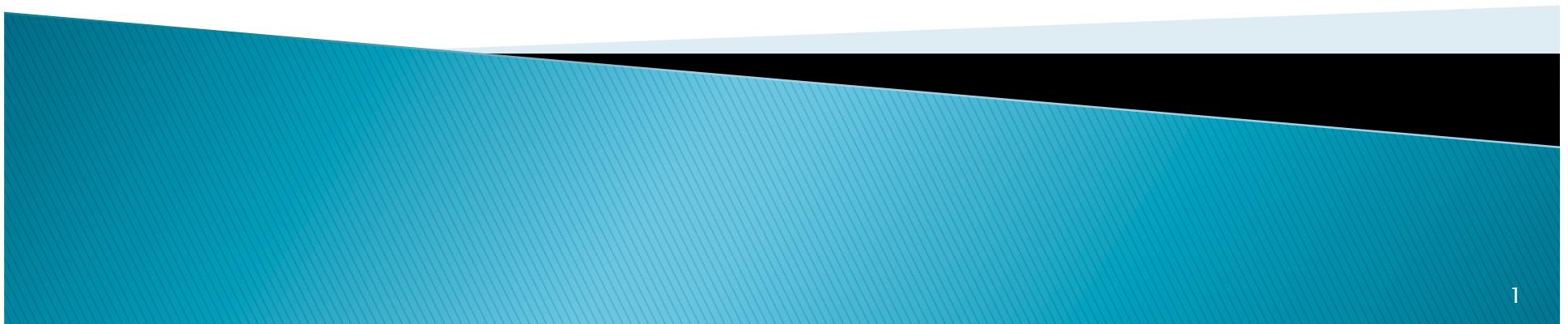


Course 4

Evaluation metrics and Question–Answering systems

Slides offered by Adrian Iftene



Content

- ▶ NLP Systems Evaluation
- ▶ Information Retrieval
- ▶ Information Extraction
- ▶ Question Answering
 - Introduction
 - System components
 - Background knowledge indexing
 - Index creation and information retrieval
 - Answer extraction
 - Results
 - Error analysis
- ▶ Conclusions

NLP Systems need evaluation

- ▶ “An important recent development in NLP has been the use of much more rigorous standards for the evaluation of NLP systems”
Manning and Schutze

- ▶ To be published, all research must:
 - establish a **baseline**, and
 - quantitatively show that it improves on the baseline **and the state-of-the-art**



NLP Systems – ways of doing evaluation

- ▶ “How well does the system work?”
- ▶ Possible domains for evaluation
 - Processing time of the system
 - Space usage of the system
 - Human satisfaction
 - Correctness of results
- ▶ Measures: (Accuracy, Error), (Precision, Recall, F-measure)



Accuracy and Error

- ▶ Comparing the output of the system with a gold standard we can verify what is correct
- ▶ The results of a system are marked as:
 - Correct: matches the gold standard
 - Incorrect: otherwise

$$\text{accuracy} = \frac{\text{number correct}}{\text{number of results}}$$

$$\text{error} = \frac{\text{number wrong}}{\text{number of results}}$$



Accuracy and Error – Example

“مساء الخير” → Arabic ✓

“Bon soir!” → French ✓

“Good evening!” → Spanish ✗

- ▶ Accuracy = 66.66 %
- ▶ Error = 33.33 %



Precision and Recall

- ▶ Precision and Recall are set-based measures
- ▶ They evaluate the quality of some set membership, based on a reference set membership
- ▶ **Precision:** what proportion of the *retrieved documents* is relevant?

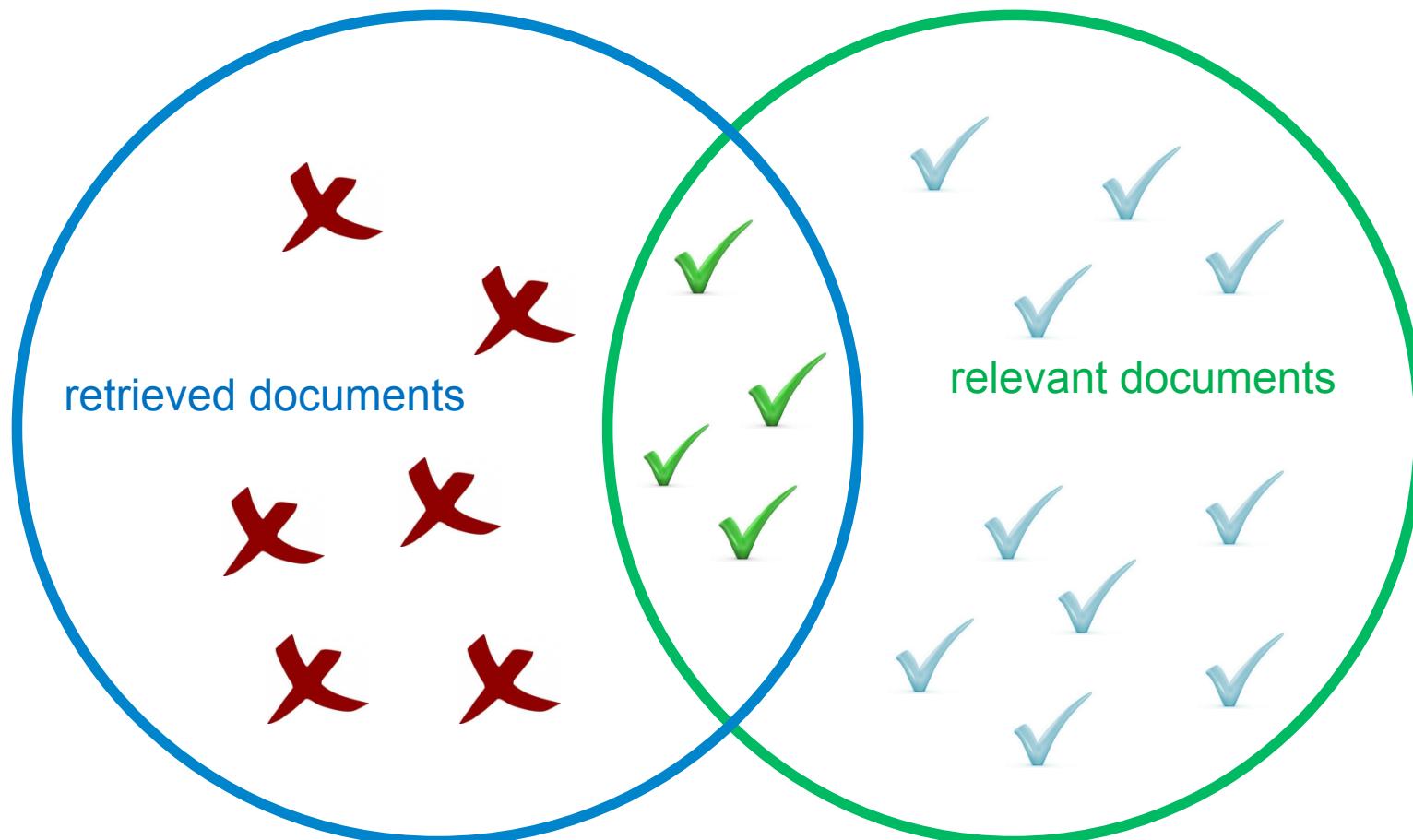
$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

