



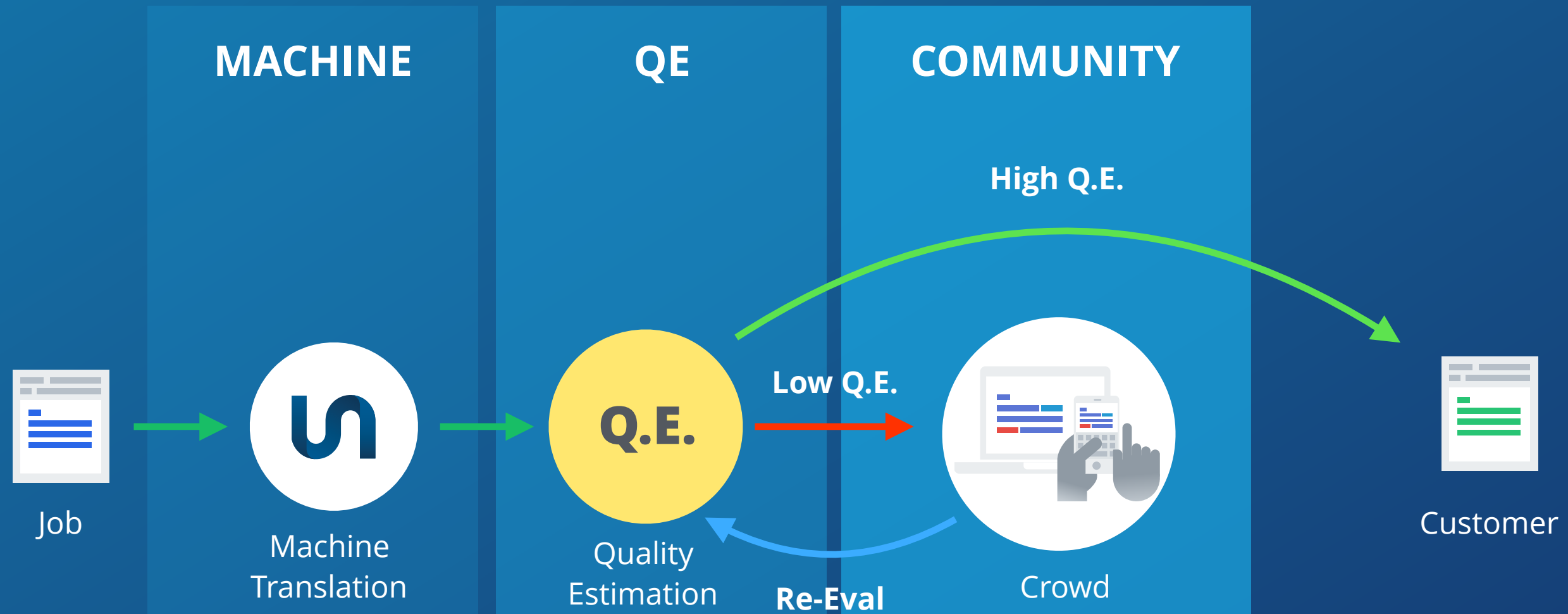
Unbabel's workflow: Crowdsourcing  
model and scalable quality  
management

**Helena Moniz**  
**Computational Linguist Researcher**

# Quality in a AI + Crowd pipeline

- What is quality and why does it matter in a AI and Crowd pipeline?
- What are the quality processes to firstly assess then assure?
- How can we create trust in a crowd ecosystem?

# Unbabel Pipeline



# Distributed Pipeline



# Dynamic Factors on Quality

- Initial Text (MT)
- Editors Pool and Assignment
- Content types
- Interfaces
- Language pairs (arguable, but true!)
- Periodic Quality Assessment for Quality Assurance:
  - Crowd evaluation
  - Jobs annotation

# Editors Pool



Evaluators

Annotators

# Editor's on-boarding

- **Language guidelines**
- **Usability guidelines**
- **Evaluation criteria**
- **Legal & ethical issues**

# Crowd Assessment

- **Evaluation task profile**

- Multiple choice tests with usually no feedback provided (not us!)

