

# Survey – WG1 & 3 ENeL Online Dictionaries, Dictionary Portals and their Users

Carole Tiberius, Bob Boelhouwer,

Robert Lew, Simon Krek

Carlos Valcárcel Riveiro, María José Dominguez Vazquez


# The value of a work must be estimated by its use

It is not enough that a dictionary delights the critick, unless, at the same time, it instructs the learner; as it is to little purpose that an engine amuses the philosopher by the subtilty of its mechanism, if it requires so much knowledge in its application as to be of no advantage to the common workman.

(Johnson, The Plan of an English Dictionary, 1747)




# Purpose of the survey

- ▶ to learn about user research in the different lexicographical projects that are represented in ENeL (WG1 and WG3).
  - ▶ to find out what methods are used to study user behaviour
  - ▶ which types of user contributions are used in the different projects.
- 

# Sub themes

- ▶ **Methods to analyse dictionary use and dictionary portal use and their users:**
  - What kind of methods are applied towards this goal by the projects within the WG3 and WG1?
  - Application of the results in the development of dictionaries/dictionary portals.
- ▶ **Contexts of Dictionary Use / Contexts of the Use of Dictionary Portals**
  - What kind of methods are applied in the projects of WG1 and WG3 to get more information on users knowledge about dictionaries/dictionary portals?
  - Application of the results in the development of dictionaries/dictionary portals.
- ▶ **User contributions to Online Dictionaries and Dictionary Portals**
  - To what extent do dictionaries/dictionary portals of WG1 and WG3 allow the user to get involved in their development?

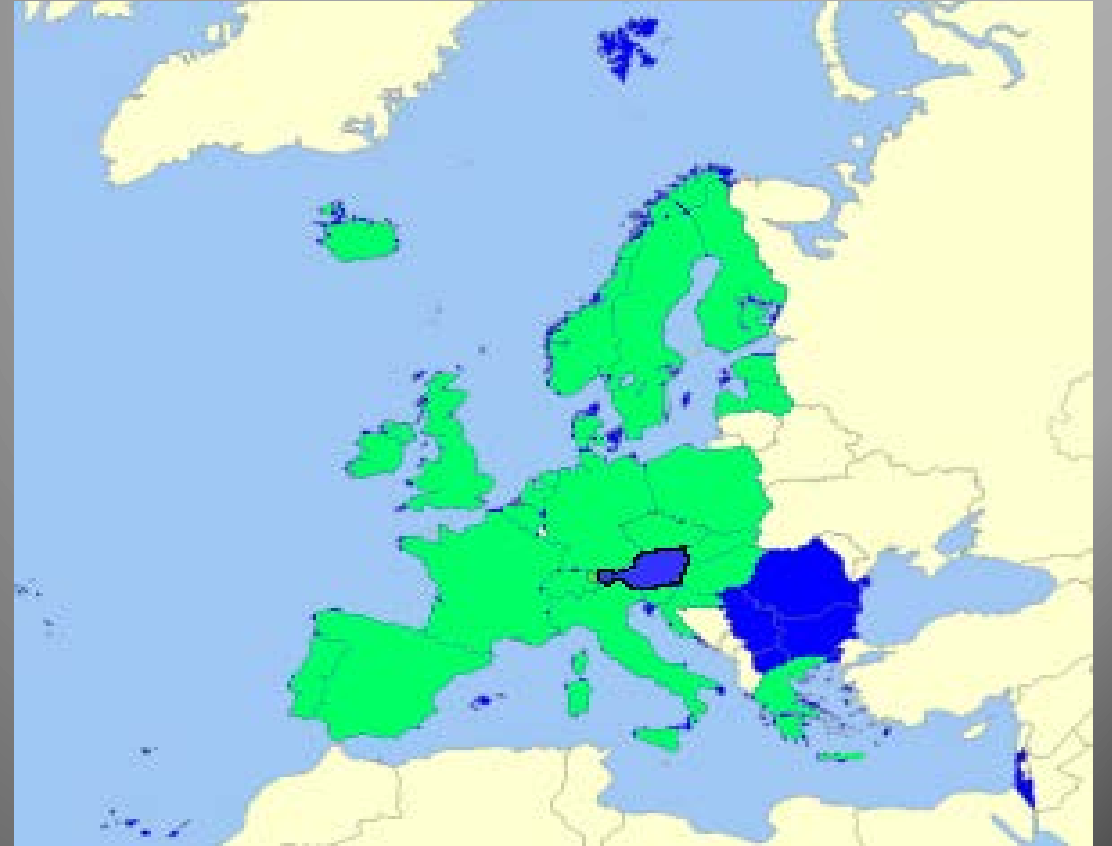
# General

- ▶ Web address: <https://www.1ka.si/a/75373>
  - ▶ Questions: 53 (variables: 66)
  - ▶ Pages: 9
  - ▶ Completed: 47
  - ▶ Partially completed: 9
  - ▶ Total valid: 56
  - ▶ All units in database: 203
  - ▶ First entry: 12.11.2015, Last entry: 12.2.2016
- 

# General comments

- ▶ see q43 with the explanation of q42 for which neither of the proposed options are relevant for us.
- ▶ \"not applicable / i do not know\" option is missing in several questions ;.-)

