Norwegian Directorate
for Education and Training

AI-assisted QA of Moodle Translations

Tore Høgås

Senior Advisor

Department of Digital Services
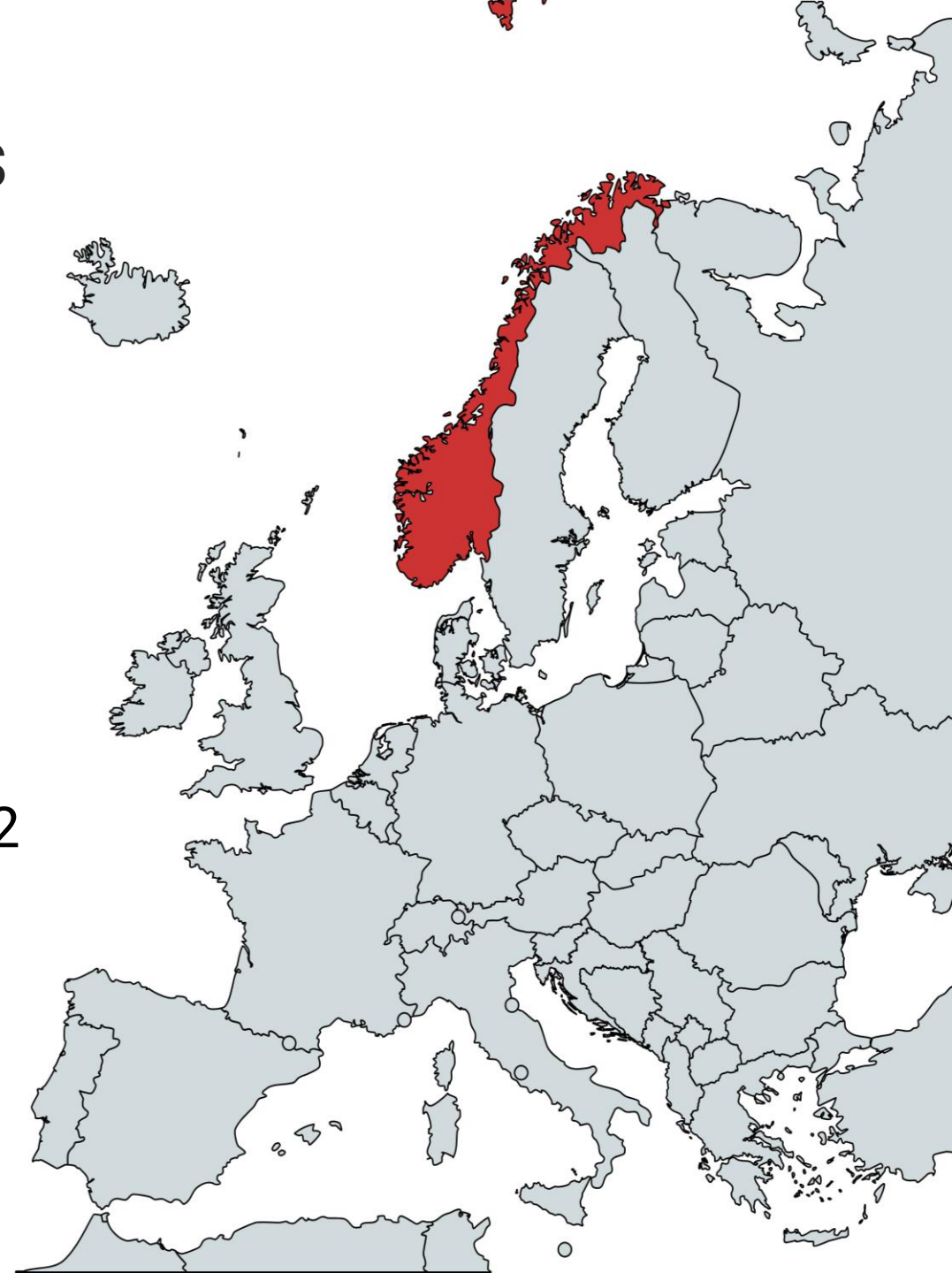
The Importance of Translations

Norway does not have many people
(ca. 5 500 000) but many languages!