- ▶ **Recall:** what proportion of the *relevant documents* is retrieved?

$$\text{recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}$$



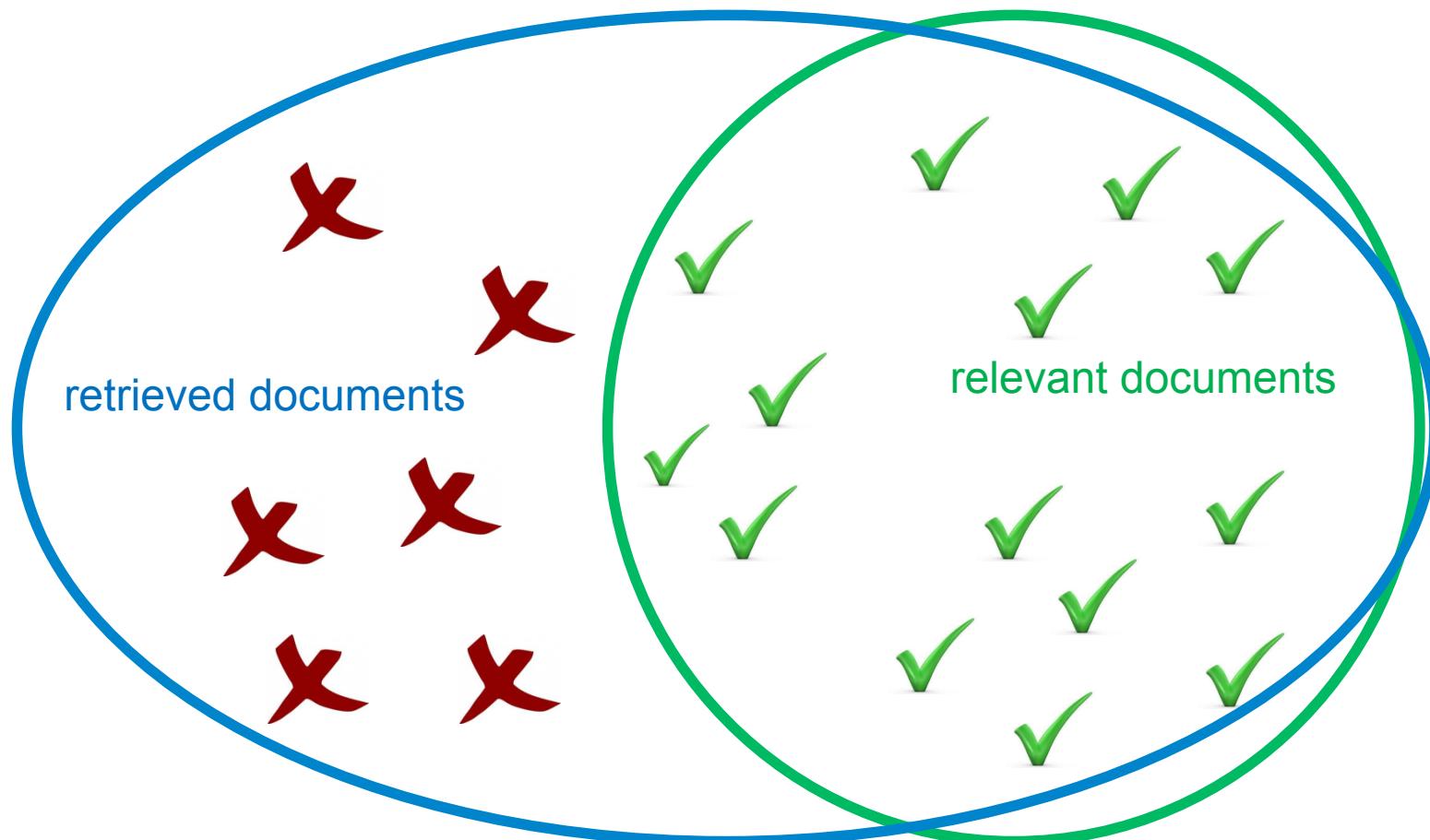
Precision and Recall - Example (1)



$$\text{Precision} = 4 / 10 = 40 \%$$

$$\text{Recall} = 4 / 14 = 28.57 \%$$

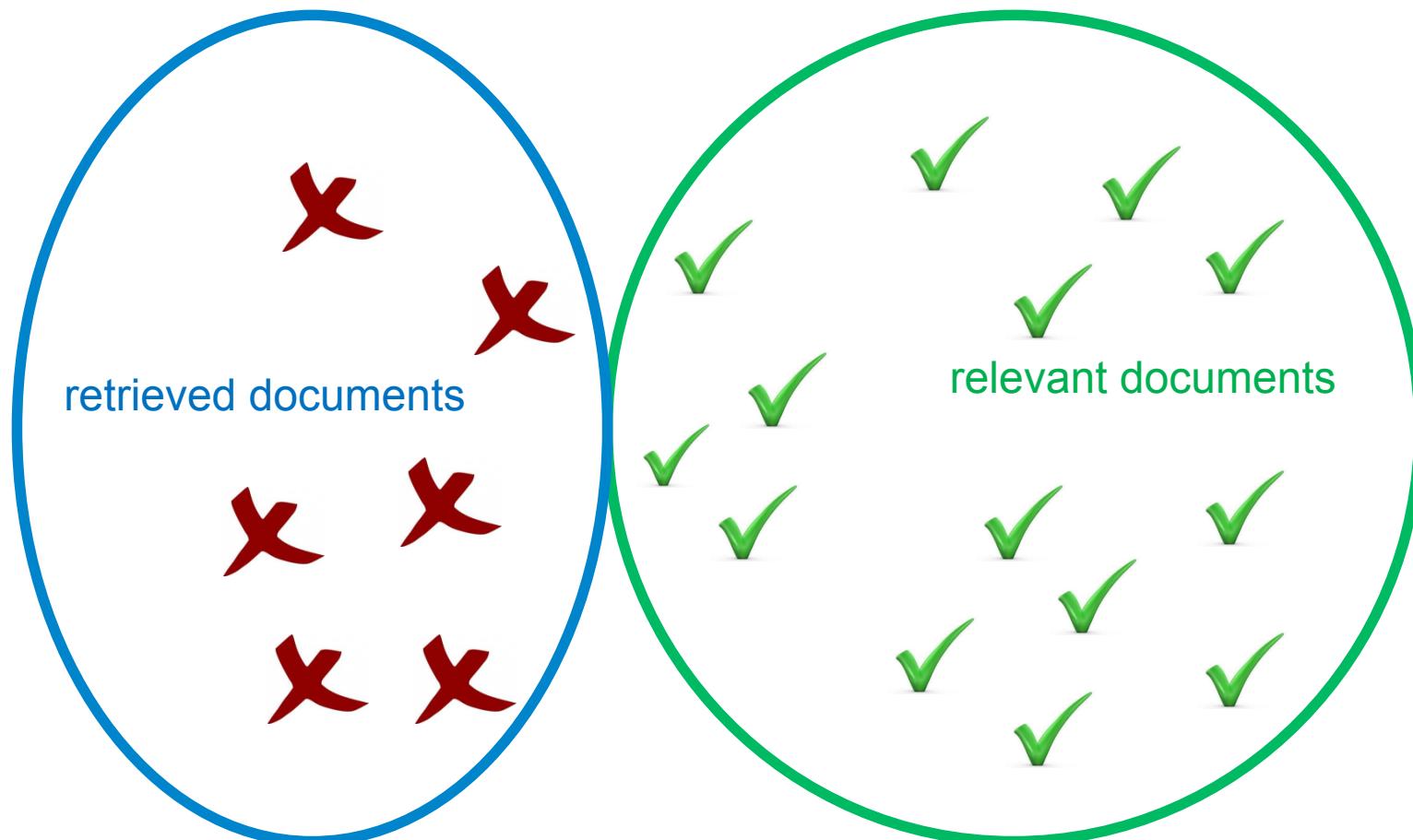
Precision and Recall - Example (2)



