# GikiCLEF overview

Diana Santos
Luís Miguel Cabral
Nuno Cardoso
**Linguateca**

POS_Conhecimento
Programa Operacional Sociedade do Conhecimento

União Europeia
FEDER

UMIC
Agência para a Sociedade do Conhecimento

FCCN
Fundação para a Computação Científica Nacional

# GikiCLEF in a nutshell

- Asking open list questions to Wikipedia
- A difficult task for people and for computers
- A realistic situation where crosslingual and multilingual skills may be of real interest

- **http://www.linguateca.pt/GikiCLEF**

- A merger of QA and GIR, a follow-up of the GikiP pilot
- 10 Wikipedia collections, 50 culturally motivated topics
- 8 participants in 2009

# Topic titles in GikiP 2008:
# 3 different language cultures, but maybe artificial

| ID | English topic title |
|---|---|
| GP1 | Which waterfalls are used in the film "The Last of the Mohicans"? |
| GP2 | Which Vienna circle members or visitors were born outside the Austria-Hungarian empire or Germany? |
| GP3 | Portuguese rivers that flow through cities with more than 150,000 inhabitants |
| GP4 | Which Swiss cantons border Germany? |
| GP5 | Name all wars that occurred on Greek soil. |
| GP6 | Which Australian mountains are higher than 2000 m? |
| GP7 | African capitals with a population of two million inhabitants or more |
| GP8 | Suspension bridges in Brazil |
| GP9 | Composers of Renaissance music born in Germany |
| GP10 | Polynesian islands with more than 5,000 inhabitants |
| GP11 | Which plays of Shakespeare take place in an Italian setting? |
| GP12 | Places where Goethe lived |
| GP13 | Which navigable rivers in Afghanistan are longer than 1000 km? |
| GP14 | Brazilian architects who designed buildings in Europe |
| GP15 | French bridges which were in construction between 1980 and 1990 |

# GikiCLEF task to the topic group

- Please find realistic questions that make sense in your language / in your culture (9 non-English languages) and hopefully have a better coverage in that particular Wikipedia
- Hopefully even difficult to translate or render in other languages

- Thanks to the topic group (14): Corina Forascu, Pamela Forner, Danilo Giampiccolo, Katrin Lamm, Thomas Mandl, Petya Osenova, Anselmo Peñas, Erik Tjong Kim Sang, Diana Santos, Julia Schulz, Yvonne Skalban, Alvaro Rodrigo Yuste (+Paula Carvalho, Christian-Emil Ore)
- Thanks to Fred Gey for English nativization

# Examples of GikiCLEF-like systems in use

- In a mountain trip in the middle of Norway, an Italian family...

- **In which European countries is the bidet usual?**
  - Portugal
  - Spain
  - Andorra
  - Italy
  - ...

# Examples of GikiCLEF-like systems in use

- In the middle of a football match on TV in a foreign station


- **What south American countries have yellow in their football team?**
  - Brazil
  - Colombia
  - Ecuador

# Examples of GikiCLEF-like systems in use

- German youth gathering (or chatting) discussing the future: where should we move to study and still have fun?

- **Which German cities have more than one university?**
  - Berlin
  - Bonn
  - Cologne
  - Aachen
  - Bremen
  - Augsburg
  - Hamburg
  - ...

# Examples of GikiCLEF-like systems in use

- Where was this American museum which had a Picasso painting on which we saw that program last Winter?


- **Which American museums have Picasso works?**
  - Museum of Modern Art (MOMA)
  - Museum of Fine Arts (Boston)
  - Solomon R. Guggenheim Museum
  - Metropolitan Museum of Art
  - National Gallery of Art
  - Art Institute of Chicago
  - Denver Art Museum
  - (...)

# Examples of GikiCLEF-like systems in use

- Which Romanian poets published volumes with ballads before 1931? ... which may have influenced Mircea Eliade? Preliminary research for a MA Thesis in Romanian literature...

Dimitrie Bolintineanu

Vasile Alecsandri

George Coşbuc

Elisabeta de Neuwied

# Results from the topic group work

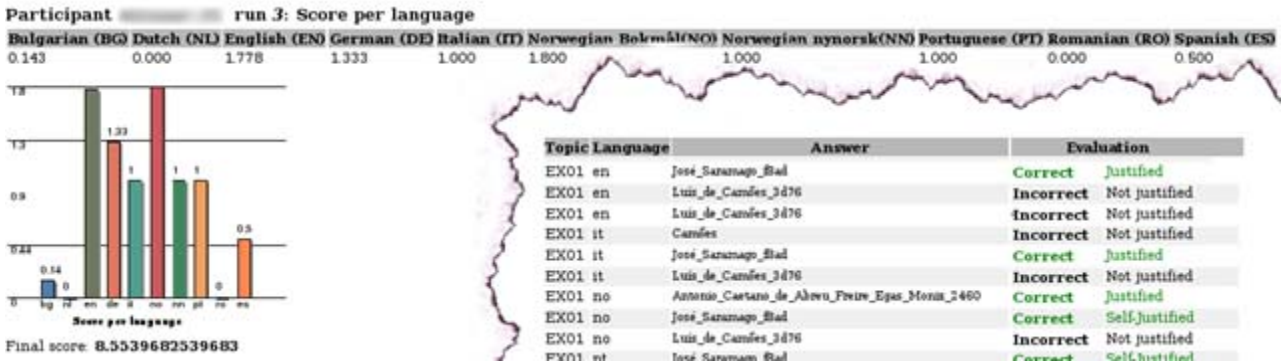
- We gathered a lof of more or less culturally-laden topics (70 -> 50)
  - Some of them **cross-cultural**, in fact (like the influence of Italy in Hemingway or the spread of Picasso's influence in the American continent)
  - Most around themes such as thinkers/writers/famous people, and places
- Not really too geographical for the most part
  - Still, Alps are **subdivided differently** in different countries/languages
  - **Flemish towns** turn out to be a not clear concept internationally
  - We allowed **visual** clues to decide (for snow, colours and/or left in a map)
- A lot of little-developed and inaccurate Wikipedia pages were found a posteriori, showing that it may not be so obvious as initially thought that searching Wikipedia crosslingually is a good idea

# The task as seen by a participant system

- In order to provide a correct answer, a system had to produce **justification** in at least one of the languages returned
- Even if correct, an answer would not be rewarded in GikiCLEF if it were not possible to assess by a human reader
- Justification could be in the page itself, or in a set of other pages provided together
  - Example question: Name places where Goethe fell in love.
  - One correct answer is Leipzig. But in the article about Leipzig this is not stated ☺
  - Where to find the justification? In "Johann Wolfgang von Goethe" page we find the following excerpt: "In **Leipzig, Goethe fell in love** with Käthchen Schönkopf and wrote cheerful verses about her in the Rococo genre."
  - Correct and justified GikiCLEF answer: **Leipzig,** Johann Wolfgang von Goethe

# GikiCLEF evaluation measures computation

- C: Number of correct answers
- N: Total number of answers provided by the system
- GP: score per language: precision vs. C: C*C/N (so the score of Romanian will be $C_{RO}$*$C_{RO}$/$N_{RO}$)
- Total score: Sum of all scores per languages: $\sum GP_{Lang}$