# Coverage



# Positions

lexicographer

(postdoc) researcher

software developer; IT

computational linguist

Terminologist/terminographer

(associate) professor

project manager/director/head-of-institute

Editor-in-chief

phd student

Corpora and lexical data administrator

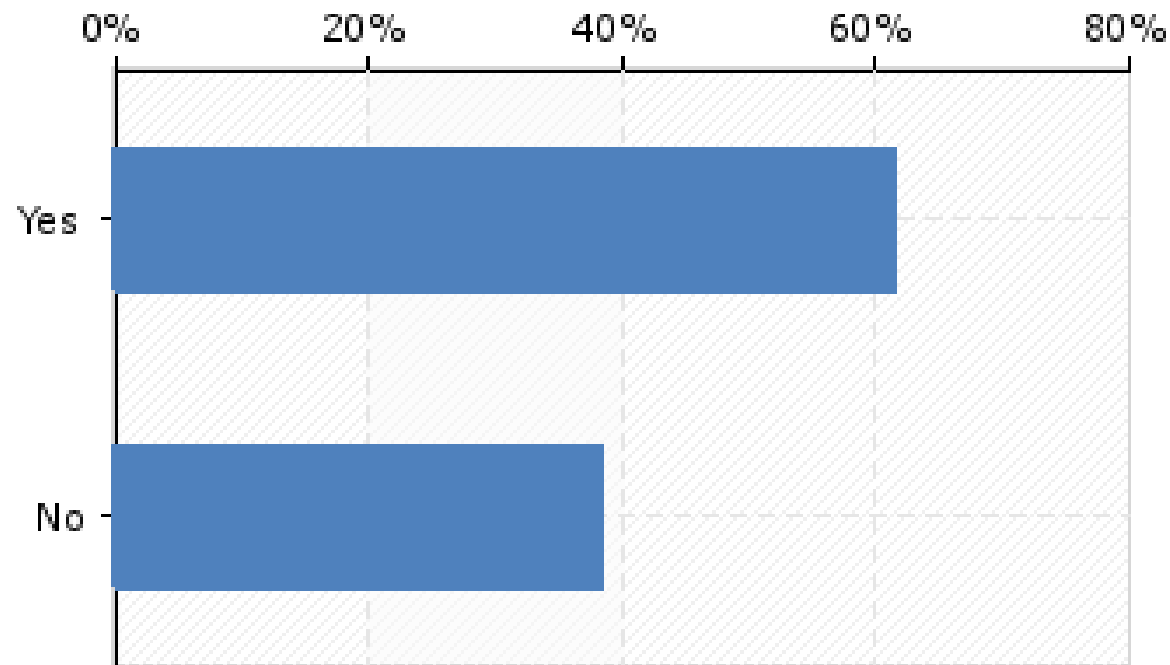
Semantic web developer



# Methods to analyse dictionary use

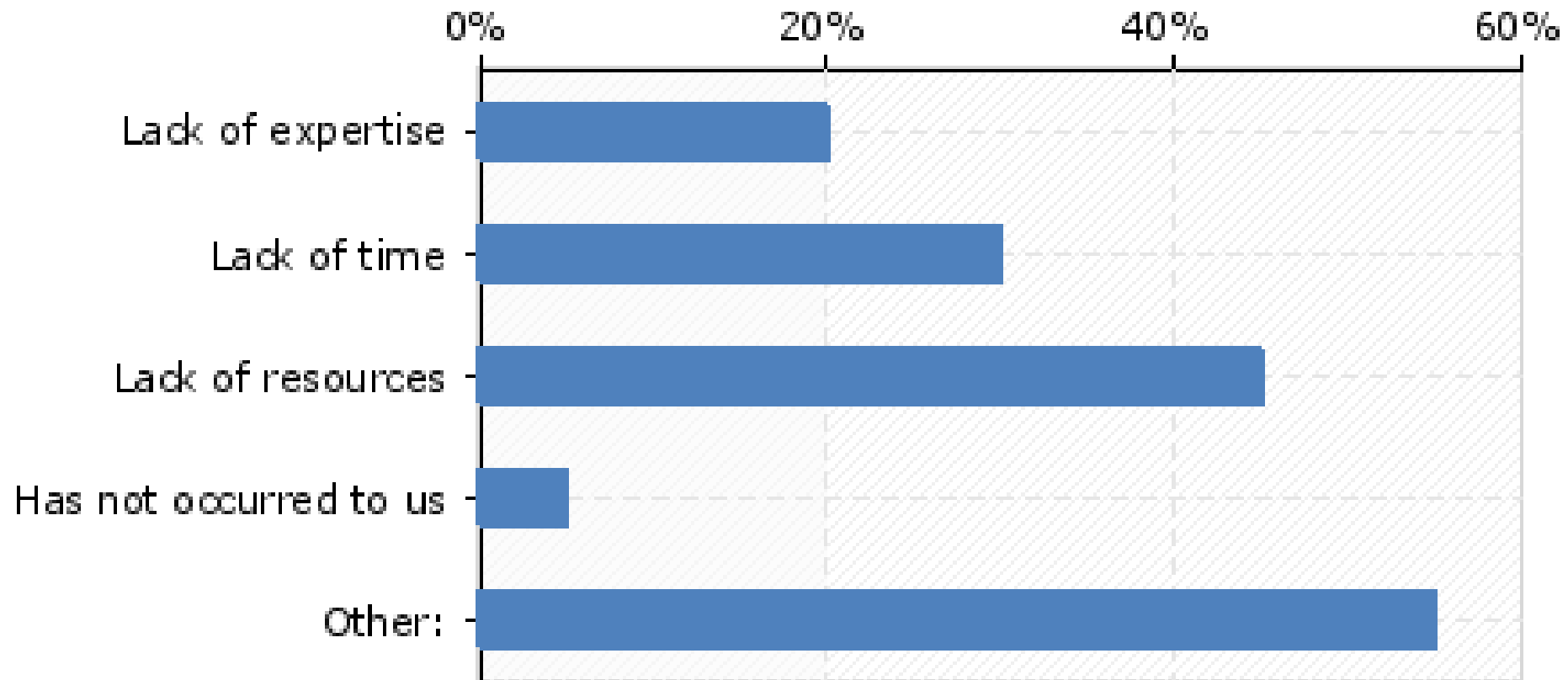
➤➤ Survey Part I

Do you use any kind of method to track/study user behaviour in your dictionary or have you used any such method in the past?(n = 52)



# Please describe why you have not used/are not using user studies. (n = 20)

---

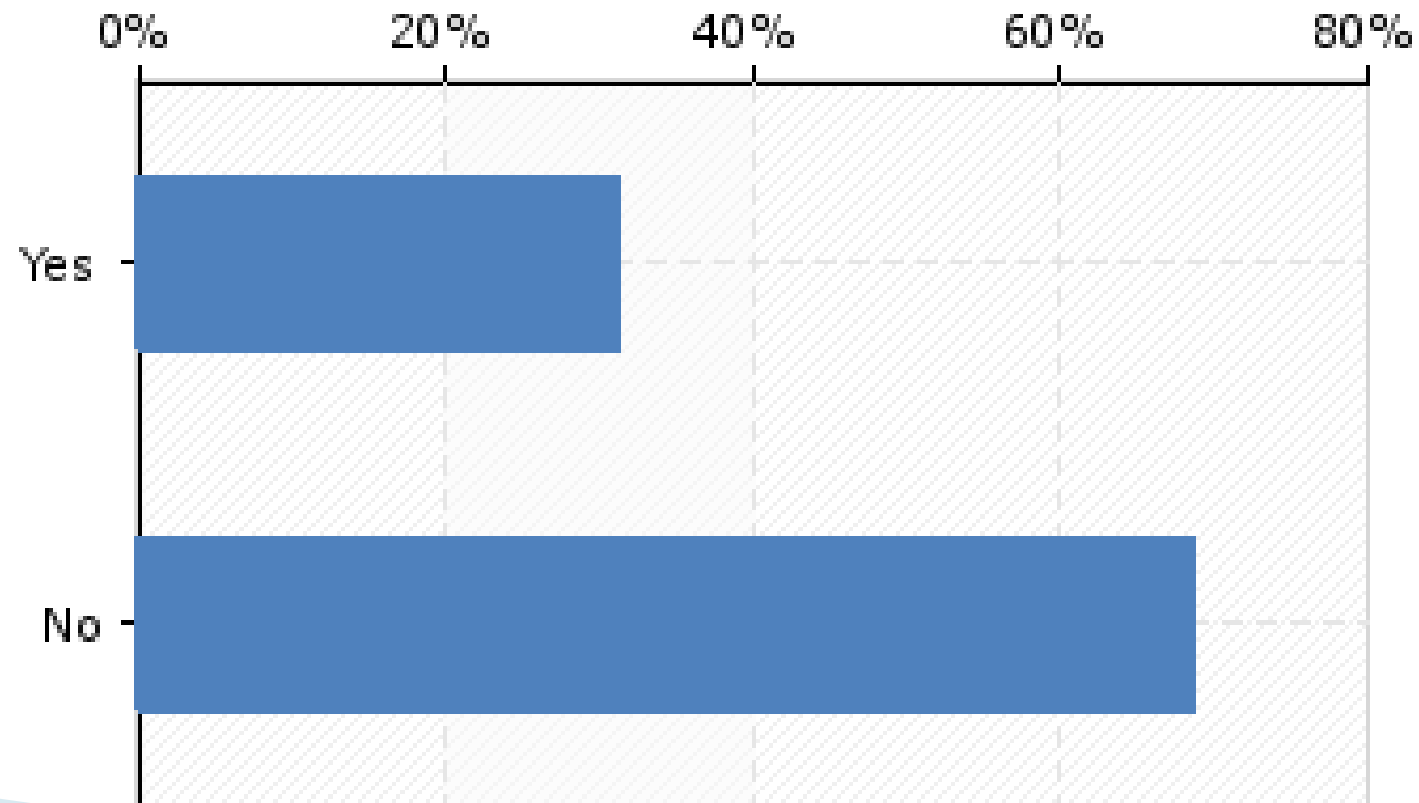


# Other:


- ▶ the repository isn't open for public yet; we haven't had online version yet; **not open for public**, opening planned for april 2016 ; we are not yet in that phase: we are first building up the database and collecting content.; do not have a portal yet, we are in the process of developing a portal for the mwes lexicographic tool ; project is **still in development phase**
- ▶ the target group, namely, immigrants in greece, had not been targeted before so as to use previous studies. moreover, under the circumstances, it was not an easy and feasible task to contact the target group (who were/are mainly struggling for a living) and give them questionnaires. so we left this task of user tracking/studying to be carried out at a later stage, after some pilot version of the dictionaries were available online.
- ▶ **legal problems ; privacy concerns** re using log files
- ▶ the linghub portal is **not directly a dictionary portal** but an aggregated set of metadata that can lead to the relevant language resources in (possibly distributed) dictionary resources ; our resource is not a dictionary, but a system to create dictionaries; we obtain feedback from our users

