



CLEF 2007

# Multilingual Question Answering Track

---

**Anselmo Peñas, UNED**  
**Danilo Giampiccolo, CELCT**

# Main Task QA 2007

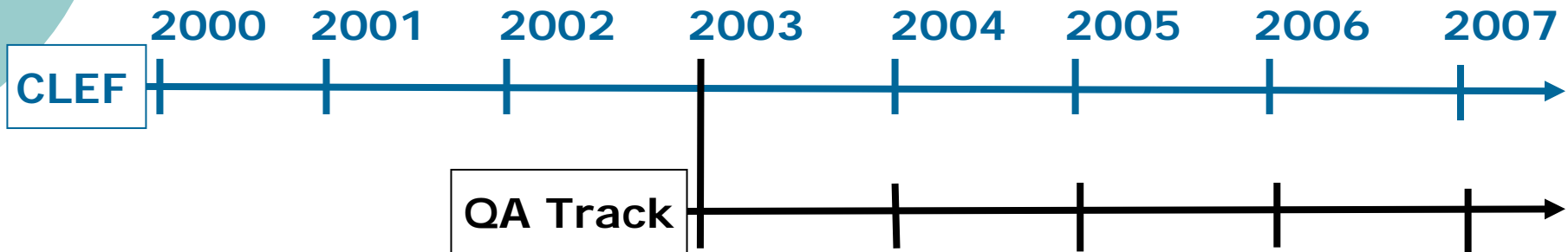
## Organizing Committee

---

- **CELCT** (D. Giampiccolo, P. Forner): Italian
- **UNED** (A. Peñas): Spanish
- **U. Amsterdam** (V. Jijkoun): Dutch
- **U. Limerick** (R. Sutcliff): English
- **DFKI** (B. Sacalenau): German
- **ELDA/ELRA** (C. Ayache): French
- **Linguatca** (P. Rocha): Portuguese
- **Bulgarian Academy of Sciences** (P. Osenova): Bulgarian
- ◆ **IASI** (D. Cristea): Romanian
  
- ◆ Only Source Languages:
  - ◆ **Depok University of Indonesia** (M. Adriani): Indonesian

# Time goes...

---



# Evolution of the Track

	2003	2004	2005	2006	2007
Target languages	3	7	8	9	10
Collections	News 1994		+News 1995		+Wikipedia Nov. 2006
Type of questions	200 Factoid		+ Temporal restrictions + Definitions	- Type of question + Lists	+ Linked questions + Closed lists
Supporting information	Doc.	Doc.	Doc.	Snippet	Snippet
Pilots and Exercises		<ul style="list-style-type: none"> <li>○Temporal restrictions</li> <li>○Lists</li> </ul>		<ul style="list-style-type: none"> <li>○AVE</li> <li>○Real Time</li> <li>○WiQA</li> </ul>	<ul style="list-style-type: none"> <li>○AVE</li> <li>○QAST</li> </ul>

# 200 questions

---

- FACTOID
  - (loc, mea, org, per, tim, cnt, obj , oth)
- DEFINITION
  - (per, org, obj, oth)
    - Person: *Who is Josef Paul Kleihues?*
    - Object: *What is a router?*
    - Other: *What is a tsunami?*
- CLOSED LIST
  - *Who were the components of The Beatles?*
  - *Who were the **last three** presidents of Italy?*
- ◆ Temporal restrictions by date, by period, by event
- ◆ NIL questions (without known answer in the collection)

New!

New!

# Linked questions

---

- TOPIC: Otto von Bismarck
  - Who was called the “Iron-Chancellor”?
  - When was he born?
  - Who was his first wife?
  
- Topics
  - Person or Event
  - Not provided to participants
  - Only a portion of the questions (from 15% depending on languages)



# Activated Tasks

	MONOLINGUAL	CROSS-LINGUAL	TOTAL
CLEF 2003	3	5	8
CLEF 2004	6	13	19
CLEF 2005	8	15	23
CLEF 2006	7	17	24
<b>CLEF 2007</b>	<b>8</b>	<b>29</b>	<b>37</b>



# Participants

	Newcomers	Veterans	TOTAL	Registered
CLEF 2003	-	-	8	-
CLEF 2004	13	5	18 (+125%)	22
CLEF 2005	9	15	24 (+33%)	27
CLEF 2006	10	20	30 (+25%)	36
<b>CLEF 2007</b>	<b>8</b>	<b>14</b>	<b>22</b> <b>(-26%)</b>	<b>29</b>

# Submitted runs

---

	Submitted runs	Monolingual	Cross-lingual
<b>CLEF 2003</b>	<b>17</b>	<b>6</b>	<b>11</b>
<b>CLEF 2004</b>	<b>48 (+182%)</b>	<b>20</b>	<b>28</b>
<b>CLEF 2005</b>	<b>67 (+39.5%)</b>	<b>43</b>	<b>24</b>
<b>CLEF 2006</b>	<b>77 (+13%)</b>	<b>42</b>	<b>35</b>
<b>CLEF 2007</b>	<b>37 (-52%)</b>	<b>20</b>	<b>17</b>

# Lower (not low) participation

---

- New collection to be indexed
  - Wikipedia
- More difficult questions
  - Linked questions
  - Closed lists
- Big surprise
  - Guidelines too late
  - Evaluate developers time reaction?