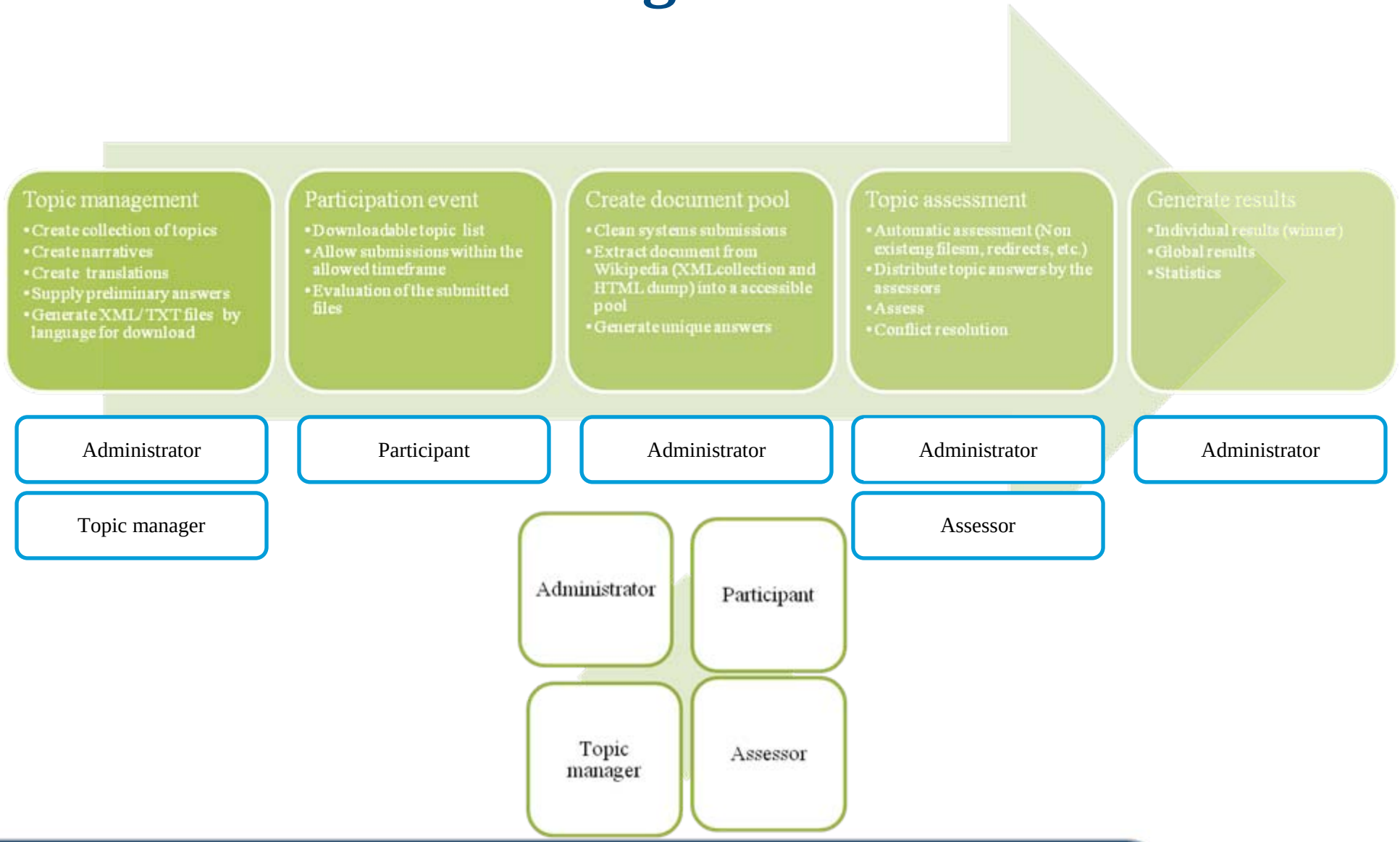


# Behind the scenes: SIGA

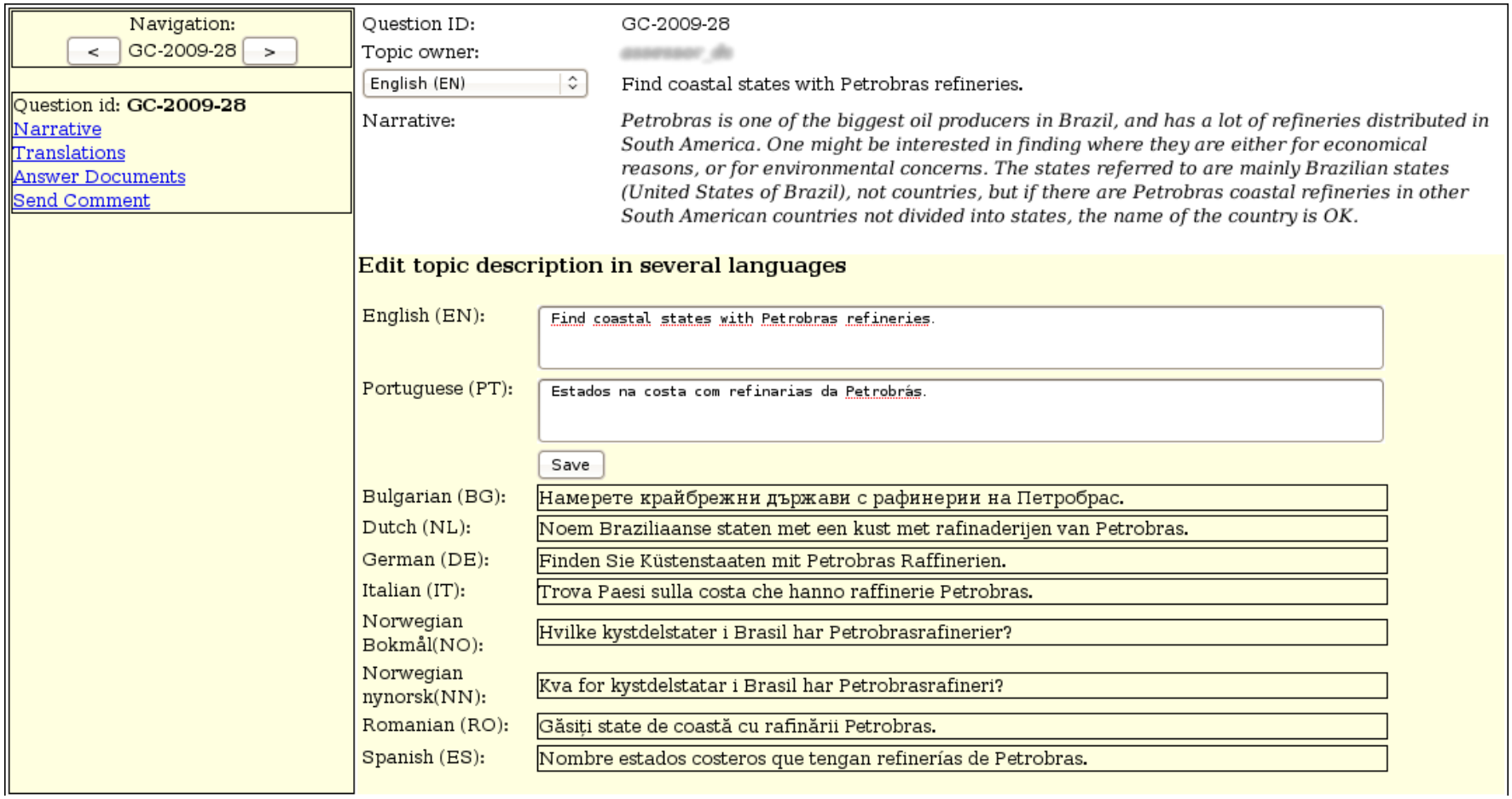
A system to help manage this evaluation track

SIGA: SIstema de Gestão e Avaliação do GIKICLEF

# SIGA: GikiCLEF organization workflow



# GikiCLEF topic management: edit topic



Navigation: < GC-2009-28 >

Question id: **GC-2009-28**
Narrative
Translations
Answer Documents
Send Comment

Question ID: GC-2009-28
Topic owner: [redacted]
English (EN)
Find coastal states with Petrobras refineries.

Narrative: *Petrobras is one of the biggest oil producers in Brazil, and has a lot of refineries distributed in South America. One might be interested in finding where they are either for economical reasons, or for environmental concerns. The states referred to are mainly Brazilian states (United States of Brazil), not countries, but if there are Petrobras coastal refineries in other South American countries not divided into states, the name of the country is OK.*

**Edit topic description in several languages**

English (EN): Find coastal states with Petrobras refineries.

Portuguese (PT): Estados na costa com refinarias da Petrobrás.

Save

Bulgarian (BG): Намерете крайбрежни държави с рафинерии на Петробрас.
Dutch (NL): Noem Braziliaanse staten met een kust met rafinaderijen van Petrobras.
German (DE): Finden Sie Küstenstaaten mit Petrobras Raffinerien.
Italian (IT): Trova Paesi sulla costa che hanno raffinerie Petrobras.
Norwegian Bokmål(NO): Hvilke kystdelstater i Brasil har Petrobrasrafinerier?
Norwegian nynorsk(NN): Kva for kystdelstatar i Brasil har Petrobrasrafineri?
Romanian (RO): Găsiți state de coastă cu rafinării Petrobras.
Spanish (ES): Nombre estados costeros que tengan refinerías de Petrobras.

# GikiCLEF topic management: find answers

Question ID: GC-2009-28

Topic owner: [redacted]

English (EN)

Find coastal states with Petrobras refineries.

Narrative: *Petrobras is one of the biggest oil producers in Brazil, and has a lot of refineries distributed in South America. One might be interested in finding where they are either for economical reasons, or for environmental concerns. The states referred to are mainly Brazilian states (United States of Brazil), not countries, but if there are Petrobras coastal refineries in other South American countries not divided into states, the name of the country is OK.*

**Add or remove answer documents**

| Document | Size | Assessor | Self justified |
|---|---|---|---|
| de/b/a/h/Bahia | 28.43 KB | *assessor_ytjk* | No |
| de/r/i/o/Rio_de_Janeiro_(Bundesstaat)_cd89 | 32.1 KB | *assessor_ytjk* | No |
| de/s/e/r/Sergipe | 23.3 KB | *assessor_ytjk* | No |
| ro/s/e/r/Sergipe | 14.13 KB | *assessor_cf* | Unset |