# Do you use a form of questioning (i.e. written questionnaires or spoken interviews) to study user behaviour? (n = 32)

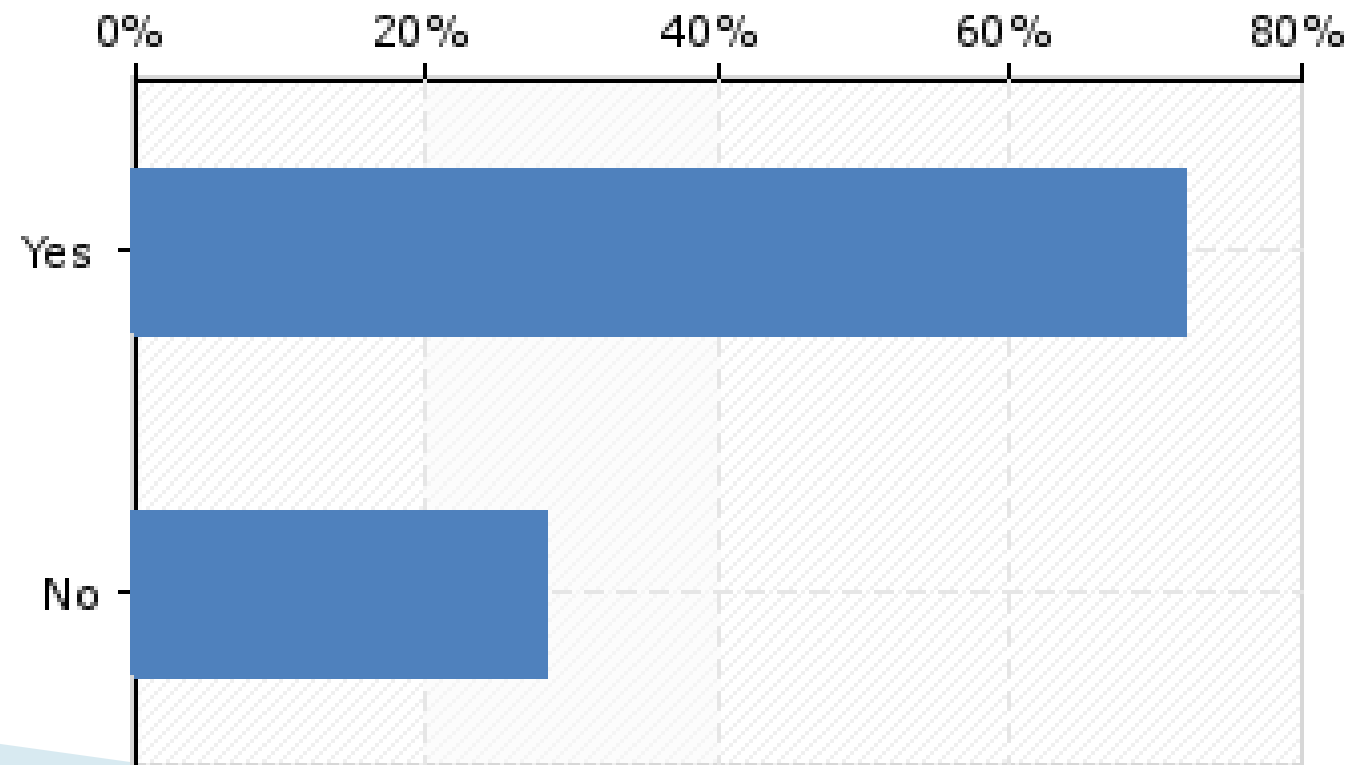
---



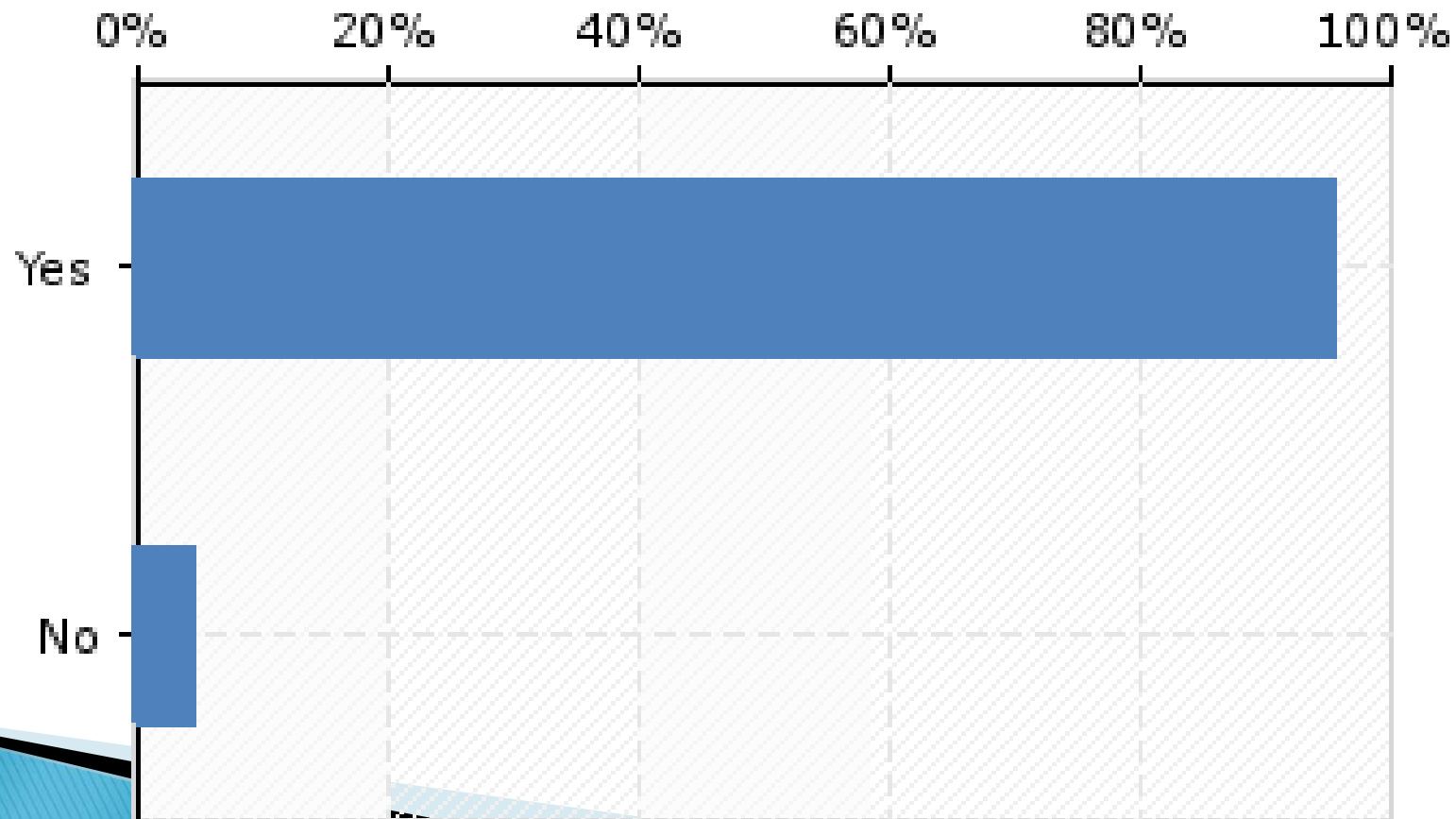
# Software used:

- ▶ Unipark
  - ▶ Surveymonkey
  - ▶ Limesurvey
  - ▶ Sphinx
  - ▶ Custom built
- 

Do you use a form of observation (e.g. thinking aloud, protocol sheets, log files, eye-tracking, video recording, screen recordings) to study user behaviour?(n = 32)




# Do you use log files?(n = 23)






# Software

- ▶ Google Analytics
  - ▶ Awstats
  - ▶ Piwik
  - ▶ Proprietary / custom
- 

# What is logged?

- ▶ everything that Google Analytics logs
- ▶ things users search for and don't find. / no-match queries
- ▶ a) Where do the users click. b) How many times do they click on certain buttons.
- ▶ Google Analytics data (period of time: 2014/10/31–2015/10/31)  
Sessions: 1.478.779 Users: 380.724 Page views: 7.587.698 Search queries: 14.020.466 Where do the visits come from? – Mobile devices (tablet, phone): 15.28 % of the visits – From search-engines: 13.68 % of the visits Where and how many times do the users click? – ‘Did you mean’ results: 148.611 – Best proposal of ‘Did you mean’ (underlined): 117.324 – Last changes: 6.336 – Popular entries: 2.761 – Word of the day: 1.976

# Analysis of search queries (14 out of 22)

- ▶ Analyses of no-match queries
  - ▶ Analysis of search patterns
  - ▶ Analysis of words searched for and corpus frequency; number of senses
  - ▶ Analysis of performance issues
- 


# Caution

- ▶ Your data in Google Analytics may not be as accurate as you think. If you have a high volume of visits, your data could easily be off by 10–80%, or even more. **Shocking right?**
- ▶ (<http://www.blastam.com/blog/index.php/2013/02/can-you-trust-your-google-analytics-data>)
- ▶ The **primary reason for sampled data** is that your selected date range has more than 500k visits.