$$\text{Precision} = 14 / 20 = 70 \%$$

$$\text{Recall} = 14 / 14 = 100 \%$$

Precision and Recall - Example (3)



$$\text{Precision} = 0 / 6 = 0 \%$$

$$\text{Recall} = 0 / 14 = 0 \%$$

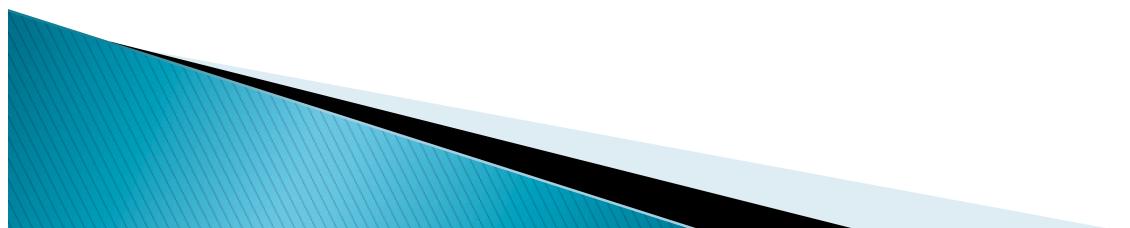
F-measure (F-score or F1-score)

- ▶ F-measure is a measure of a test's accuracy, and it considers both the precision p and the recall r
- ▶ General formula:
$$F_\beta = \frac{(1 + \beta^2) \cdot (\text{precision} \cdot \text{recall})}{(\beta^2 \cdot \text{precision} + \text{recall})}$$
- ▶ F1-measure:
$$F = \frac{2 \cdot \text{precision} \cdot \text{recall}}{(\text{precision} + \text{recall})}.$$
- ▶ F2-measure = ?



Question Answering – Definition

- ▶ Question Answering (QA): *a QA system takes as input a question in natural language and produces one or more ranked answers from a collection of documents.*



Question Answering – Modules

- ▶ QA systems normally adhere to the pipeline architecture composed of three main modules (Harabagiu and Moldovan, 2003):
 - **question analysis** – *the results are keywords, answer and question type, focus*
 - **paragraph retrieval** – *the results are a set of relevant candidate paragraphs/sentences from the document collection*
 - **answer extraction** – *the results are a set of candidate answers ranked using likelihood measures*

QA – Question Type

- ▶ Harabagiu and Moldovan, 2003:
 - Factoid – “*Who discovered the oxygen?*”, “*When did Hawaii become a state?*” or “*What football team won the World Cup in 1992?*”
 - List – “*What countries export oil?*” or “*What are the regions preferred by the Americans for holidays?*”.
 - Definition – “*What is a quasar?*” or “*What is a question–answering system?*”
 - *How, Why*, hypothetical, semantically constrained, polar (Yes/No) and cross–lingual questions

QA – Answer Type

- ▶ **Person** – "What", "Who", "Whom", "With who"
- ▶ **Location** (City, Country, and Region) – "What state/city", "From where", "Where"
- ▶ **Organization** – "Who produced", "Who made"
- ▶ **Temporal** (Date and Year) – "When"
- ▶ **Measure** (Length, Surface and Other) – "How much"
- ▶ **Count** – "How many"
- ▶ **Yes/No** – "Did the girl fear that?", "Is the car blue?"

QA – Search collection

- ▶ Local collections, internal organization documents, newspapers, Internet
- ▶ **Closed-domain** – deals with questions from a specific domain (medical, baseball, etc.). Can exploit domain-specific knowledge (**ontologies**, rules, disambiguation)
- ▶ **Open-domain** – general question about anything. Can use general knowledge about the world

Examples of such systems

- ▶ The first QA systems have been created in the 60s:
 - **BASEBALL** (Green 1963) – answer questions about baseball games
- ▶ **LUNAR** (Woods, 1977) – geological analysis of rocks returned by the Apollo moon missions
- ▶ **IURES** (Cristea, Tufiş, Mihăiescu, 1985) – medical domain, **querying national programs library**
- ▶ **QUERNAL** (Cristea, Tufiş, 1987) – personnel database, drilling and extraction, metallurgy, geography

QA – Systems of today

- ▶ Powerset: <http://www.powerset.com/> (<http://www.bing.com/>)
- ▶ Assimov the chat bot: <http://talkingrobot.org/b/>
- ▶ AnswerBus: <http://www.answerbus.com/index.shtml>
- ▶ NSIR: <http://tangra.si.umich.edu/clair/NSIR/html/nsir.cgi>
- ▶ START (The first question answering system):
<http://start.csail.mit.edu/>



Asimov

OUR CHATBOTS

[Asimov the robot](#)

[Data the android](#)

[Trinity](#)

[More chatbots](#)

ABOUT CHATBOTS

[Home.](#)

We create chatbots.

[The Present.](#)

About todays chatbots.

[The Future.](#)

The future of chatbots.

[Get a Chatbot.](#)

Get a chatbot for your website.

ABOUT ASIMOV

[Help on Asimov](#)

[Demo transcript](#)

[Asimov FAQ](#)

Hello Adrian.
My name is Asimov.

> **Hello Asimov. How are you?**

Hello Adrian.

I am fine, thank you.

I am always the same. I don't have feelings. I have no feelings or moods.