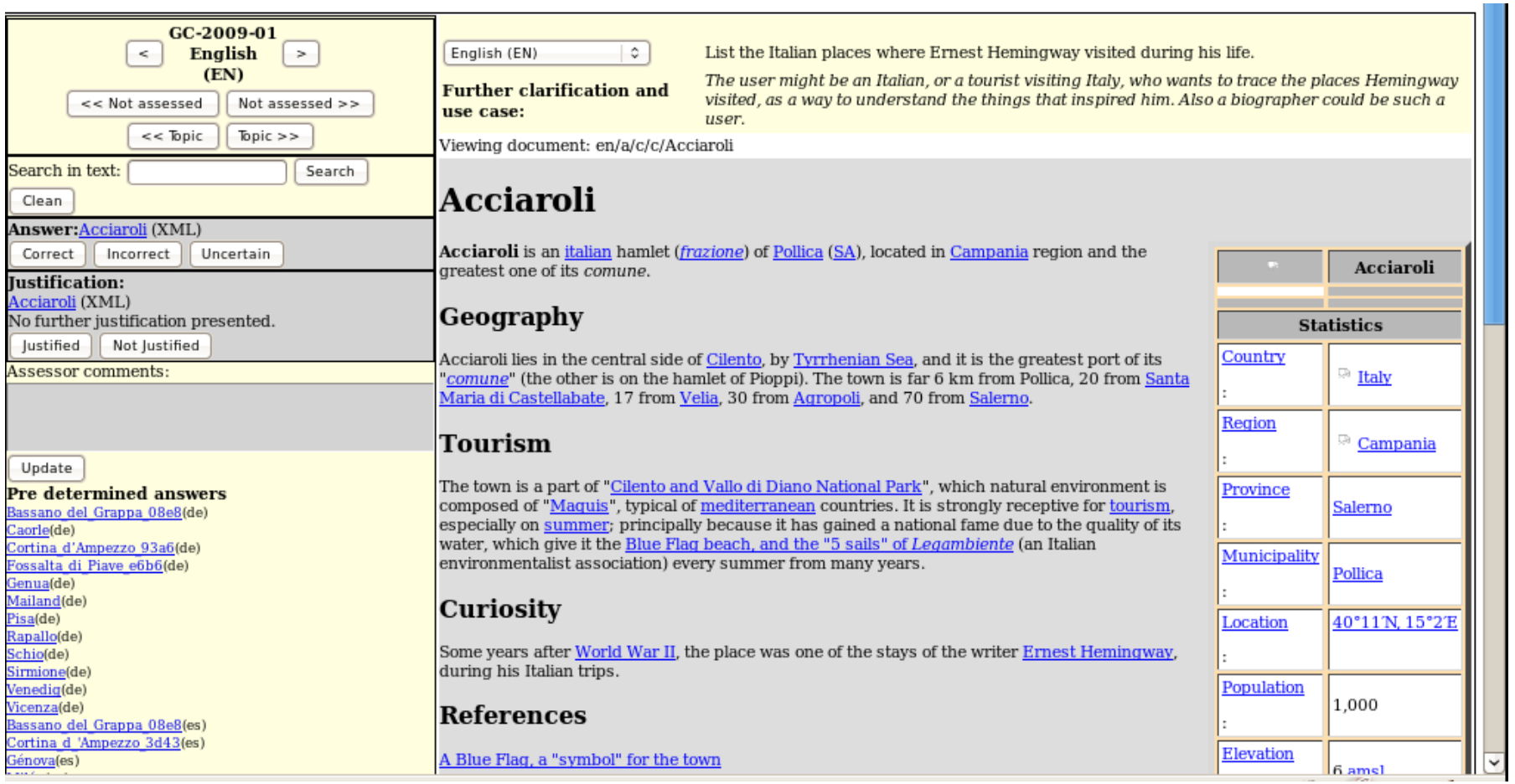
To find a document in the collection, insert a Wikipedia URL (**e.g.** "*http://en.wikipedia.org/wiki/FC_Porto*") or a Wikipedia title (**E.g.** "*Sky Tower*"). To restrict the search by language you can place the language code (bg|de|en|es|it|nl|nn|no|pt|ro) at the beginning of the box (**E.g.** "*en Olimpic_games*" )

Bahia | Search | Search alphabetically

**Pick an available document (found 1342) [Sort by score]:**

| | | |
|---|---|---|
| es/b/a/h/Bahia | 657 B | Add Preview |
| it/b/a/h/Bahia | 3.97 KB | Add Preview |
| nl/b/a/h/Bahia | 4.84 KB | Add Preview |
| no/b/a/h/Bahia | 25.81 KB | Add Preview |
| nn/b/a/h/Bahia | 9.9 KB | Add Preview |
| ro/b/a/h/Bahia | 15.14 KB | Add Preview |
| pt/b/a/h/Bahia | 295.81 KB | Add Preview |
| de/b/a/h/Bahia | 28.43 KB | Add Preview |
| en/b/a/h/Bahia | 146.42 KB | Add Preview |

# GikiCLEF assessment system: assess answers



# GikiCLEF conflict resolution

- To check assessment coherence and understand better the whole assessment process, we assigned some overlap in the answers that the assessors should evaluate
- This caused (at least for some assessors) a total repetition of their assignments, but was deemed very useful

- Thanks to all **30** assessors: Sören Auer, Anabela Barreiro, Gosse Bouma, Luís Miguel Cabral, Nuno Cardoso, Leda Casanova, Luís Costa, Iustin Dornescu, Ana Engh, Corina Forascu, dPamela Forner, Fredric Gey, Danilo Giampiccolo, Sven Hartrumpf, Katrin Lamm, Ray Larson, Laska Laskova, Johannes Leveling, Thomas Mandl, Cristina Mota, Constantin Orasan, Petya Osenova, Anselmo Peñas, Erik Tjong Kim Sang, Diana Santos, Julia Schulz, Yvonne Skalban, Rosário Silva, Kiril Simov, Alvaro Rodrigo Yuste

# GikiCLEF assessment system: conflict resolution

GikiCLEF answer assessment

[Main page | Participants list | Users | Logs | Tools | List answers | Profile | Role | Logout]
You are logged in as
Show 529 of 529 answers
[ Show all | Show only conflits | Show only unassessed | Show Correct & Justified | Show Language issues ]

[01|02|03|04|05|06|07|08|09|10|11|12|13|14|15|16|17|18|19|20|21|22|23|24|25|26|27|28|29|30|31|32|33|34|35|36|37|38|39|40|41|42|43|44|45|46|47|48|49|50]

| # | Topic | Language | Answer | Justification | Correct | Justified | Result | Comment | Info |
|---|---|---|---|---|---|---|---|---|---|
| 1 | GC-2009-09 | pt | 1728_Goethe_Link_b9c0 | | No | No | INCORRECT | | Systems: incorrect; - Re-Assess - |
| 2 | GC-2009-09 | pt | 1729_Beryl_154d | | No | No | INCORRECT | | Systems: incorrect; - Re-Assess - |
| 3 | GC-2009-09 | pt | 3047_Goethe_fda9 | | No | No | INCORRECT | | Systems: incorrect; - Re-Assess - |
| 4 | GC-2009-09 | bg | Adolf_Meschendörfer_329f | | No | No | INCORRECT | Document does not exist | Systems: auto:; - Re-Assess - |
| 5 | GC-2009-09 | de | Adolf_Meschendörfer_329f | | Conflict | No | INCORRECT | | Systems: incorrect; - Re-Assess - / correct; - Re-Assess - / incorrect; |
| 6 | GC-2009-09 | en | Adolf_Meschendörfer_329f | | No | No | INCORRECT | Document does not exist | - Re-Assess - |
| 7 | GC-2009-09 | es | Adolf_Meschendörfer_329f | | No | No | INCORRECT | Document does not exist | Systems: - Re-Assess - |

- Re-Assess -
Uncertain
Correct
Correct & justified
Correct & unjustified
Incorrect
--- Assess by Language ---
Override assess(Correct & Justified)
Override Correct(Correct only)
------
Delete assessment

# GikiCLEF result propagation

- Only the minimum number of answers were assessed, so we had to propagate the scores to the runs (to the other languages)
- As well as confirm whether there were true multilingual conflict results -> another round of conflict resolution
- For those, language propagation was inhibited
  - The language of the Vatican
    - Italian and/or Latin?


  - Height of a Norwegian waterfall
    - EN: Rjukanfossen is a waterfall of **104 meter**...
    - NO: Fossen har en total fallhøyde på **238 meter** og høyeste loddrette fall er 104 meter

# Overview of GikiCLEF 2009

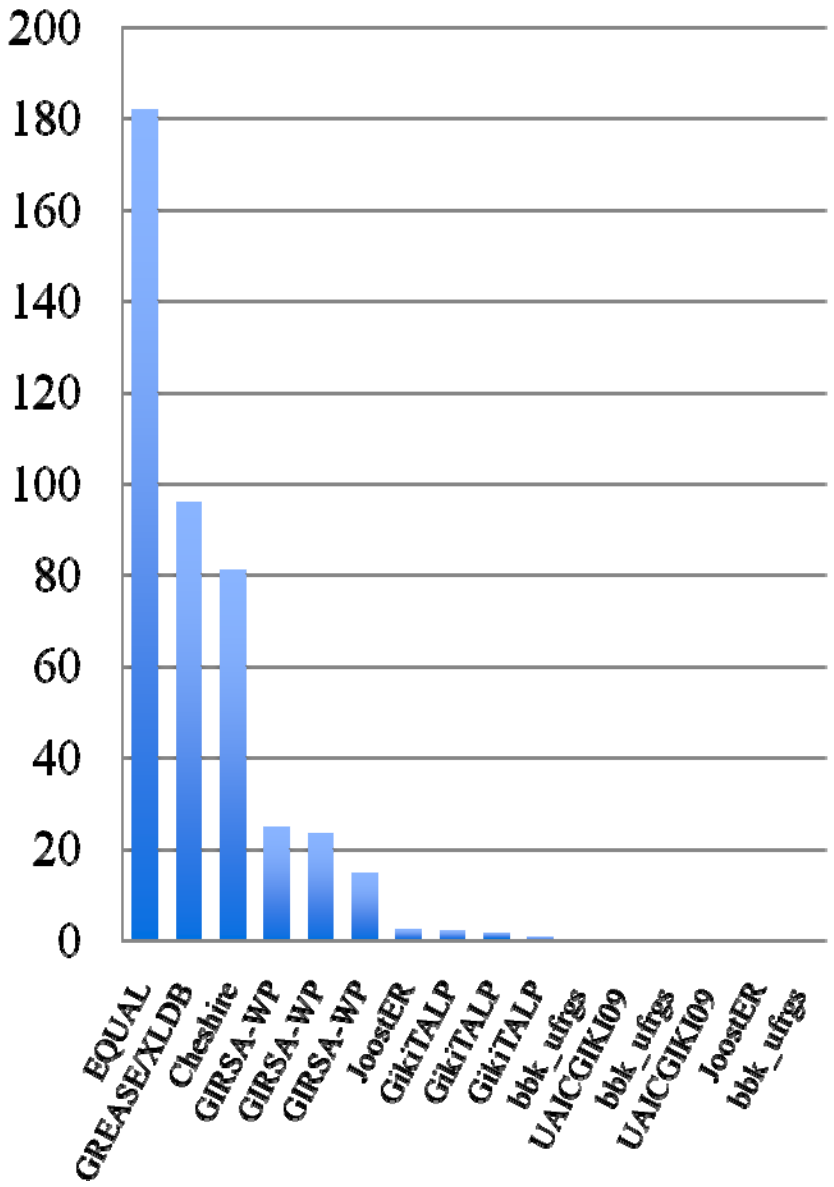
# Participants (8)