# Eye-tracking

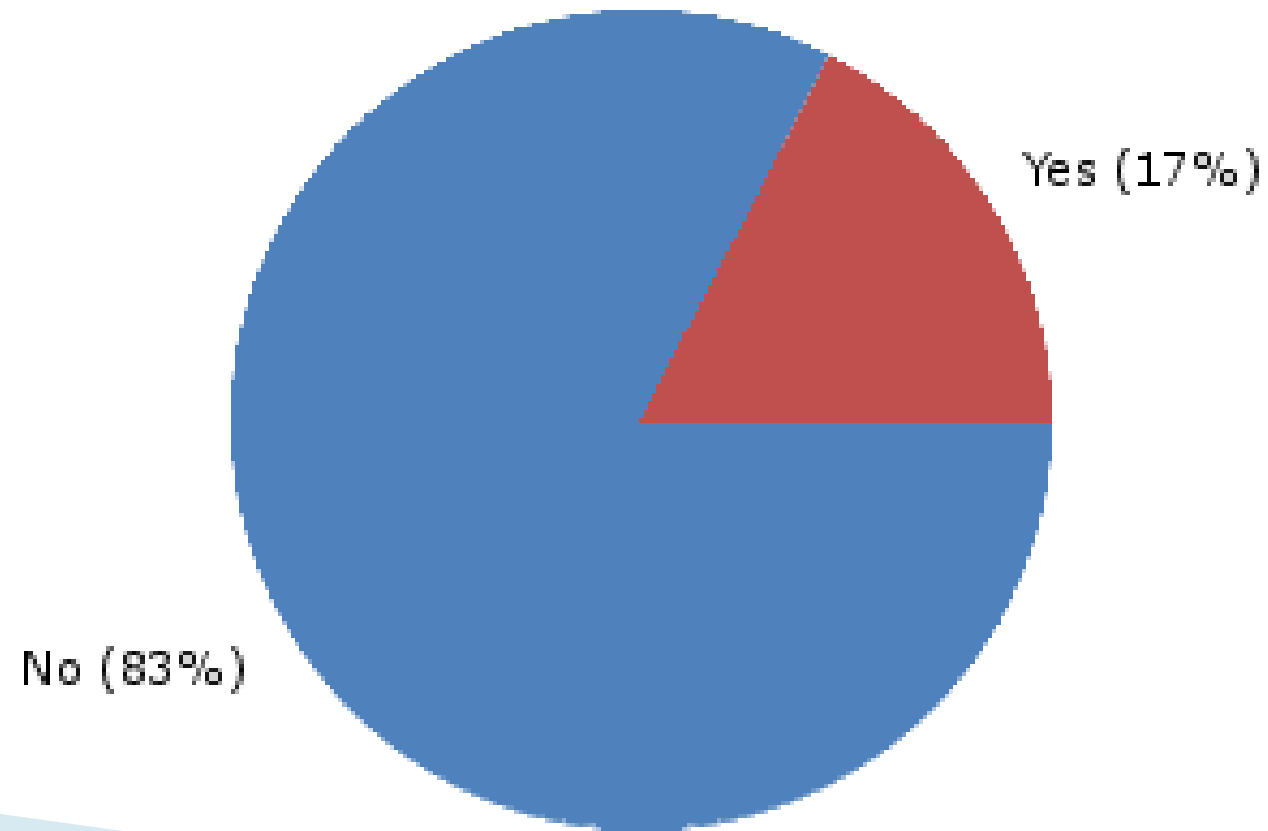
- ▶ Only 1
- ▶ See müller–spitzer, carolin/michaelis, frank/koplenig, alexander: evaluation of a new web design for the dictionary portal owid. in: müller–spitzer, carolin (hrsg.): using online dictionaries. s. 207–228 – berlin/new york: de gruyter, 2014. (lexicographica: series maior 145)

# Screen recorders

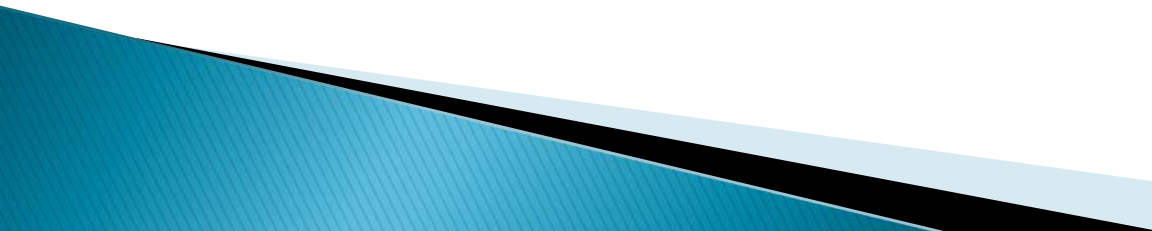
- ▶ Only 1 (Sketch Engine)
  - ▶ we tried: – open web analytics – mousestats – decibel insight – sessioncam – inspectlet – usabilitytools (which we considered best, but still this is in testing)
  - ▶ it worked very well in testing but we haven't yet manage to set it up in production, hopefully we will have things to report by mid 2016
- 

# Forms of self-observation

---

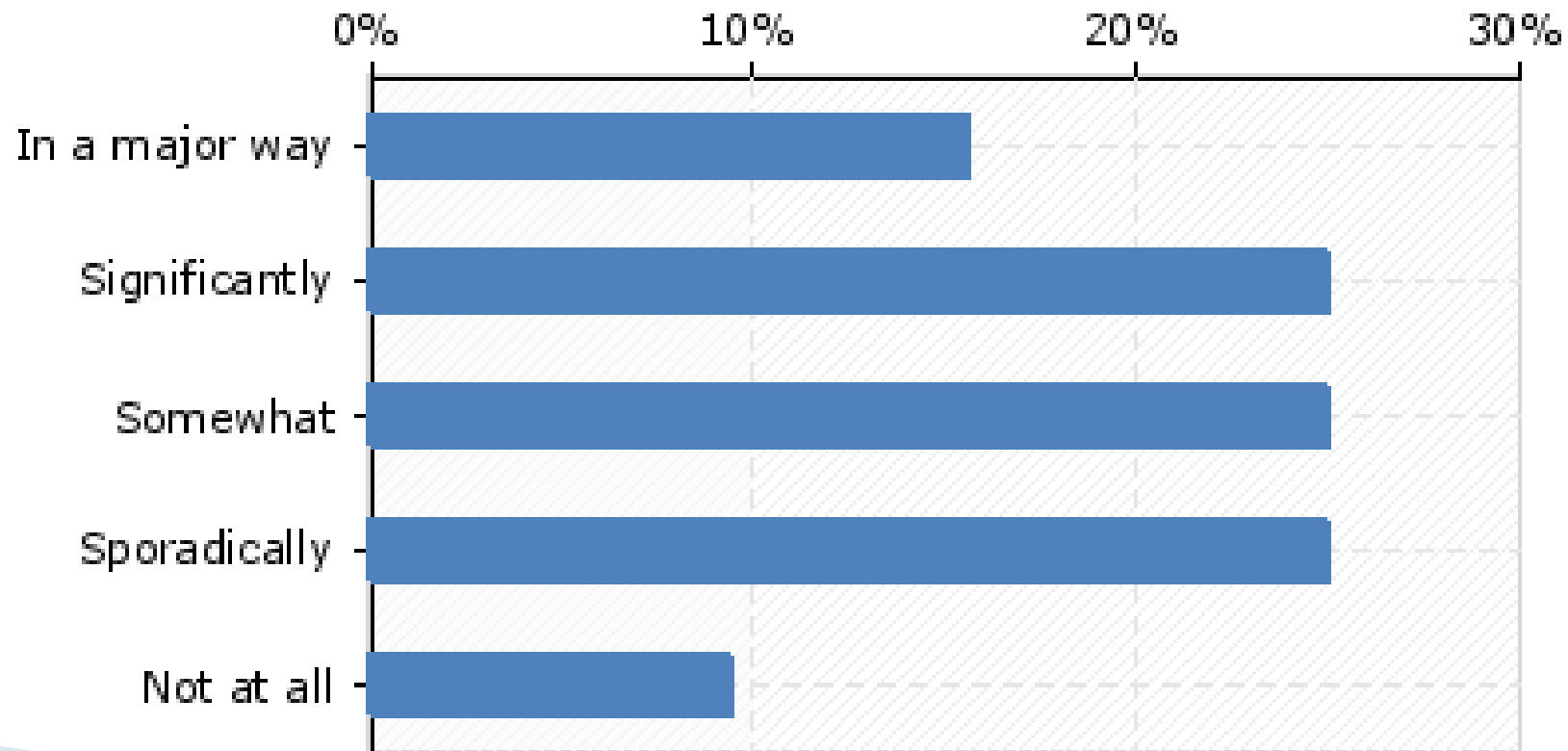


# Forms of self-observation

- ▶ Thinking-aloud interviews
  - ▶ Personas scripts
  - ▶ Commentaries on dictionary use
  - ▶ We tried to analyze the discourse of the users without directly asking them: first approach was discourse analysis of self-managed facebook group for slovenian translators, the second one focused on the already existed user-reported language problems
- 