OR

- Scoring with feedback, so editors may learn and be evaluated again

- **Inter-rater agreement**

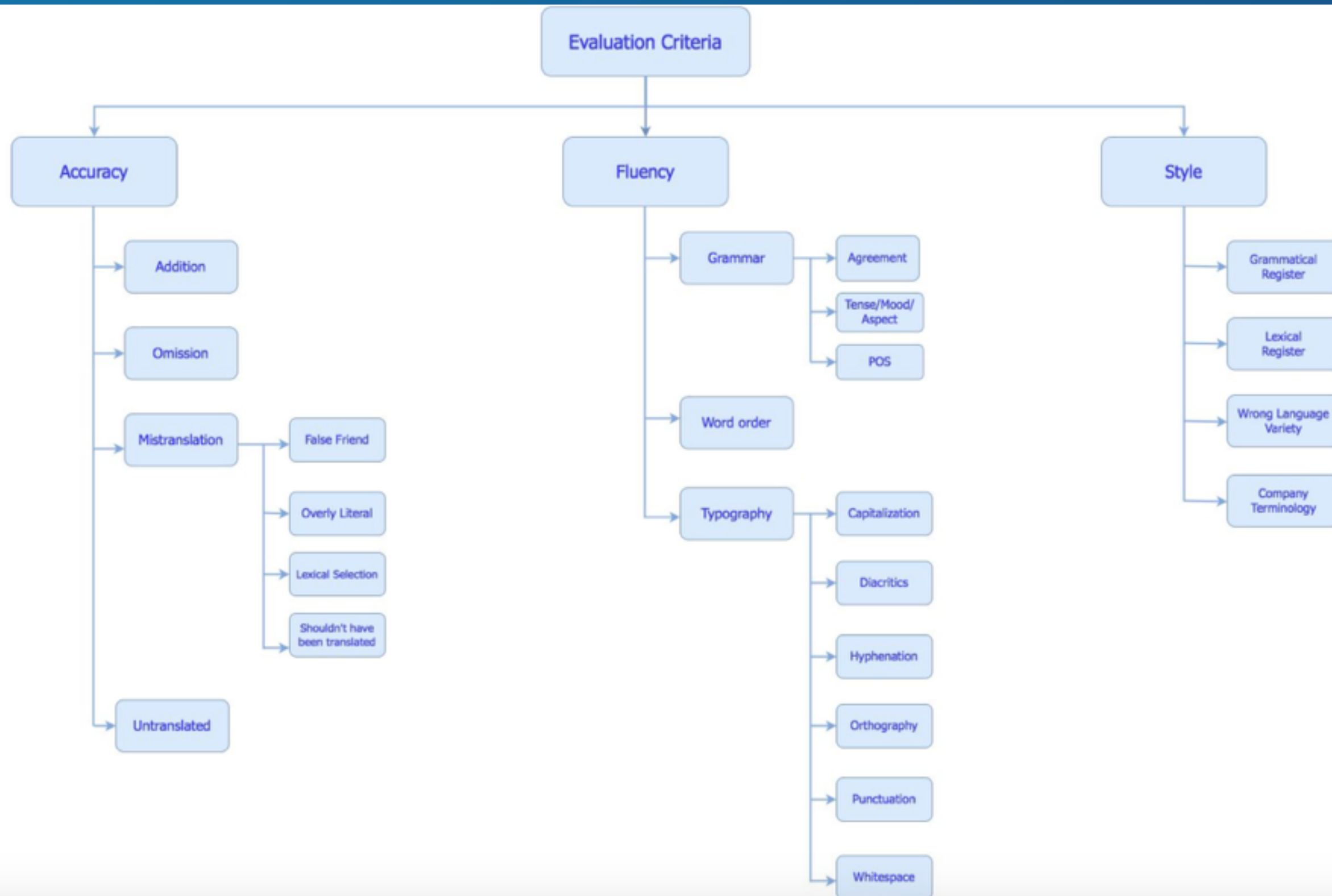
- several metrics
- Editors automatic evaluation
- Gold tasks for both evaluators and editors



# Crowd Assessment

- **Evaluators profile and training**
  - Usability / interface training
  - Templates to provide constructive feedback
  - Forum to discuss problematic issues

# Evaluation Error Criteria



# How Editors are Evaluated

The screenshot displays the 'Evaluation Tool' interface in a browser window. The URL is 'evaluation-tool-staging.unbabel.com'. The page header includes the logo, the text 'Evaluation Tool', and a user greeting 'Hi pg27042!'. Below the header, there is a navigation bar with a flag icon, the word 'Evaluation', and details: 'Language Pair: English to Portuguese(BR)', 'Tone: Not available', and 'Editor: blackshadowbr'. A 'Report' button is located on the right side of this bar.

