



Experience success. Together.

# QUANTO Transliteration Center

QUANTO Solutions GmbH

Stand: 08.03.2021

## Basics

In the process of globalization and internationalization, companies must take into account many legal, cultural and other challenges.

To remain competitive, companies must address international markets and communicate with business partners as well as customers worldwide. Language and cultural differences present companies with new challenges.

In the IT sector, the following aspects in particular must be taken into account when it comes to internationalization:

- **Encoding:** ASCII character encoding is sufficient for texts in most languages with the Latin alphabet. Other languages such as Cyrillic or Chinese usually require a larger character encoding such as Unicode.
- **Hardware:** Software designers should keep in mind that not all hardware equipment may be available in all countries.
- **Usability:** Texts in English are usually quite short. Already in German, one notices that expressions are longer and require more space. This should be taken into account when designing the user interface. Furthermore, it should be noted that many languages are written from left to right, but other languages such as Arabic or Hebrew are written from right to left.

## Transliteration & Transcription

There are more than 7,000 languages worldwide. What do you think are the most widely spoken languages?

Perhaps it's no surprise that at the top of the list of the world's most widely spoken languages is Mandarin Chinese (more than 1.2 billion native speakers), followed by Spanish (400 million native speakers) and English (350 million native speakers).

To represent content, words and texts in languages with different alphabets, experts make use of transliteration, transcription and romanization. But how do these terms differ and how are they relevant for the globalization and internationalization of companies?

### Transliteration

Transliteration means the letter-for-letter conversion of words from the alphabet of a source language to the alphabet of the target language. Transliteration also helps the reader understand how a particular word is pronounced in the target language. This is achieved by converting the alphabet of the original text into similar sounding and corresponding characters of another alphabet in a foreign language. Diacritical

marks are often used for this purpose as additional information to reproduce the exact pronunciation. As a rule, there is only one official transliteration in Chinese. This is pinyin, i.e. the transcription of Chinese characters into sounds, which also indicates the tones or pronunciation together with the diacritics.

Transliteration is mainly used for personal and company names as well as addresses and products. Thus transliteration solves the problem of different characters in different alphabets.

It also allows words to be entered on different laptop keyboards and word processors, especially in languages that do not use the Latin alphabet, such as Chinese, Russian, or Arabic.

In addition, transliteration is useful when a person does not know the alphabet of a language, but understands and can speak the language.

### **Transcription**

In comparison, transcription can be described as faithful transcription using a phonetically defined phonetic transcription. Even if someone does not know a language, he/she can pronounce a word or text through it.

Here is an illustrative example:

Greek	Transliteration	Transcription
Μπερλίν, Ντακάρ, Γκέντ	Mperlín, Ntakár, Gként	Berlin, Dakar, Gent

## **Romanization**

Romanization is a transcription/transliteration into Latin letters. So what does all this have to do with globalization and internationalization?

## **Relevance of transliteration for international companies**

Transliteration and transcription can help to present the following information in the best possible way:

- Personal names
- Brand names
- Company names
- Addresses

Translated are:

- Titles, departments
- Slogans
- Advertising material such as brochures
- Online information such as websites

German tech company Jungheinrich is a good example of how transliteration helps brands position themselves. In Chinese, Jungheinrich becomes "永恒力." The transliteration results in "Yonghengli," which means "eternal, constant force" and perfectly reflects Jungheinrich's products and services, for example, its forklifts.

The challenge here is that readers cannot know whether transliteration or translation has taken place. And if transliterated, which characters were used.

So when using software for transliteration, it is important to have a well-structured database for company names and also parts of addresses, and match them to identify what other information needs to be translated.