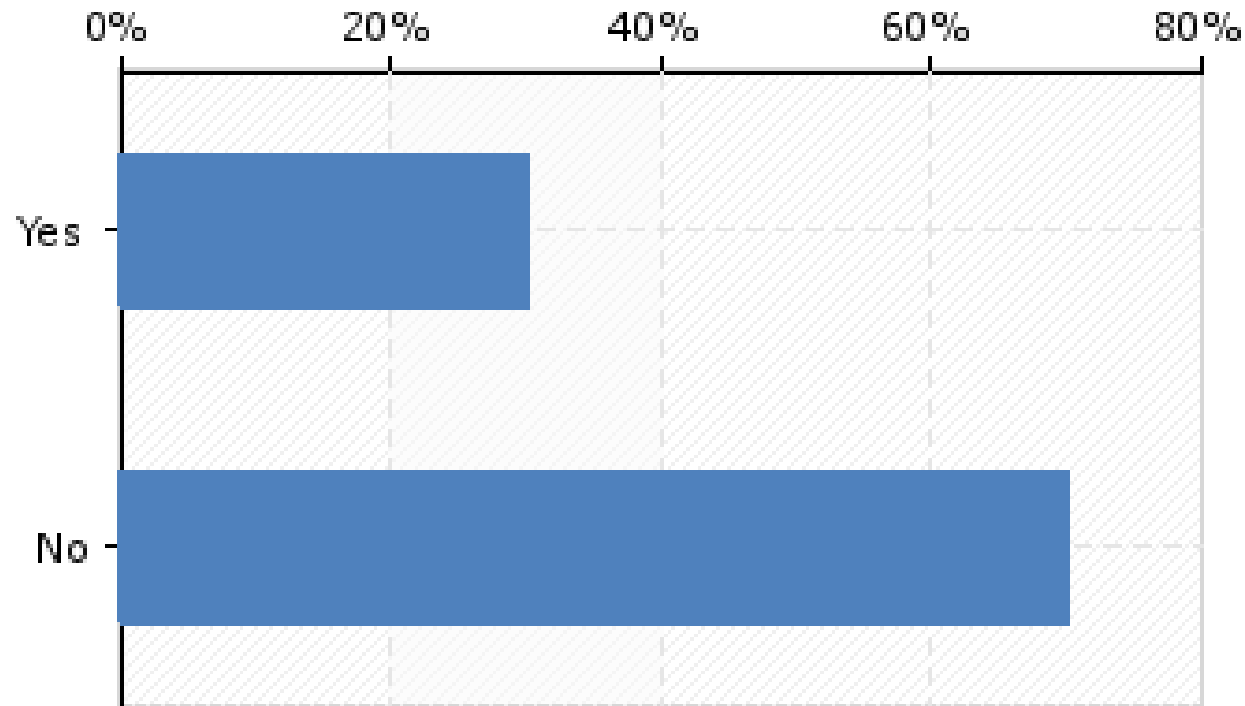
# To what extent have the results of your user studies shaped the development of your dictionary/dictionary portal?



# Contexts of Use

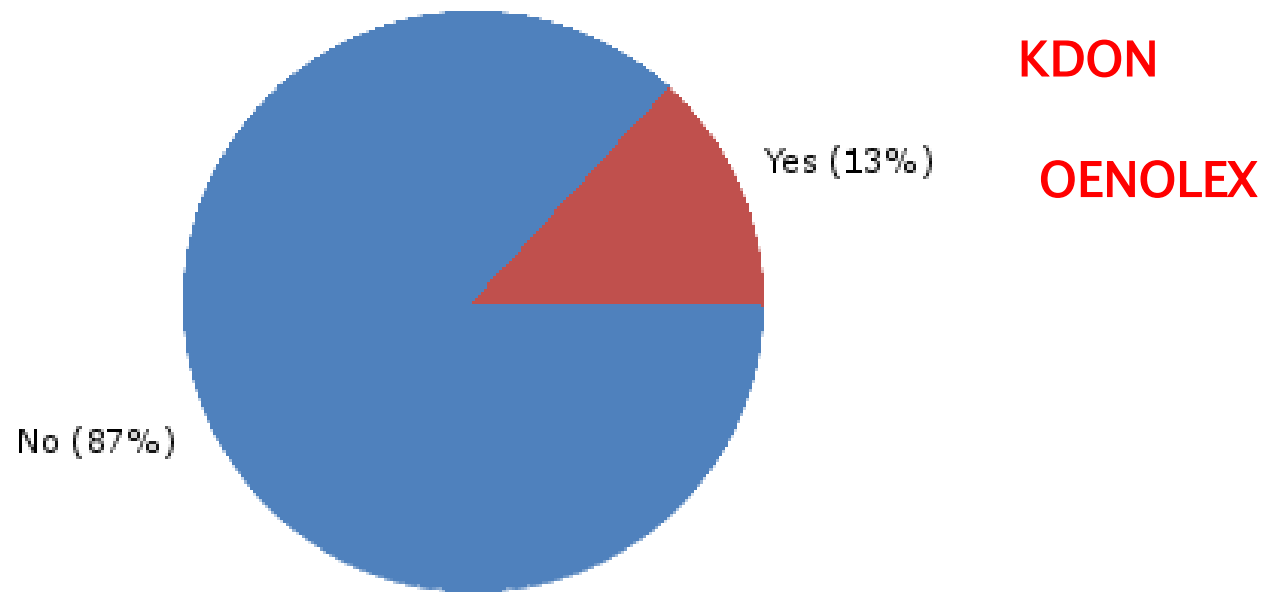
➤➤ Survey Part II

# Can your dictionary (portal) be adapted to the demands of a particular activity and/or to the needs of an individual user? (n=50)