- Richard Flemmings *et al*. (BBK_UFRGS)
- Gosse Bouma & Sérgio Duarte (JoostER)
- Nuno Cardoso *et al*. (GREASE/XLDB)
- Iustin Dornescu (EQUAL)
- Ray R. Larson (CHESHIRE)
- Sven Hartrumpf & Johannes Leveling (GIRSA-WP)
- Daniel Ferrés & Horácio Rodríguez (GIKITALP)
- Adrian Iftene *et al*. (UAICGIKI09)

**17 runs submitted**

# Final score

| System | RunScore | #answers | #Corrects | Precision | Score |
|---|---|---|---|---|---|
| EQUAL | 1 | **813** | **385** | **0.4736** | **181.9329** |
| GREASE/XLDB | 1 | 1161 | 332 | 0.2860 | 96.0070 |
| Cheshire | 1 | 564 | 214 | 0.3794 | 80.9247 |
| GIRSA-WP | 1 | 38 | 31 | 0.8158 | 24.7583 |
| GIRSA-WP | 3 | 985 | 142 | 0.1442 | 23.3919 |
| GIRSA-WP | 2 | 994 | 107 | 0.1076 | 14.5190 |
| JoostER | 1 | 638 | 36 | 0.564 | 2.4053 |
| GikiTALP | 3 | 356 | 26 | 0.0730 | 1.9018 |
| GikiTALP | 2 | 295 | 20 | 0.0678 | 1.3559 |
| GikiTALP | 1 | 526 | 18 | 0.0342 | 0.6964 |
| bbk_ufrgs | 1 | 726 | 8 | 0.0110 | 0.0882 |
| UAICGIKI09 | 2 | 6420 | 8 | 0.0012 | 0.0156 |
| bbk_ufrgs | 2 | 734 | 3 | 0.0041 | 0.0123 |
| UAICGIKI09 | 1 | 1133 | 2 | 0.0018 | 0.0062 |
| JoostER | 2 | 272 | 0 | 0.0000 | 0.0000 |
| bbk_ufrgs | 3 | 686 | 0 | 0.0000 | 0.0000 |
| UAICGIKI09 | 3 | 4910 | 0 | 0.0000 | 0.0000 |

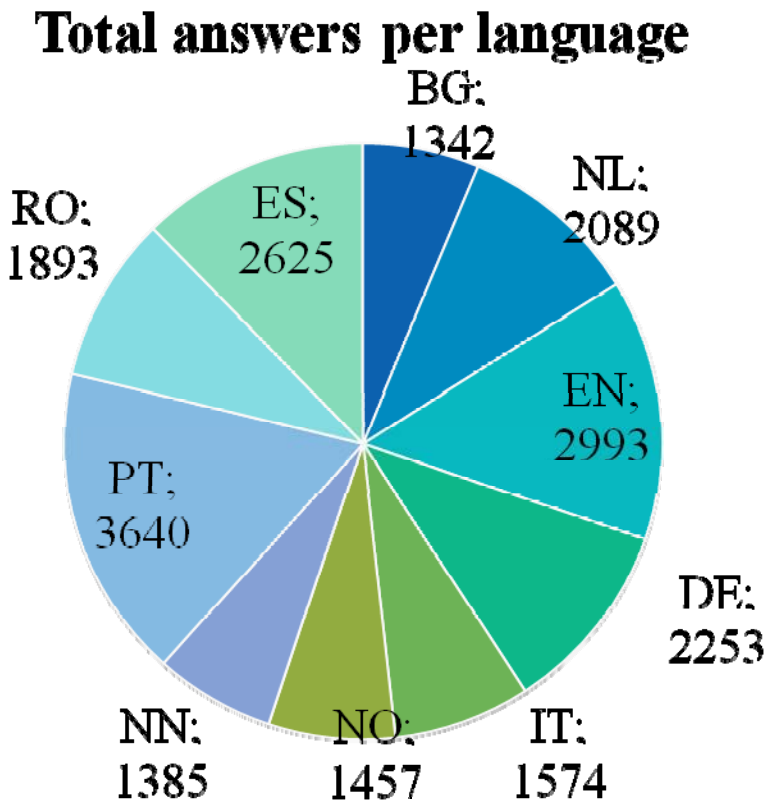


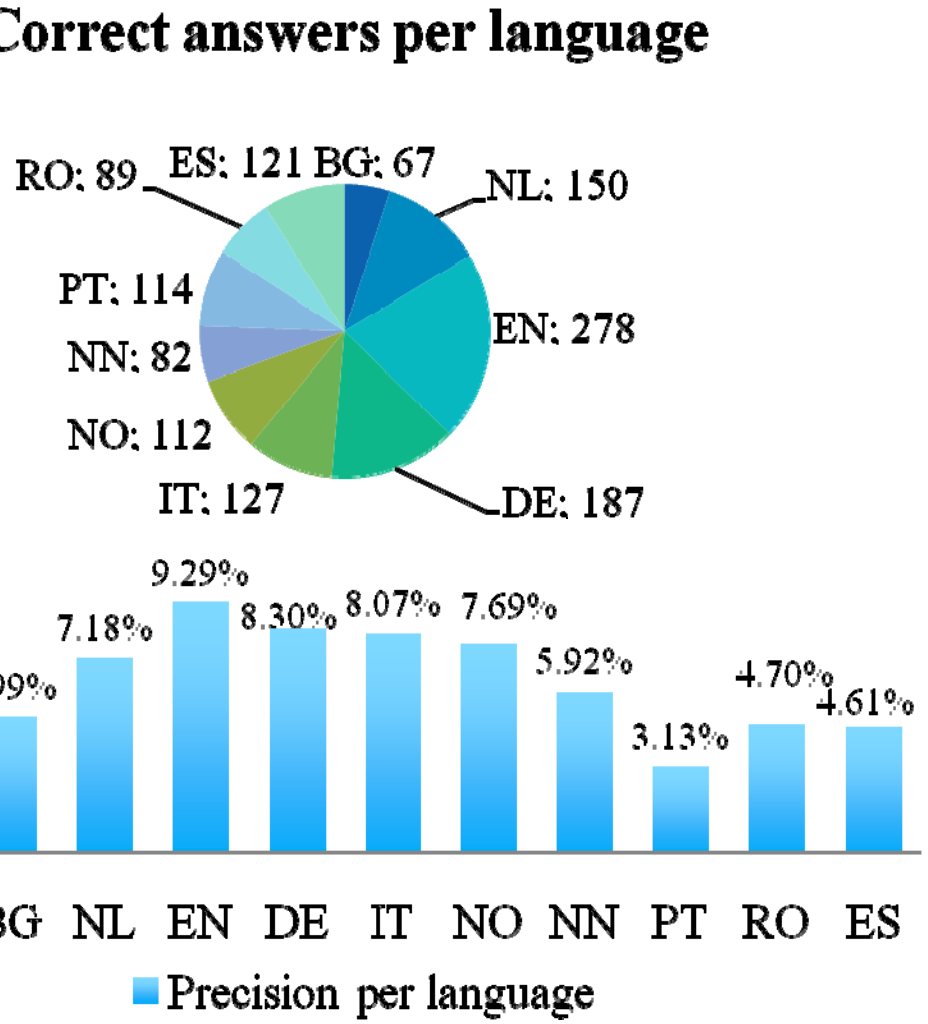
# Scores per language

| System / run | BG | NL | EN | DE | IT | NO | NN | PT | RO | ES | Total | Languages |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **EQUAL 1** | **9.757** | **18.980** | **34.500** | **25.357** | **17.391** | **17.254** | **9.308** | **15.515** | **14.500** | **16.695** | **181.932** | **10** |
| GREASE/XLDB 1 | 6.722 | 8.258 | 13.657 | 12.007 | 8.533 | 11.560 | 9.557 | 7.877 | 6.720 | 11.115 | **96.007** | **10** |
| Cheshire | 1.091 | 9.132 | 22.561 | 9.000 | 11,200 | 7.043 | 3.368 | 4.891 | 7.714 | 4.923 | **80.924** | **10** |
| GIRSA_WP 1 | 1.333 | 2.250 | 1.800 | 1.125 | 2.250 | 3.000 | 2.000 | 3.000 | 3.000 | 3.000 | **24.758** | **10** |
| GIRSA_WP 3 | 3.030 | 1.798 | 1.390 | 3.661 | 1.988 | 2.526 | 3.064 | 2.250 | 1.684 | 2.000 | **23.391** | **10** |
| GIRSA_WP 2 | 2.065 | 1.299 | 0.496 | 1.540 | 1.429 | 1.723 | 1.841 | 1.350 | 1.029 | 1.306 | **14.519** | **10** |
| JoostER 1 | | 0.964 | 1.441 | | | | | | | | **2.405** | **2** |
| GikiTALP 3 | | | 1.635 | | | | | | | 0.267 | **1.901** | **2** |
| GikiTALP 2 | | | 1.356 | | | | | | | | **1.356** | **1** |
| GikiTALP 1 | | | 0.668 | | | | | | | 0.028 | **0.696** | **2** |
| bbk_ufrgs 1 | | | | | | | | 0.088 | | | **0.088** | **1** |
| UAICGIKI09 2 | 0.000 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.000 | 0.002 | 0.000 | 0.006 | **0.015** | **10** |
| bbk_ufrgs 2 | | | | | | | | 0.012 | | | **0.012** | **1** |
| UAICGIKI09 1 | | | | | | | | | 0.000 | 0.006 | **0.062** | **2** |
| JoostER 2 | | | | | | | | | | 0.000 | **0.000** | **1** |
| bbk_ufrgs 3 | | | | | | | | 0.000 | | | **0.000** | **1** |
| UAICGIKI09 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | **0.000** | **10** |
| **Total language score** | **23.998** | **41.717** | **78.340** | **52.690** | **31.602** | **43.106** | **29.138** | **34.883** | **34.647** | **39.039** | **427.193** | |
| **Total Participations** | **8** | **9** | **12** | **8** | **8** | **8** | **8** | **11** | **9** | **12** | | |