The main content area is divided into two columns. The left column contains a yellow banner with the text 'Translate this article to train your skills.' Below this, there are sections for 'Source' and 'Previous Version'. The 'Source' section contains the text: "I just started fucking crying," says Drozd. "It's my most cherished musical possession." And the old Flickinger console, which Linkous had struggled to restore, is fully functioning at yet another old friend's studio in Richmond, capturing the sounds of songs awaiting to be heard. The 'Previous Version' section contains a Portuguese translation of the source text.

The right column contains a 'Translation' section with the Portuguese text: "Eu simplesmente comecei a chorar," diz Drozd. "É a minha propriedade musical mais querida." E o velho console Flickinger, que Linkous teve dificuldade para restaurar, está funcionando muito bem no estúdio de um outro velho amigo em Richmond, gravando os sons de músicas que estão esperando para serem escutadas." Below this is an 'Edits' section with a 'Hide edits' button. The 'Edits' section shows the same Portuguese text with several words highlighted in yellow, indicating corrections or suggestions. At the bottom of the 'Edits' section, there is a 'Rate this translation' section with a star rating system (None, ☆, ☆☆, ☆☆☆, ☆☆☆☆).

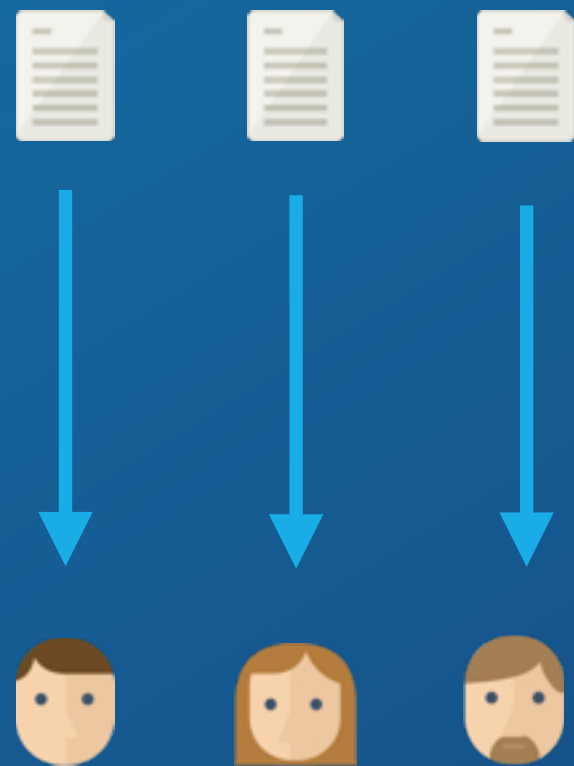
On the right side of the interface, there is a 'Comments' section with a text input area and a 'Submit Evaluation' button. Below the input area, there is a link to 'guidelines' and a 'Take a break' button. A message at the bottom of the right side says 'This will save your progress'.

# Editors Profiling



# Editor Assignment

Regular distribution



★★★★★ 3.8

old rating

Smart distribution



★★★★★ 4.6

Improved rating

# Editors Pool



Evaluators

**Annotators**

# Error typologies

- Localization Quality Evaluation (LQE, Google)
- Dynamic Quality Framework (DQF, TAUS), normalised with MQM
- Multidimensional Quality Metric (MQM, QT21)