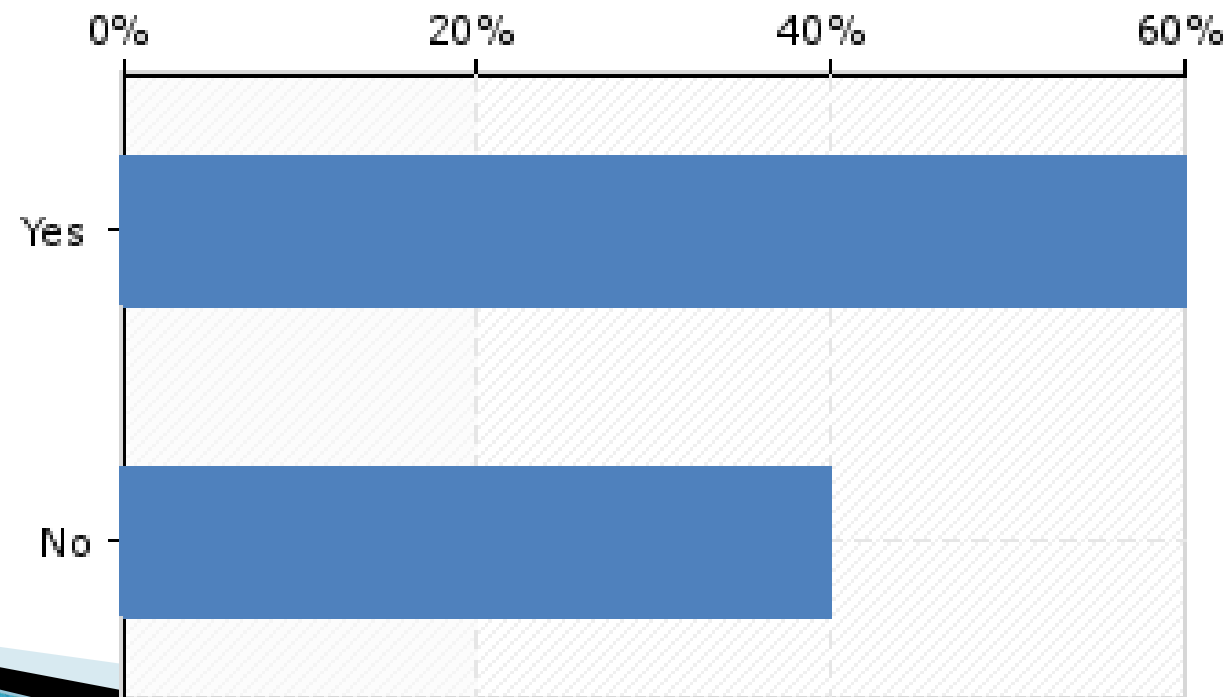


# Do you have a customized user interface ?


(to facilitate access to relevant personal information, the user interface automatically adapts to the user's preferences depending on the item classes used in the previous search requests, cf. homepages of mail-order companies such as Amazon, which change according to the user and his previous shopping preferences)



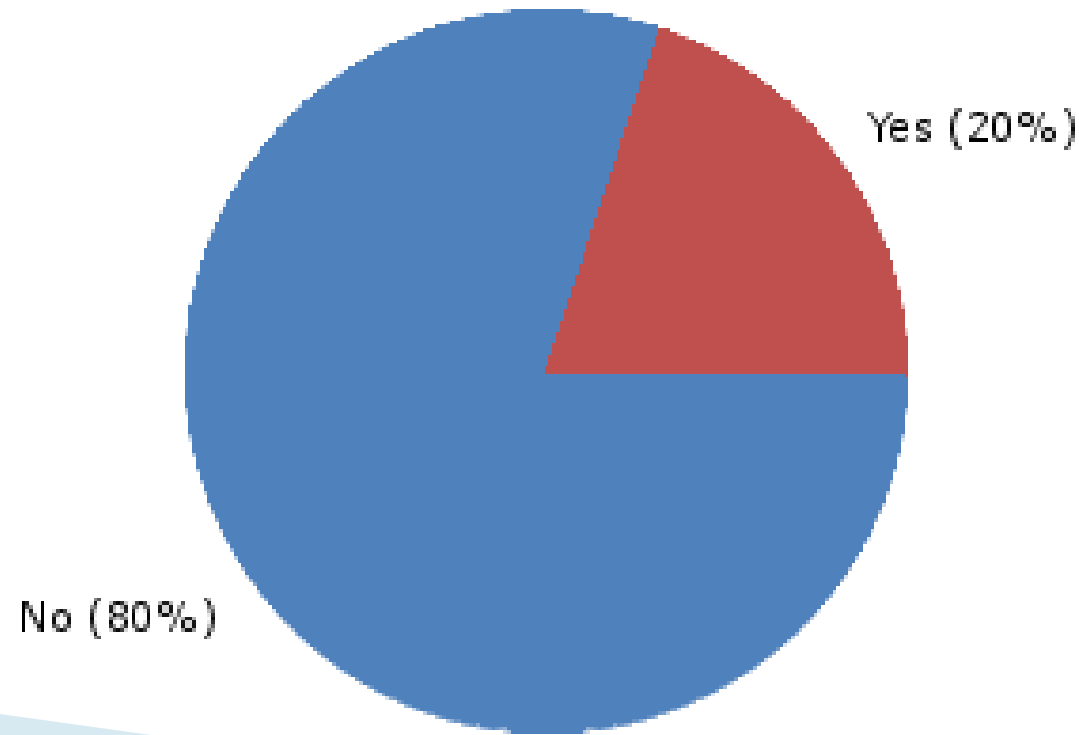
Does your dictionary (portal) allow users to create a personalized user view of the dictionary (portal), e.g. by choosing between different item classes (definitions, sense relations etc.) to be displayed? (n=15)



# Details:

- ▶ it is possible to select the language combinations if different dictionaries.
  - ▶ user can choose from several predefined views of each dictionary
  - ▶ the user can create personal lexicon, personal translations and store them, and have personal more features
  - ▶ two general presentations (long and short)
  - ▶ users can filter search results according to a range of criteria.
- 

Does your dictionary (portal) allow users to choose between different profiles that optimally adjust the content according to the user's needs? (n=15)

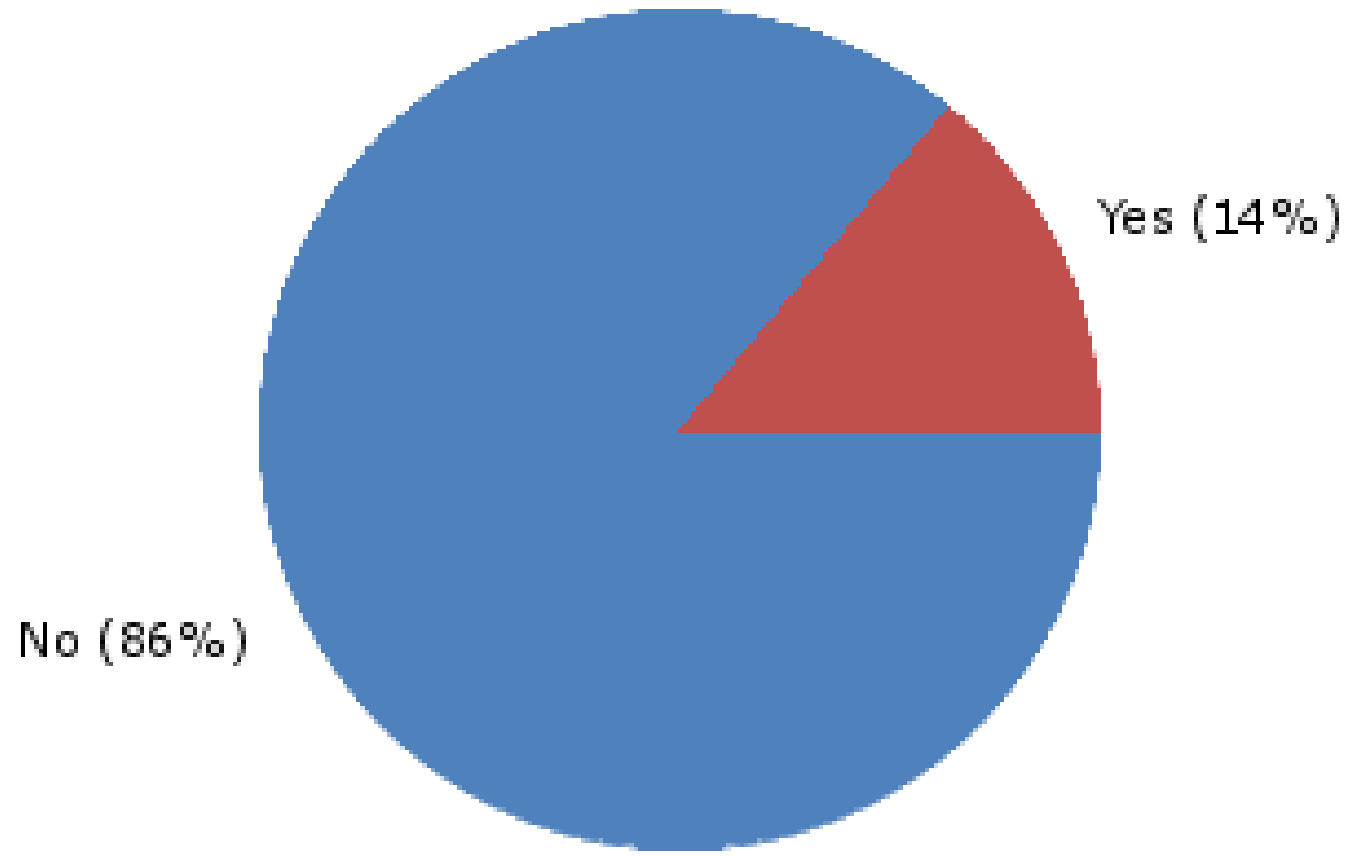


# Details


- ▶ user interface is adapted to different user roles (reader, editor...)
- ▶ filtering of search results and choice of data presentation according to personal preference
- ▶ the dictionary is planned to be customisable to meet users' profiles according to three aspects:
  - language variety (Brazilian or European Portuguese),
  - language status (speaker of Portuguese as a mother tongue or as an additional language)
  - disciplinary groups ( for now, four broad areas – as in BAWE)



