
Crowdsourcing terminology: harnessing the potential of translator's glossaries

Ivanka Rajh

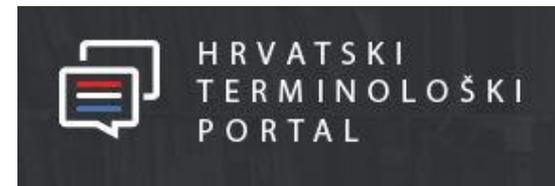
Zagreb School of Economics and Management

Siniša Runjaić

Institute of Croatian Language and Linguistics



ZZP



Translators and Interpreters Interest Group (TIIG)

- an interest group within the **Croatian Chamber of Economy**
 - in 2009 - *Translators Group* was formed within the *Foreign Languages Association of the CCE*
 - in 2013 – became autonomous
 - **umbrella organisation**
 - around 220 members (free-lancers, companies, associations)
-

National Terminological Infrastructure group within TIIG

- Terminology - one of the most important issues common to translators of all specializations
 - Main idea: „to support translators in their efforts to **exert influence** on the creation of terminological infrastructure and **adapt that infrastructure** to their practical needs...”
 - First task: „**to assess the current situation** and needs of the translators' community...”
 - **survey** conducted among the members of TIIG in **early 2016** (Gracin et al., 2016). The survey collected answers to 37 questions from 99 translators within a period of around 40 days.
-

The main conclusions of the survey

Respondents:

- freelancers (59.6%), translators employed in companies (27.3%), and part-time translators (13.1%).

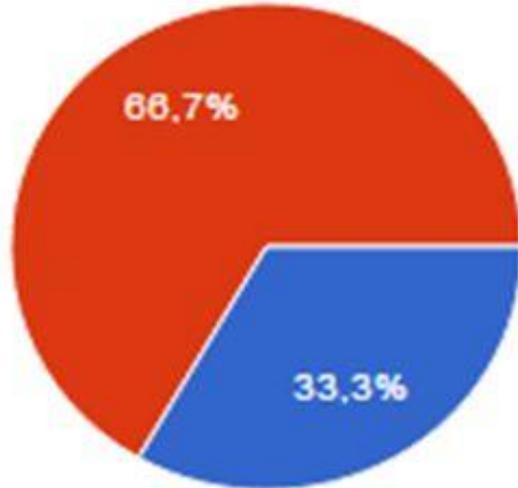
Particular resources they're using during the translation work:

- the internet (73%)
 - online scientific and professional literature (65%)
 - multilingual texts (51%)
 - their own termbases and translation memories (40%)
 - inquiries among colleagues (32%)
 - print dictionaries (21%).
-

Illustrative answers and distribution...

Question:

- Are you satisfied with the quality of terminology in available resources?



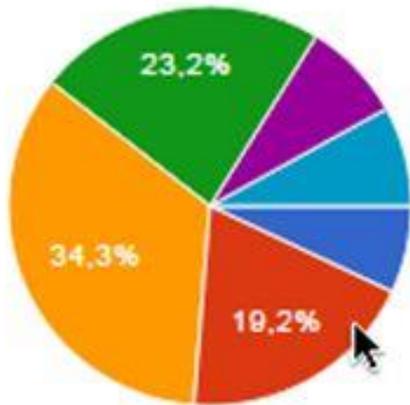
Da	33	33.3%
Ne	66	66.7%

- So, 2/3 are not satisfied with the quality of the resources, and...
-

...

Question:

- How much of the translation time do you spend on the terminological research?



do 10%	7	7.1%
od 10 do 20 %	19	19.2%
od 20 do 30 %	34	34.3%
od 30 do 40 %	23	23.2%
od 40 do 50 %	8	8.1%
više od 50 %	8	8.1%

- So, most of the respondents do spend a lot of time for the terminology research during their work, but on the other hand...

About individual and online resources

Individual resources:

- 78.8% - create their own termbases (for particular subject)
- 65.7% - share that material inside the translators community

IATE:

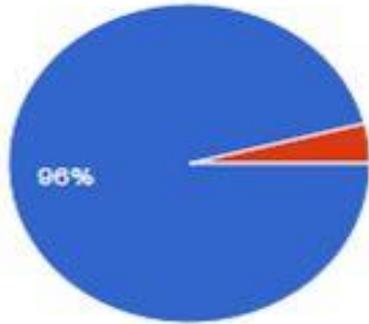
- 57% of respondents are familiar with the biggest European terminological resource, they use it sometimes, and most of them think it **doesn't fulfil the translators' needs**

Croatian Terminology Portal:

- 58% of respondents are familiar with that freely available resource, but they use it rarely or very rarely, and most of them think it **doesn't fulfil their needs** for various reasons
-

Solutions from translators' community perspective

95% of them would be interested in using similar resource built from the within TIIG:



