

# eTranslation

The **Automated Translation (eTranslation) Building Block** helps European and national public administration exchange information across language barriers in the EU by providing machine translation capabilities. Public administrations, citizens and businesses in the EU will thus be able to use digital services in the language they choose. For example, users of the **Online Dispute Resolution Portal** are able to access information in languages they do not understand and communicate with their counterparts in their own language. While eTranslation is mainly intended to be integrated into such digital services, it also offers useful stand-alone services for the translation of documents and snippets of text.

The Core Service Platform offers automatic text translation based on the eTranslation service and provides an EUwide coordination and networking mechanism for collecting and sharing language resources, the European Language Resource Coordination (ELRC). The ELRC was launched by the European Commission to identify and gather language and translation data relevant to national public services, administrations and governmental institutions across all 30 European countries participating in the CEF programme. These resources are needed in order to improve the quality and the coverage of the machine translation engines in eTranslation. Data resources gathered in this ELRC initiative will therefore be used to develop a high-quality machine translation service.

The eTranslation service builds on the previous machine translation service of the European Commission, MT@EC, developed by the Directorate-General for Translation (DGT). MT@EC translation engines were trained using the vast Euramis translation memories, comprising over 1 billion sentences in the 24 official EU languages, produced by the translators of the EU institutions over the past decades.

Generic Services Projects focus on supporting the collection of language tools and resources for specific sectors and for under-resources languages, as well as to stimulate the integration and take-up of eTranslation services.

[LEARN MORE](#)

## Highlights



### Generic Services Projects

**€24.2 million**

**29 projects / 26 Member States,**  
Iceland, Norway and the United Kingdom



### eTranslation service

**843 million** translation requests

**410 million** translated pages.  
*The service launched in 2017.*



### Top translation engine

eTranslation **ranked amongst top engines** in a **competition** with over 150 machine translation systems.

More info on the Generic Services:

[HaDEA](#)

[Success Story](#)

## Indicators

As the CEF funding of the Core Service platform came to an end in Q4 2021 for ELRC and in Q4 2022 for eTranslation, these are the cut off dates for the last monitoring activities linked to this work strand. The monitoring of the Generic Services projects presents the state of play by Q4 2022. 8 Generic Services projects are still under implementation and are foreseen to be completed by 2023.

## Uptake

**843,970,348**

translation requests processed by  
eTranslation

**97**

information systems connected to  
eTranslation

**4,715**

language resources collected by the  
European Language Resource Coordination  
initiative

**656**

updates to eTranslation engines

## Service Availability

**99.73%**

Average uptime of the eTranslation service

## Financial Monitoring

**€ 28,507,368**

CEF Core Service Platform funding

**€ 24,203,771**

CEF funding through Generic Services  
Projects

## Reuse of eTranslation

eTranslation is being reused by Digital Skills and Jobs Platform, EESSI, EU e-Justice, eProcurement, EDMO, Europeana, ODR, Public Open Data, Safer Internet and Wifi4EU

For more information, click [here](#)