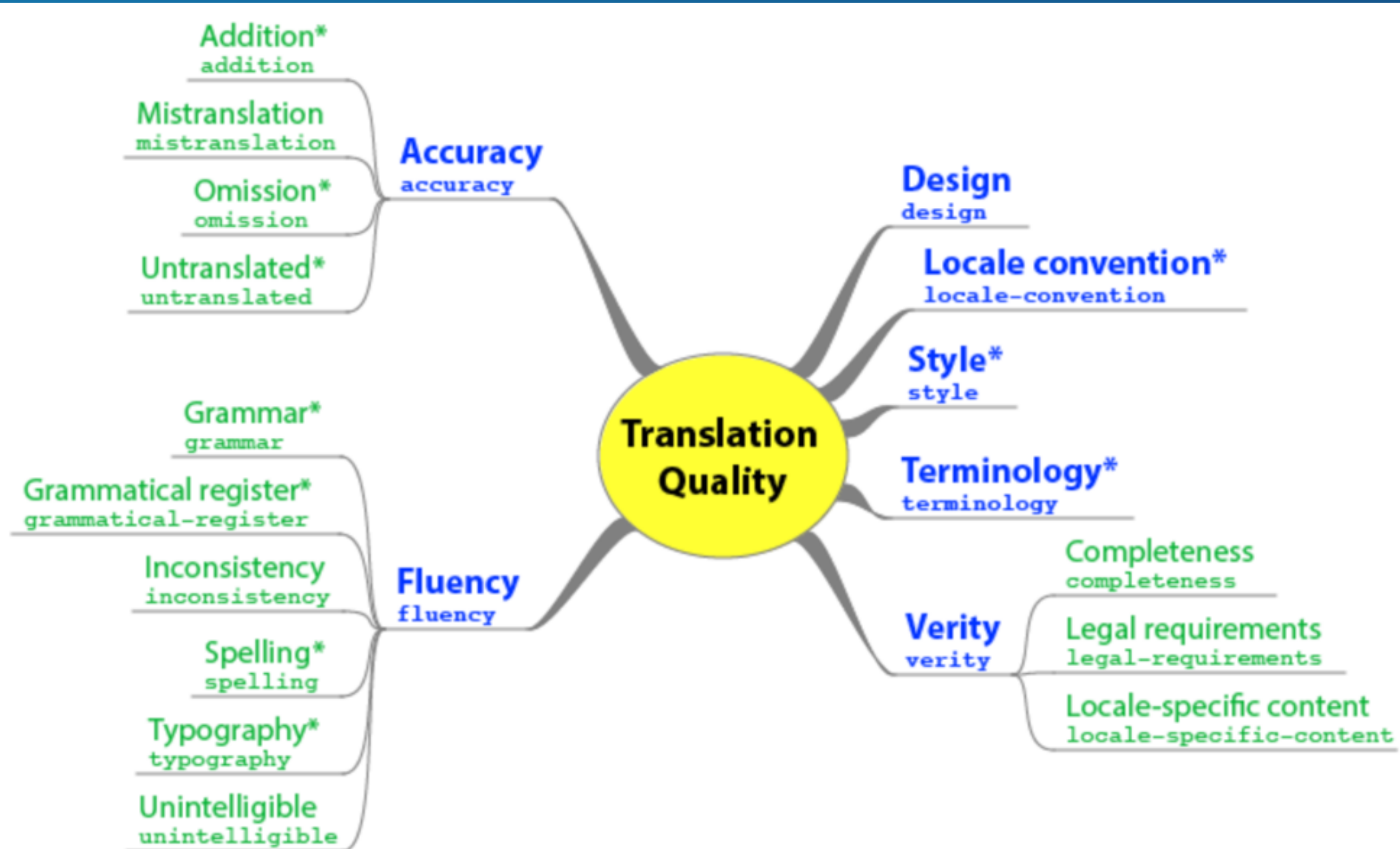


# MQM (QT21)

- 7 major categories:
  - Accuracy
  - Fluency
  - Style
  - Design
  - Locale
  - Verity
    - Fine-grained subcategorisation into 120 parameters
    - Subset of 60 parameters is used at Unbabel