# Final list of participants (random order)

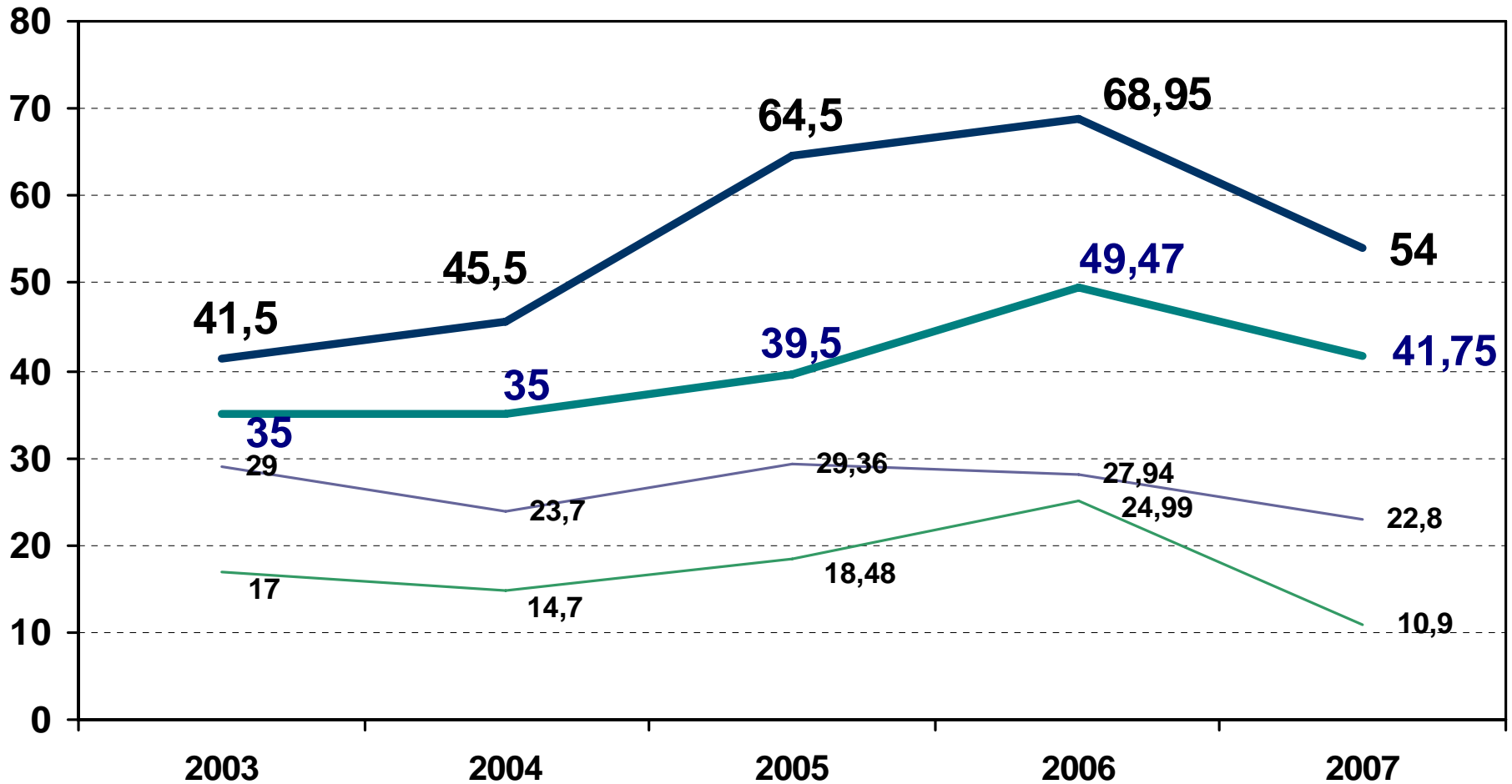
NAME	COUNTRY
SYNAPSE Developpement	France
INESC-ID	Portugal
Universiteit van Amsterdam	Netherlands
FernUniversität in Hagen	Germany
University of Évora	Portugal
RACAI	Romania
INAOE	Mexico
University of Indonesia	Indonesia
DFKI	Germany
MIRACLE	Spain
Linguateca-SINTEF	Norway
Priberam Informatica	Portugal
Universidade do Porto	Portugal
Rijksuniversiteit Groningen	Netherlands