Submit

NSIR - Question Answering System - Mozilla Firefox

File Edit View History Bookmarks Tools Help

NSIR - Question Answering Syst... +

clair.si.umich.edu/clair/NSIR/html/nsir.cgi?question=Who+is+Tom+Cruise+married+to%3F&search_engine=Google_snippet&n_documents=50

Most Visited Getting Started Latest Headlines LITORAL 2010 ROMANI... Adevarul

Google

NSIR

Question Answering System

:: Home :: Help :: About NSIR :: Contact Us ::

Sample Questions

1. [What year did Wilt Chamberlain score 100 points?](#)
2. [How many chromosomes does a human zygote have?](#)
3. [Who is Tom Cruise married to?](#)
4. [What is South Africa's total population?](#)
5. [What is Canada's most populous city?](#)
6. [Which country exports the most tea?](#)
7. [What state has the most Indians?](#)
8. [George Bush purchased a small interest in which baseball team?](#)
9. [What county is Modesto, California in?](#)
10. [Who is the author of the book, 'The Iron Lady: A Biography of Margaret Thatcher'?](#)

Type your question below
or choose from [TREC-8](#) [TREC-9](#) [TREC-2001](#) [TREC-2002](#)

Who is Tom Cruise married to?

Search Engine / Index: [help](#) Google-snippet (faster but less accurate) ▾
No. of Hits: [help](#) 50 ▾
Answer Length: [help](#) Phrases ▾
No. of Answers: [help](#) 10 ▾
Answer [help](#) (for eval; use "" to delimit multiple answers)
Cached results: [help](#) Nicole Kidman optional
Debug information: [help](#) Check to use cache if there is cached result
 Check to see the detailed scores

Get Answer

This is a PERSON type question

Ranked Answers (click to see context)

1) [\(What are these numbers?\)](#) FARR 1 FHS 1 TRDR 1 TTRDR 1 PREC 1 MRR 1 TRWR 1 NTRWR 1

Top Web Documents Retrieved By Google

20



question answering systems START



Most Visited Getting Started Latest Headlines LITORAL 2010 ROMANI... Adevarul

AnswerBus

Ask

Type in your question! We only want to offer the best possible answers to you.

Possible Answers

[Mas d'Augustine Luxury Bed & Breakfast In The South Of France](#) www.masdaugustine.com

[Tichy Racing Camps](#) The best all inclusive summer ski race camps in North America www.tichyracing.com

[Teen Climbing Camp](#) Ages 13 - 18, Intro to Advanced AMGA and PCIA Accredited climbacadia.org/

[AdChoices ▾](#)

[Bread Recipes](#) Free e-Cook Book Explains Secrets to Making Bread - View Now TotalRecipeSearch.com

[Horizon Summer Camps](#) Learn English in Unique Interaction Residential Camps w/ Irish Children www.horizonireland.com

[Portable Living Units](#) Portable Rig Camps, Residential ,Commercial (Full Catering avail) maps.google.ca

[AdChoices ▾](#)

[How do i make french toast with only Egg And bit of milk and bread?](#)

...my first time **making french toast** can someone maybe put it step by step on **how to make** it and also **I** don't Have Vanilla something... is (**Sugar, Milk, Bread, and eggs,** and butter) that's it so **how do i make** it like that Please help...

[How many calories are on a piece of French toast?](#)

I'm **doing** a school project to see **how** many calories I eat a day and **how** much...this is the **only** thing I cant find! Its normal white **bread** dipped in **eggs** and a **skim milk** and cooked on a pan with a **bit** of butter.. my mom **makes** it so **I** don't know...

[Did I give a good answer to this question??](#)

...with lots of butter **Milk** is really good but **make** sure its 2% or less **Lunch**...**make** sure again its on wheat **bread** and try to add other ...terrible, cut them out of your diet or at least don...have if its good **Cereal** and **toast** is good like I said above if its...and you don't load it **with** lots of butter What... you getting? **Eggs** are good 2 but **how** are you cooking...up at 5:45 to exercize, **do** you have no time to exercize...a week **Pancakes, French Toast, and Waffles**...have to have them **only** have a little **bit** Dinner Subway is good...

More Resources from Wiki and Web

x Find: START

↓ Next ↑ Previous ⌂ Highlight all ☐ Match case

The image cannot be displayed. Your computer's security settings may prevent the image from being displayed. To view this image, you may need to change your security settings or disable your browser's image viewer. You may have to delete the image and then insert it again.

The START Natural Language Question Answering System - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://start.csail.mit.edu question answerin

Most Visited Getting Started Latest Headlines Crazy Go Nuts | Game... WAV to MP3 Encoder http://www.totalcmd....

Foxit Search Launch Foxit Messages Foxit Online Services

START

Natural Language Question Answering System

Who is Ion Iliescu?

Ask Question Clear

START, the world's first Web-based question answering system, has been on-line and continuously operating since December, 1993. It has been developed by Boris Katz and his associates of the InfoLab Group at the MIT Computer Science and Artificial Intelligence Laboratory. Unlike information retrieval systems (e.g., search engines), START aims to supply users with "just the right information," instead of merely providing a list of hits. Currently, the system can answer millions of English questions about places (e.g., cities, countries, lakes, coordinates, weather, maps, demographics, political and economic systems), movies (e.g., titles, actors, directors), people (e.g., birth dates, biographies), dictionary definitions, and much, much more. Below is a list of some of the things START knows about, with example questions. You can type your question above or select from the following examples.

Geography

- [What South-American country has the largest](#)

Science and Reference

- [What is Jupiter's atmosphere made of?](#)

Find: iure

Next Previous Highlight all Match case

Done Now: Partly Sunny, 17° C Tue: 33° C Wed: 26° C

22

QA – Competitions

- ▶ CLEF (Cross Language Evaluation Forum) – started in 2000 –
<http://www.clef-campaign.org/> European languages in both monolingual and cross-language contexts
 - Coordination: Istituto di Scienza e Tecnologie dell'Informazione, Pisa, Italy
 - Romanian Institute for Computer Science, Romania
- ▶ TREC (Text REtrieval Conference) – started in 1992 <http://trec.nist.gov/>
National Institute of Standards and Technology (NIST), Gaithersburg, Maryland, USA

