Words in Flux: How AI is Rewriting the Language Industry

Gábor Bessenyei

Globalese

memoqfest

memoQ Internal

Two souls alas! are dwelling in my breast.

memoqfest

memoQ Internal

Trends in the language industry

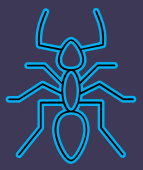
memoqfest

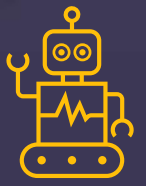
memoQ Internal

Trends in the language industry

Increasing demand for translation

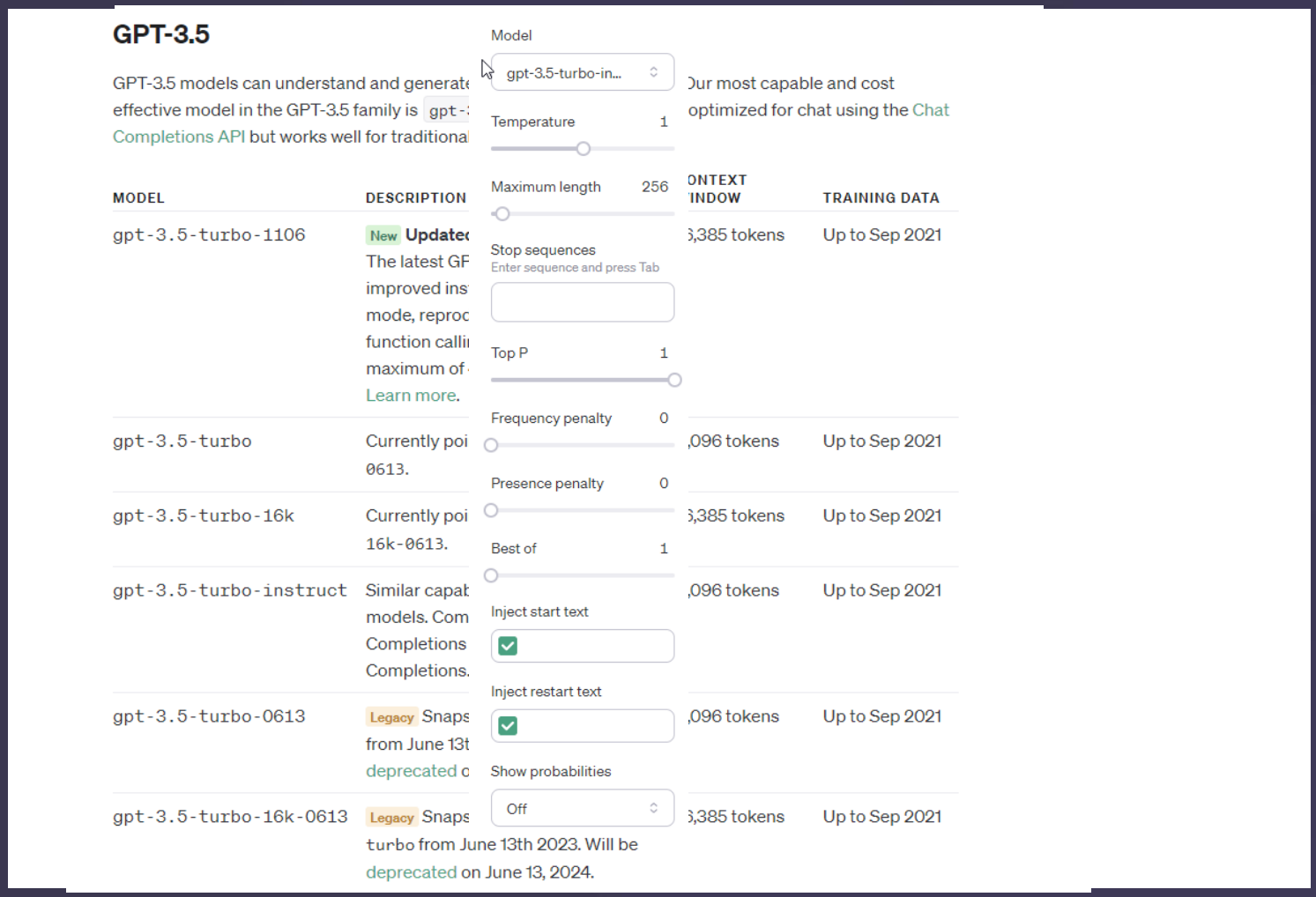
Significant decrease of turnaround times