NAME	COUNTRY
LCC	USA
Universitatea "Alexandru Ioan Cuza"	Romania
Universidad Politécnica de Valencia	Spain
FBK-irst	Italy
Universitat Politècnica de Catalunya	Spain
University of Wolverhampton	UK
Cindi Group	Canada
Macquarie University	Australia



**Industrial Companies**

# Results: Best and Average scores



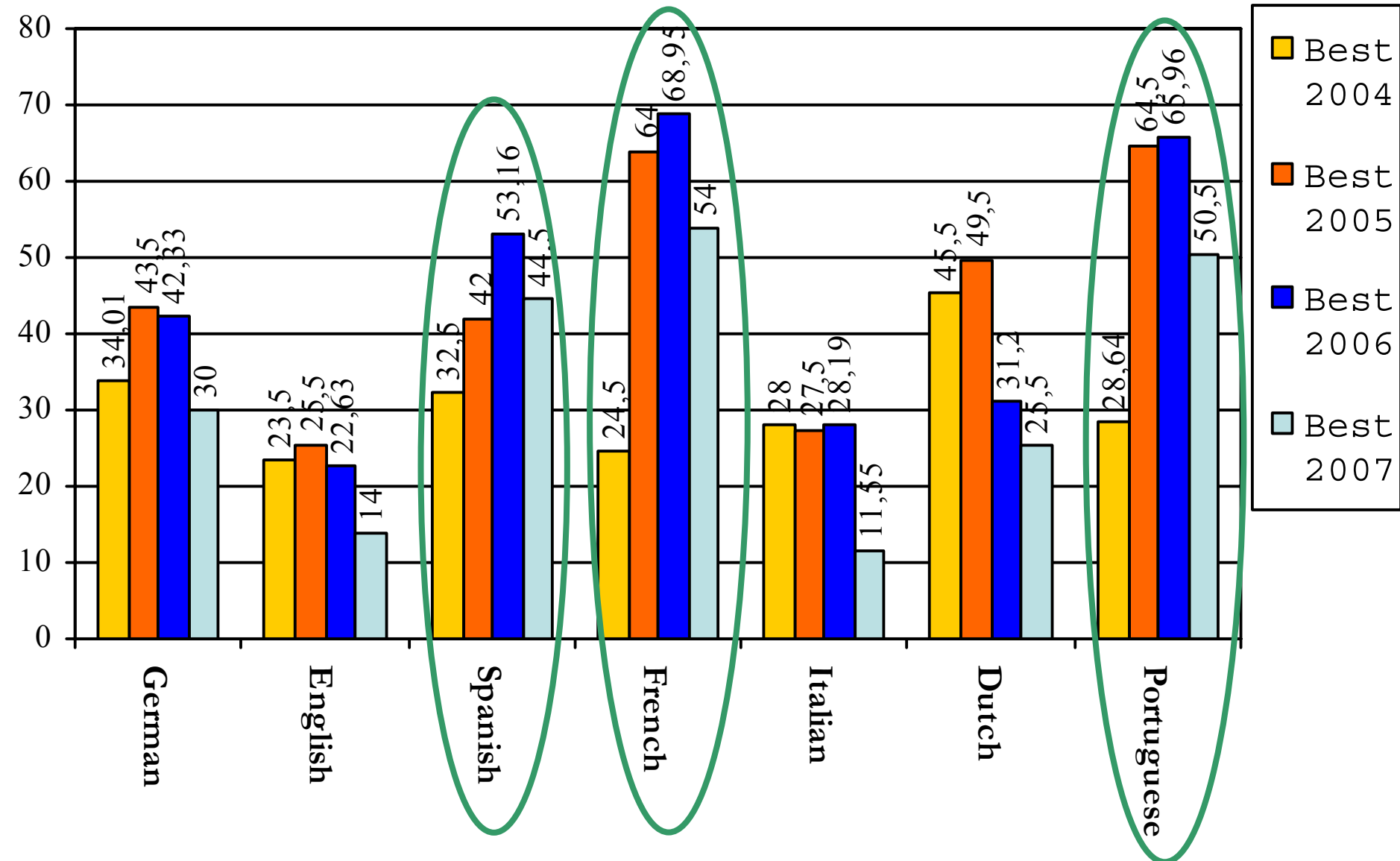
Best Bilingual

Average Bilingual

Best Monolingual

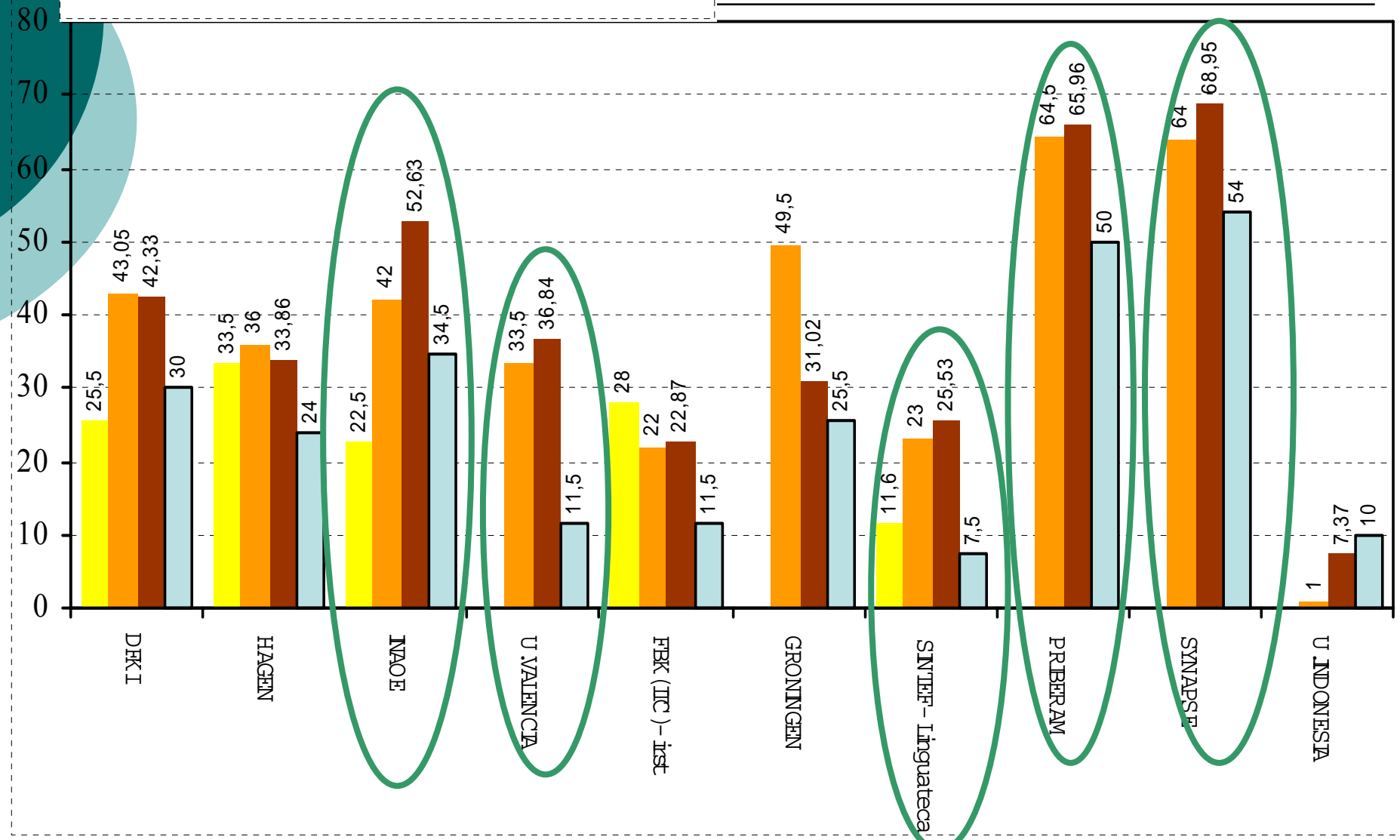
Average Monolingual

# Best scores by language



# Best scores by participant

Best 2004 Best 2005 Best 2006 Best 2007



# Lower results

---

- Some answers only in wikipedia
- Closed lists
  - Almost no answers
- Temporal restrictions
  - Still very difficult
- Linked questions
  - Topic not provided
  - Fail the first, fail the rest
  - Co-reference resolution



# Conclusion

---

- Much more difficulty
  - Less participants
  - Poorer results
- But
  - New challenges
  - New collections
  - 10 languages
  - 37 activated subtasks
  - 22 participants
  - 37 runs

# Conclusion

---

- QA Track continues its evolution
  - Although we are a big heterogeneous community
- Trying to find a compromise between
  - Real world application
  - Interest for research
  - User needs / model
  - Systems ability
  - Available collections
  - Replication of experiments
  - Components evaluation
  - Newcomers
  - Natural progress
  - ...

# Questions for breakout

---

- Repeat task (second chance)
  - Simplification
- Components evaluation
  - Question classification
  - Passage retrieval
  - Answer extraction
- Pilots
  - Repeat existing?
  - New exercises
- 2007 exercises -> 2008?
- Multilinguality
- NILs, types of questions
- Vision
- ...