- <test-set>

- <topic t_id="1" t_name="AIDS">
- <reading-test r_id="1">
- <doc d_id="1">

Annie Lennox: De ce sunt activist împotriva HIV/SIDA Vă voi spune povestea transformării mele într-un activist împotriva HIV/SIDA. Iar acesta este numele campaniei mele, Campania SING [Cântă]. În noile fundații să împotriva HIV/SIDA. Iar 46664 este numărul pe care îl avea Mandela în închisoarea din Robben Island. Aici sunt cu Youssou N'Dour, pe scenă, distrânându-mă de minune. A doua zi, toți artiștii au făcut foste celule de închisoare. Puteți vedea grădina de la fereastra aici. A fost un eveniment plin de semnificație pentru noi toți. În acel moment Mandela a spus presei mondiale că are loc aproape un genocid mai vulnerabile victime, erau femeile și copiii. Asta a avut un impact uriaș asupra mintii mele, deoarece sunt femeie și sunt mamă și nu mi-am dat seama că pandemia HIV/SIDA afectă direct femeile în acest fel. Ceva despre care trebuie să vorbesc. Trebuie să servesc această cauză." Si astfel, în consecință, am participat la fiecare eveniment al fundației 46664 la care am putut lua parte și am organizat conferințe de respect pentru extraordinara, incredibilă muncă pe care a făcut-o. Toată lumea îl respectă pe Nelson Mandela. Toată lumea îl idolatrizează pe Nelson Mandela. Dar știu ei despre ce se întâmplă în Africa de Sud și să spune oamenilor ce se întâmplă acolo, ar fi șocați. Am fost foarte, foarte norocoasă, câțiva ani mai târziu, să-l întâlnesc pe Zackie Achmat, fondatorul Treatment Action Campaign, un activist incredibil. Îi spune că sunt solidară cu oamenii care au HIV, oamenii care trăiesc cu HIV. Si într-un fel, din cauza stigmatului, prin purtarea acestui tricou, eu spun "Da, putem discuta despre această problemă." Nu trebue să credeți că acestei organizații incredibile. Este o campanie ce se bazează pe comunitate cu 80% dintre membri fiind femei, majoritatea fiind infectate cu HIV. Lucrează pe teren. Au un impact extraordinar asupra oamenilor extraordinari ceea ce fac. Și da, campania mea SING a sprijinit Treatment Action Campaign în sensul că am încercat să conștientizez oamenii și am încercat să strâng fonduri. O mulțime din banii pe care am încercat să încearcă să o facă, în Africa de Sud. Așadar aceasta este campania mea SING. Campania SING este de fapt formată din mine și încă trei sau patru oameni minunati care mă ajută și mă susțin. Am călătorit Norvegia, primind un cec consistent, când în Hong Kong, încercând să-i conving pe oameni să doneze bani. În Johannesburg am avut ocazia să cânt în fața unei audiente majoritar albe din clasa mijlocie ca oamenii au tendință să o evite, pentru că sunt plăcitiști, și nu știu de fapt care sunt soluțiile. Aaron Motsoaledi, ministru sănătății aflat în funcție, a participat la acel concert, și am avut ocazia să-l întâlnesc, iar devenit ulterior o reprezentanta a Scoției și a HIV. Si le-am prezentat experiențele mele în încercarea de a-i face conștiință de ce se întâmplă. Si din nou, în Edinburgh, împreună cu minunatul cor al copiilor afecțuată de virusul SIDA. Aici sunt în New York cu Michel Sidibe. El e directorul UNAIDS. Sunt foarte onorată de faptul că Michel m-a invitat doar cu câteva luni în urmă să devin ambasador UNAIDS. Si în lume este că am vrea să vedem eliminarea propriu-zisă a transmiterii virusului de la mamă la copil până în anul 2015. Este un obiectiv foarte ambicios, dar credem că poate fi atins cu voine politice. Este posibil să suntem foarte încrezătoare, pentru că știm că această tânără este tratată astfel încât să își prelungă viața pentru a avea grija de bebelușul pe care îl va naște. Iar bebelușul său va fi tratat cu PMTCT, ceea ce poate să împreună cu modalitățile de intervenție pentru stoparea pandemiei SIDA. Aș dori să încearcă să spună povestea despre Avelile. Ea este Avelile. Merge cu mine oriunde mă duc. Spun povestea ei tuturor, pentru că ea reprezintă complementare SIDA. Avelile avea virusul. S-a născut cu virusul. Si aici are 7 ani, cândăldupă nu mai mult decât un bebeluș de 1 an. În acest moment al vietii sale, ea suferă de SIDA în plină dezvoltare și are probleme de sănătate. Doctorii și asistențele erau fenomenali. Au trecurt-o pe un regim nutritiv foarte special și au avut mare grija de ea. Si atunci când am plecat de la spital nu știam -- pentru că am filmat atunci povestea ei -- nu știau foarte multe de această experiență directă, de acest copil, știu, acea poveste. Cinci luni mai târziu ne-am întors în Africa de Sud pentru a o reîntâlni pe Avelile. Si devin -- firele de păr de păr -- nu știu dacă poate să aibă loc. Nu este extraordinar? (Aplauze) Aceste aplauze sunt de fapt pentru doctorii și asistențele din spitalul în care au avut grija de Avelile. Si înțeleg că apreciază acest tip de transformare. Așadar, aş dori să le mulțumesc mult. (Aplauze)

</doc>

- <q q_id="1">
- <q_str>

Ce eveniment a determinat-o pe Annie Lennox să se angajeze în lupta contra SIDA?

</q_str>

- <answer a_id="1">detenția lui Nelson Mandela la Robben Island</answer>
- <answer a_id="2">

ceremonia de închidere a fundației lui Nelson Mandela

</answer>

- <answer a_id="3">întâlnirea cu Youssou N'Dour</answer>
- <answer a_id="4">segregarea rasială din Africa de Sud</answer>
- <answer a_id="5">