Best results for that system | Second best results for that system

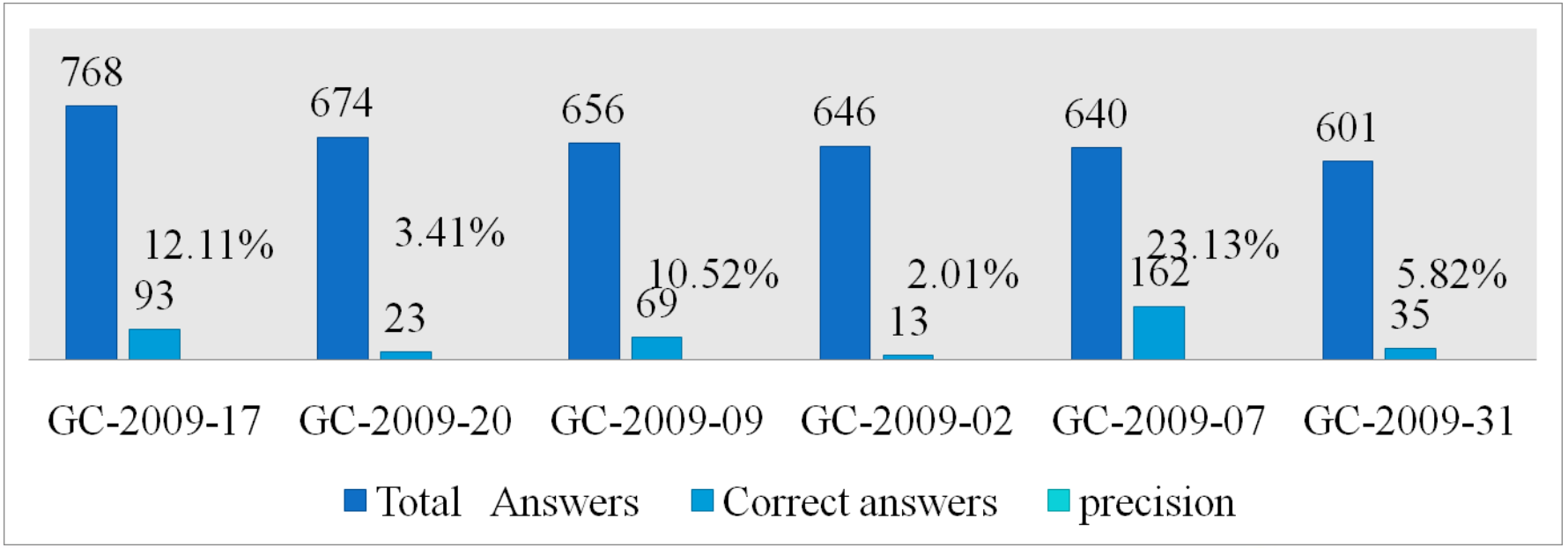
# Results overview per language



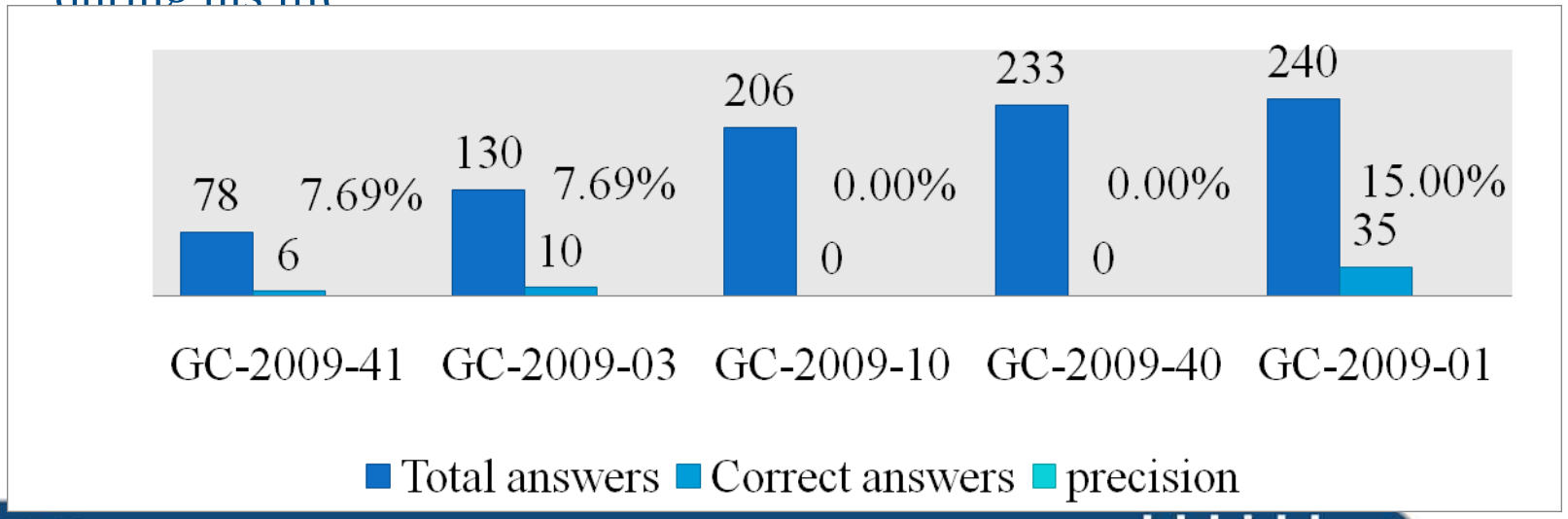
# Topics with most answers (by the participant set)

- GC-2009-17 (768): List the 5 Italian regions with a special statute
- GC-2009-20 (674): List the name of the sections of the North-Western Alps
- GC-2009-09 (656): Name places where Goethe fell in love.
- GC-2009-02 (646): Which countries have the white, green and red colors in their national flag?
- GC-2009-07 (640): What capitals of Dutch provinces received their town

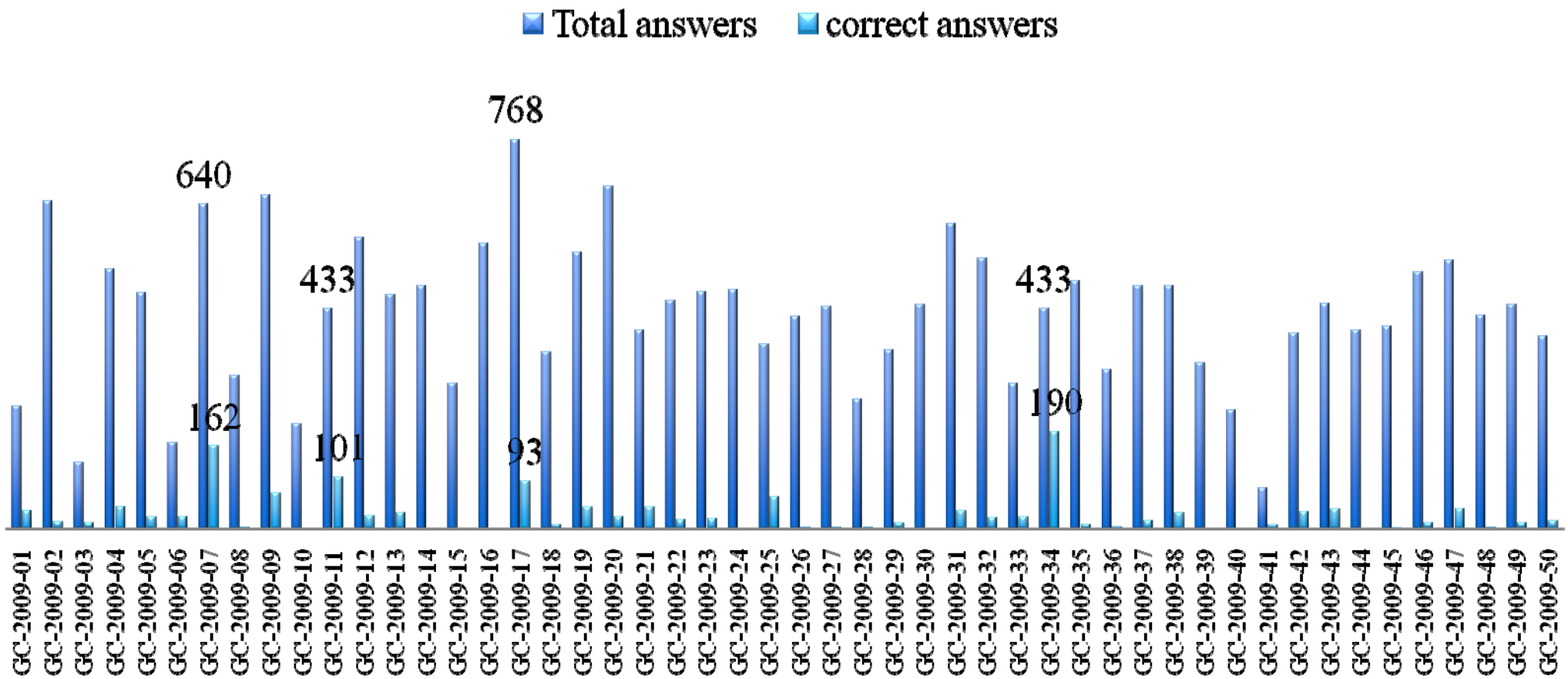


# Topics with least answers (by the participant set)

- GC-2009-41 (78): Chefs born in Austria who have received a Michelin Star.
- GC-2009-03 (130): In which countries outside Bulgaria are there opinions on Petar Dunov's (Beinsa Duno's) ideas?
- GC-2009-10 (206): Which Flemish towns hosted a restaurant with two or three Michelin stars in 2008?
- GC-2009-40 (233): Which rivers in North Rhine Westphalia are approximately 10km long?
- GC-2009-01 (240): List the Italian places which Ernest Hemingway visited during his life