Da	95	96%
Ne	4	4%

And the most important:

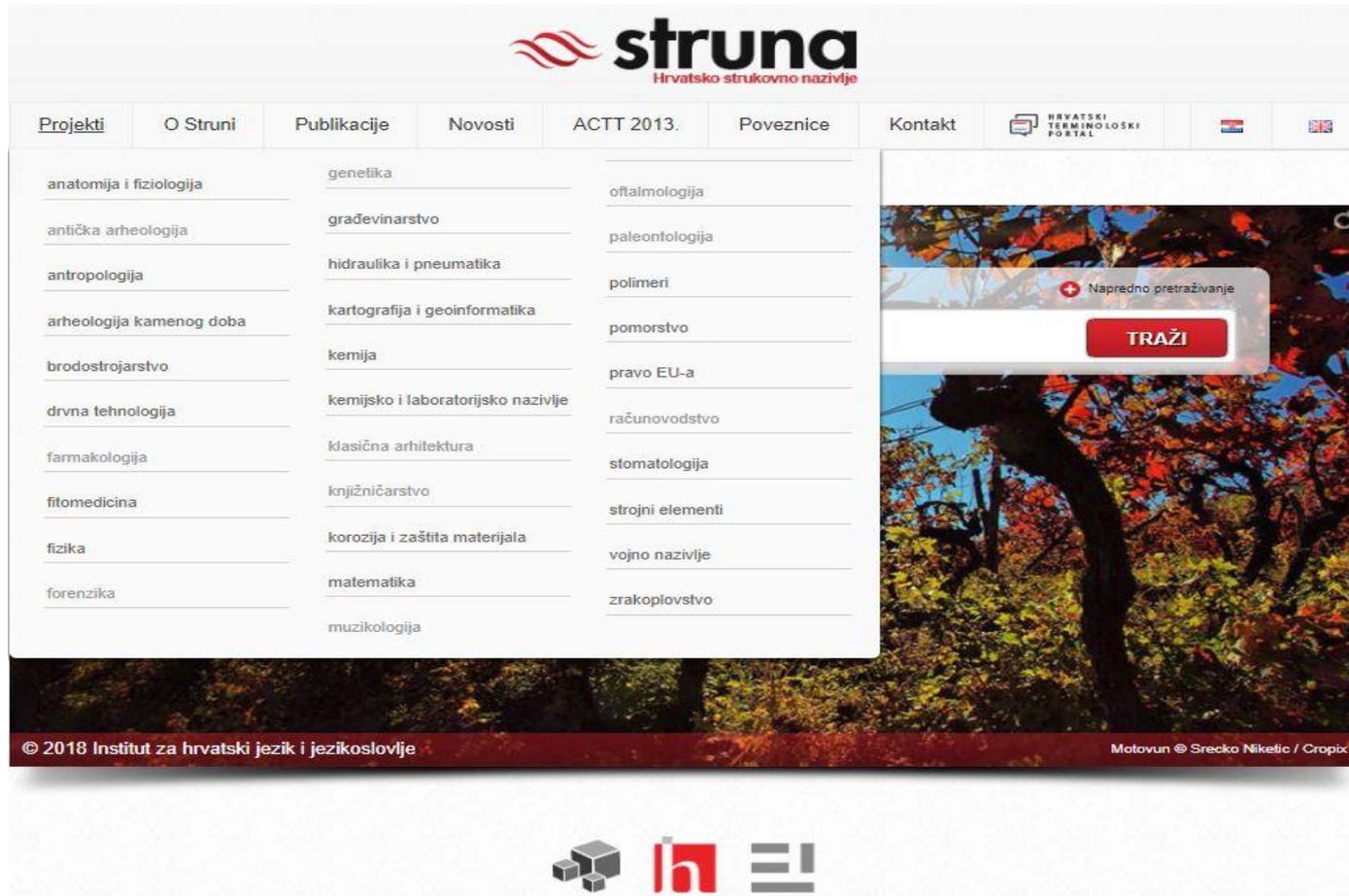
- 77% of respondents would be **ready to contribute** to such a project/resource by sharing their individual termbases, translation memories, glossaries, etc. on the basis of free data access and crowdsourcing

Choosing the partner for the solution...