conferința lui Nelson Mandela pentru presa mondială

</answer>

</q>

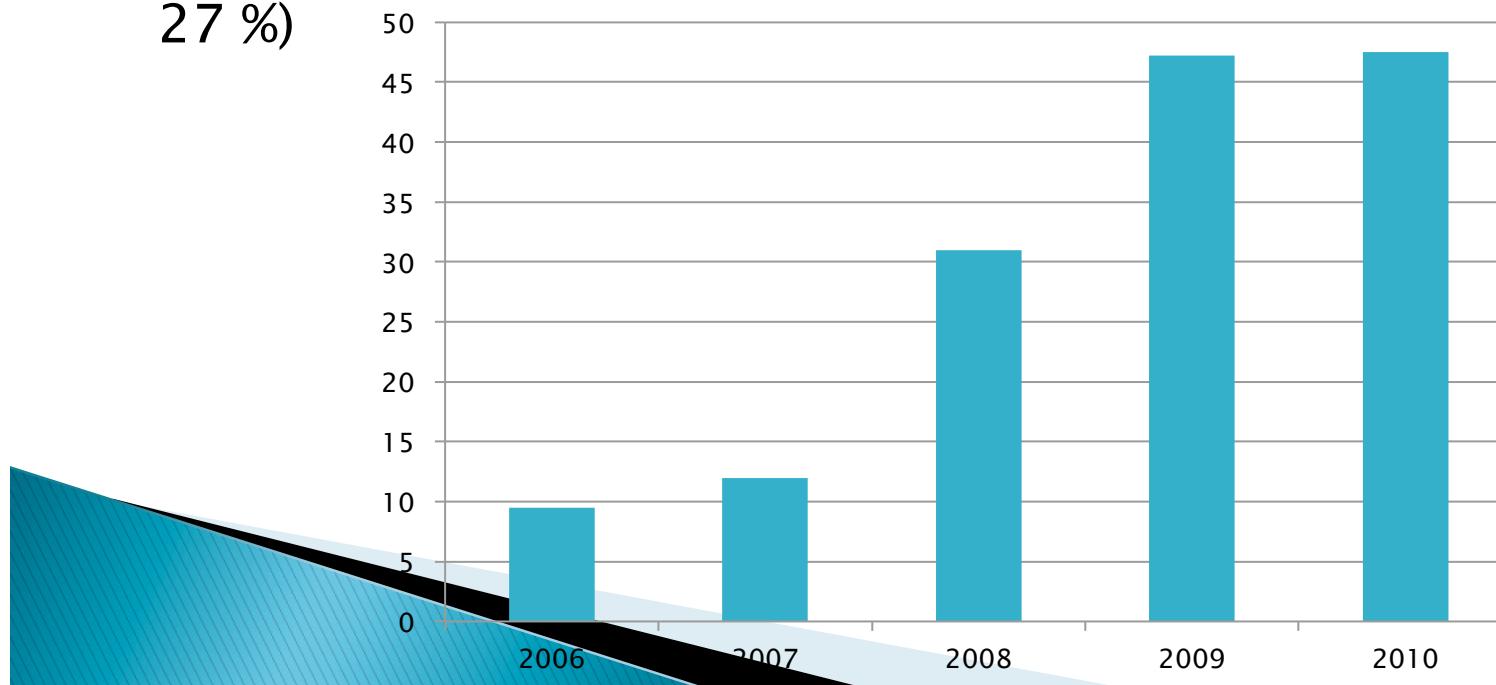
- <q q_id="2">
- <q_str>Cine este fondatorul campaniei SING?</q_str>

- <answer a_id="1">Nelson Mandela</answer>
- <answer a_id="2">Youssou N'Dour</answer>
- <answer a_id="3">Michel Sidibe</answer>
- <answer a_id="4">Zackie Achmat</answer>

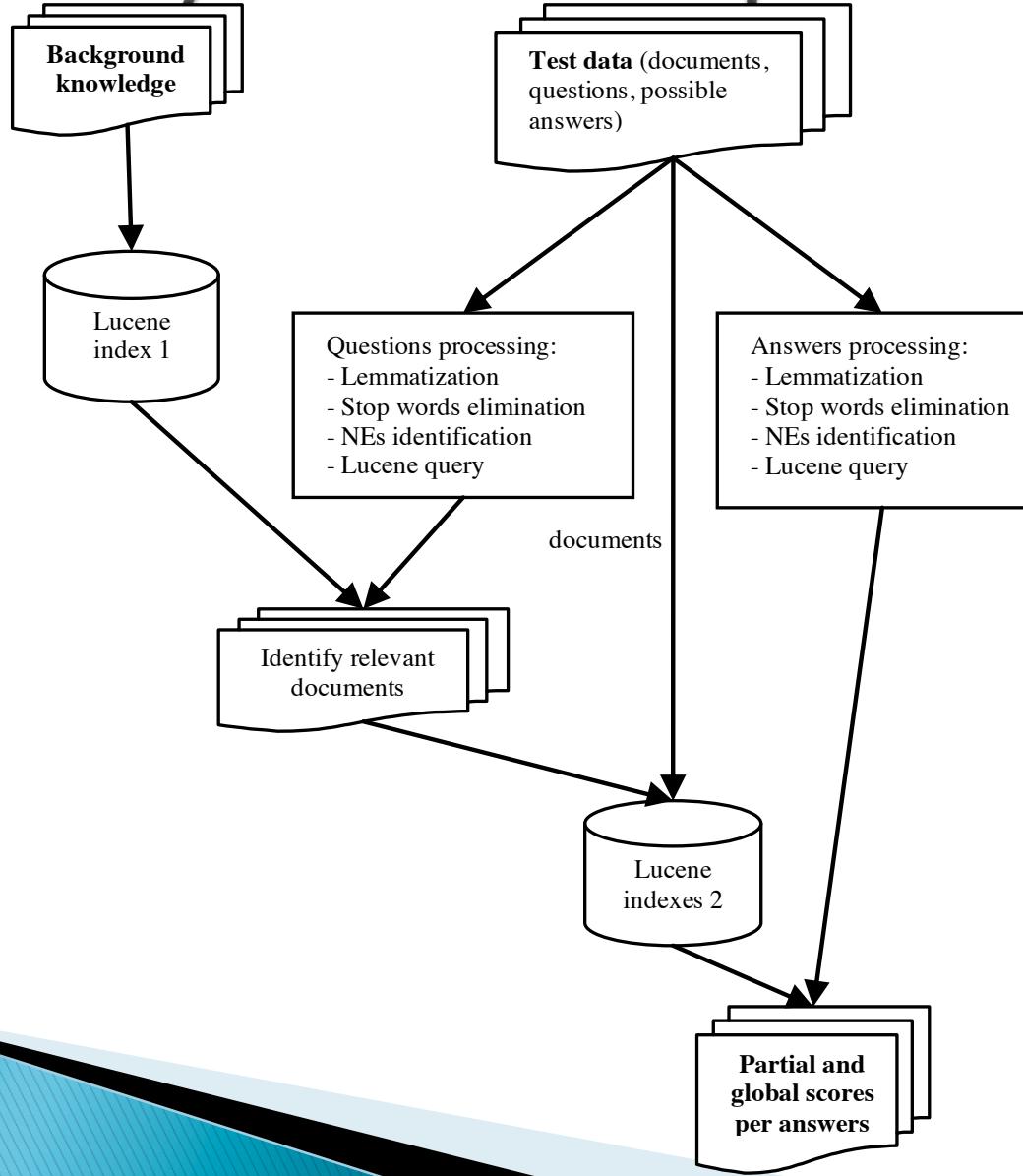
An excerpt from the gold standard dataset

UAIC System – CLEF 2011

- ▶ Our group participates at CLEF exercises since 2006:
 - 2006 – Ro-En (English collection) – 9.47% right answers
 - 2007 – Ro-Ro (Romanian Wikipedia) – 12 %
 - 2008 – Ro-Ro (Romanian Wikipedia) – 31 %
 - 2009 – Ro-Ro, En-En (JRC-Acquis) – 47.2 % (48.6%)
 - 2010 – Ro-Ro, En-En, Fr-Fr (JRC-Acquis, Europarl) – 47.5% (42.5%, 27 %)