Smaller, but more frequent projects 

Booming technology (MT & AI) 

memoQ Internal

Behind AI, there is more than just ChatGPT



GPT-3.5

GPT-3.5 models can understand and generate
effective model in the GPT-3.5 family is `gpt-`
Completions API but works well for traditiona

Our most capable and cost
optimized for chat using the Chat

MODEL	DESCRIPTION	CONTEXT WINDOW	TRAINING DATA
gpt-3.5-turbo-1106	New **Update** The latest GF improved ins mode, reproc function calli maximum of Learn more.	,385 tokens	Up to Sep 2021
gpt-3.5-turbo	Currently poi 0613.	,096 tokens	Up to Sep 2021
gpt-3.5-turbo-16k	Currently poi 16k-0613.	,385 tokens	Up to Sep 2021
gpt-3.5-turbo-instruct	Similar capab models. Com Completions Completions.	,096 tokens	Up to Sep 2021
gpt-3.5-turbo-0613	Legacy Snaps from June 13t deprecated o	,096 tokens	Up to Sep 2021
gpt-3.5-turbo-16k-0613	Legacy Snaps turbo from June 13th 2023. Will be deprecated on June 13, 2024.	,385 tokens	Up to Sep 2021

Model
gpt-3.5-turbo-in...

Temperature 1

Maximum length 256

Stop sequences
Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

Show probabilities
Off

memoqfest

memoQ Internal

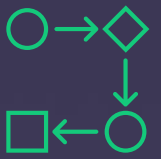
Consequences of the new trends

memoqfest

memoQ Internal

Consequences of the new trends

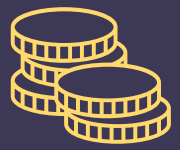
Changing workflows & processes



Changing tools

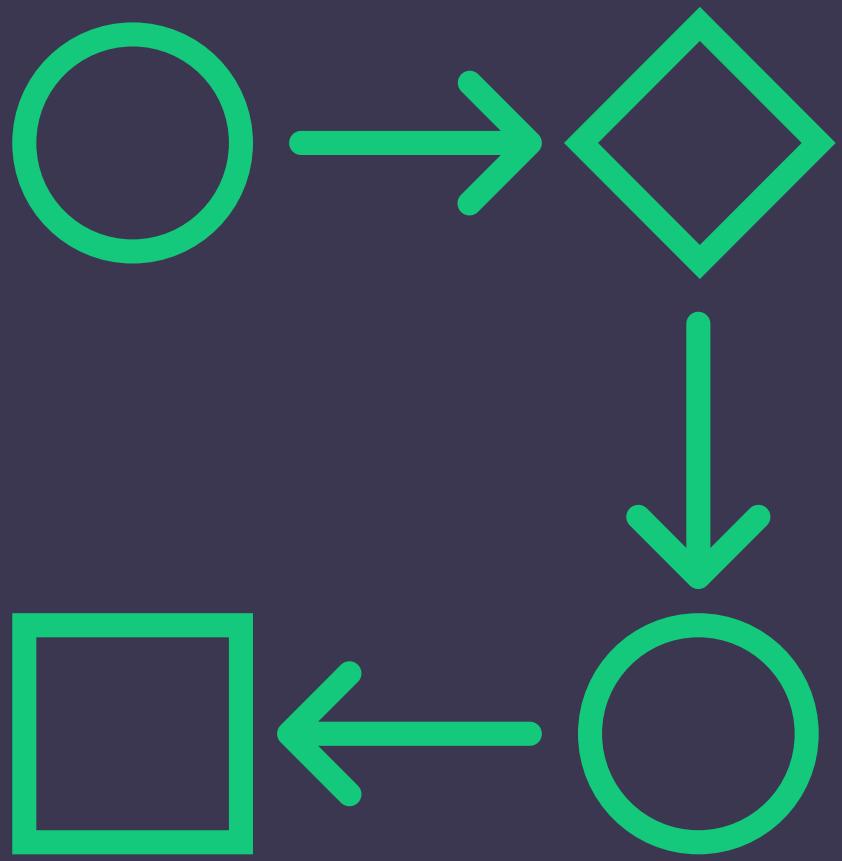
Changing tasks & roles

Changes in pricing



memoQ Internal