Institute of the Croatian Language and Linguistics:

- Reached out to TIIG
 - presented the national terminology project *Struna* in 2010
 - presented the *Croatian Terminology Portal* in April 2016
 - Team of educated and experienced terminologists
 - technological support for a database with customized CMS
 - and a future search engine of translators' dictionaries
-

Struna – Croatian Special Field Terminology



The screenshot displays the Struna website interface. At the top, the logo "struna" is accompanied by the tagline "Hrvatsko strukovno nazivlje". Below the logo is a navigation menu with items: Projekti, O Struni, Publikacije, Novosti, ACTT 2013., Poveznice, and Kontakt. To the right of the menu are icons for the "HRVATSKI TERMINOLOŠKI PORTAL" and flags for Croatian and English. The main content area features a grid of 30 specialized terminology fields, each with a horizontal line underneath. A search bar is overlaid on the right side of the page, containing the text "Napredno pretraživanje" and a red button labeled "TRAŽI". The footer of the page includes the copyright notice "© 2018 Institut za hrvatski jezik i jezikoslovlje" and the text "Motovun © Srećko Niketić / Cropix".

anatomija i fiziologija	genetika	oftalmologija
antička arheologija	građevinarstvo	paleontologija
antropologija	hidraulika i pneumatika	polimeri
arheologija kamenog doba	kartografija i geoinformatika	pomorstvo
brodostrojarstvo	kemija	pravo EU-a
drvena tehnologija	kemijsko i laboratorijsko nazivlje	računovodstvo
farmakologija	klasična arhitektura	stomatologija
fitomedicina	knjižničarstvo	strojni elementi
fizika	korozija i zaštita materijala	vojno nazivlje
forenzika	matematika	zrakoplovstvo
	muzikologija	

- 21 finalized projects – 10 more in the pipeline.

Struna – Croatian Special Field Terminology

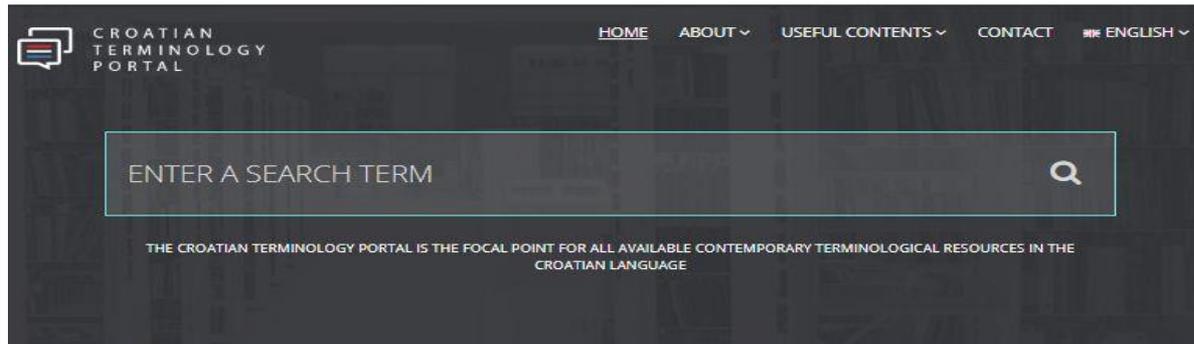
Underlying methodology:

- Normative termbase
 - organized according to the principles of the *General Theory of Terminology* and relevant ISO/TC 37 terminological standards
 - TBX compatible, 46 fields
 - collaborative terminology work

Since February 2012 – publicly available

- more than 30,000 concepts, i.e. standardized, preferred terms, accompanied by different synonyms (admitted, deprecated, obsolete and jargon terms)
 - source of information for translators
 - every term must have at least an English equivalent
 - also equivalents in other European languages
 - the search engine includes both simple and advanced types of search
-

Croatian Terminology Portal



- user- and translator-oriented addition to the *Struna* system – a central place of gathering diverse terminologies.

Croatian Terminology Portal

Publicly released after two years of preparatory work in July 2015:

- a metasearch engine (aggregate search engine)
 - simplified and made to simultaneously search all available terms in any language for a string of characters entered into a search bar
 - the results can be further narrowed down and sorted depending on the user needs and preferences
 - Portal is simultaneously searching through four separate resources:
 - *Struna*
 - donated terminological dictionaries transformed into a terminological database *Terminology dictionaries and glossaries*
 - digitally accessible resources of the *Miroslav Krleža Institute of Lexicography*
 - and donated terminology collections of the *Croatian Standards Institute*.
 - Approx. 95,000 Croatian terms available on the portal with roughly 150,000 foreign equivalents from 60 different sources – before the new coordination with TIIG
-

Terminology dictionaries and glossaries database

- Unlike *Struna* database, significantly simplified in structure
 - **adapted for the import of ready-made terminology manuals**, with only minor adjustments
 - The database
 - made up of basic metadata for individual terminological resource
 - the minimum terminological entry consists of at least one Croatian term and an equivalent in one of the European languages
 - **The idea was born!**
 - negotiations with TIIG and preparations for the creation of the database and the search engine for translator's glossaries took off
 - From the beginning of 2017 – common efforts
 - Specifications for the internal content management system (CMS)
 - Simple procedure of importing glossaries in classic table formats
-