# Indicators

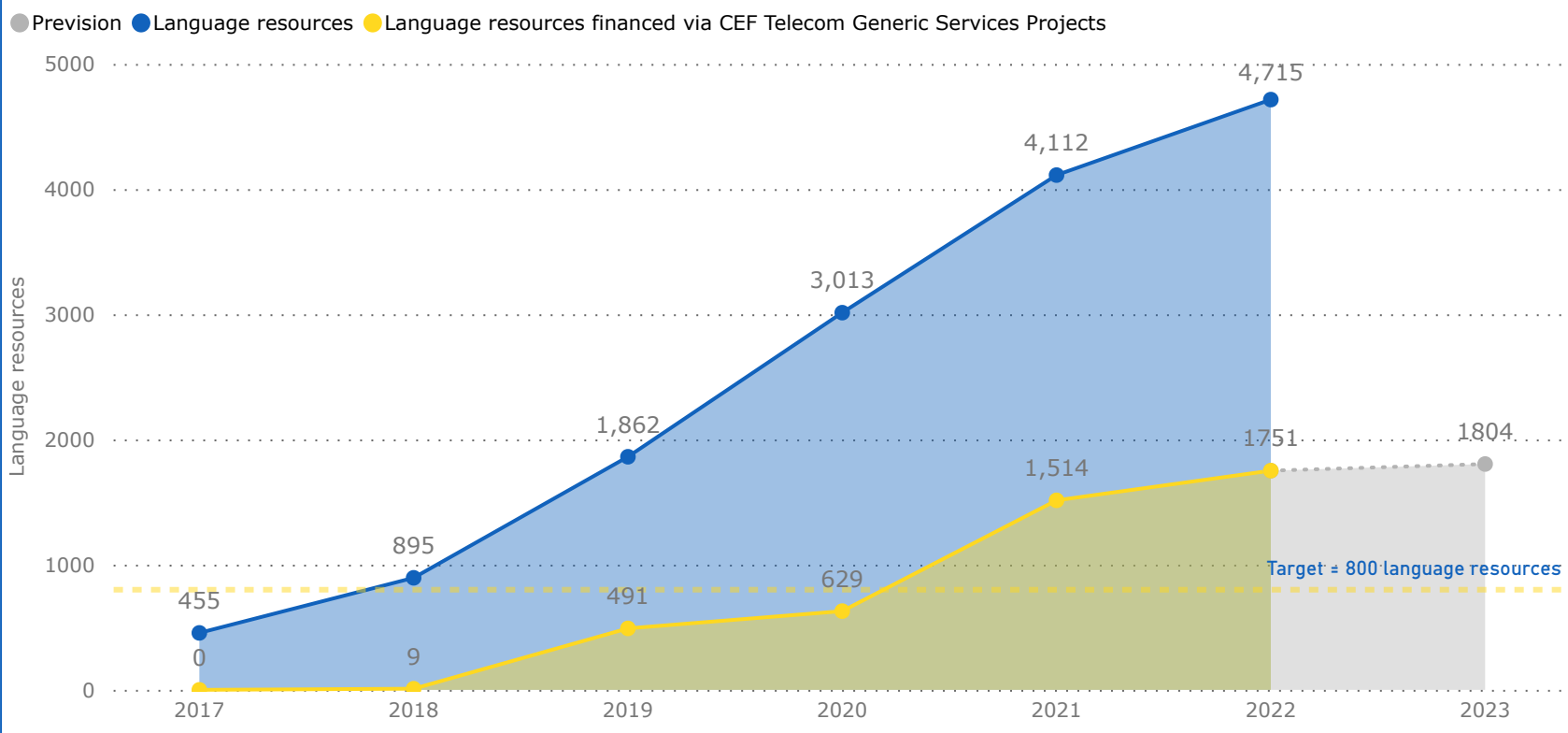
## eTranslation > Uptake > Language resources generated by the ELRC

This indicator measures the number of **language resources collected by the European Language Resource Coordination (ELRC) initiative**. The baseline for this number is the operational launch of the Data collection and curation service in Q2 2017 (245 language resources).

Language resources refer to sets of language data and descriptions in machine readable form, including written and spoken corpora, grammars, and terminology databases. A language resource is considered as collected by the ELRC initiative when made available to the Commission freely and openly through the [ELRC-SHARE repository](#).

### Target

**800 language resources** generated by the European Language Resource Coordination (ELRC) initiative, financed as Generic Services Projects, **by Q4 2023**.