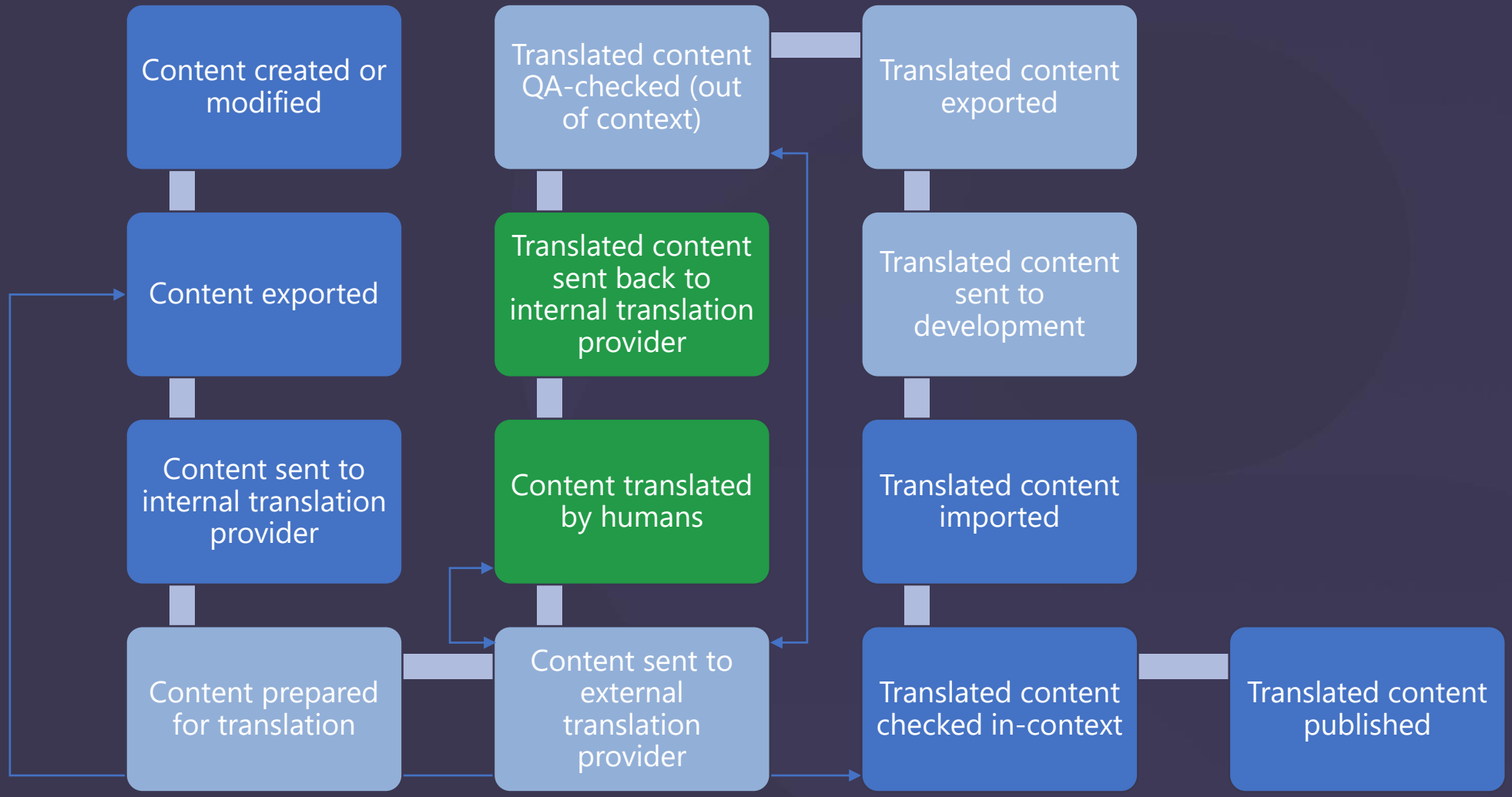
Changing workflows & processes



memoqfest

memoQ Internal

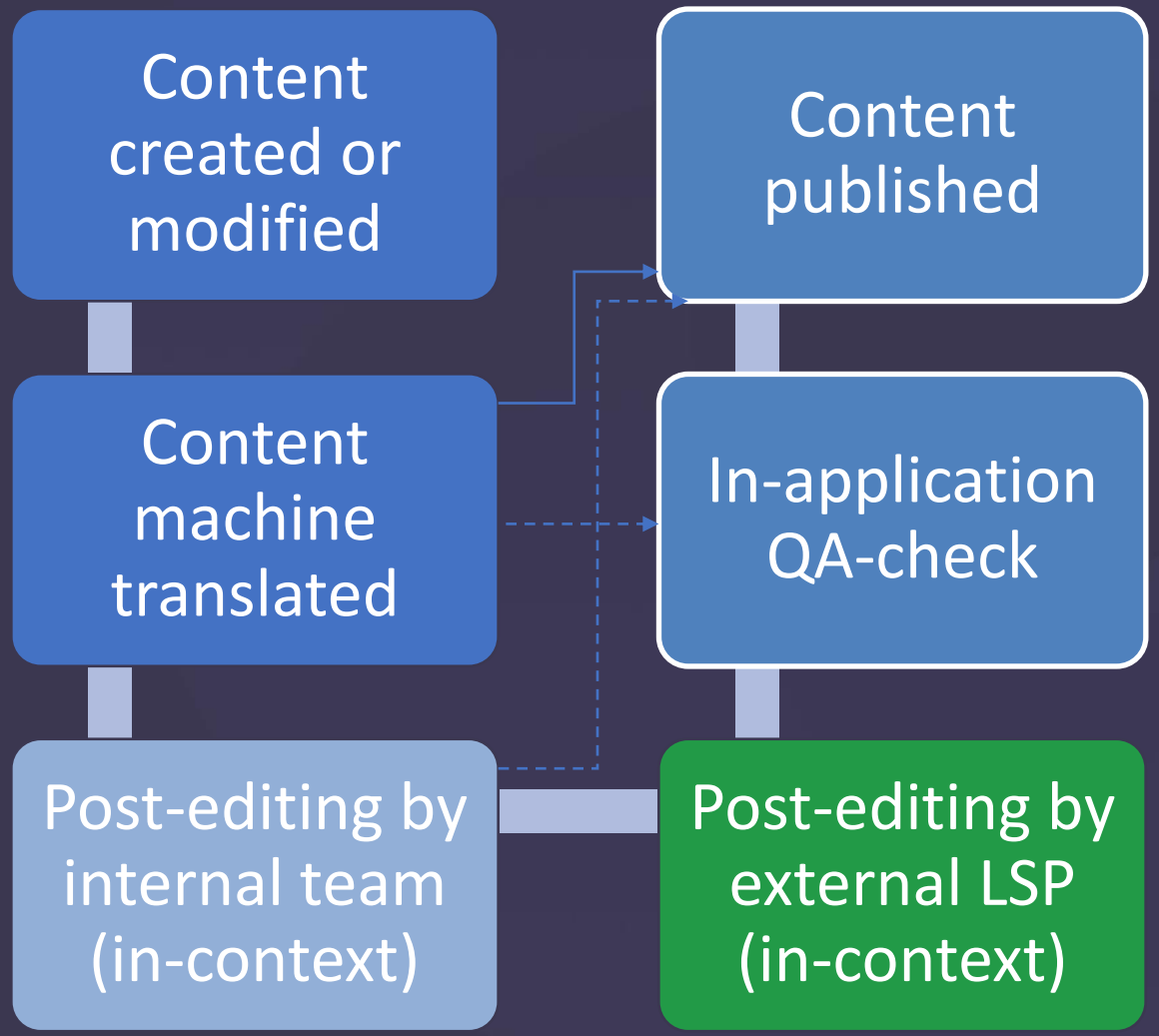
Changing workflows & processes



memoqfest

memoQ Internal

Changing workflows & processes



memoqfest

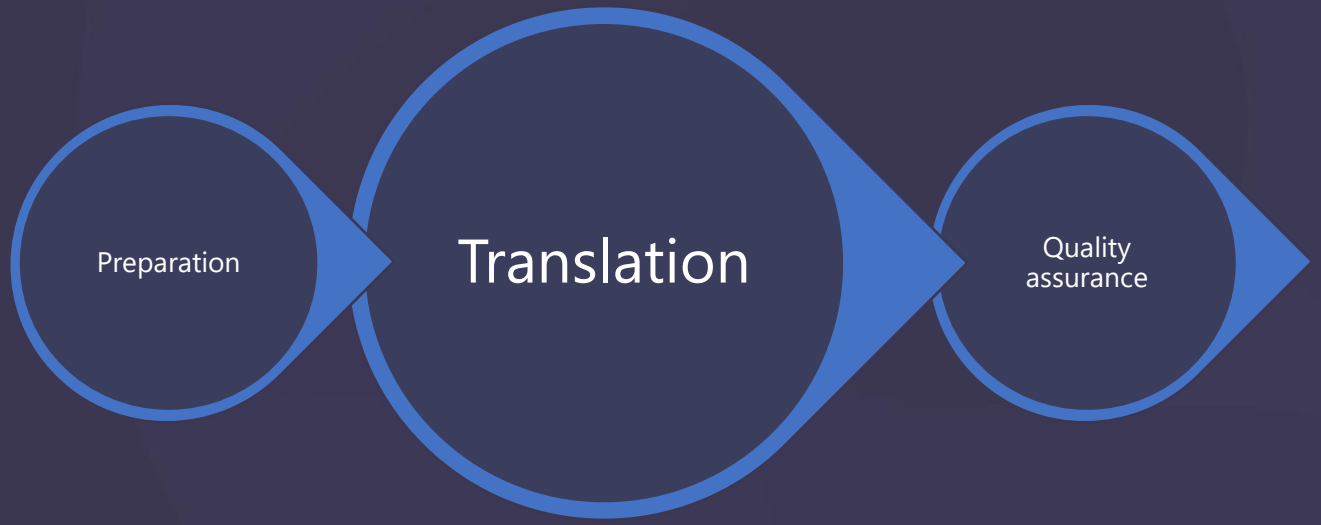
memoQ Internal

Changing tasks & roles

memoqfest

memoQ Internal

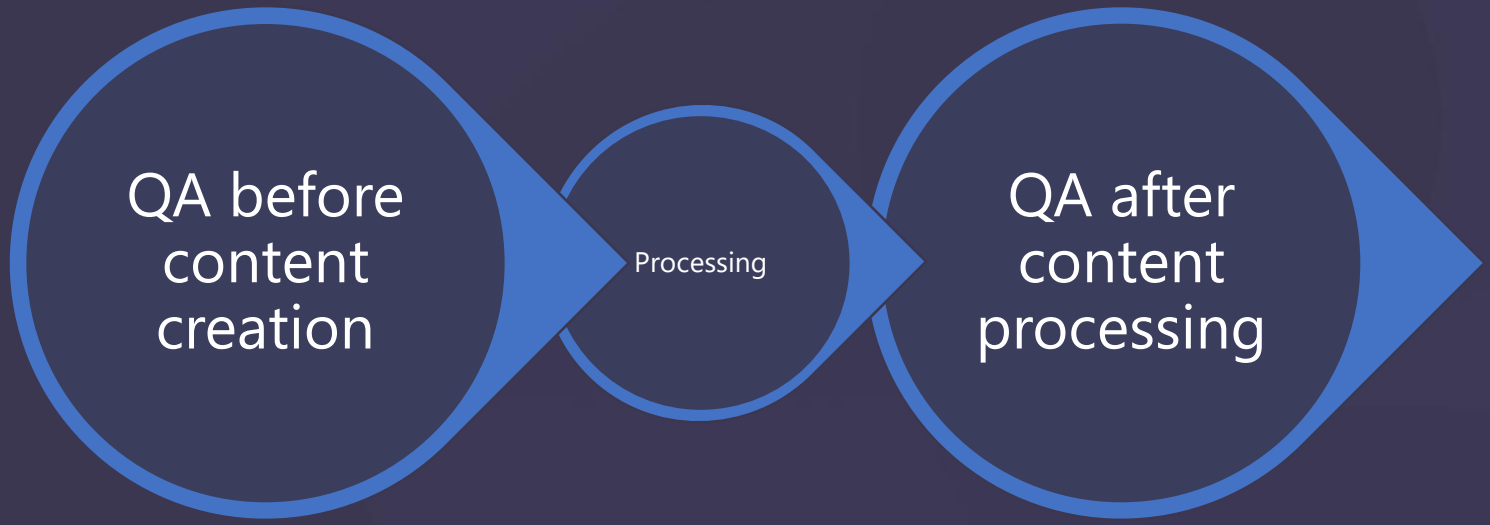
Changing tasks & roles



memoqfest

memoQ Internal

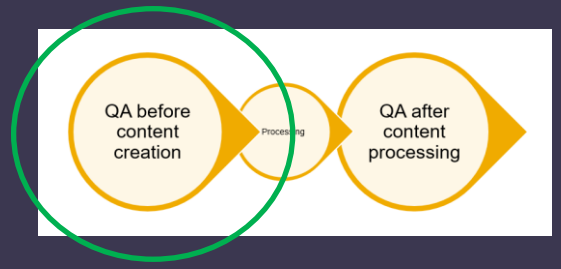
Changing tasks & roles



memoQ Internal

Changing tasks & roles

QA before content creation

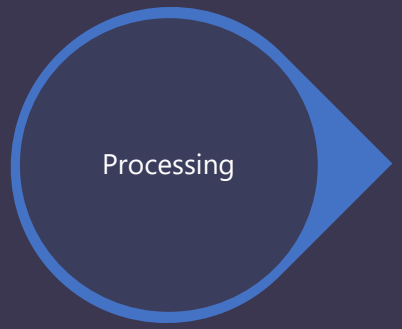


- Terminology work
- Translation training data set maintenance
- Translation reference data set maintenance
- Data labelling
- Content qualification
- Metadata creation and maintenance
- MT quality evaluation
- Prompting, prompt testing

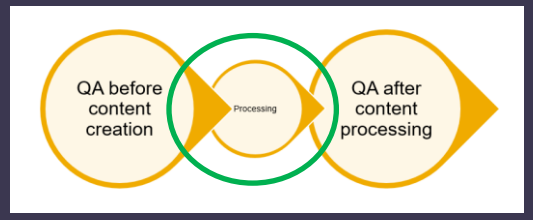
memoqfest

memoQ Internal

Changing tasks & roles



- QA on source
- QA on target (aka post-editing)