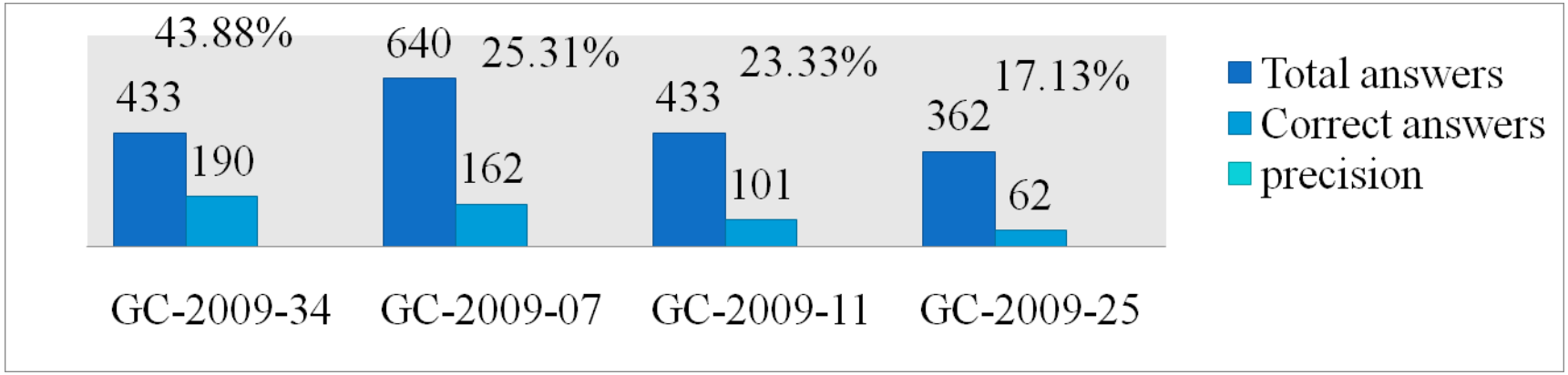


# Number of answers per topic



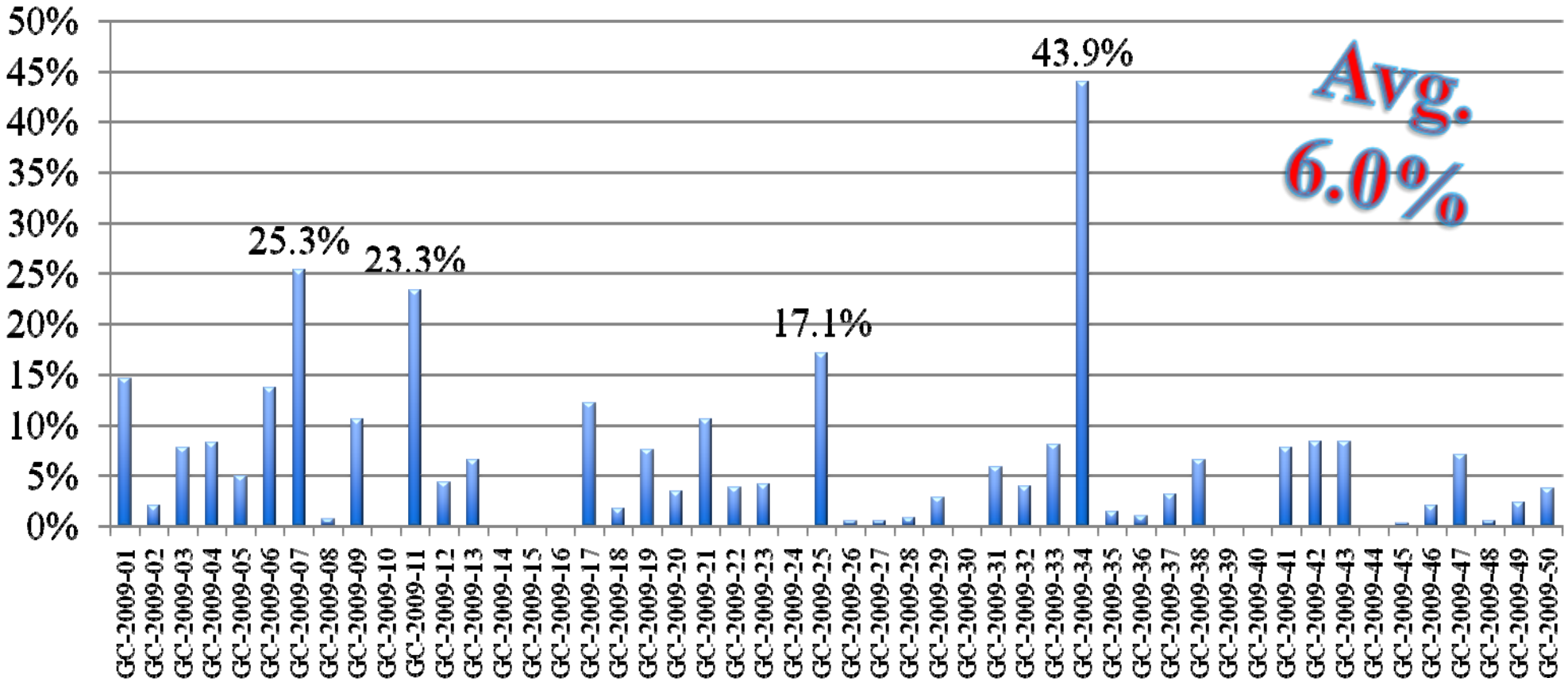
# Topics result overview: by precision

- Highest precision
  - GC-2009-34 (43.9%): What eight thousanders are at least partially in Nepal?
  - GC-2009-07 (25.3%): Which capitals of  Dutch provinces received their town privileges before the fourteenth century?
  - GC-2009-11 (23.3%): What Belgians won the Ronde van Vlaanderen exactly twice?
  - GC-2009-25 (17.13%):  Name Spanish drivers who have driven in Minardi.



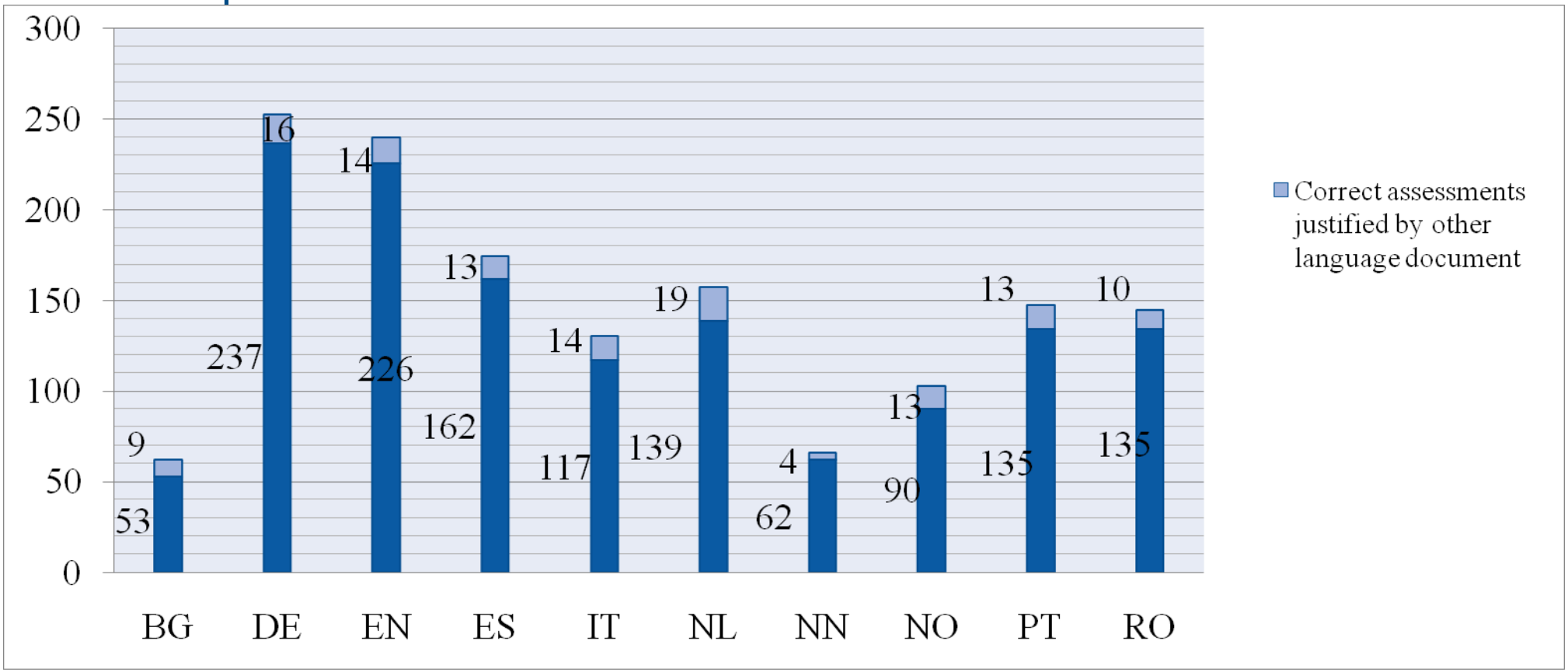

- 9 topics without any correct answer by the systems

# Precision by topic



# Authoritative languages: where is the justification

- For each language, was the justification found in the corresponding Wikipedia?



