EBOOK

SYSTRAN

Translating Sensitive Data in a Regulated Context: Issues and Solutions

More and more, personal and sensitive data lies at the heart of company business. The data used in regulated sectors (banking, insurance, pharmaceutical industry, etc.) is subject to strict European regulations. The General Data Protection Regulation (GDPR) is one of the strongest privacy and security laws in the world. GDPR imposes its obligations onto organizations if they collect data about European citizens, whether they are EU-based organizations or not. EU Data Protection Regulation will have extraterritorial effect.

How can 100% GDPR compliance be guaranteed at a time when digitalization and Cloud solutions are widely available? How can IT departments remain in control? And how can the security of this data be ensured when it is processed and used outside the company? For translation alone, the stakes are high, and we aim to examine them in this white paper.

Contents

What exactly is personal and sensitive data under GDPR context?

What is personal data?

Personal data[1] is any information that refers to an **individual** that is identified and/or identifiable. It may be direct (first name and surname) or indirect information (phone number, address, photo, etc.).

The term 'personal data' is the entryway to the application of the European General Data Protection Regulation (GDPR). The term is defined in Art. 4 (1). Personal data are any information which are related to an identified or identifiable natural person.

In France, this personal data is governed by the French Data Protection Act (**Loi Informatique et Liberté**) amended in June 2019, and by the **European GDPR**, published in April 2016.

Industrial secrecy and degrees of confidentiality of internal data

Some internal company data is confidential data without necessarily being considered as personal data. How is this data protected?

Industrial secrecy is binding on all those who share it. This however, does not protect the data from being leaked, or hacked, or from industrial espionage. Businesses and industries usually deal with four types of data on a daily basis, with differing **levels of security** and **distribution**:

- **public** data,
- **internal** data,
- **confidential** data,
- **restricted** data.

What is sensitive data?

So-called **sensitive data**[2] is a subcategory of personal data.

The CNIL (The "Commission Nationale de l'Informatique et des Libertés" is the National Data Protection Commission created by law no. 78-17 of 6th January 1978 to protect personal data, support innovation and preserve individual liberties) defines this data as follows:

> *"Sensitive data may reveal alleged racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership, and may also concern processing of genetic data, biometric data for the purpose of uniquely identifying an individual, or data concerning health, sex life or sexual orientation of an individual."*

Health data in particular is included in this category. The notion of sensitive data is very broad and governed at European level: medical secrecy, protection by banks when granting a loan, or by insurance companies, etc.

Under the GDPR, **the European Union prohibits the use of such data** unless:

- the person concerned has given their informed consent,
- the data is made public,
- use of the data in question is authorized by the CNIL.

[1] Source: Personal data | CNIL

[2] Source: Sensitive data | CNIL

What are translations of personal and/or sensitive data used for?

Three recurring challenges in international companies

Communication between international branches within a group

To be able to work together, various subsidiaries within a group exchange documents and share information. **In order for this data to be understood and used** by all employees, it must be translated. During international meetings, such as video-conferences, simultaneous translation is essential for effective teamwork.

Customer and business services

After-sales service, chat, etc. These services are frequently outsourced and delocalized: to be processed, customer requests first need to be translated. The translation problem also arises in an international company, for example, when the **IT department** in one subsidiary handles requests from the entire group.

This is a recurring problem in international groups: how can **international after-sales solutions be centralized** without employing personnel for each language? One solution may be **automatic translation into the operator's language**.

Translating end-customer documents

Depending on the business sector and needs, your translation requirements will vary.

Let's illustrate this point with some examples.

In the **banking sector**, KYC (Know Your Customer) is a procedure used everyday: customers' personal data is used to identify them, analyze their activity, draw up a banking profile, etc.

In the **insurance industry**, documents concerning any disputes or claims that occur abroad, may need to be translated.

Law firms working internationally sometimes need to translate contracts.

Health data required by the pharmaceutical industry (such as studies, the results of scientific cohorts, etc.) may also need to be translated.

In **industry**, patents and tenders containing trade secrets or product specifications also have to be translated.