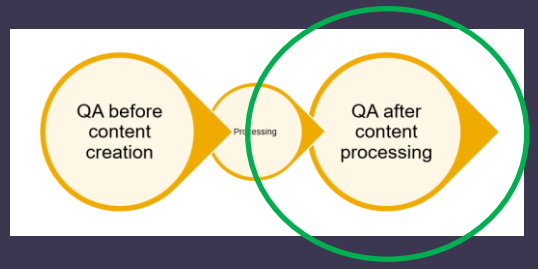
memoqfest

memoQ Internal

Changing tasks & roles

QA after content processing

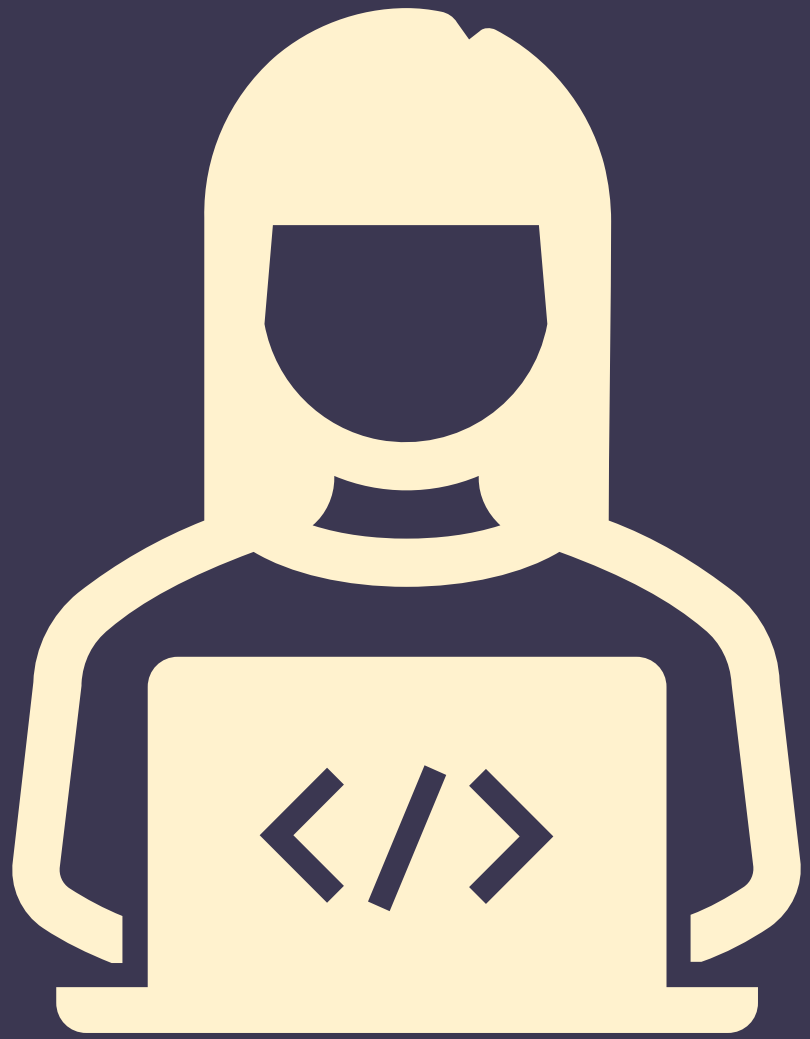
- In-context review
- In-application testing
- + all the tasks from the preparation phase



memoqfest

memoQ Internal

Changing tools



memoqfest

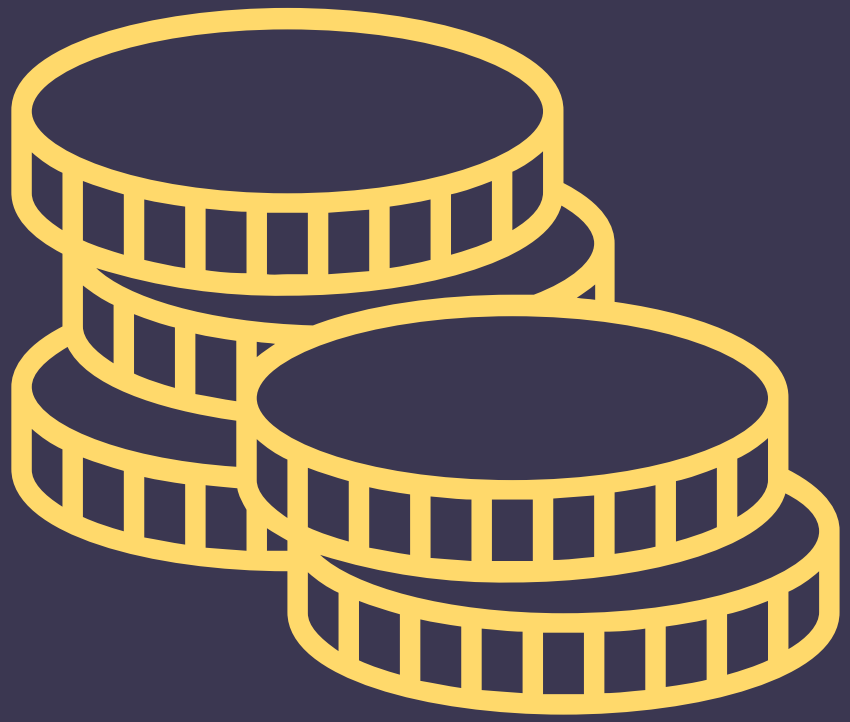
memoQ Internal

Changing tools

- Focusing on content analysis and
- Focusing on segments manipulation
- Editor heavy
- In-context editing
- QA tools are only additions
- QA tools are the heroes
- Prompting tools

memoQ Internal

Changes in pricing



memoqfest

memoQ Internal

Changes in pricing

- Volume-based pricing is obsolete
- Fuzzy-match and Edit Distance type calculations are obsolete
- Hourly-based, project-based, individual flat price models will come back