The idea was realized

Novi naziv Početna > Nazivi DODAJ JEZIK

Naziv

Glosar

Ekologija

 Ekologija

- Glosar kanalizacije, strojarskih i zemljanih radova
- Glosar lokalne samouprave
- Glosar SEPA
- Glosar statistike osigurateljnih usluga
- Glosar termina iz područja osiguranja
- Glosar Zakona o izvršavanju državnog proračuna

Sinonim

Napomena

Kontekst

Izvor konteksta

Faza

JEZIK 1

Jezik

Istovrijednica

Kontekst

Izvor konteksta

UNESI NAZIV

Categories of data – possible for import into CMS

- name of a glossary (and its creator)
 - terms' classification (subject fields, even branches)
 - Croatian term (and possible synonyms)
 - equivalents in foreign languages (depending on the language specialization of the translator who is the author of the glossary)
 - context (term and equivalent)
 - additional fields for possible information on the context
 - any additional remarks
-
- very robust so as to accommodate different formats of glossaries
 - **mandatory categories:** the term in Croatian, its equivalent in any foreign language, the title of the glossary and the subject field.
-

An illustration

ZZP

Upišite pojam za pretraživanje...

TRAŽI

[O projektu](#)

[Kontakt](#)

radna snaga

NAPOMENA

U glosaru MVPEI-ja: radna snaga.

KONTEKST

Anketa o radnoj snazi najopsežnija je anketa o obilježjima tržišta rada provedena na uzorku kućanstava u Hrvatskoj.

Statistički ljetopis 2010., www.dzs.hr

ISTOZNAČNICE

ENGLESKI

labour force

SLOVENSKI

delovna sila

GLOSAR

[Pojmovnik ECB-a](#)

POLJE

[Ekonomija](#)

GRANA

the record radna snaga '*labour force*' in the search engine of TIIG database.

Status

- April 2018 – the searchable version <http://struna.ihjj.hr/zzp/> was released to public, i.e. TIIG members, on the website of the Croatian Chamber of Economy
- July 2018 – the TIIG database was linked through the application programming interface (API), thus becoming searchable on the *Croatian Terminology Portal* (further improvement of visibility and quality of results for all interested publics)
- Currently – 19 glossaries with 11,124 Croatian terms and 14,854 equivalents in different foreign languages, and there are a few more glossaries in the pipeline
- Glossaries characteristics:
 - vary in structure, length and overall quality
 - require additional editing performed by the terminologists from the Institute

Example from Croatian Terminology Portal

Rezultati pretraživanja za "MACKEREL"

10 rezultata ▾

PRIKAZI SAMO REZULATE IZ

STRUNA STRUKOVNI RJEČNICI I GLOSARI LEKSIKOGRAFSKI ZAVOD MIROSLAV KRLEŽA HRVATSKE NORME

Sužavanje rezultata...

NAZIV ▲	ISTOVRIJEDNICA ◆	IZVOR ◆
plavica STRUKOVNI RJEČNICI I GLOSARI	Chub mackerel	Morske ribe
skuša STRUKOVNI RJEČNICI I GLOSARI	Atlantic mackerel	Morske ribe
šarun STRUKOVNI RJEČNICI I GLOSARI	Atlantic horse mackerel	Morske ribe
šarun STRUKOVNI RJEČNICI I GLOSARI	Mediterranean horse mackerel	Morske ribe
šnjur golemi STRUKOVNI RJEČNICI I GLOSARI	Blue jack mackerel	Morske ribe

Prikazano 1 - 5 od ukupno 5 rezultata

< Prethodna 1 Sljedeća >

Search result for 'mackerel' in the Croatian Terminology Portal search engine.

Limitations of translator's glossaries

- glossary is the results of a **unique translation project**
 - reflects the process of a **terminology search** translators find themselves in
 - translators are forced to **compile their glossaries from a variety of sources** of different quality and origin
 - contain **terms found in the text that was being translated** and not the terminology of an entire subject field
 - they may contain **terms specific to a particular project** or client requirements
 - However, the terms in the database, even though they may not be the ones they are looking for, may help skilled translators **turn their search in the right direction**
-

Future – plans, goals, cooperations...

- **promotional activities** among the translator community in Croatia
 - reaching out to **government institutions**
 - detecting internal terminology resources
 - persuading the owners to enable the access to the general public (either by their inclusion into the translator's glossary database or directly into the *Croatian Terminology Portal*)
 - cooperation with the translation studies – Department of Romance Languages (Faculty of Arts and Humanities, University of Zagreb, Croatia)
 - collection of **10 students' French-Croatian glossaries**
 - more than **2000 terms** from various subject fields
 - cooperation with researchers Simon Krek and Andraž Repar from the Jožef Stefan Institute
 - inclusion of the translator's glossaries database into the **eTranslation Termbank project**, co-financed by the European Union's Connecting Europe Facility
-

Thank you for the attention!

ivanka.rajh@zsem.hr

srunjaic@ihjj.hr