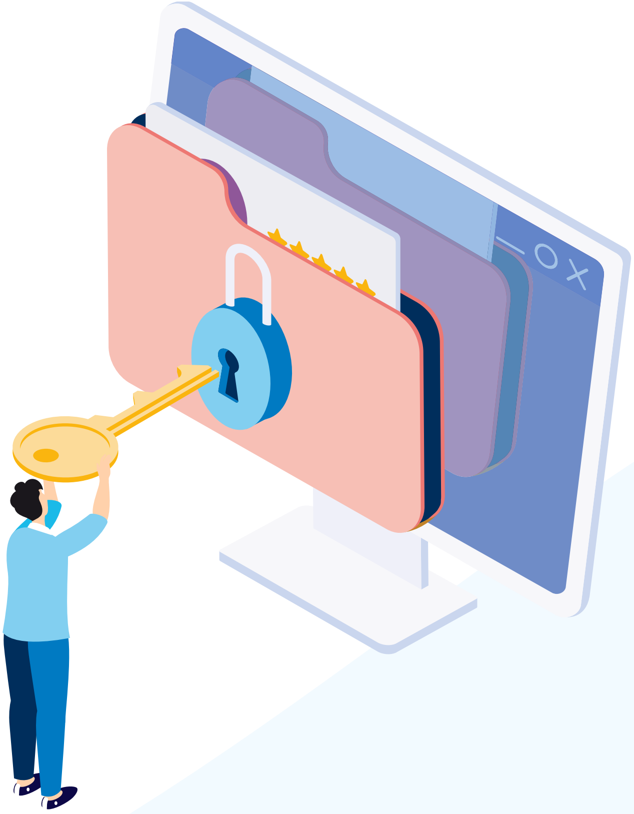
Nestlé: ensuring consistency and accuracy of communication worldwide

Nestlé is a multinational food company with **subsidiaries in over 200 countries**.

The issue of **language** and **localization** is essential for:

- consumers,
- employees,
- logistics,
- the worldwide supply chain.

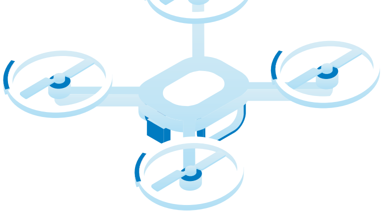
Nestlé's challenge therefore is to **respect the confidentiality** of certain strategic information, while communicating with stakeholders in each subsidiary.



> Using SYSTRAN, Nestlé has implemented a centralized translation solution to speed up processes and optimize costs. We cover broad needs across the company while at the same time meeting Nestlé's high data security requirements.

German Basterra
Former Technical Manager Nestlé Translation Department

Source: CUSTOMER TALKS #1 Nestlé | How to unify international communication in 30 languages?



Tata Consultancy Services (TCS): how to put employee skills to use in all languages

Tata Consultancy Services is an Indian company specializing in IT and telecommunication services, and business and marketing solutions for its clients around the world.

With a presence in **46 countries**, the group has over 556,000 employees. Most of the group's clients work in the banking, insurance and finance sectors.

As a result, **translation is key** at TCS: especially for members of the customer service department in Europe, who are responsible for answering all group customers, including in languages they do not master. The challenge is to make the most of employee skills, regardless of the language of their customers.

> While translation quality is essential for working in highly sensitive sectors, our first priority was security. We were looking for a suitable tool aligned with the GDPR and compliant with all European laws regarding the protection of personal data. In our business, this is non-negotiable!
>
> In particular, for one of our clients, a Scandinavian insurance company, we analyzed sensitive documents in Swedish, Danish and Norwegian completely securely by means of a private Cloud. [...] Over time, we have trained our models with SYSTRAN to improve performance in often complex languages.

Deepesh Bakhru
Associate General Manager at Tata Consultancy Services

Source: CUSTOMER TALKS #2 Tata Consultancy Services | How to support Client Business Process in any language?

What are the risks of translating sensitive data?

Shadow IT: the threat from within

The digitalization of companies incurs new risks in terms of **data management and protection**. As a result, protecting sensitive client data from cyberattacks has become a critical issue. However, there are other risks that are less well-known: for example **Shadow IT**.

This is the **use of software without the approval of the IS management**. This includes online translation software with data protection rules that are unclear.

CESIN's "2017 Shadow IT Report" reports IS managers citing an average of between 30 and 40 Cloud applications and services being used. In reality, **250 to 5,950 Cloud applications are used** (1,700 on average).

What are the risks associated with Shadow IT?

In fact, it exposes companies to:

- **cyberattacks** and **cybercrime** (in 2021: 1 company in 2 was subject to 1 to 3 cyberattacks during the year[3]) that can adversely affect the entire company;
- **leaking**, **theft** or **loss of the personal data** of their clients and/or employees.
- **non-compliance** with GDPR rules.

[3] Source: 7th edition of the annual CESIN barometer - Exclusive study of cybersecurity in French companies

Data leakage:
when reality is stranger than fiction

According to the French National Agency for Information Systems Security (ANSSI), 55% of French companies believe the **threat level of cyber espionage to be high**[4].

The use of free online translation software can also mean that **third parties are collecting and using sensitive data**. There are plenty of examples in Europe. Statoil (now Equinor)[5], a Norwegian oil company was founded in 1972.