memoQ Internal

Summary

memoqfest

memoQ Internal

Real-life example

- Software UI translation
- 50 target languages
- 800k characters in total
- Domain: law enforcement
- Raw MT, in-context review
- 1 Project Manager/linguist, 1 prompt engineer

memoQ Internal

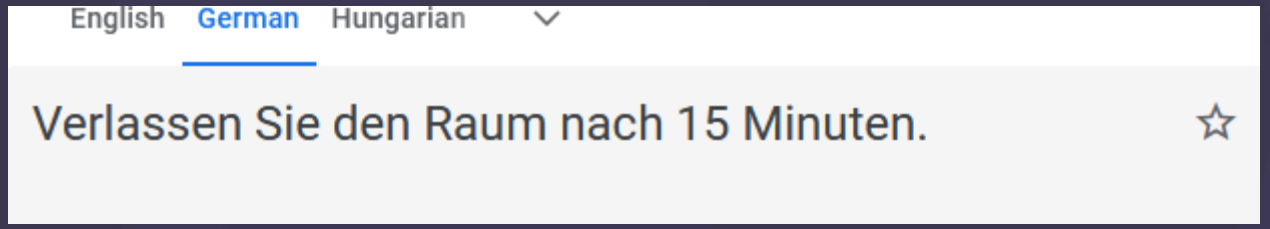
Challenges (all about prompting)

- Abbreviations
- Keywords/terminology
- Context window
- Validation

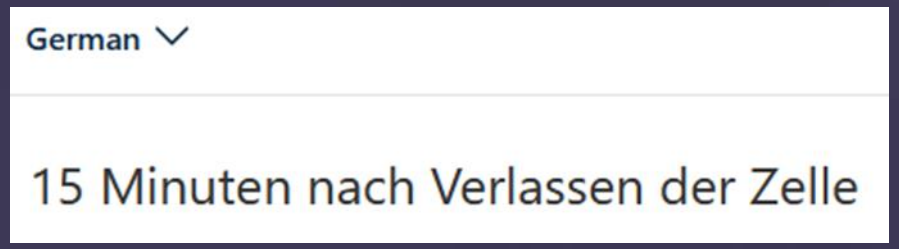
memoQ Internal

Challenges (all about prompting)

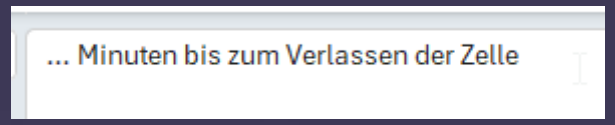
- Keyword/terminology



Generic MT 1



Generic MT 2



LLM with customized prompt

memoQ Internal

Challenges (all about prompting)

- Keywords + Abbreviations

Hungarian	German
Jogerúsen elítélt - fogház	Verurteilt nach dem Gesetz - Gefängnis

Hungarian	German
Jogerúsen elítélt - börtön	Rechtskräftig verurteilt - Gefängnis

Hungarian	German
Jogerúsen elítélt - fegyház	Verurteilt nach dem Gesetz - Gefängnis

Jogerősen elítélt - Fogház	Convicted with final judgment - Detention Center
Jogerősen elítélt - Börtön	Convicted with final judgment - Prison
Jogerősen elítélt - Fegyház	Convicted with final judgment - Penitentiary
Jogerősen elítélt - Pbh. Fogház	Convicted with final judgment - detention replacing a fine
Jogerősen elítélt - Km. Fogház	Convicted with final judgment - detention replacing community service
Jogerősen elítélt - Fk. Fogház	Convicted with final judgment - Juvenile Detention Center
Jogerősen elítélt - Fk. Börtön	Convicted with final judgment - Juvenile Prison

Generic MT 1 or Uncustomized LLM

LLM with customized prompt

memoQ Internal

Summary

- AI has reached already the level of human quality in translation
- Translation as a task / process step will disappear or at least shrink to a minimum
- The focus of the tools, process and roles will switch from translation to QA, data manipulation and prompt engineering

memoQ Internal

memoqfest

Thank you!

Any questions?

gabor.bessenyei@globalese.ai

memoQ Internal