# MQM Core



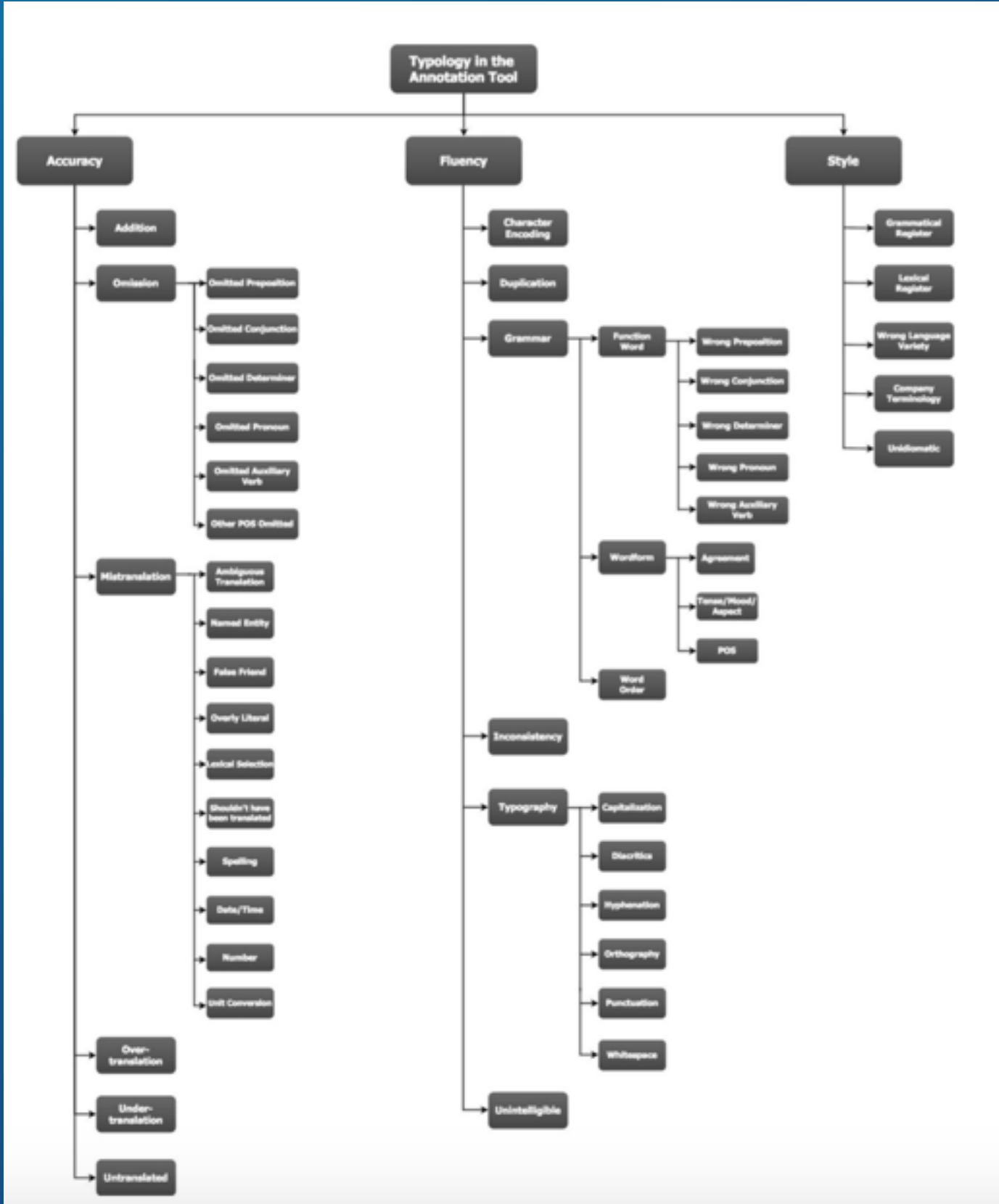
# MQM (QT21)

- 4 levels of severity:
  - neutral, minor, major, and critical.
- Formula:
  - $=100 - \text{SUM}((1 * \text{MINORS}) + (5 * \text{MAJORS}) + (10 * \text{CRITICALS})) / \# \text{Words} * 100$
  - *95% as professional level translation*

# Annotation at Unbabel

- **Profile of the annotator:**
  - PhD on Linguistics and/or Translation Studies
- **Inter-annotator agreement**
  - training usability / parameters
  - gold tasks
  - feedback
- **Recent research studies:**
  - positive impact of Decision Trees (tested with EN > DE)
  - severity levels harder for agreement than the type of errors

# Error Typology for Annotation



# Deep Annotations

Annotate. powered by Unbabel. Admin Statistics Hi Annotatoradmin

Job 1/25 (Annotator: Ingrid.zesta1, batch completed. Batch: HQ\_Batch\_20-08-2017\_to\_27-08-2017 QT21 Score: 105.93 Tone: Formal)

Client Instructions: These are Customer Support tickets for Rovio, game company that developed Angry Birds and others. Leave game titles and characters in English.

