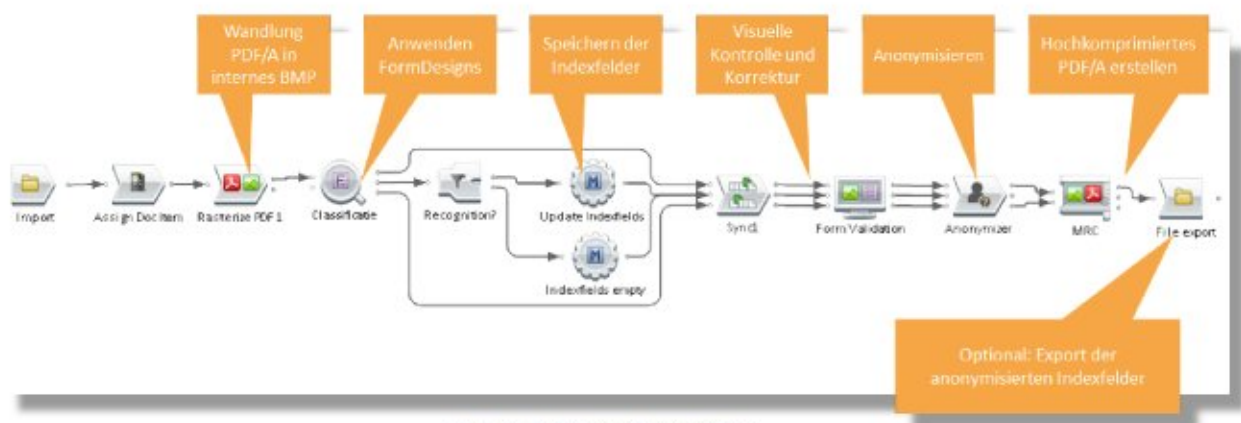
Picture 2: DocYard-Workflow

Result

By setting up smart workflows with the DocYard software, organizations can irreversibly anonymize large amounts of documents and data in an easy way. All person-related data that is processed will be removed which guarantees compliance with the new European Law GDPR (DSGVO).

Please contact your DocYard partner who can anonymize a document of your choice for you. The details of the nearest DocYard partner can be found on the website: www.docyard.eu

On our site is also a sample document deposited (original document and anonymous version).

Implementation

Decide for yourself how you want to use DocYard. The platform can be implemented in your organization, but you can also use it through a cloud version.

Visit the website for more information and the contact details of your nearest DocYard partner.

About DocYard GmbH

DocYard GmbH delivers production software solutions offering total document management and document and data conversion solutions, along with customized services and outstanding support. DocYard solutions are easy to handle and require a manageable level of customization and individual programming. This to implement and integrate every step of production into configurable, centrally manageable workflows to achieve the highest level of automation and integration in document processing. Using the deeply integrated MRC (Mixed Raster Content) Compression Technology of Foxit and other tools for Classification and Data Extraction DocYard covers all requirements for Mail Room Solutions and Scan Services.

We create proven software

We create professional software like our proven DocYard suite.
By delivering great software our clients can work faster by automating common, repetitive tasks. Our products make it easy for you to take care of quality during all stages of production.

We help you to optimize your process

We don't just ship software but we help you make great use of it.
We support and assist our customers until the software is running productively and a return on investment has been achieved. We don't consider a project successful until our customer feels the same way.

Work with humans, not robots

At the heart of every successful project is an agile team (of humans) fully equipped to deliver results from start to finish. Our interdisciplinary teams are scalable for projects of all shapes and sizes, allowing for efficient collaboration with the client and each other while never losing sight of what really works for your business and your customers.

We believe in measurable results

We strive to start producing a return on investment from the very first project, while significantly increasing productivity and ensuring lasting competitiveness.
We know from our own experience that investments with a given budget have to lead to bespoke solutions – and with our solutions, you can still achieve that goal while keeping customizing and individual programming work to a minimum. And the best thing? It's all measurable!

We're flexible and always ready to act

Your business can change, which means the way you use our products can change too.
Our products and people are flexible enough to meet your needs and your scope of work. This way, we're always ready to act.

For more information visit our website: www.docyard.eu or contact us.

Your DocYard partner: