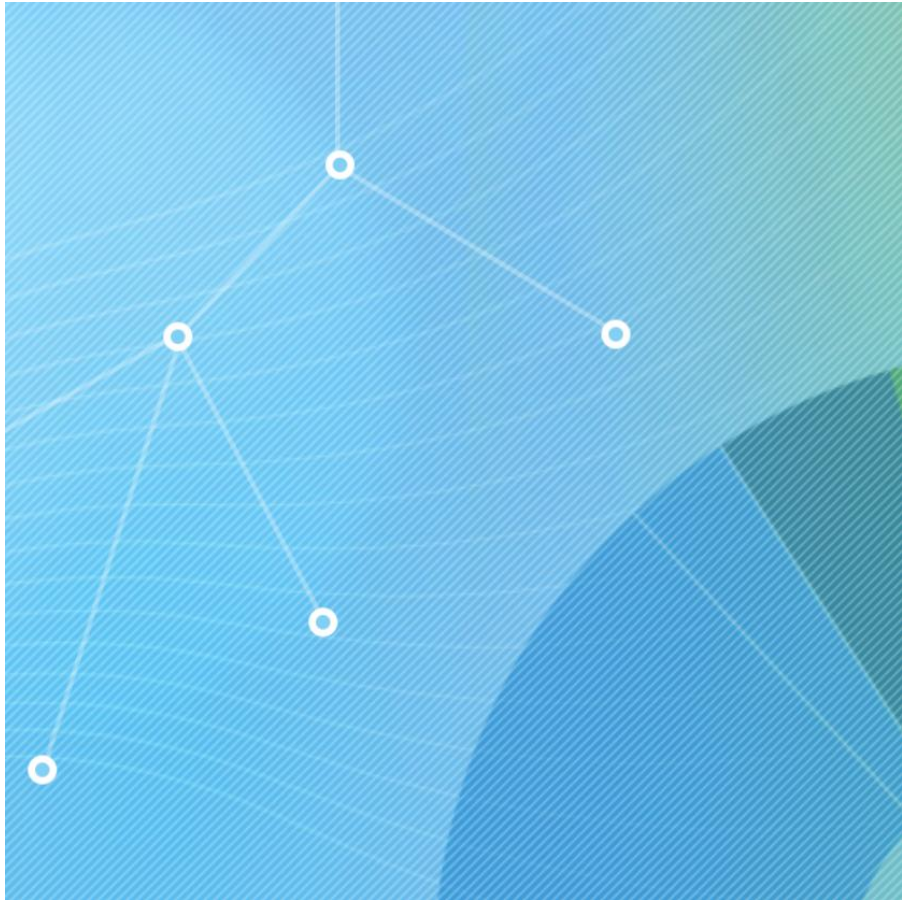


Sunsetting CAT



Coreon GmbH



- ✓ Berlin-based software company
- ✓ Leader in Multilingual Knowledge Systems
- ✓ Software-as-a-Service
- ✓ Management owned – CEO Jochen Hummel
- ✓ Makes search, machine learning, and IoT applications interoperable
- ✓ Deployed in industry, eGov, and education

Topics



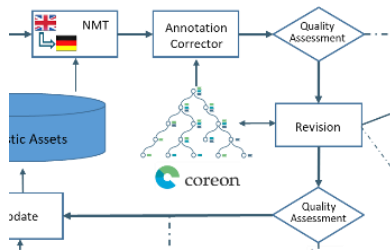
NMT disruptes translation

Industry getting it wrong

NMT Workflow sunsets CAT

Multilingual Knowledge System

New roles for linguists



Disruption



image credit: mindjet.com

“Nearly Indistinguishable From Human Translation”—Google Claims Breakthrough

by Florian Faes on September 27, 2016



Human Parity

Intelligent Machines



The Google machine translation service is coming out of its beta phase and will be available to all users in 2016.

Google’s New Service Translates Languages Almost as Well as Humans Can

Chris Wendt @Tian500

Follow

Replying to @Jeromobot @tarfandy @allisonlinn

Thanks. It is not a career path I can reasonably recommend to a young person. Not for that language pair and without a really deep specialization. We can debate the human parity claim endlessly. There is no doubt about the fact that machines are getting better at it every day.

5:55 PM - 15 Mar 2018

Microsoft | The AI Blog | The Official Microsoft Blog | Microsoft On the Issues

Microsoft Research Achieves Human Parity For Chinese English Translation

Written by Sue Gee

Wednesday, 21 March 2018

Researchers in Microsoft’s labs in Beijing and in Redmond and Washington have developed an AI machine translation system that can translate with the same accuracy as a human from Chinese to English.

jump in efforts to

y Tom Sin



Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

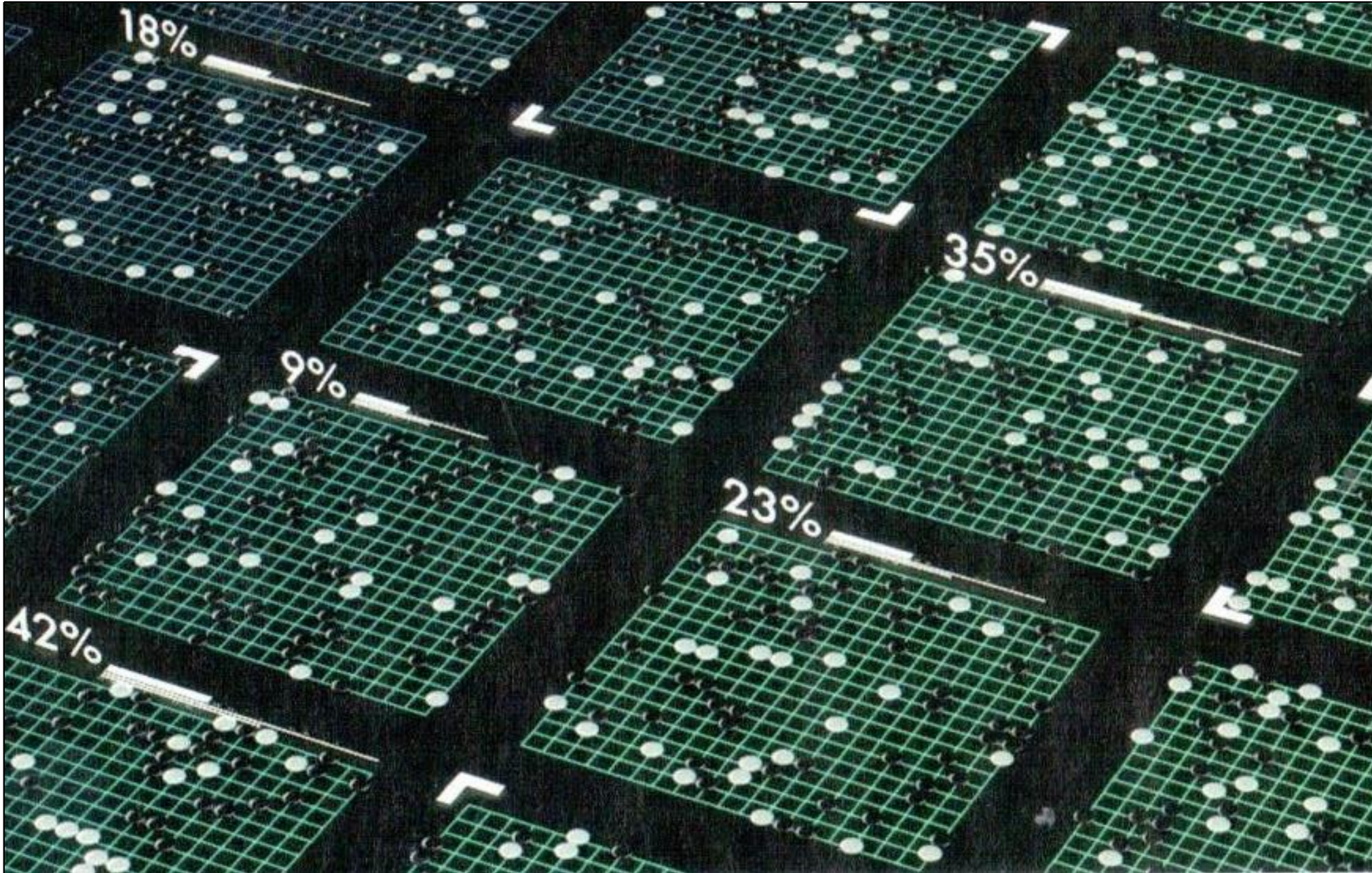


Jagwar [grrr...] @terakasorotany · 29 Aug 2017

I've just realised that the quality of #DeepL machine translations is mind-blowingly good! 🤯 Holy sh*t, the future is now! #FuturesNow

SDL Cracks Russian to English Neural Machine Translation

Global Enterprises to Capitalize on Near Perfect Russian to English Machine Translation as SDL Sets New Industry Standard

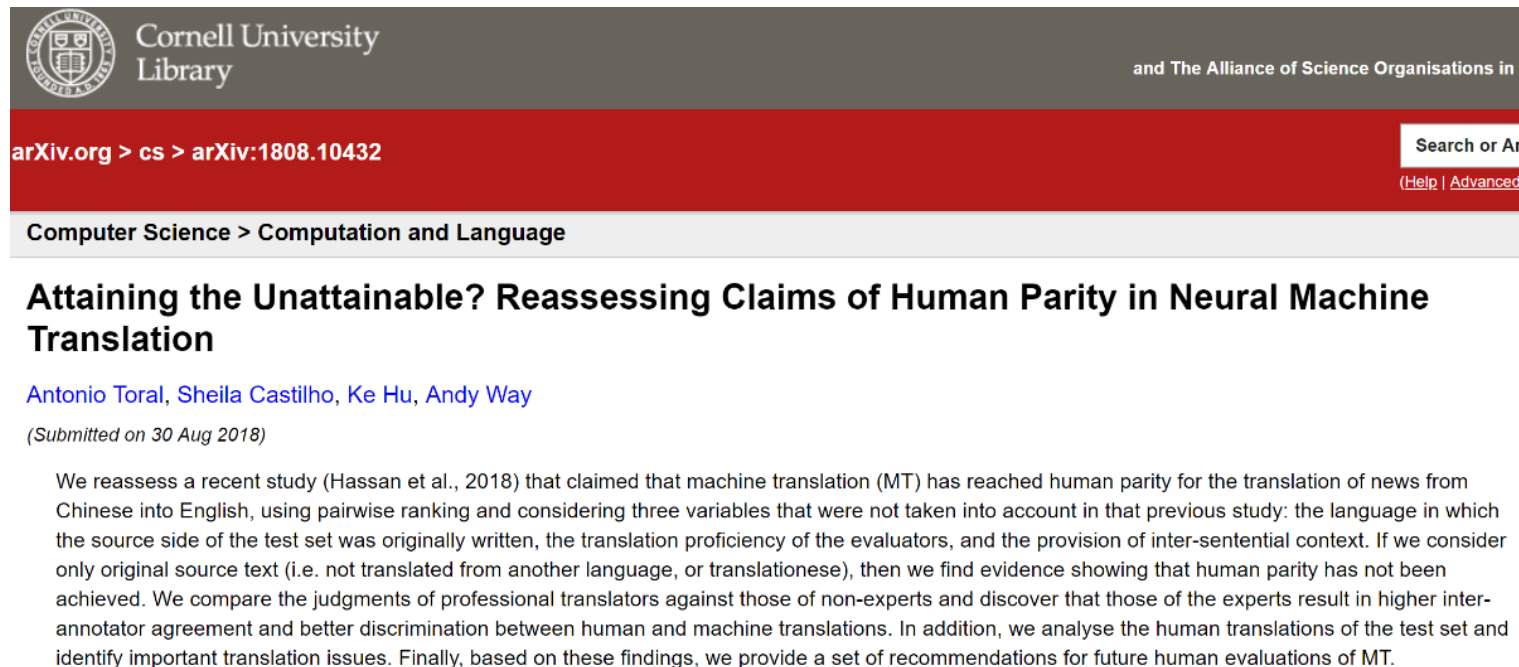


Intuition

Why now?



“...evidence showing that human parity has not been achieved.”

A screenshot of an arXiv paper page. At the top left is the Cornell University Library logo. The page title is "Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation" by Antonio Toral, Sheila Castilho, Ke Hu, and Andy Way. The submission date is 30 Aug 2018. The abstract text is visible below the title.

Cornell University Library and The Alliance of Science Organisations in C

arXiv.org > cs > arXiv:1808.10432 Search or Art (Help | Advanced)

Computer Science > Computation and Language

Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation

Antonio Toral, Sheila Castilho, Ke Hu, Andy Way
(Submitted on 30 Aug 2018)

We reassess a recent study (Hassan et al., 2018) that claimed that machine translation (MT) has reached human parity for the translation of news from Chinese into English, using pairwise ranking and considering three variables that were not taken into account in that previous study: the language in which the source side of the test set was originally written, the translation proficiency of the evaluators, and the provision of inter-sentential context. If we consider only original source text (i.e. not translated from another language, or translationese), then we find evidence showing that human parity has not been achieved. We compare the judgments of professional translators against those of non-experts and discover that those of the experts result in higher inter-annotator agreement and better discrimination between human and machine translations. In addition, we analyse the human translations of the test set and identify important translation issues. Finally, based on these findings, we provide a set of recommendations for future human evaluations of MT.

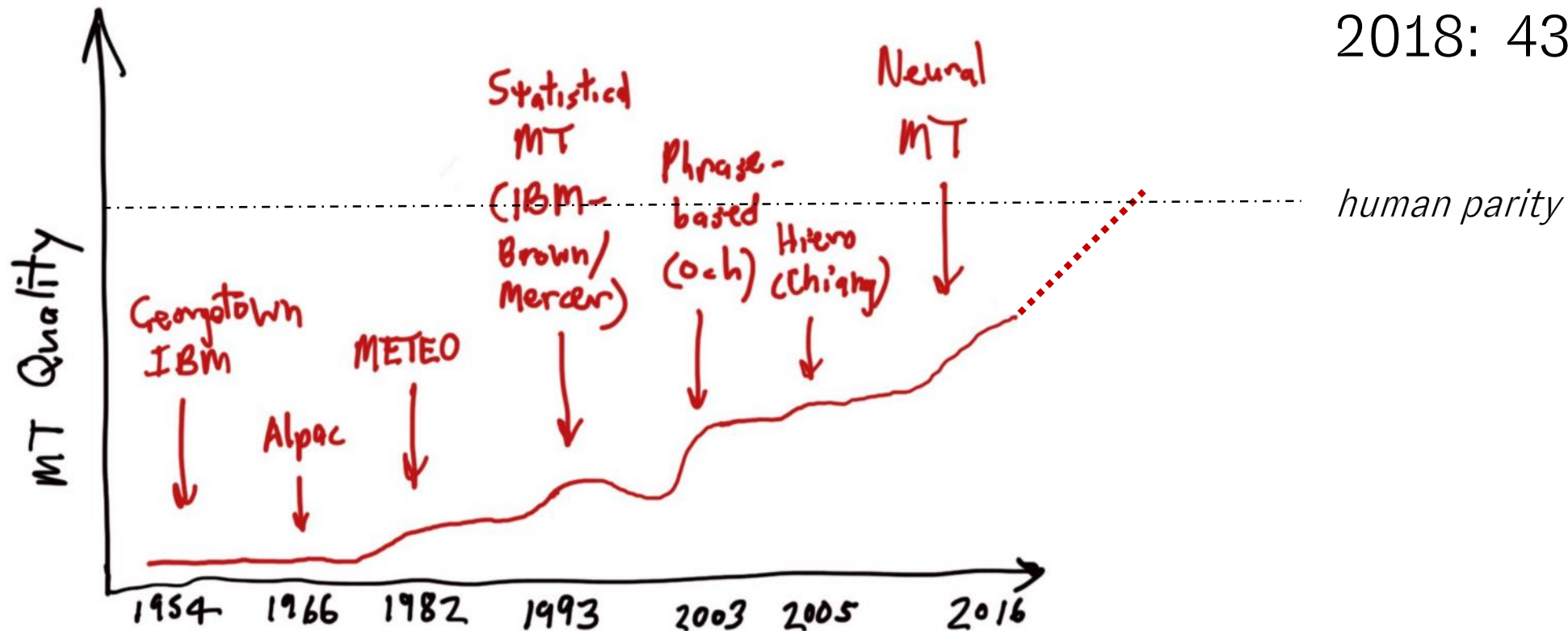
1. originally written in source language
2. take the whole document into account
3. evaluations be done by professional translators

4. time and money?



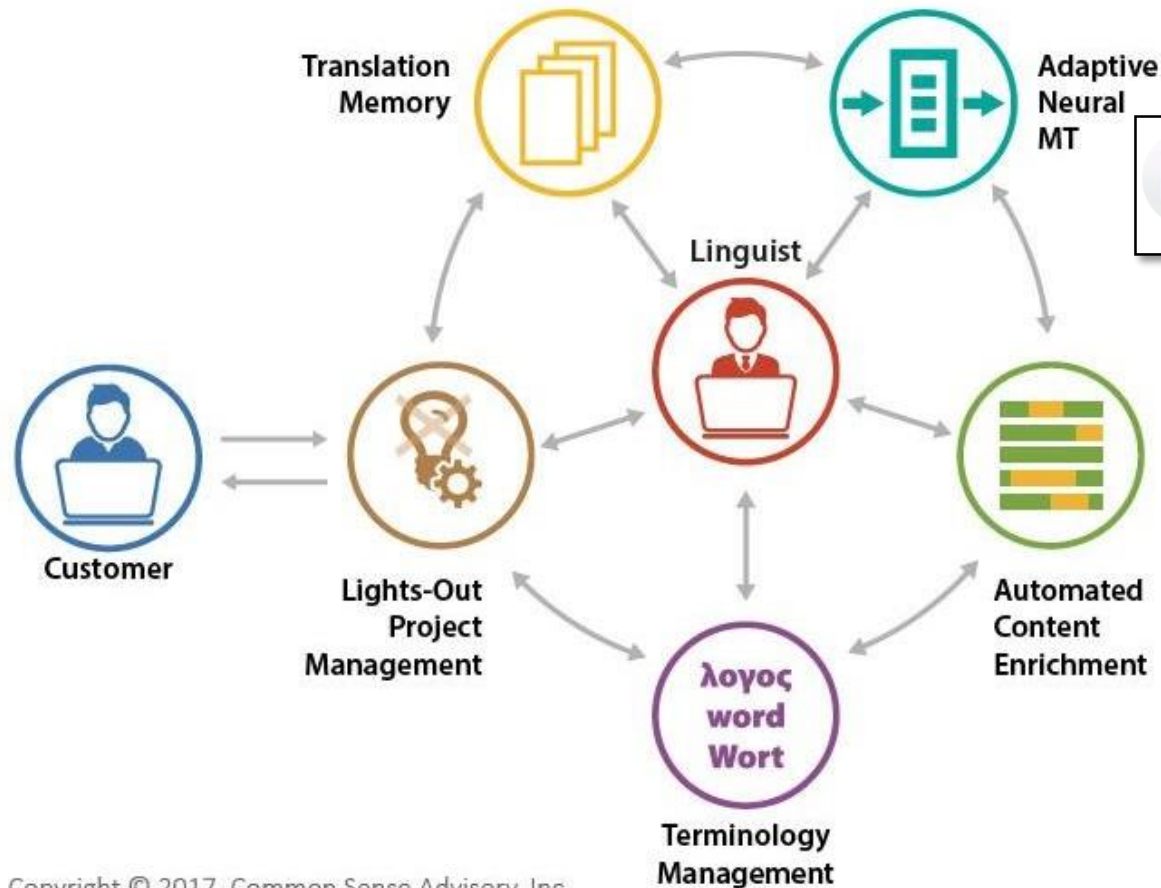
NMT Crossing the Rubicon

Progress in MT




2017: 199 papers
2018: 430 papers


Augmented Translation?



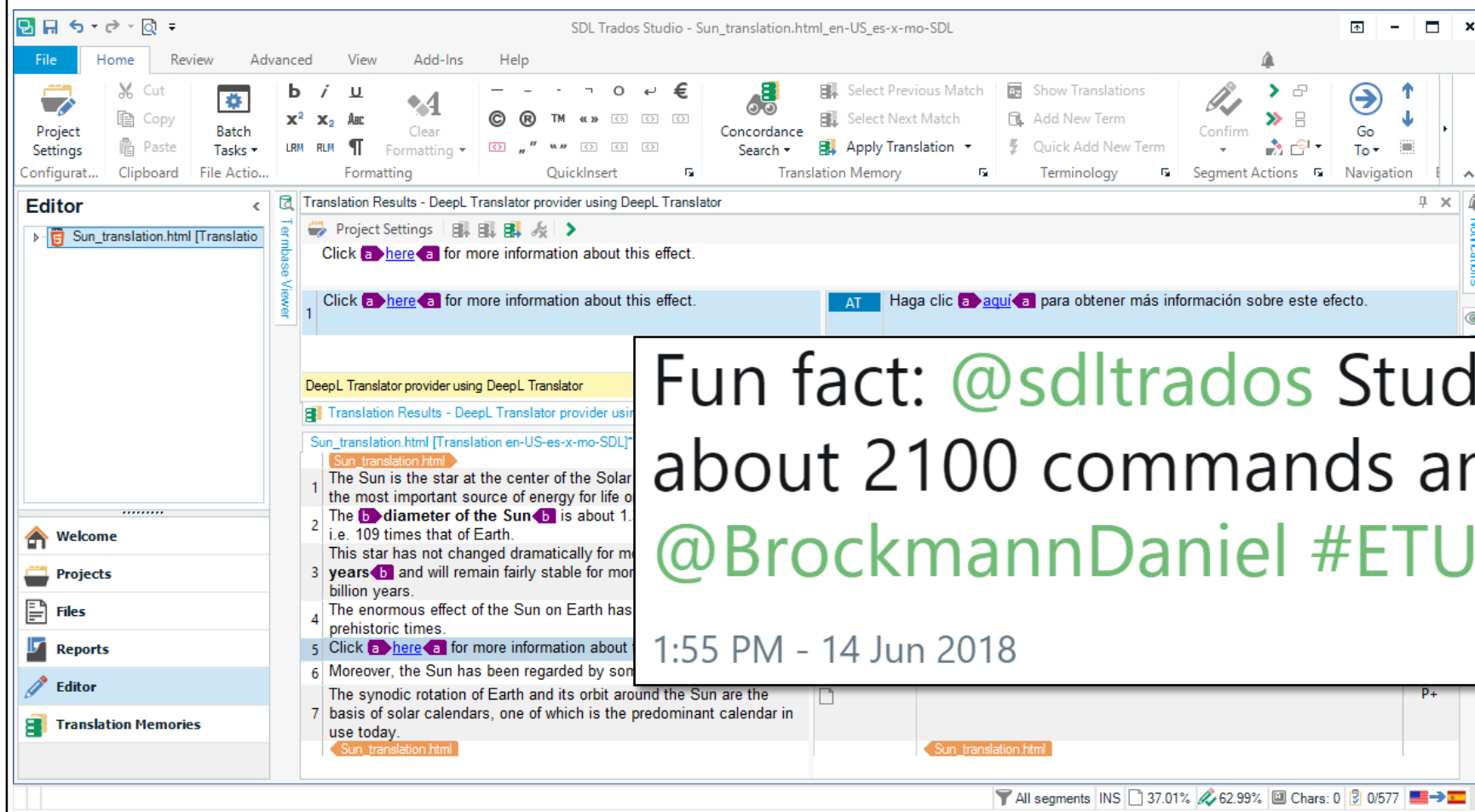
Copyright © 2017, Common Sense Advisory, Inc.

 **Lilt** @LiltHQ · 18 Feb 2017
Adaptive MT from **Lilt** one of four fundamental technologies for the "augmented translator" #xl8 @CSA_Research buff.ly/2ITm5j2

Jeromobot @Jeromobot
Replying to @ProfessorOlsen @adrechsel and 3 others
Yes, I think this is true. The positive aspect of "augmented translation" is that the translator is the agent who is being supplied with data by technology as s/he drives the process. Post-editing of MT would therefore not be augmented translation.

 **Véronique Özkaya** @vozkaya Following
With NMT, the future is happening faster. Augmented translation at our door.
[#SlatorConSF18](#)

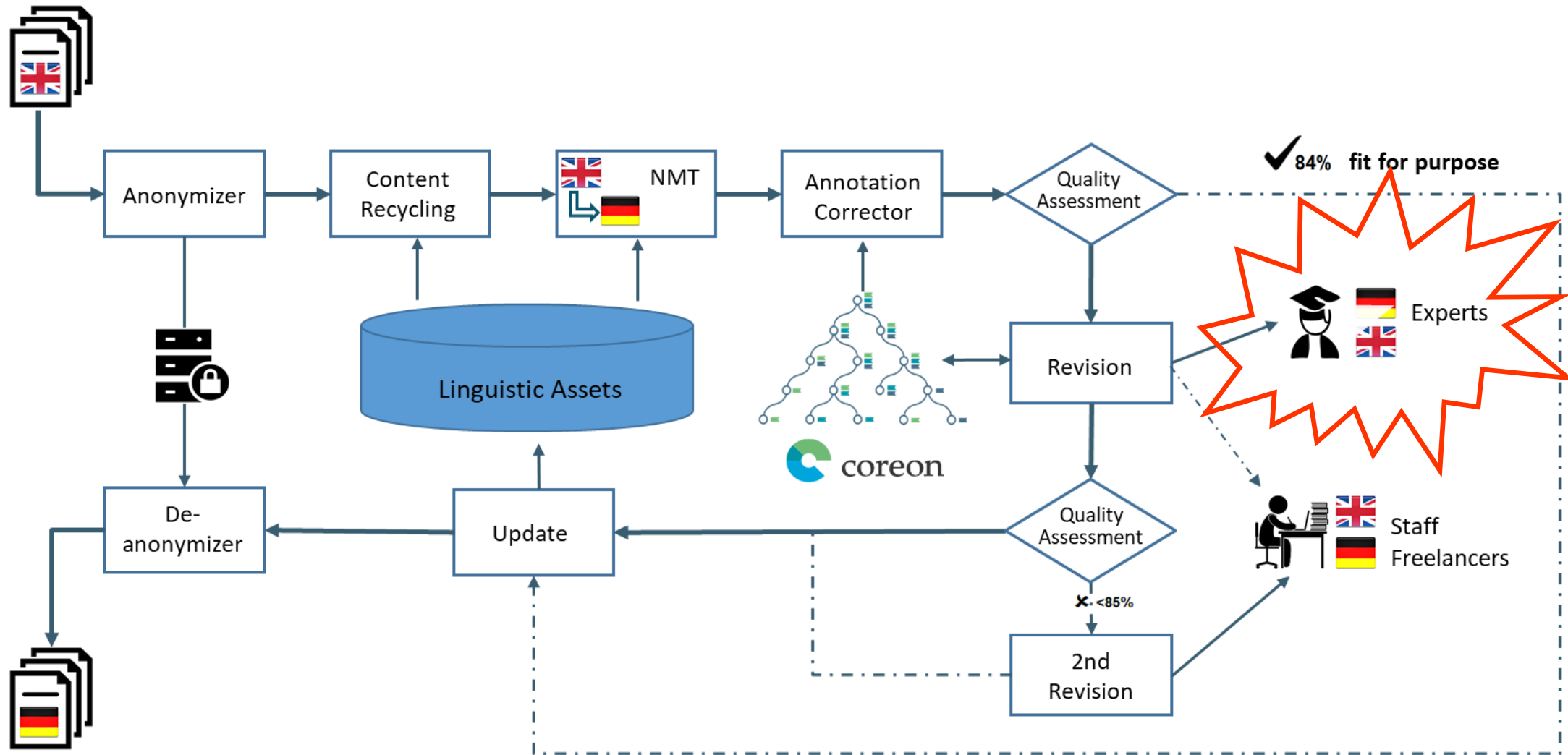
Fat, Complex Clients



Fun fact: @sdltrados Studio currently has about 2100 commands and settings.
@BrockmannDaniel #ETUG2018

1:55 PM - 14 Jun 2018

New NMT Workflow Sunsets CAT

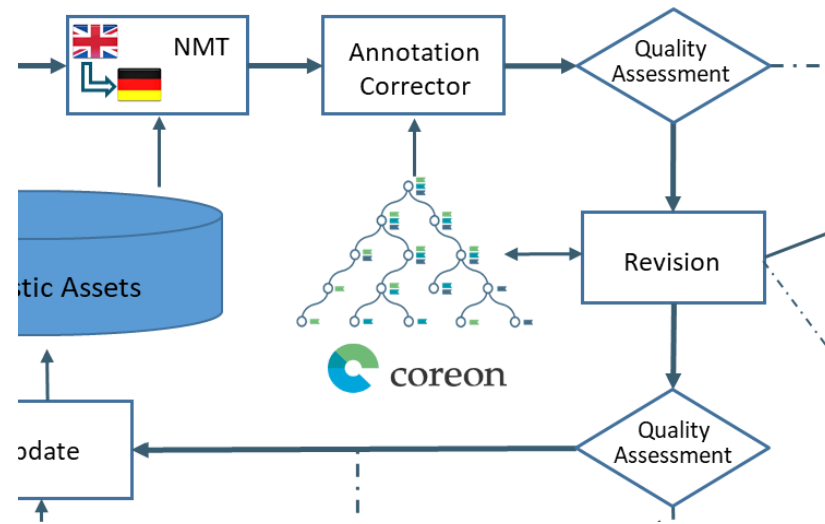


NMT Workflow Questions



- ¿ Content to recycle?
- ¿ Translations to train the NMT?
- ¿ NMT engine to pick?
- ¿ Terms correct or wrong?
- ¿ Style rules to apply?
- ¿ Criteria for fit-for-purpose?
- ¿ Best experts?
- ¿ If busy, second best?

Key Enabler Multilingual Knowledge System



An MKS classifies linguistic assets, human resources, workflow rules, and projects in a unified system which is expandable, dynamic, and provides fallback paths.

Leverage Existing Assets

Multilingual Terminology List



Knowledge Graphs



*Structure
semantically*

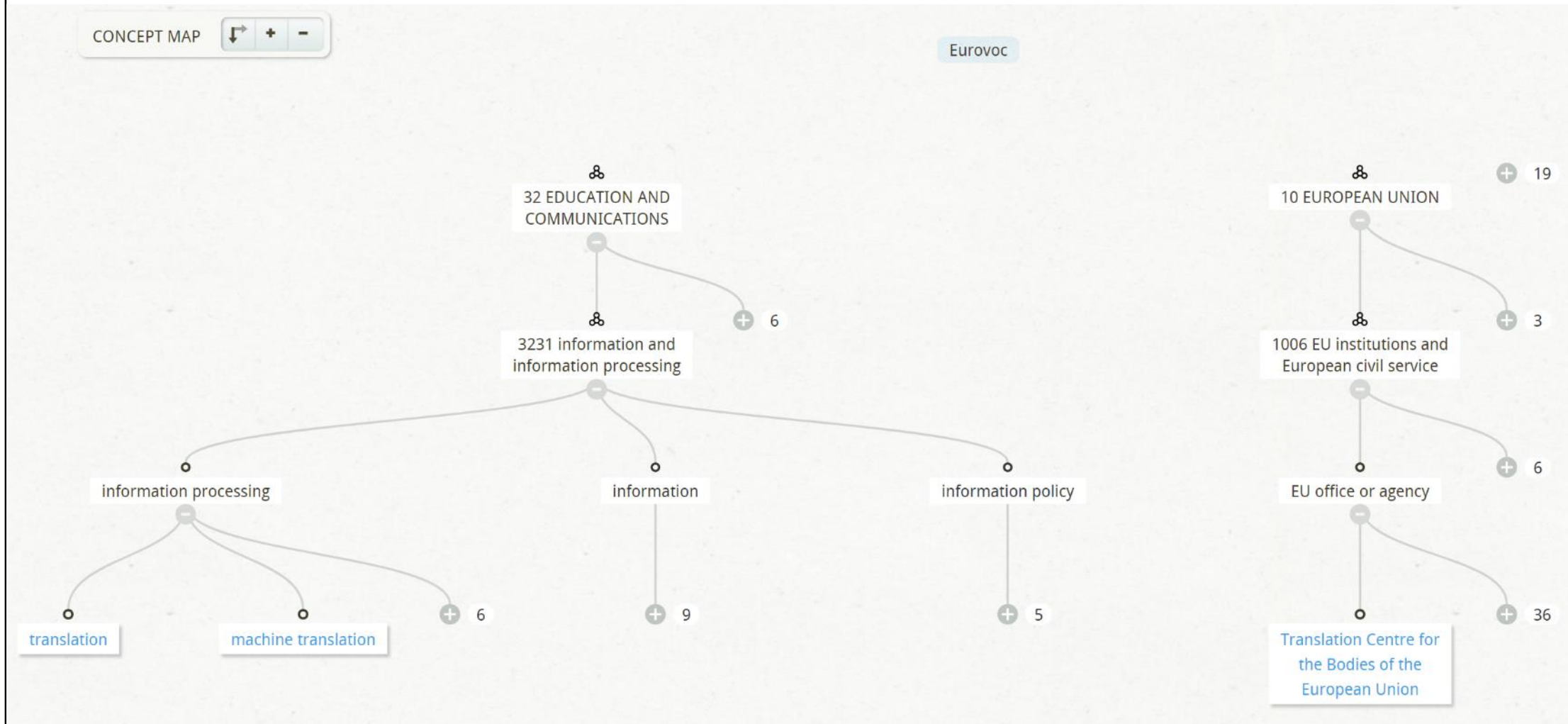


*Fill-up with languages,
variants, control*

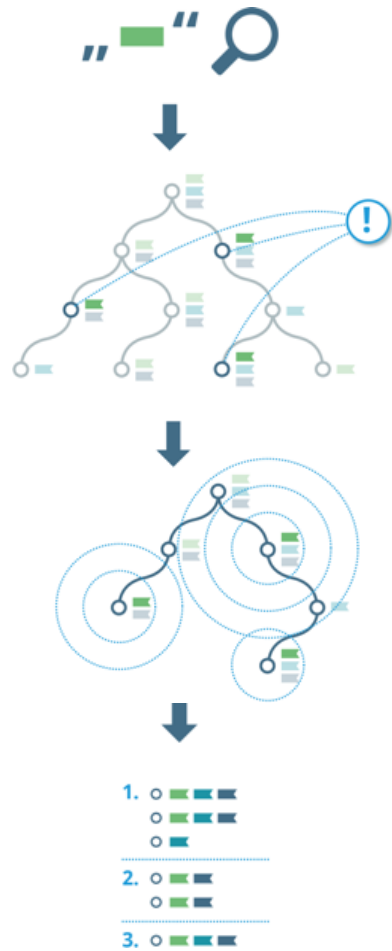


Multilingual Knowledge System

Multilingual Knowledge System



From Strings to Meaning



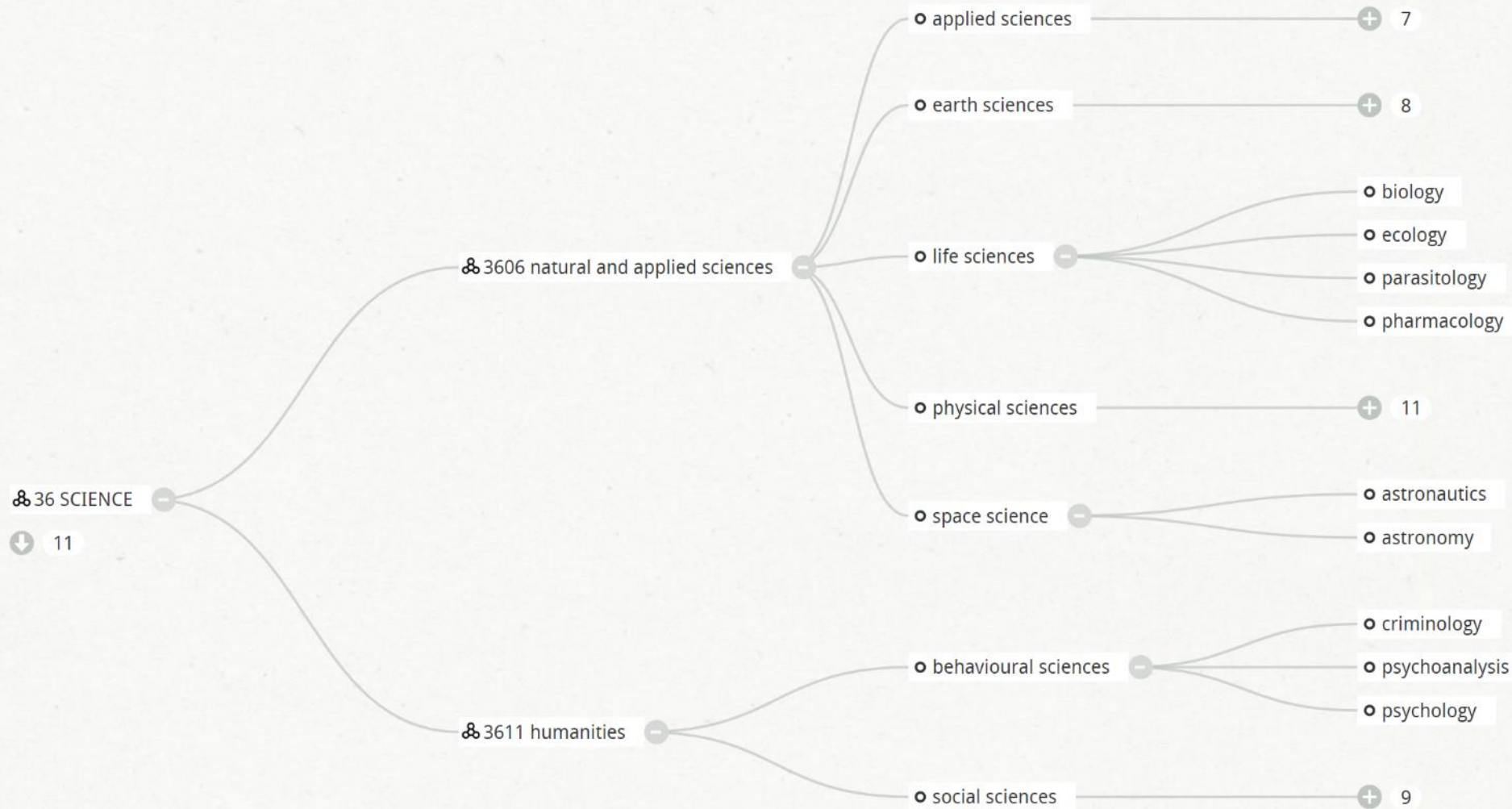
Find terms in text using Advanced Linguistic Search.

Locate them in the taxonomy.

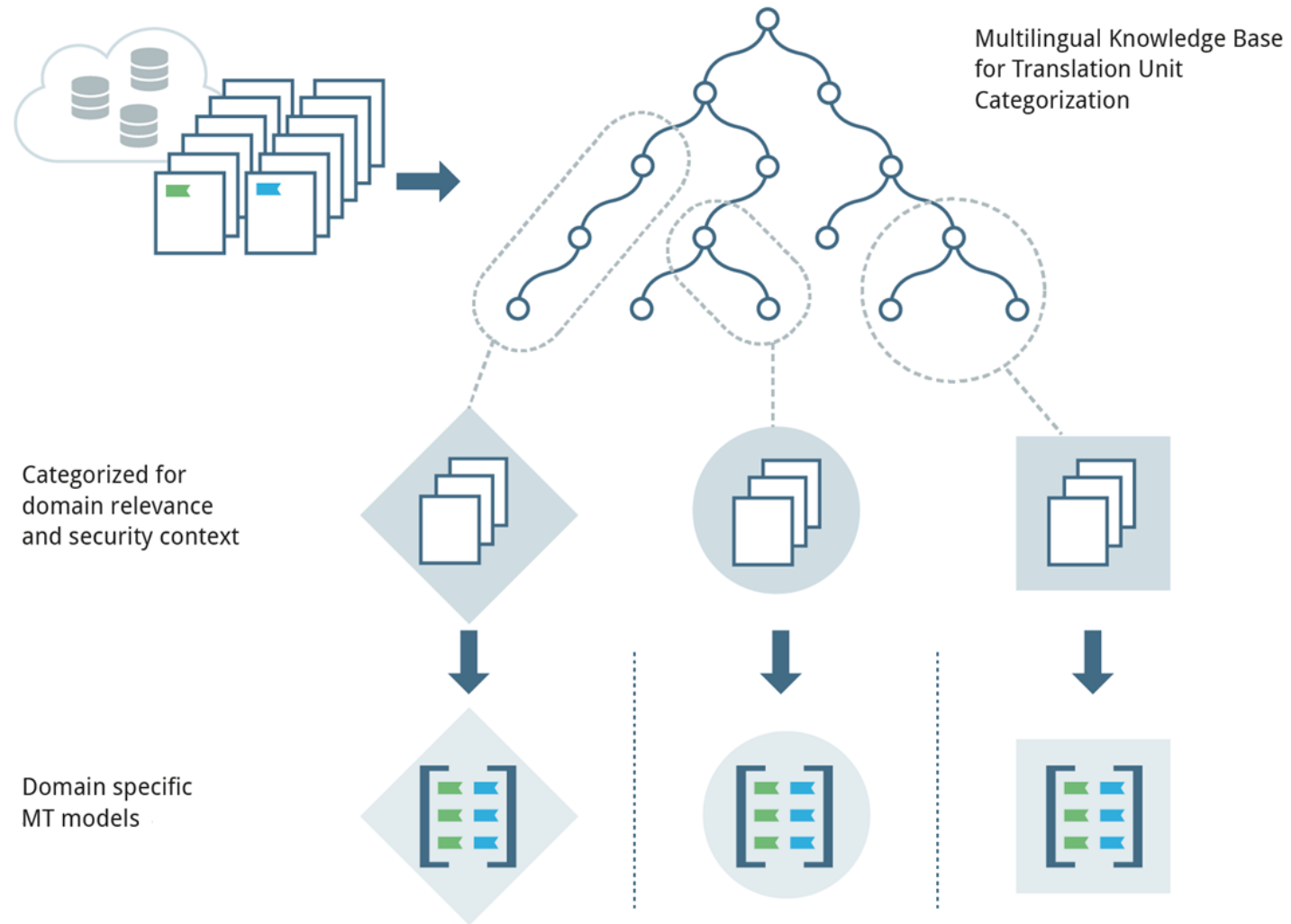
Create domain heat maps.

Link to other datasets (HR, DMS, QA rules, etc.).

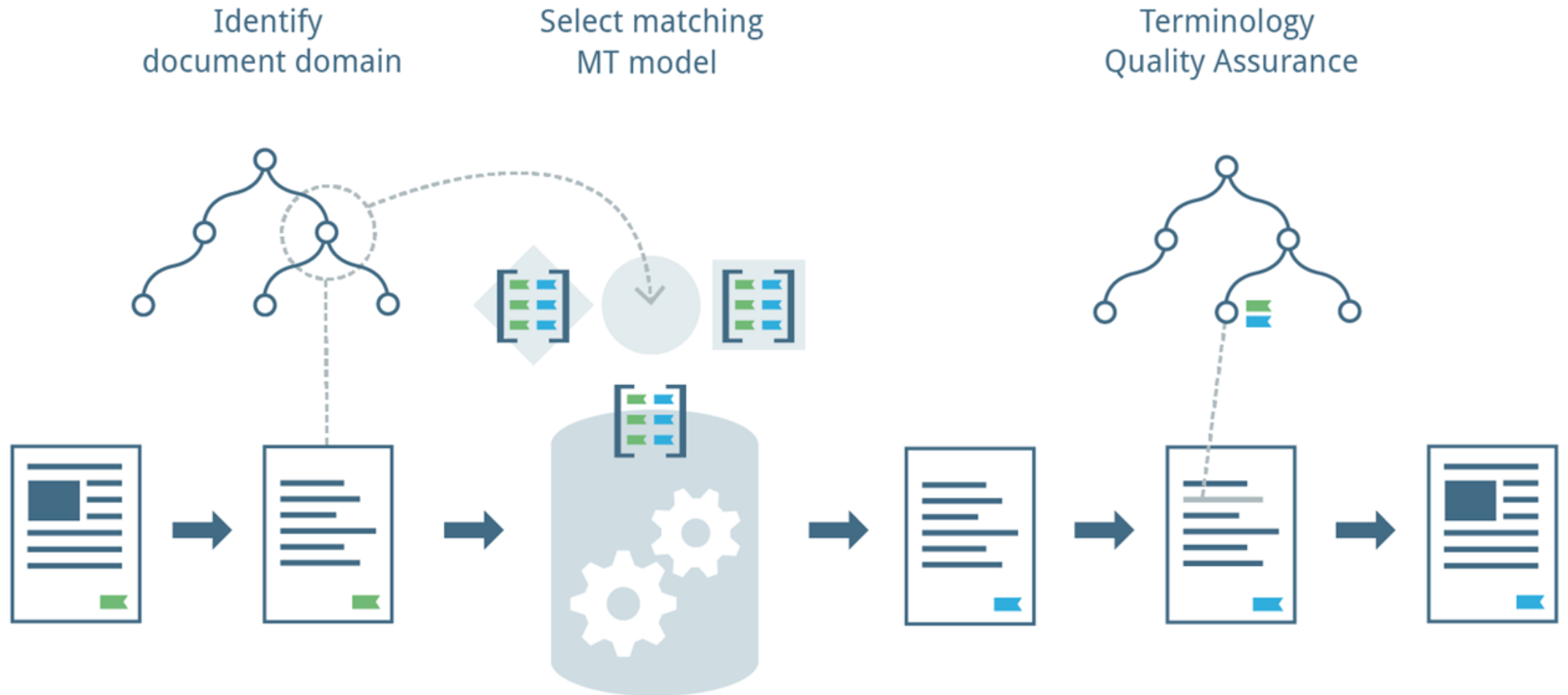
Classification with a Multilingual Taxonomy



Knowledge-based MT Training



Knowledge-based MT Workflow



New Roles for Linguists



- ✓ Revisor
- ✓ Multilingual Knowledge Manager
- ✓ Linguistic Assets Curator
- ✓ Translation Workflow Engineer

Ask me two Questions!



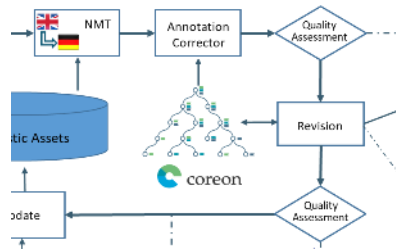
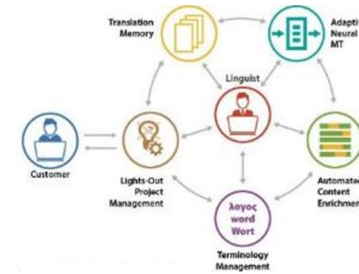
NMT disruptes translation

Industry getting it wrong

NMT Workflow sunsets CAT

Multilingual Knowledge System

New roles for linguists





coreon

Knowledge meets language.

m jochen@coreon.com

s Jochen.Hummel

t @JochenHummel

l Berlin-Mitte

Discover it now:
app.coreon.com
www.semiox.com