# Public resources

http://www.linguateca.pt/GikiCLEF/

- Topic lists per language, in XML and text format
- Topic descriptions (in English, together with assessment decisions)
- Lists of correct answers
  - Correct and justified in some language: GikiCLEF_answers_correct_justified.txt
  - Correct (because we know): GikiCLEF_answers_correct.txt
- Results (global, per language) and general statistics
- SIGA system, open source
- GikiCLEF collections

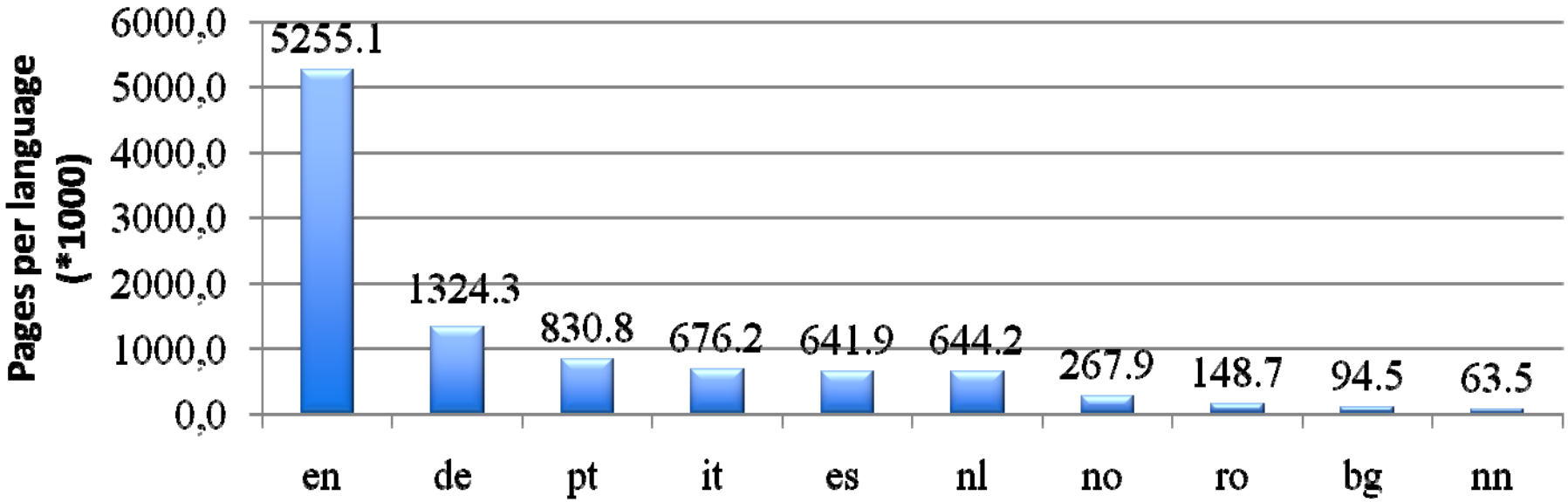
# Final discussion: Was GikiCLEF worthwhile?

- Recurrent comment: **too difficult!**
- Often not clear which (Wikipedia) pages were sought
  - Flag pages or country pages?
  - Team pages or country pages?
- Most (organizers) were probably not expecting so much work

- It is possible to create a system capable of answering many languages
- It pays to process English! The amount provided by each language, even in culturally-aware topics, is negligible
- The quality of other languages' Wikipedias is in strong need of improvement