Original	Translation
Hello there,	Ciao.
and thanks for contacting us!	e grazie per averci contattato!
We're sorry to hear that the game is causing trouble.	Siamo spiacenti di sapere che il gioco le sta causando problemi.
In order to help you we need some additional information about the problem please.	Per <u>aiutarti</u> abbiamo bisogno di ulteriori informazioni sul problema per favore.
Which game are you referring to?	Quale <u>gioco</u> <u>stai</u> facendo riferimento?
What exact model is the device you are using and what operating system version does it have?	Quale <u>modello esatto</u> è il <u>dispositivo</u> che <u>stai</u> utilizzando e quale versione del sistema operativo ha?
What happens exactly when you are trying to play the game?	Cosa succede esattamente quando si sta cercando di <u>giocare il gioco</u> ?
Do you get an error message?	Riceve un messaggio di errore?
Have you connected the game to your Rovio account or Facebook?	<u>Hai</u> connesso il gioco al <u>tuò</u> account Rovio o Facebook?
If yes, what's your login email address?	Se sì, qual è il <u>tuò</u> indirizzo di posta elettronica di accesso?
We apologize for any inconvenience and look forward to your response.	Ci scusiamo per gli eventuali disagi e restiamo in attesa di una sua risposta.
Best regards,	Cordiali saluti,
Olga Rovio Support Team	Olga Rovio <u>Support Team</u>
More about Rovio Newsletter - URL-0 Facebook - URL-1 Twitter - URL-2 URL-3	<u>Scopri</u> di più su <u>Rovio Newsletter</u> - URL-0 Facebook - URL-1 Twitter - URL-2 URL-3

Annotations	
Type	Context
Register	Scopri
Register	aiutarti
Register	Hai
Overly Literal	Cosa succede esattamente quando si sta cercando di giocare il gioco
Word Order	Rovio Newsletter
Prepositions	
Overly Literal	Quale modello esatto è il dispositivo
Register	Ciao
Untranslated	Support Team
Register	tuò
Register	stai
Register	stai
Register	tuò

# Annotation at Unbabel

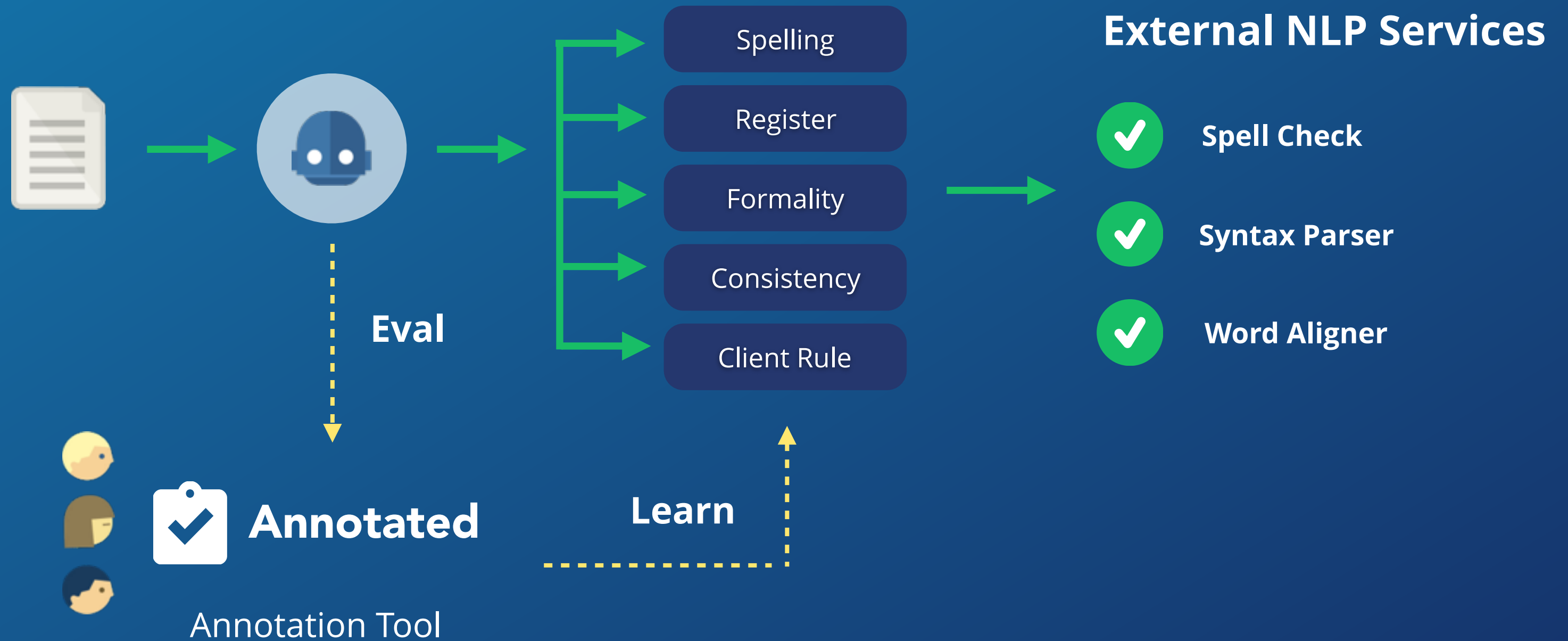
- **Annotation output**
  - monitor quality in all steps / customer / domain /interfaces / QE and MT produced
  - feedback for editors (linguistic tips)
  - correlation with evaluation scores
  - feed NLP modules