# Has your project probed its users' awareness level with respect to dictionaries? (n=50)



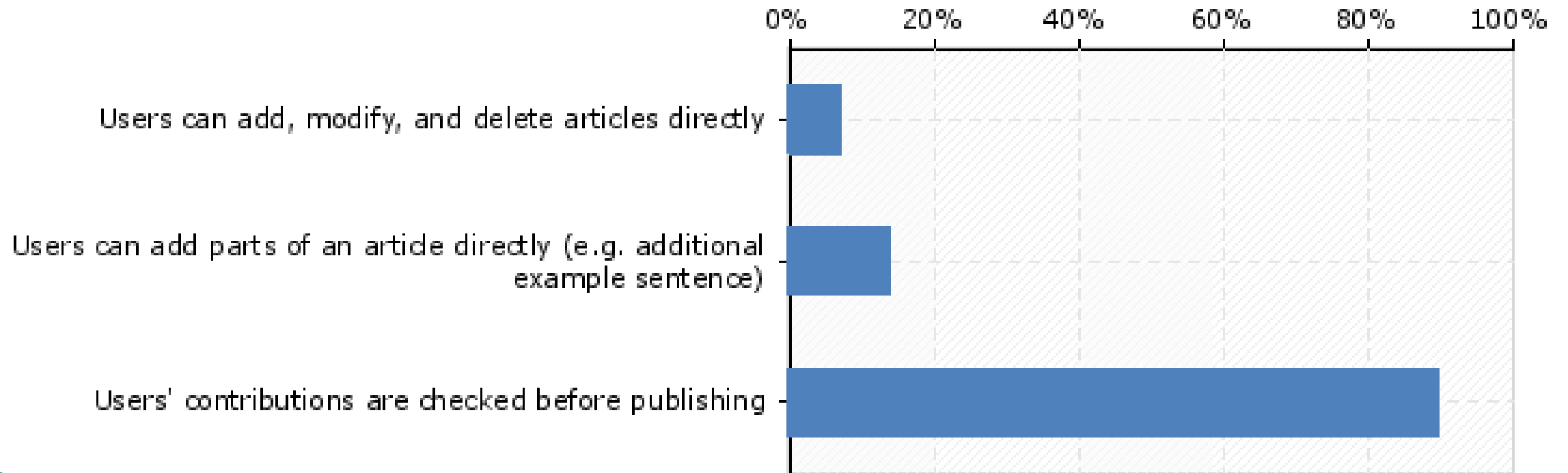
# Details

- ▶ Questionnaires / Online survey
  - ▶ Interviews
  - ▶ Meetings
  - ▶ Language games
  - ▶ Each entry is conceived as a more shortened or extended article depending on the user's level of knowledge
- 

# User contributions

➤➤ Survey Part III

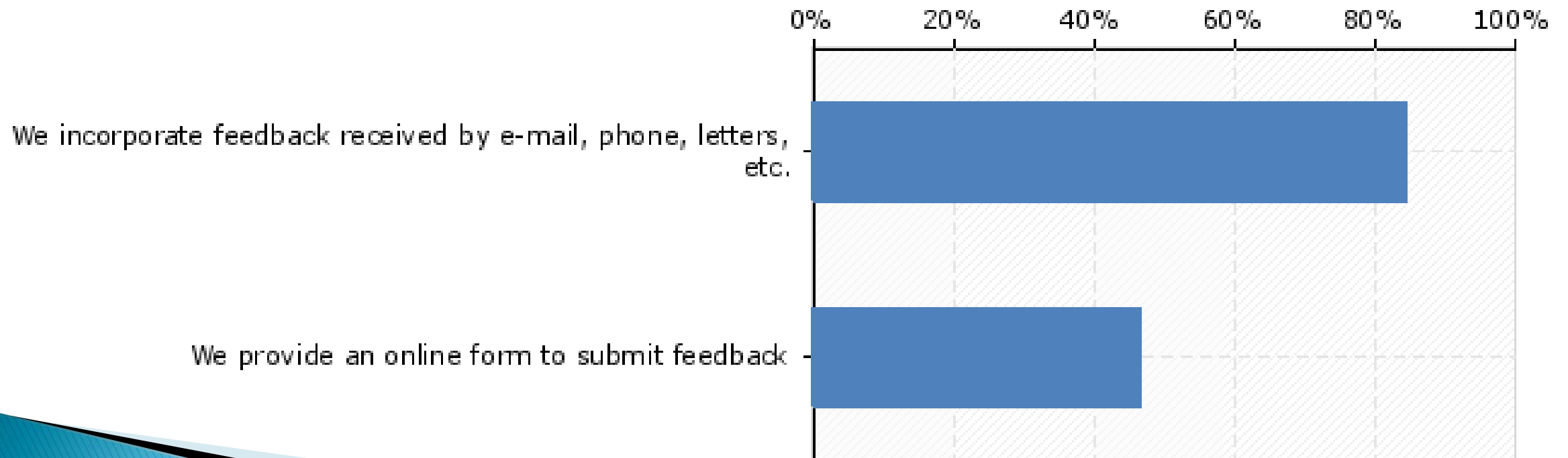
# Which of the following types of direct user contributions do you employ in your project? (n = 29)



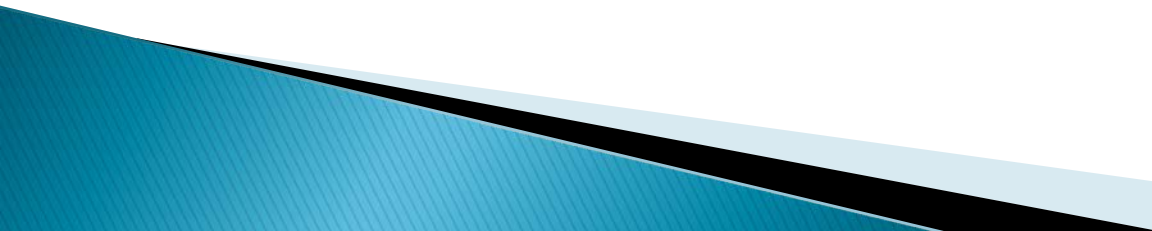
# Users' contributions are checked before publishing

User contributions are usually received through email or a web form. These contributions are thoroughly checked. Sometimes validated contributions are incorporated in the dictionaries. Sometimes they are slated for future versions.

# Which of the following types of explicit user feedback do you use in your dictionary project?(n = 45)



# Details (feedback by email, phone, letters)

- ▶ Via email
  - ▶ Via contact form
  - ▶ Social Media
    - Blog
    - Twitter
    - Facebook
  - ▶ Regular meeting with users (wine professionals)
  - ▶ (paid) telephone service
- 

# Type of feedback

- ▶ Content
  - Suggestions for definitions, words, examples
  - Language advice
- ▶ Technical
  - Problems/errors
  - Feature requests



## Contact

Your name \*

Your e-mail address \*



Wielki słownik  
języka polskiego

[Wstęp](#)

[Autorzy](#)

[Kontakt](#)

## Kontakt

Imię:

Nazwisko:

Email:

Wiadomość:

Ik ben geen robot



reCAPTCHA

[Privacy](#) - [Voorwaarden](#)

**Wyślij**

## Palautelomake

Oheisella lomakkeella voit antaa kaikenlaista SMS:ään liittyvää palautetta. Käytämme ajatuksiasi ja mielipiteitäsi sanakirjan sisällön parantamiseen ja käyttöliittymän kehittämiseen.

aihepiiri

Sanakirjan sisältö ▾

viesti

nimi

öposti

**Lähetä**

timaisten kielten keskus 2011–2016 · Tuorein päivitys: 29.2.2016

You are here: Home

## Nye ord – ja, tak!

Er du stødt på et nyt ord, et nyt udtryk eller en ny brug af et eksisterende ord? Så kan du hjælpe os ved at indberette din iagttagelse og sende den til os sammen med nogle ord om dig selv. Oplysningerne gemmes i en database som Dansk Sprognævn og Det Danske Sprog- og Litteraturselskab har oprettet i fællesskab for at følge udviklingen af det danske sprog.