## QUANTO Transliteration Center

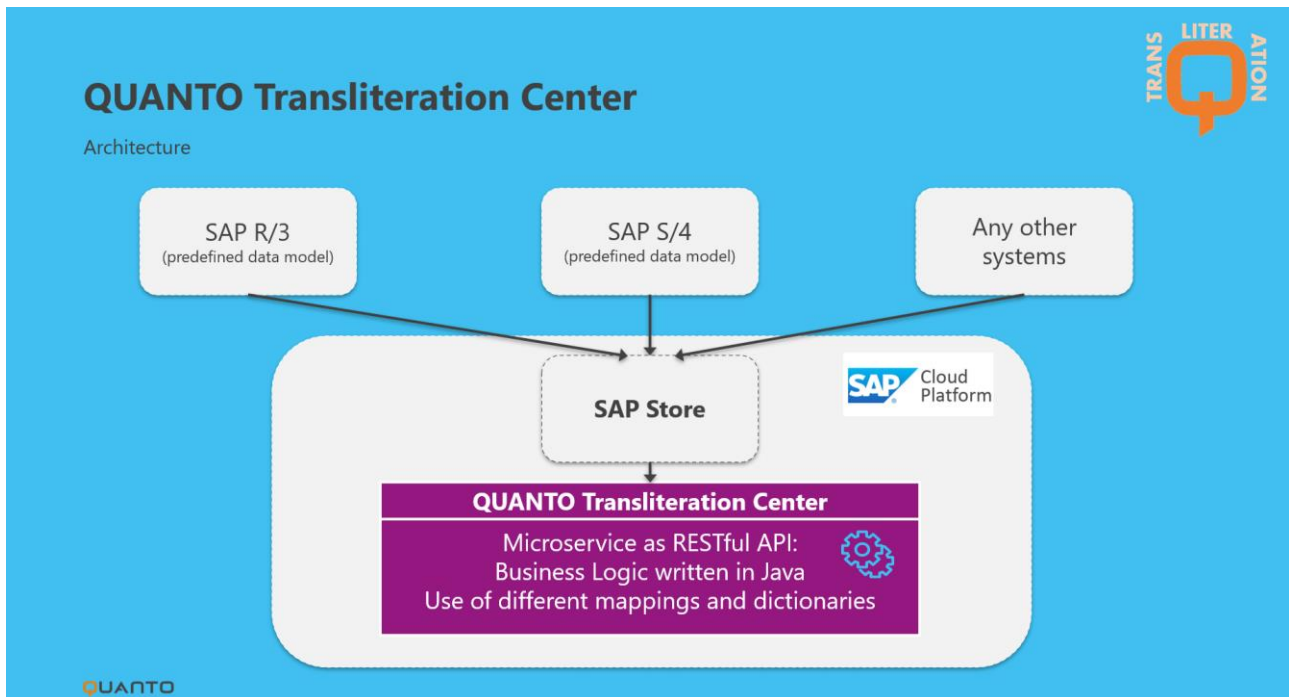
QUANTO Transliteration Center enables transliteration and romanization of business partner data (proper names, company names and addresses of database extracts) and incorporates localized translation of relevant parts of the texts like legal entities or geographical names. a REST API to transliterate data.

**QUANTO Transliteration Center** provides the following **core features**:

- Transliteration and transcription of entire business partner and address objects in high quality
- Choose from different scripts
  - Chinese to Latin
  - Cyrillic ↔ Latin
  - Hangul ↔ Latin
  - Greek ↔ Latin
  - Thai ↔ Latin
- Different variants
  - Transliteration only
  - Transliteration combined with transcription
- High performance

### RESTful API

Our cloud-based RESTful API performs character-based transcription at Unicode level. To achieve high quality results, additional complex logic has been implemented. For example, our product uses word dictionaries to transcript certain words and characters. This becomes important when addresses are transcribed. There are dictionaries for city and province names, legal entities, companies and many more. The cloud-based REST API was developed with state-of-the-art technology, such as Java Spring-Boot and Cloud Foundry.



Users can reach the REST API using common HTTP methods. An unlimited amount of business partner / address objects can be sent in a POST request. The REST API will process the data and expose the results to the user. The results can be retrieved by GET requests.

Our product can be used in many ways. The most basic way is to send data via tools like Postman or Swagger UI. Data must be entered manually in this case. A better solution is to implement the API into existing systems, such as ERPs. Depending on the subscription, for example repeatedly transliteration jobs can be performed.

A transliteration from Chinese to Latin letters could look like this:

```
{
  "license_key": "1234",
  "transliteration_rule":
  "CHINESE_TO_LATIN",
  "variant": "COMPLEX",
  "datamodelType": "PLAIN",
  "addresses": [
    {
      "firstname": "景",
      "lastname": "涛",
      "companyname": "瑞仪集团上海",
      "country": "中国",
      "postcode": "30565",
      "city": "上海",
      "district": "徐汇区",
      "street": "赵家bang路",
      "housenumber": "308",
      "region": "山西"
    }
  ]
}
```



```
"transliteratedAddress": {
  "firstname": "Jing",
  "lastname": "Tao",
  "companyname": "Ruiyi Group
  Shanghai",
  "country": "China",
  "postcode": "30565",
  "city": "Shanghai",
  "district": "Xuhui",
  "street": "Zhaojiabang",
  "housenumber": "308",
  "region": "Shanxi"
}
```