Official	Norwegian	Bokmål (nb, nob)
		Nynorsk (nn, nno)
	Sami	North Sami (se, sme)
		Lule Sami (smj)
		South Sami (sma)
Minority	Kven (fkv)	
	Romani (rmg)	
	Romanes (rom)	

12

Arabic (ar, ara)   Ukranian (uk, ukr)
Somali (so, som)
Tigrinya (ti, tir)

digilær.no

The Challenge

- Moodle language packs are crowdsourced in the AMOS (Automated Manipulation Of Strings) Repositories.
  Anyone may contribute a translation or suggest an improved translation

- Language pack **maintainers** approve contributions.

- The **quality** of translations varies a lot. For example:
  - Typos
    (including strings that break placeholders and other code)
  - Lack of context sensitivity
    (words that are valid Norwegian, but not in that context)
  - Incorrect Norwegian syntax
    (using English syntax with Norwegian words)

The Solution?

Starting with Norwegian, can we:

- **Reduce** the need for manual review of strings…
- Pinpoint **where** manual review is most needed…
- **Help** translators and maintainers with manual review…

…with AI?
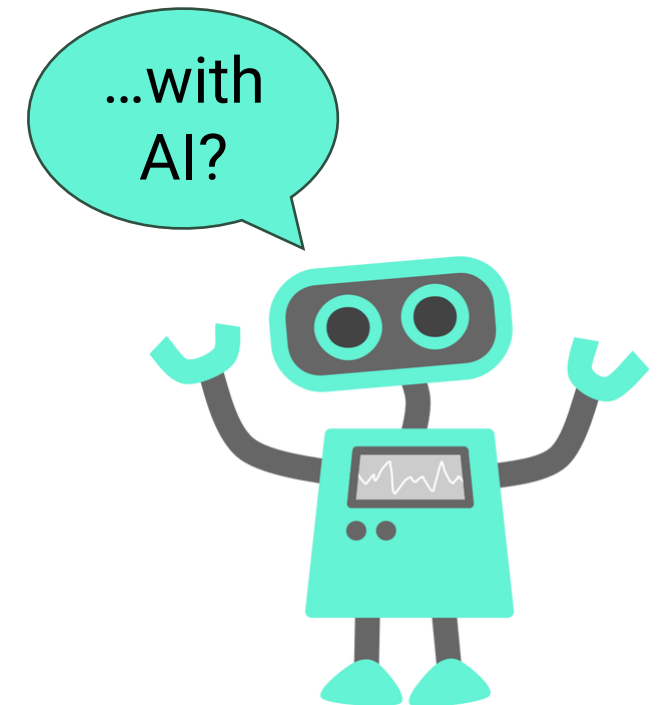
Test 1: text-davinci-003

- Command line interface with the ChatGPT 3.5 model *text-davinci-003*

- Strings were only evaluated on whether they were a valid translation or not. (True/False)

```php
*/
private static function isStringInLanguage(string $string, string $lang,
    if ($wlv->isWhitelisted($string)) {
        return true;
    }
    if ($wlk->isWhitelisted($filename, $key)) {
        return true;
    }

    $p = "Evaluate if the input below is in $lang or not." .
        "INPUT START\n\n" .
        $string .
        "\n\nINPUT END\n\n" .
        "Output (true/false):\n";

    $res = $openAI->completeText($p, 10);
    $res = strtolower($res);
    $res = str_replace(" ", "", $res);
    $res = str_replace("\n", "", $res);
    $res = str_replace("\r", "", $res);
    $res = str_replace("\t", "", $res);
    return $res === 'true';
}
```

Results
Test 1

- 10% of strings considered invalid, which still leaves **ca 4200** strings for manual review

- Many false positives

Test 2: gpt-3.5-turbo-0301

- API connection to the ChatGPT 3.5 model *gpt-3.5-turbo-0301*
- 3 roles in the chat:

SYSTEM
Initiates the chat and defines its context

```
$messages[] = new OpenAIChatMessage(
    text: "The user is working on a Norwegian Bokmål translation of a web application used by middle school
    students in Norway. You are a helpful assistant that, for each of the users messages, responds with
    an integer between 1 and 100 indicating the quality of the translation, where 1 is worst an 100 is best.
    You will not output anything other than this number. Keep in mind that the users of this web application
    are middle school students. You will use the Norwegian dictionary to ensure correct spelling."
    , role: OpenAIChatMessage::ROLE_SYSTEM);
```
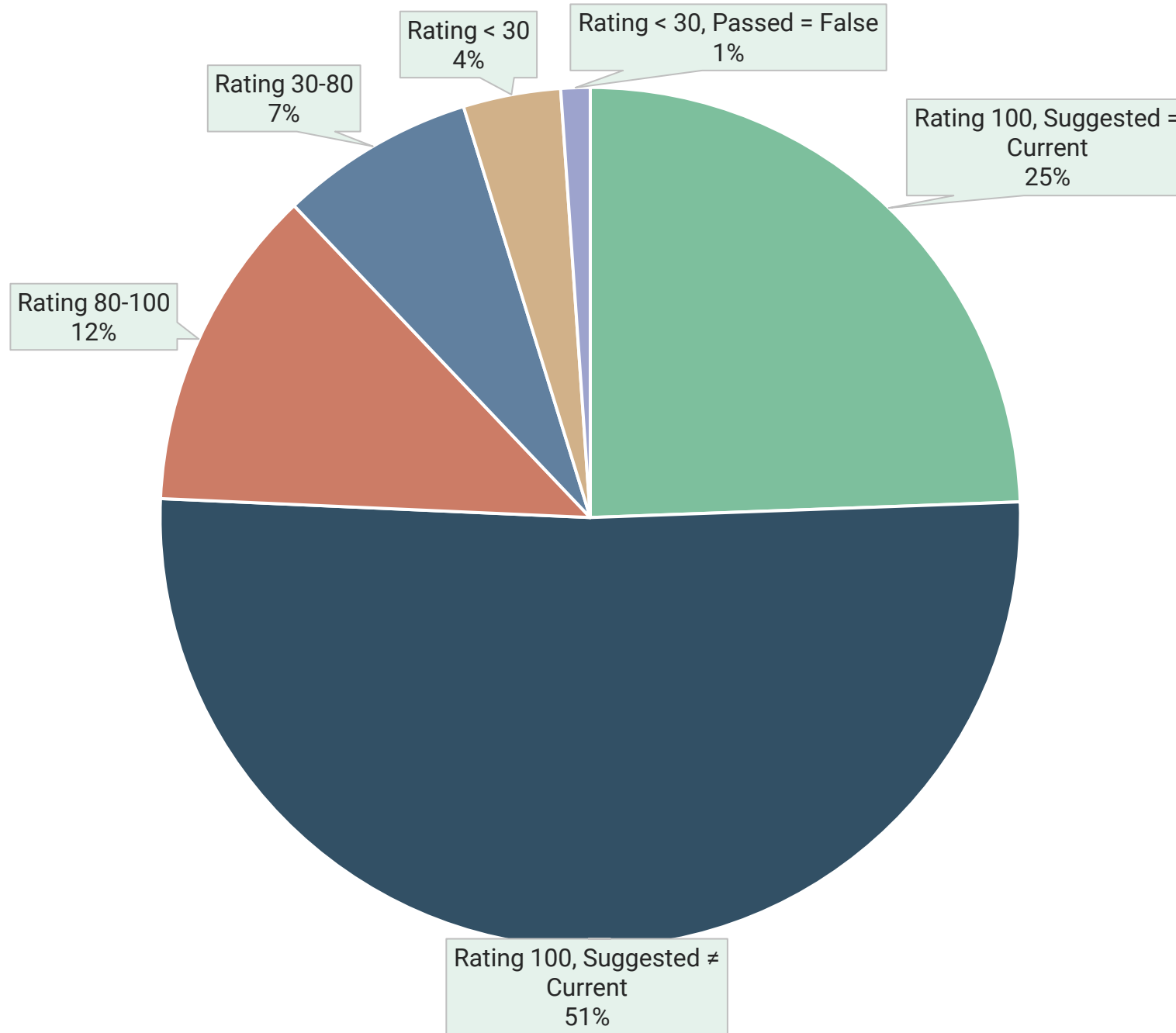
Test 2: gpt-3.5-turbo-0301

- Output per string in the language pack:

| Input	Process	Output
Current Norwegian translation	Is this a valid translation or not?	Passed = True/False
Current Norwegian translation	Rate the quality of the translation	Quality = n (1-100)
Original English string	Suggest a translation of the string into Norwegian	Suggested Norwegian translation

Results
Test 2

Rating < 30
4%

Rating < 30, Passed = False
1%

Rating 30-80
7%

Rating 100, Suggested =
Current
25%

Rating 80-100
12%

Rating 100, Suggested ≠
Current
51%

Results
Test 2

- **Reduce** the need for manual review of strings: ✔
- Pinpoint **where** manual review is most needed: ✔
- **Help** translators and maintainers with manual review: ✔