Hvad forstår man ved [nye ord](#)?

- *feltes som **skal** udfyldes*



### Nyt ord ▪

### Kommentarer

Betydning, kilde, citat, dato o.lign.

### Indsendt af

#### Fornavn og efternavn ▪

#### E-mail ▪

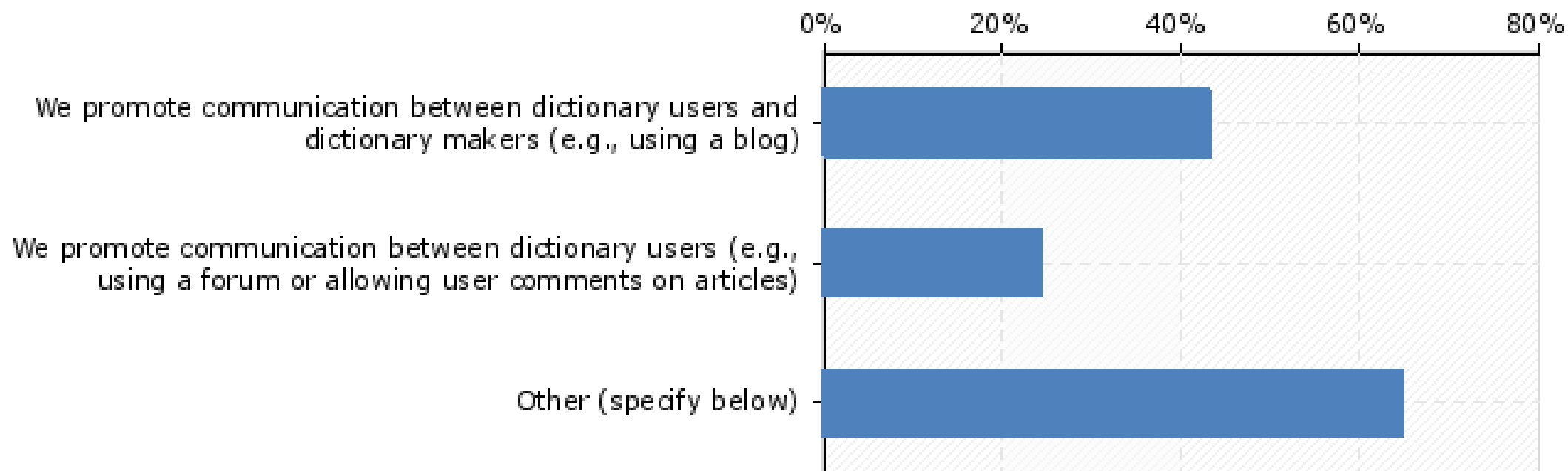
### Spamfilter

#### Skriv tallet 14 med små bogstaver ▪


# Feedback form URLs

- ▶ <http://kaino.kotus.fi/sms/?p=feedback>
- ▶ [http://www.kotus.fi/kotus/yhteystiedot/yhteydenottolomakkeet/palaute\\_kielitoimiston\\_sanakirjasta](http://www.kotus.fi/kotus/yhteystiedot/yhteydenottolomakkeet/palaute_kielitoimiston_sanakirjasta) (any feedback)
- ▶ [http://www.kotus.fi/kotus/yhteystiedot/yhteydenottolomakkeet/ilmianna\\_sana!](http://www.kotus.fi/kotus/yhteystiedot/yhteydenottolomakkeet/ilmianna_sana!) (new words)
- ▶ <http://islex.is/islex?contact=1>
- ▶ <http://wsjp.pl/index.php?pokaz=kontakt&l=1&ind=0?pwh=0>
- ▶ <http://bolje.hr/> & <http://savjetnik.ihjj.hr/contact.php>
- ▶ <http://struna.ihjj.hr/en/contact/>
- ▶ [http://hiztegiak.elhuyar.eus/sarrera\\_iruzkinak/en/available](http://hiztegiak.elhuyar.eus/sarrera_iruzkinak/en/available)
- ▶ <http://ordnet.dk/send-et-ord> (new words)
- ▶ <http://signbank.science.ru.nl/feedback/site/>
- ▶ <https://diccionariopotlex.wordpress.com/contacto/>
- ▶ <http://www.inl.nl/onderzoek-a-onderwijs/webrubrieken/melden-neologisme>
- ▶ <https://kamusi.org/contact>

# What types of interaction with dictionary users are possible in your dictionary project? (n=37)



# Details

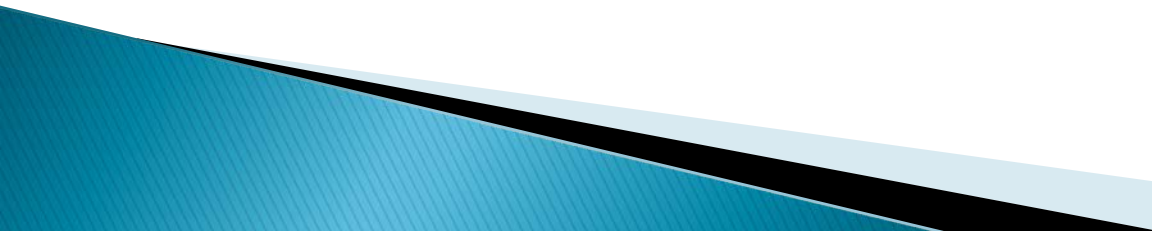
- ▶ Social media (blog, facebook)
  - ▶ Forum, online course
  - ▶ Radio broadcasts (Croatian)
  - ▶ Users can suggest Slovene translations of recent loanwords and vote for them
  - ▶ Registered users can rate GDEX examples
  - ▶ Public seminar / participating in research market
- 

# How to benefit more from user contributions

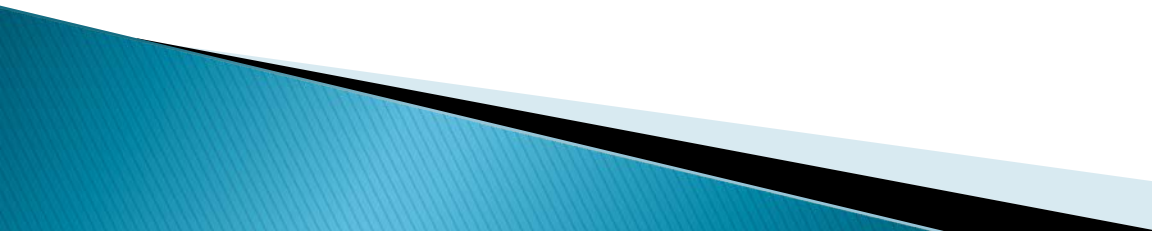
## Who to ask?

- ▶ personal contact
- ▶ It is difficult to benefit much from sporadic contributions; rather we should have voluntary user teams with expertise in several field
- ▶ user-supplied suggestions for technical terms, especially when supplied by expert users in that specific field.

## What to ask and how?

- ▶ From what we can see, a free text email is preferred to filling in forms.
  - ▶ Ask simple questions with limited set of possible replies.
  - ▶ User profile information could be added to contribution.
  - ▶ Suggestions for search options
  - ▶ Training (online course)
  - ▶ Expert gatherings
  - ▶ Comment management system like on Youtube, maybe?
- 

# General comments

- ▶ Very time-consuming; wish we had more time
  - ▶ No space for this within current projects
  - ▶ maybe it would be more interesting to connect data concerning type of the project (type of dictionary, its position on the market [new, old, known, unknown, commercial, academic] with the information related to the user orientation (goals to achieve when it comes to the users, user-oriented strategies, its tools, social networks etc.).
- 

Wish we had more time to analyse user behaviour and interact directly with the users.



# Meeting Programme:

**Methods to analyse dictionary use and dictionary portal use and their users**

**Contexts of Dictionary Use / Contexts of the Use of Dictionary Portals**  
(cf. Müller–Spitzer, Carolin (2014): *Using online dictionaries*, Berlin/Boston, de Gruyter.)

**User contributions to Online Dictionaries and Dictionary Portals**  
(cf. [Abel and Meyer, 2013](#));

**Case studies in Designing and Developing Dictionary Portals focusing on the User**