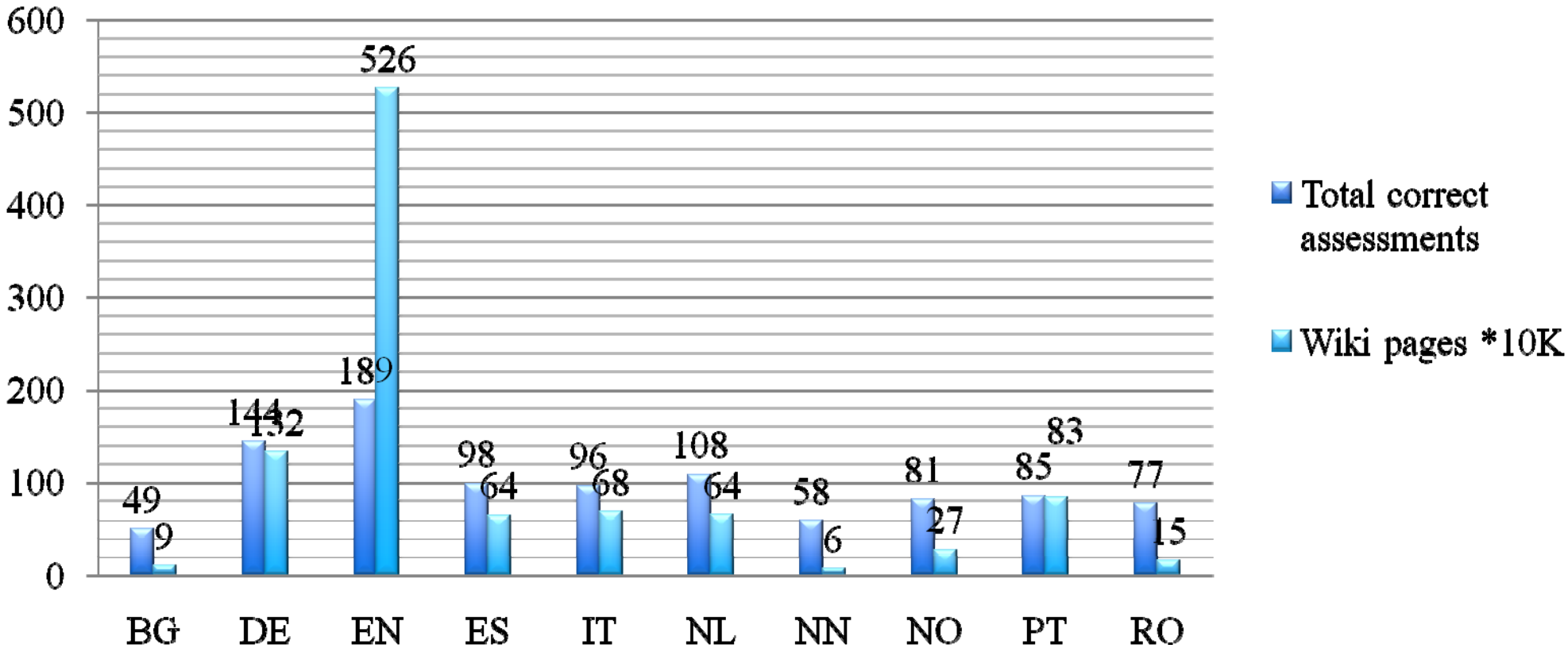
# Further details

# Size of collections (pages * 1k)

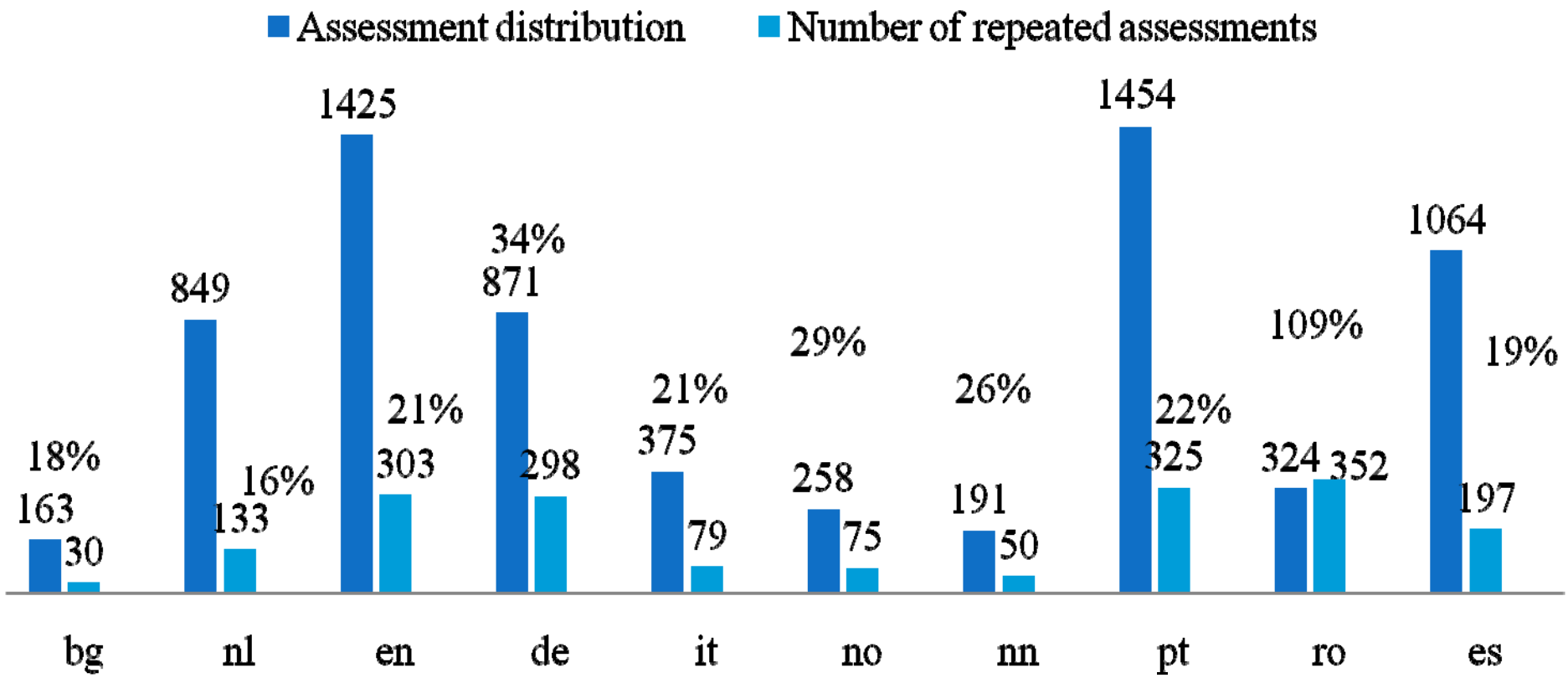



All collections made available were dated from June 2008, except for the English collection SQL dump, from May and July instead (since the SQL dump of June was not available )

# Number of authoritive answers per wiki size

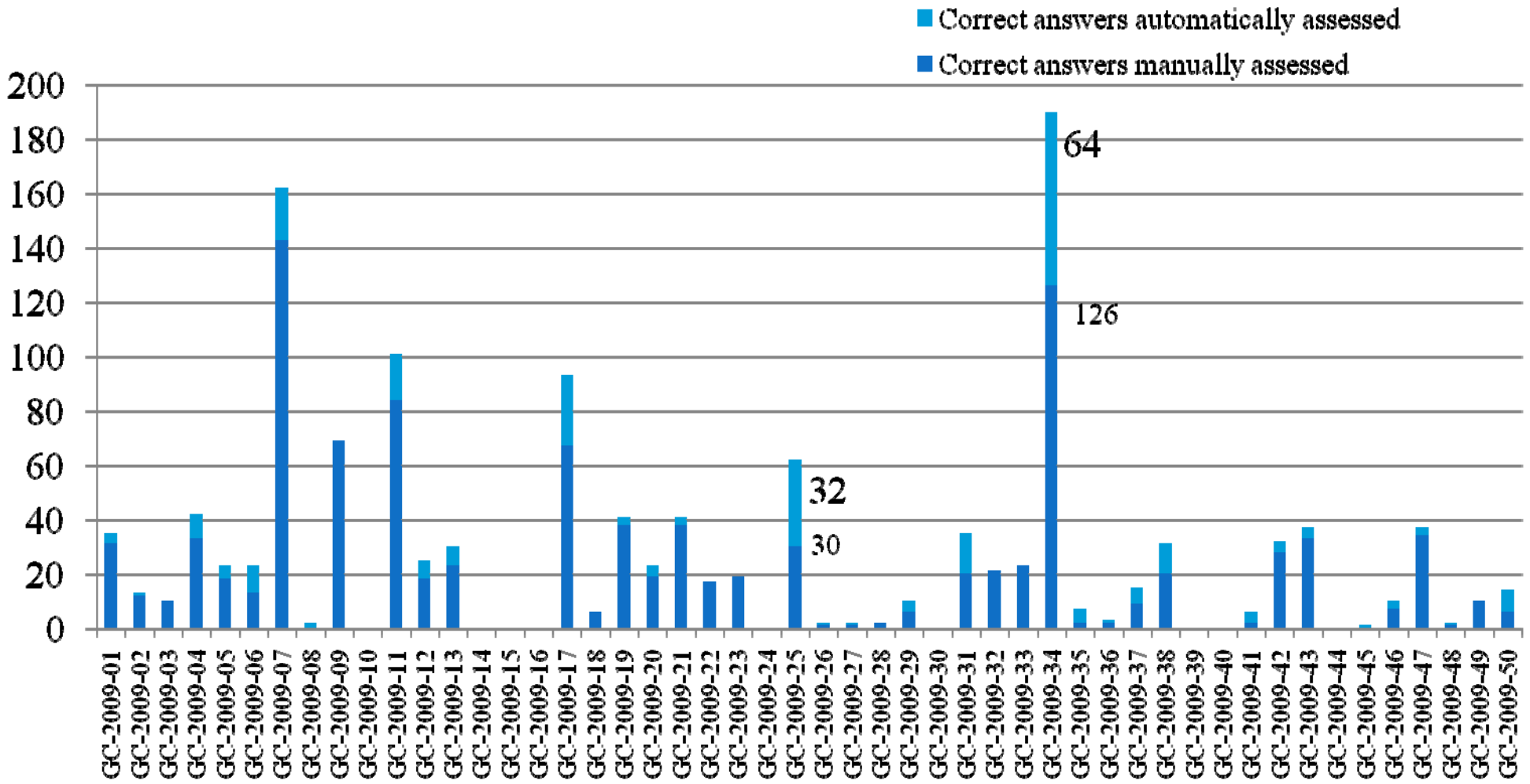


# Assessment processing

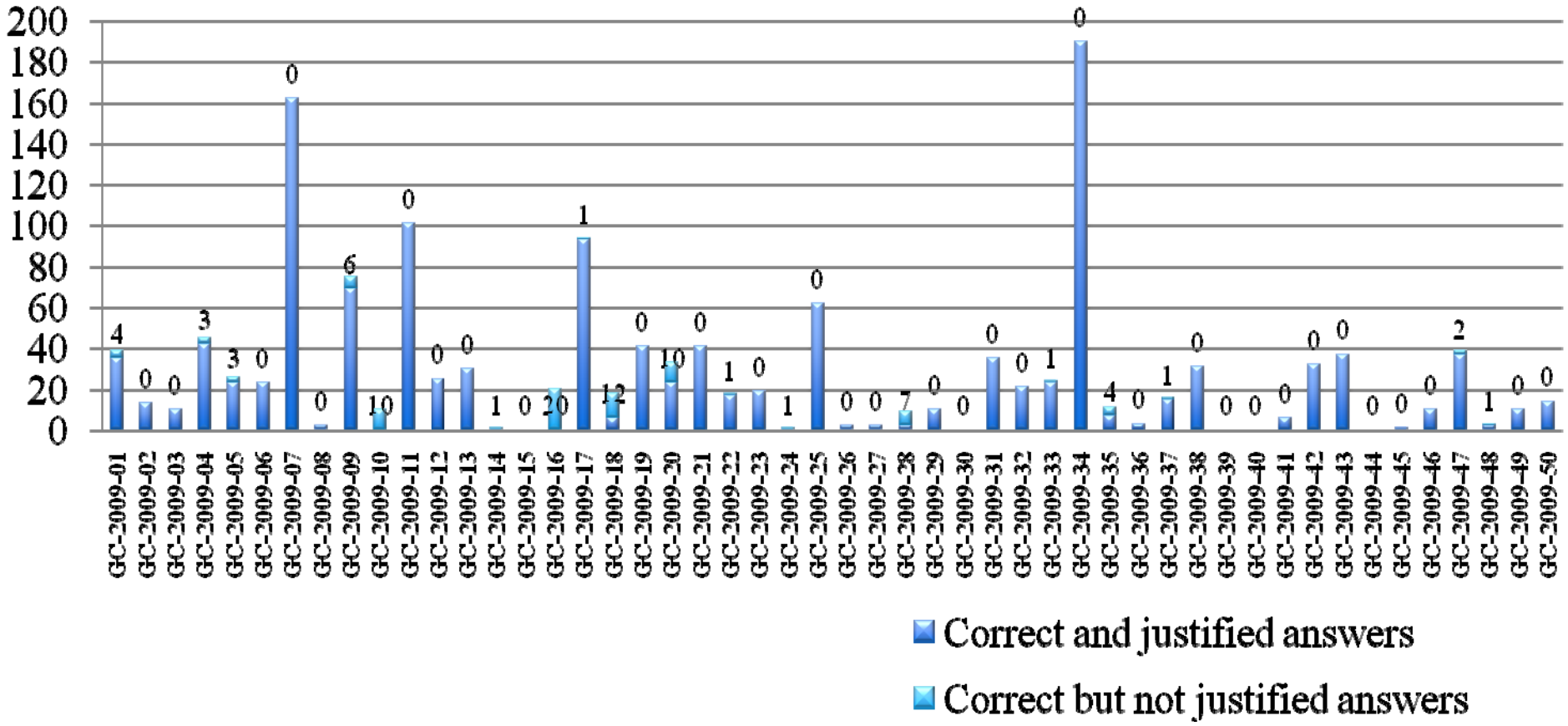



"Overlap": ratio of repeated vs. total, but some answers were assessed by three or even four assessors, since repetition was automatically computed by SIGA.

# Number of correct answers automatically assessed and manually assessed



# Correct and justified answers VS correct not justified answers



# Use of justification

- Only 2 systems made use of the justification field (4 runs) with a total of 229 answers with another justification document