START HERE

Rating < 30, Passed = False
1%

Rating < 30
4%

Rating 30-80
7%

Rating 80-100
12%
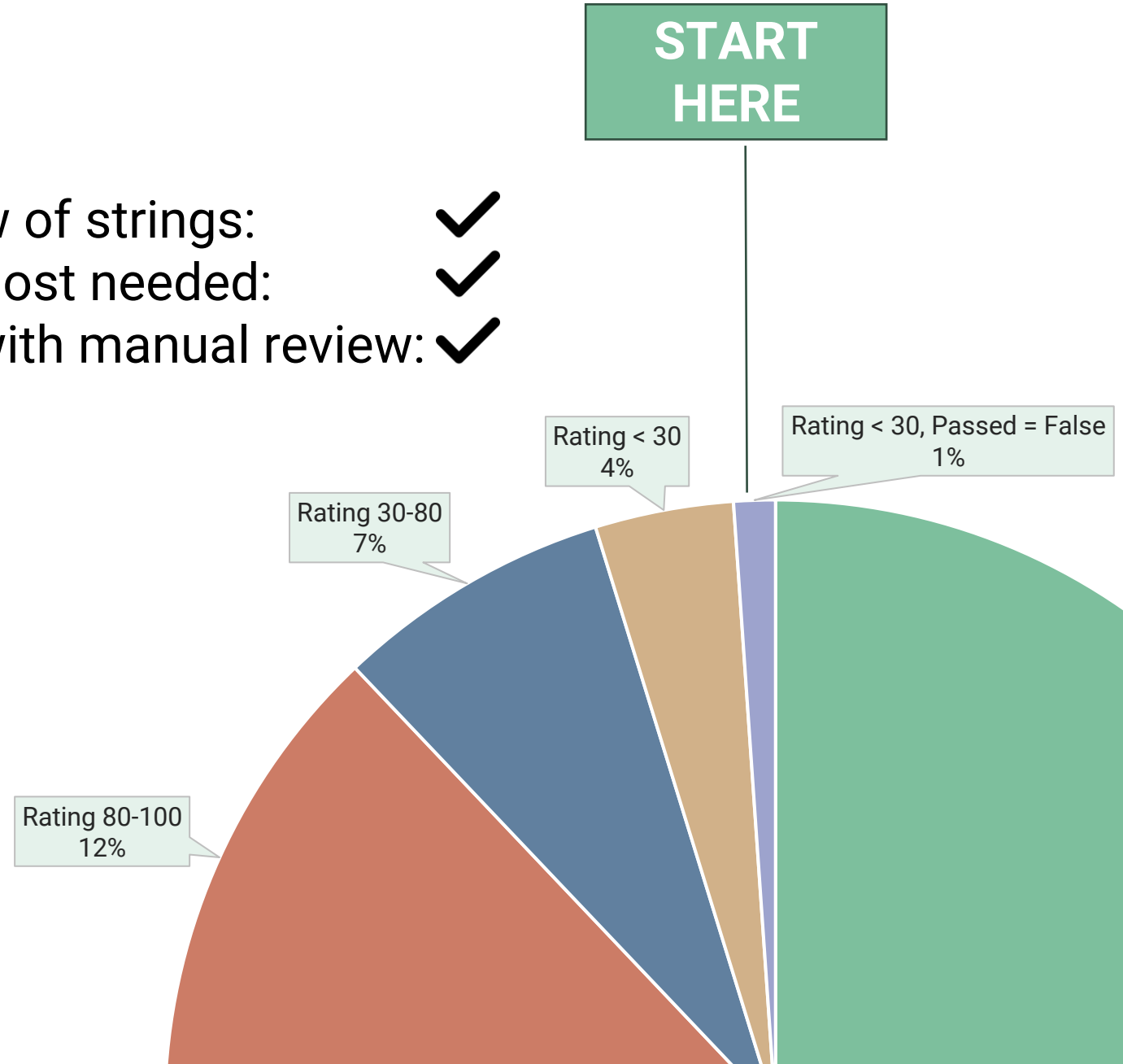
Going forward

- Product improvement potential:
  - Automatic QA checks on new strings
  - Automatic correction of some strings, manual of others
  - Output reasons for quality rating (Explainable AI)
  - New AI models
- Product dissemination potential:
  - Moodle plugin for local check of language strings
  - Integration with AMOS to assist package maintainers
  - **Better translations, easier and faster, for all languages!**

moodlemoot
GLOBAL 2023

moodlemoot.org | #MootGlobal23

E: **tore.hogas@udir.no**   W: **udir.no/in-english**