# Smartcheck

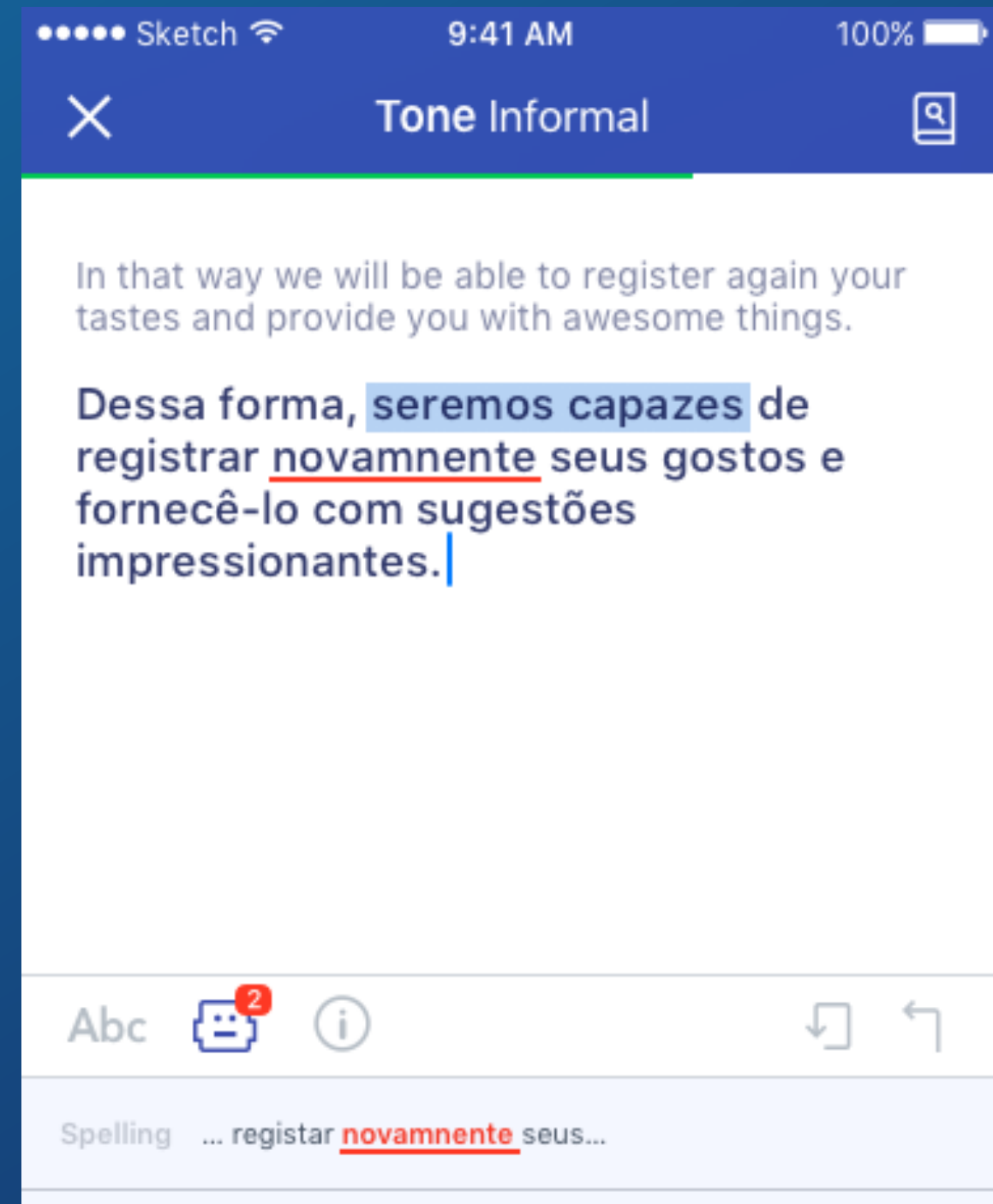
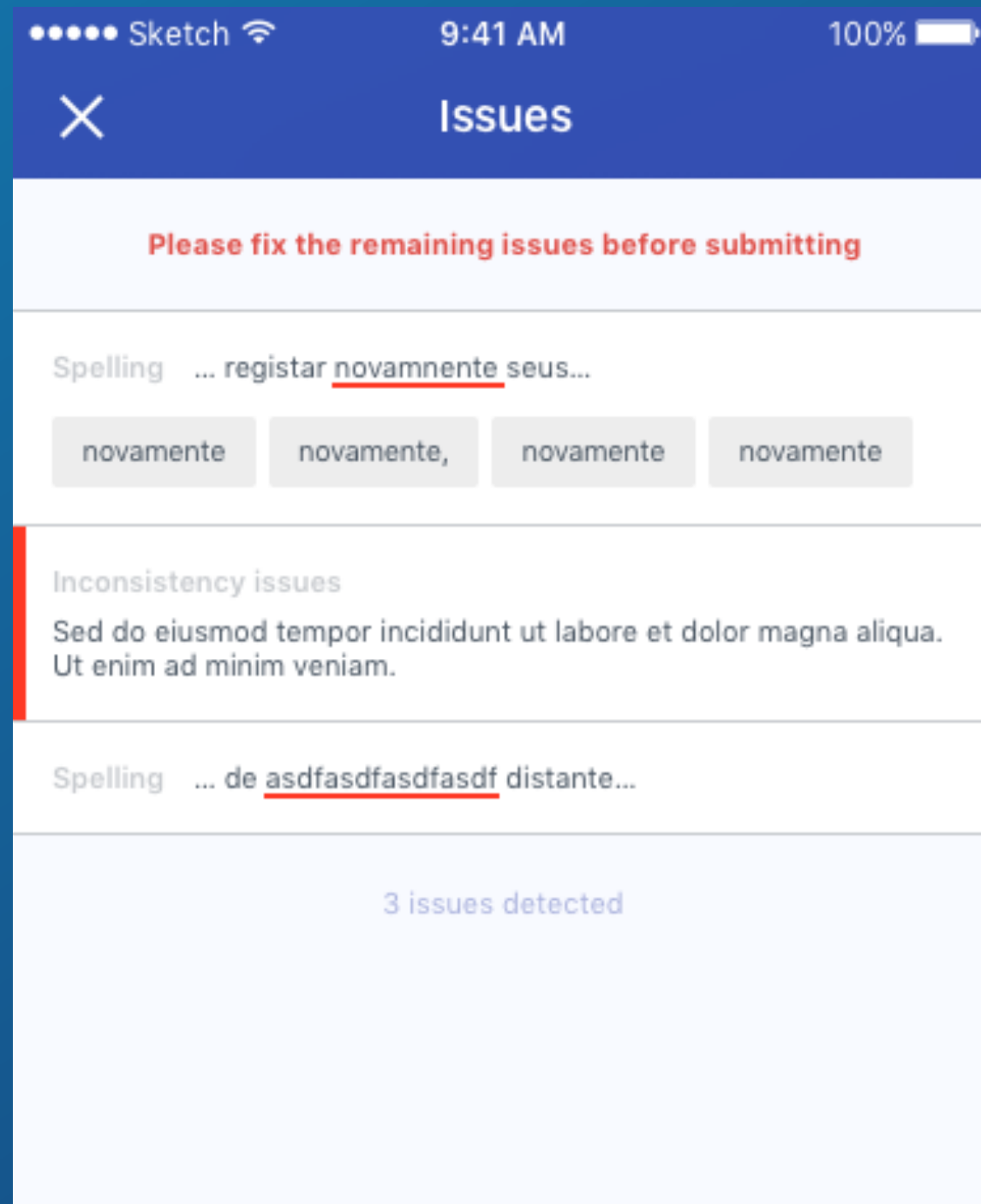


# Smartcheck





# Pos-Editing Interfaces







Unbabel

Thank you!

[helena@unbabel.com](mailto:helena@unbabel.com)