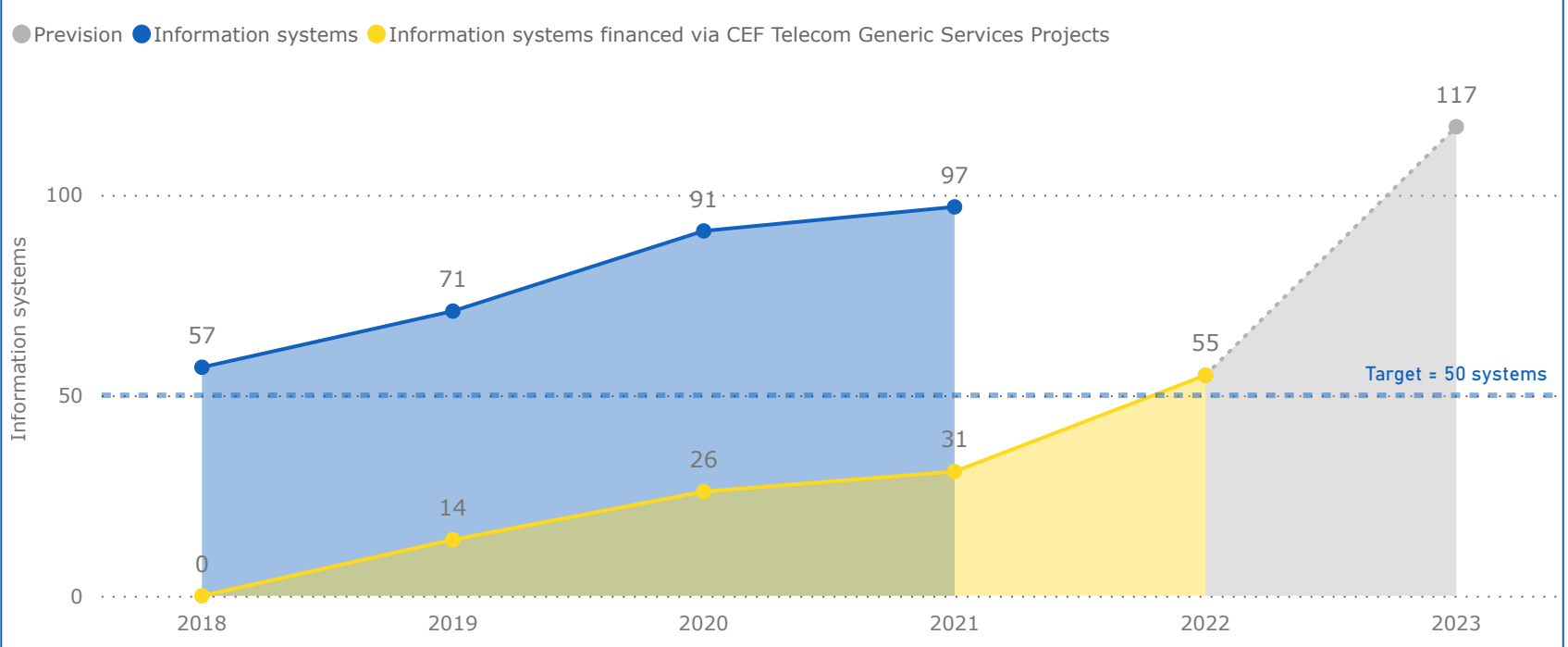
53 additional language resources are expected to be generated by the European Language Resource Coordination through Generic Services Projects by 2023.

## eTranslation > Uptake > Organization information systems connected to eTranslation

This indicator measures the number of **information systems connected to eTranslation** at the time of data collection.

eTranslation can be integrated into public administrations' information systems to make digital public services and content multilingual. eTranslation is available to public administrations in EU countries, Iceland and Norway, as well as EU institutions and agencies. A system is considered as connected to eTranslation when it has sent a text snippet or a document to be translated by the eTranslation service in the last 6 months.

**Target**  
**50 systems connected to eTranslation, by Q4 2021.**



62 additional systems are expected to be connected to eTranslation through Generic Services Projects by 2023.



## eTranslation > Uptake > Translation requests processed & pages translated by eTranslation

This indicator measures the number of **translations requests** and **translated pages** carried out **relying on the eTranslation system**. The baseline for this number is the operational launch of the eTranslation component/tool/service in Q1 2018 (1,109,672 translations).

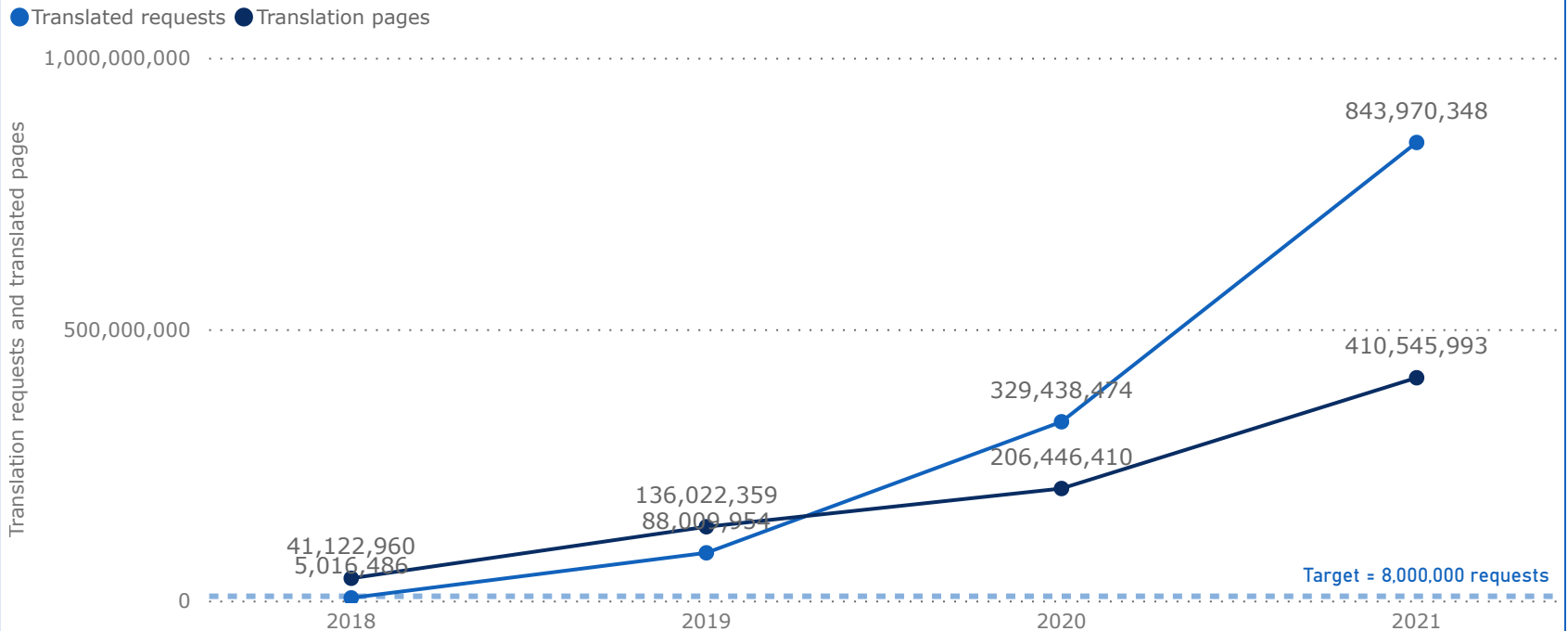
The number of translation requests refers to the sum of text snippets and documents that were translated by the eTranslation service. A user can, in a single request, translate anything from one word to more than 500 pages. While the number of translation requests reflects the frequency of use of the eTranslation service, the number of translated pages reveals how much the service actually translates.

The eTranslation service can be used in two distinct ways to process translation requests:

1. One-off translations, for which it provides [a web user interface](#) for direct use by individuals (human-to-machine use).
2. Integrated machine translation functionality, for which it provides machine translation capabilities for digital services through a web service (machine-to-machine use).

### Target

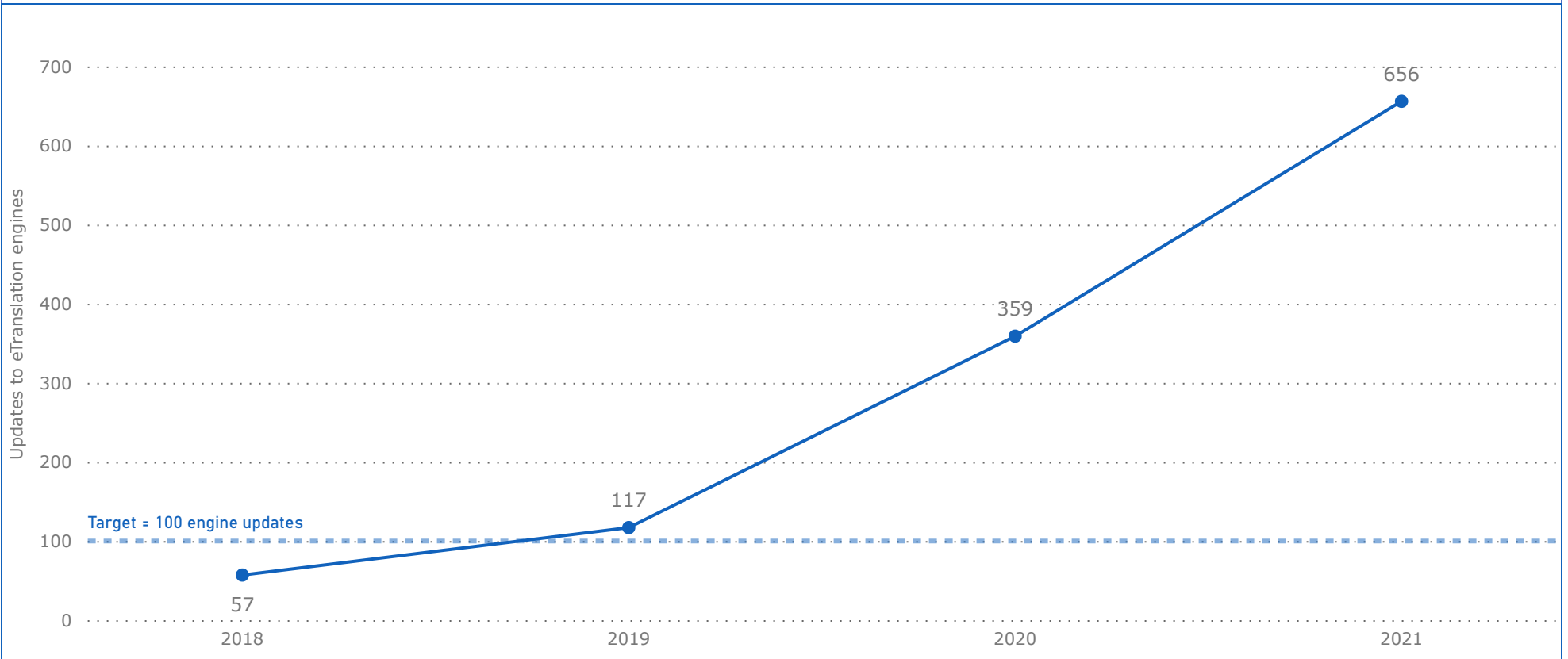
**8 million translation requests**, by the end of 2020.



This indicator measures the number of **updates to eTranslation engines**. The baseline for this number is the start of the monitoring of this indicator in Q1 2018 (0 updates to eTranslation engines).

An engine handles translation from one language into a second in one direction (e.g. French to English) for a specific domain (e.g. Public Health). An existing engine is said to be updated when a new release occurs, i.e. when the engine has an increased average score across several metrics, including the [Bilingual evaluation understudy \(BLEU\) score](#) (i.e. the most commonly used metric for Machine Translation quality).

**Target**  
**100 updates to eTranslation engines**, by the end of 2020.



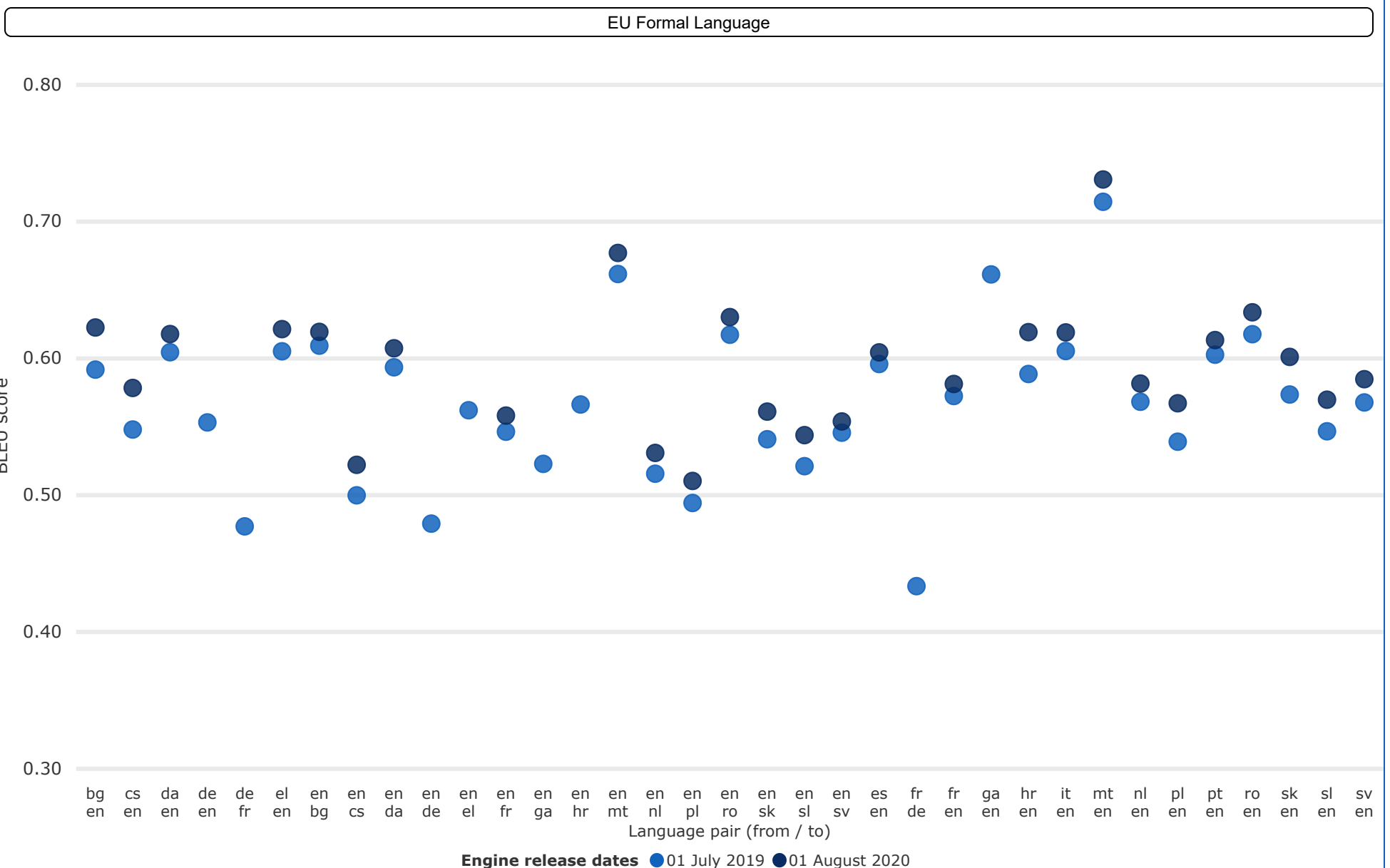
### BLEU score

The BLEU score compares the translation produced by the engine against the pre-existing human translation. This score is calculated on a 10,000 segment dataset (test set) reserved from the additional training data prior to training.

The scatter chart below displays the BLEU scores of two subsequent releases of a certain engine. These BLEU scores are computed on the same test set (i.e. the test set used for the former release), for sake of proper comparison.

### Engine Domain

This is the domain of the engine for which the scatter chart will compare the **BLEU scores** of two subsequent releases:



# eTranslation Milestones

All the **eTranslation** milestones defined within the CEF Monitoring framework were successfully achieved.

A detailed overview of the milestones can be found in the table below.

| Milestone title                                                                                              | Start date        | End date          | % complete | Status    |
|--------------------------------------------------------------------------------------------------------------|-------------------|-------------------|------------|-----------|
| Additional NLP modules                                                                                       | 01 January 2018   | 30 December 2020  | 100        | Completed |
| Batch processing for engine training                                                                         | 30 June 2019      | 30 September 2019 | 100        | Completed |
| Cloud scalability                                                                                            | 31 May 2019       | 31 March 2020     | 100        | Completed |
| Continuous security enhancements, leading to enhanced monitoring by DIGIT (SOC - security operations centre) | 30 September 2018 | 30 June 2019      | 100        | Completed |
| Data cleaning (cooperation with language depts, Euramis, WMT task, SMART workshop in December 2018)          | 31 March 2018     | 31 March 2019     | 100        | Completed |
| Data management tool (interfaced with ELRC)                                                                  | 30 June 2017      | 31 December 2018  | 100        | Completed |
| Domain specific engines                                                                                      | 30 September 2018 | 31 December 2018  | 100        | Completed |
| Dynamic resource allocation for cloud                                                                        | 31 May 2019       | 31 December 2019  | 100        | Completed |
| Dynamic resource management (cloud elasticity)                                                               | 31 October 2017   | 30 June 2019      | 100        | Completed |
| Integration of CEF eDelivery building block (secure data transmission)                                       | 30 September 2017 | 31 March 2019     | 100        | Completed |
| Integration of eID                                                                                           | 30 September 2017 | 30 June 2018      | 100        | Completed |
| Interfacing with European Language Resource Coordination (ELRC) to obtain data for domain-specific engines   | 31 January 2017   | 30 June 2017      | 100        | Completed |
| Migration to neural MT                                                                                       | 30 November 2017  | 30 June 2018      | 100        | Completed |
| Pilot data-centre engines using eTranslation workflow                                                        | 31 January 2018   | 01 June 2018      | 100        | Completed |
| Quality estimation                                                                                           | 30 June 2018      | 31 December 2019  | 100        | Completed |
| Rebuilt NMT engines to include data from other EU Institutions                                               | 31 March 2018     | 30 September 2018 | 100        | Completed |
| Render cloud secure (Limited High)                                                                           | 31 March 2017     | 30 June 2018      | 100        | Completed |
| Reporting System (eTranslation Statistics Dashboard)                                                         | 31 January 2017   | 31 March 2019     | 100        | Completed |
| System Administration Tools                                                                                  | 31 January 2017   | 30 June 2018      | 100        | Completed |
| Terminology incorporation through constrained vocabulary                                                     | 30 November 2018  | 31 March 2021     | 100        | Completed |
| Translation Quality Enhancements (Neural Networks, Domain adaptation, analysis of translator output...)      | 01 September 2017 | 31 December 2019  | 100        | Completed |
| User friendliness campaign for interface                                                                     | 31 October 2019   | 31 March 2020     | 100        | Completed |
| Website integration reference implementations                                                                | 30 September 2019 | 31 March 2020     | 100        | Completed |