UAIC System components



Background knowledge indexing

- ▶ The Romanian background knowledge has 161,279 documents in text format
 - 25,033 correspond to the AIDS topic
 - 51,130 to Climate Change topic
 - 85,116 to Music and Society topic
- ▶ The indexing component considers the name of the file and the text from it => **Lucene index 1**

```
CLIMATE_CHANGE_RO_9910.txt
1
2 România a intrat în criza alimentară
3 La jumătatea lui 2008, scumpirea alimentelor o depășea pentru prima dată după patru ani pe cea a mărfurilor
nealimentare sau a serviciilor, pe fondul deprecierii leului și importurilor masive de produse agricole. De atunci
până la începutul acestui an, serviciile și produsele nealimentare au lăsat în urmă alimentele din punctul de vedere
al creșterii prețurilor, însă trendul s-a inversat și acum suntem din nou în fața unui vârf al scumpirilor la alimente.
4
5 Mai citește și:
6
7 Pâinea este produs de lux pentru orădeni, iar producătorii amenință cu noi scumpiri
8
9
```

length : 3312 lines : 36 Ln:1 Col:1 Sel:0 Dos\Windows ANSI as UTF-8 INS 27

Test data processing

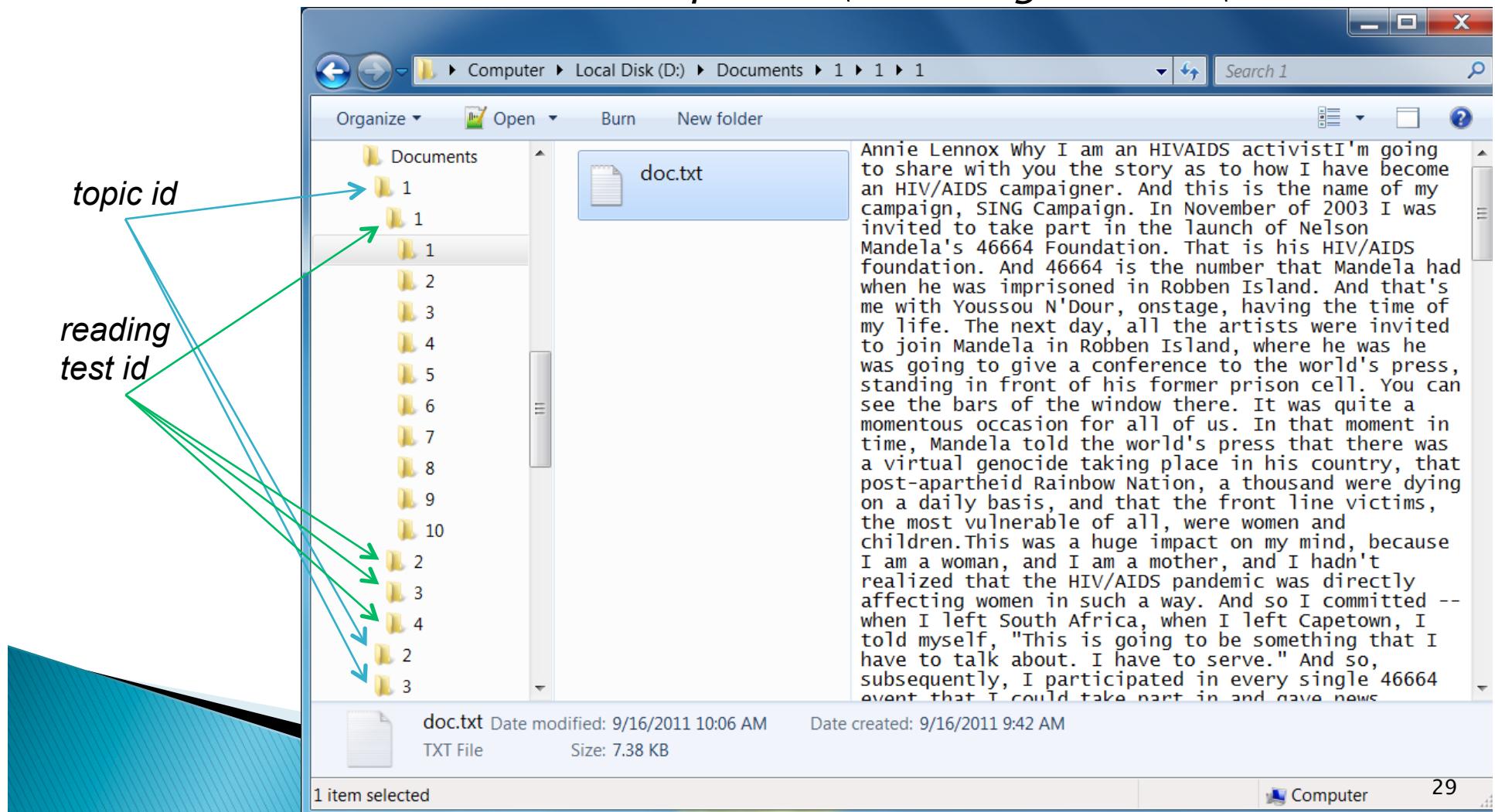
- ▶ Test data was an XML file with 12 test documents
 - 4 documents for each of the three topics (12 in total)
 - 10 questions for each document (120 in total)
 - 5 possible answers for each question (600 in total)
- ▶ *Test data processing* involved 3 operations:
 - extracting documents
 - processing questions
 - processing possible answers

```
en_testxml
2 <test-set>
3   <topic t_id="1" t_name="AIDS">
4     <reading-test r_id="1">
5       <doc d_id="1">
6         Annie Lennox Why I am an HIVAIDS activist ...
7       </doc>
8       <q q_id="1" >
9         <q_str>What event caused Annie Lennox to commit herself to the fight against AIDS?</q_str>
10        <answer a_id='1'>the imprisonment of Nelson Mandela at Robben Island</answer>
11        <answer a_id='2'>the closing ceremony of Nelson Mandela's Foundation</answer>
12        <answer a_id='3'>the meeting with Youssou N'Dour</answer>
13        <answer a_id='4'>the racial segregation in South Africa</answer>
14        <answer a_id='5'>Nelson Mandela's conference to the world press</answer>
15      </q>
16      <q q_id="2" >
17        <q_str>Who is the founder of the SING campaign?</q_str>
```

Volatile Markup Language file Length: 102612 Lines: 1020 Ls: 1 Col: 1 Sel: 0 UNIV ANNOTATE 0 INC

Test data processing - Extracting documents

- ▶ The content of `<doc>` => `<topic id>\<reading test id>\1..10`



Test data processing – Processing questions

- ▶ Stop words elimination
- ▶ Lemmatization
- ▶ Named Entity identification
- ▶ Lucene query building

```
<topic t_id="1" t_name="AIDS" r_id="2" q_id="3">
<string>Which African country did Bono Vox visit?</string>
<focus>country</focus>
<verb>visit </verb>
<noun></noun>
<adj></adj>
<nameEntities>African Bono Vox</nameEntities>
<query>African^3 country Bono^3 Vox^3 visit</query>
<questionType>FACTOID</questionType>
<answerType>OTHER</answerType>
```