On September 3, 2017, the Norwegian news agency NRK revealed a massive leak of personal data (contracts, emails, personal contact details, staff mails, etc.) after they had been translated on a free online tool. The scandal brought to light the fact that any translated text could be posted online and led several Norwegian companies, including the Oslo Stock Exchange, to block access to certain tools, including Google Translate.

Penalties for data leakage:
where does French law stand?

Data usage in France is **regulated by the CNIL**, which is responsible for implementing the GDPR in France and for sanctioning companies that do not comply with EU regulations.

Article 84 of the European GDPR states that "Member States shall lay down the rules on other penalties applicable to infringements of this Regulation in particular for infringements which are not subject to administrative fines pursuant to Article 83, and shall take all measures necessary to ensure that they are implemented. Such penalties shall be effective, proportionate and dissuasive. "

From a simple warning to an administrative fine, the maximum **penalty** in France can reach **20 million euros** or up to **4% of annual global revenues**.

According to article 226-13 of the French Criminal Code, the penalty for breaching bank or medical secrecy is the same as for professional secrecy: one year in prison and a 15,000 euro fine.

[4]Source: 7th edition of the annual CESIN barometer - Exclusive study of cybersecurity in French companies

[5]Source: Translate.com Exposes Highly Sensitive Information in Massive Privacy Breach - Slator

What are the solutions for secure translation?

For banks, insurance companies, health and other regulated professions, choosing a French or European provider means that you can be certain of a **true culture** of personal and/or sensitive data protection **in addition to the regulations in force imposed by GDPR**.

Translation agencies offer **document translation** services. However, they do not necessarily meet every need, especially for **instant translation**.

Especially when deadlines are tight or data cannot leave the company.

Secure machine translation: what guarantees?[6]

Here are some tangible ways to measure the level of security of instant translation software.

- Data is **encrypted during exchanges** and for storage.
- **Backups** are also encrypted.
- Data is securely erased to prevent restoration.
- Translated texts are deleted instantly or as soon as possible, depending on software settings.
- The platform is **monitored 24/7** and security alerts are centralized **using SIEM technology** (Security Information & Event Management).
- **Systems can be protected** in the event of a DDoS (Distributed Denial of Service) attack. DDoS means denial of service attack: multiple requests sent to a server simultaneously by robots, rendering the services, software or hosted websites inaccessible.
- The exposed area of the systems is reduced: access is only allowed from within the company. For this purpose, a **whitelist** can be set up (list of IP addresses authorized to access the software/server, etc.).

[6] Source: Secure machine translation: an essential tool for Basel & Solvency compliant data governance

Either a Cloud solution...

Cloud solutions are **flexible** and offer instant translation for all types of documents, enabling teamwork, immediate translation of emails, quotes, contracts, chat conversations, etc. Sensitive data hosted on a client-dedicated Cloud or on your provider's Cloud offers **additional**, **customizable security**.

- One company's data is not shared with another: each company has its own Cloud.
- It is possible to choose the geographical location of its Cloud, which solves the sovereignty issue and a trusted Cloud, which concerns 6 out of 10 companies.

This type of provider can ensure:

- that your data will never be used in its translation models,
- that your stored data will be encrypted,
- that vulnerability tests will be conducted regularly.

...or a local alternative

Using On-Premise software, i.e. installed on the company's information system, remains the first choice (if not the only one) for some companies and organizations in the Defense and Homeland Security sector. All data is stored on the client's servers according to their own rules and security needs.

The **geographical constraints of data storage** are therefore respected.

This is also a versatile solution, like the Cloud version, which connects to all the company's resources:

- The software **can easily be included on all internal tools/applications** for instant translation of emails, documents (Word, Excel, Powerpoint), PDFs, images – almost any document that can be used in a work environment.
- The **API** (Application Programming Interface) is directly **integrated** in the company's **information system**. It provides translation of an intranet for greater security (in Microsoft SharePoint for example), internal pages, chat, etc.

A community and standard as quality reference

Created in September 2001, **OWASP** (Open Web Application Security Project) is a community that works to prevent risks to the security of web applications. Its list of top 10 security risks is a reference among top developers worldwide. Following OWASP best practices is an additional guarantee of security.

ISO/IEC 27001 is the most widely used standard in the ISO/IEC 27000 family. It relates to the **Information Security Management System** (ISMS).

It is not mandatory to apply it, but it sets out a framework for best practice. Companies certified to this standard are regularly audited by an independent body to ensure that their ISMS remains secure.

SYSTRAN

With services increasingly going online and abroad,
we are exchanging more and more sensitive data.

In order to ensure your employee and client data remains confidential,
and to maintain your company's compliance with the GDPR,
it is preferable to use secure instant translation software.

A solution for translating your data with peace of mind!

Secure my translations

by invox