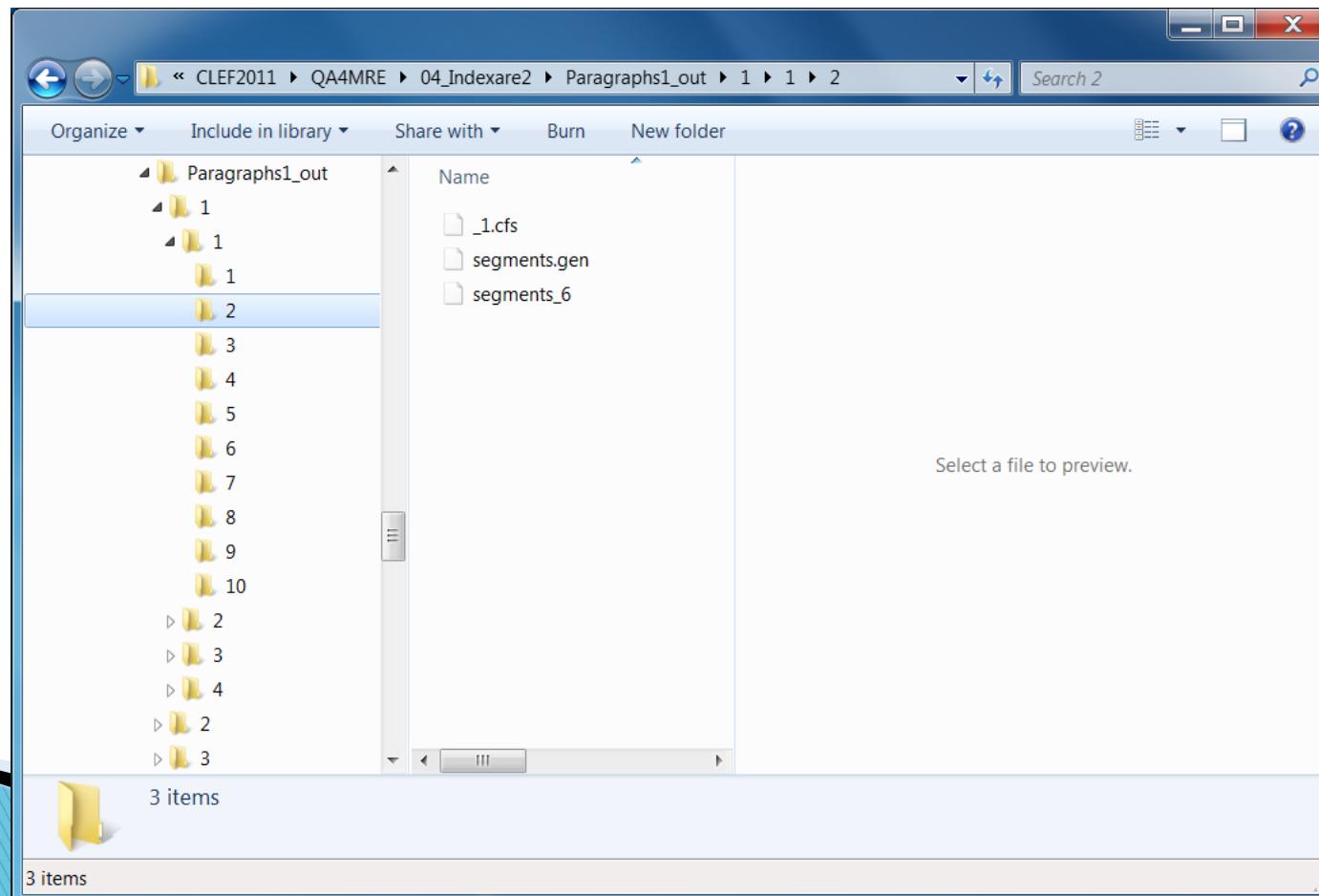
Test data processing – Processing possible answers

- ▶ Similar to processing questions +
- ▶ We use ontology (Iftene and Balahur, 2008) for elimination of possible answers with low probability to be final answer (relation [is_located_in])
- ▶ *In which European cities has Annie Lennox performed?*
- ▶ We eliminate from the list of possible answers the answers with non-European cities (we replace non-European cities with the value XXXXX)

```
<answer a_id='1'>London^3 Paris^3</answer>
<answer a_id='2'>XXXXX^3</answer>
<answer a_id='3'>XXXXX^3</answer>
<answer a_id='4'>Edinburgh^3 Oslo^3</answer>
<answer a_id='5'>XXXXX^3</answer>
```

Indexing relevant documents for questions

- For every question we index relevant documents returned by Lucene at the previous step and relevant documents saved from the initial test file



Searching using relevant documents for questions

- ▶ Then in every index, we performed searches using Lucene queries associated to possible answers
- ▶ For every answer, we obtained a list of documents with Lucene relevance scores
- ▶ $Score2(d, a)$ is the relevance score for document d when we search with the Lucene query associated to the answer a

```
<question q_id="1" r_id="1" t_id="1" a_id="1">
    <paragraph doc_id="doc.txt" score2="0.11814499"/>
</question>
<question q_id="1" r_id="1" t_id="1" a_id="2">
    <paragraph doc_id="doc.txt" score2="0.068914026"/>
</question>
<question q_id="1" r_id="1" t_id="1" a_id="3">
    <paragraph doc_id="doc.txt" score2="0.018048882"/>
</question>
<question q_id="1" r_id="1" t_id="1" a_id="4">
    <paragraph doc_id="doc.txt" score2="0.04478345"/>
</question>
<question q_id="1" r_id="1" t_id="1" a_id="5">
    <paragraph doc_id="doc.txt" score2="0.1216592"/>
    <paragraph doc_id="AIDS_RO_13800.txt" score2="0.0015860073"/>
</question>
```

Results 1

- ▶ Results of UAIC's runs at question answering level

	Ro-Ro			En-En
answered right	30	11	19	25
answered wrong	85	19	43	47
total answered	115	30	62	72
unanswered right	0	19	11	12
unanswered wrong	0	66	42	34
unanswered empty	5	5	5	2
total unanswered	5	90	58	48
Overall accuracy	0.25	0.09	0.16	0.21
C@1 measure	0.26	0.16	0.23	0.29

Links

- ▶ Yes-no question: [http://en.wikipedia.org/wiki/Yes%E2%80%93no_question](http://en.wikipedia.org/wiki/Yes/no_question)
- ▶ Question Answering: http://en.wikipedia.org/wiki/Question_answering
- ▶ Lecture 13: Evaluation: Precision and Recall
<http://courses.washington.edu/ling473/Lecture13.pdf>
- ▶ Precision and Recall of Five Search Engines for Retrieval of Scholarly Information in the Field of Biotechnology:
<http://www.webology.org/2005/v2n2/a12.html>