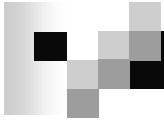




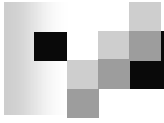
iClef-2002 at Universities of Alicante and Jaen



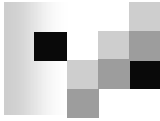
University of Alicante (Spain)



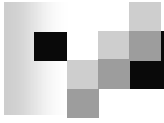
- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions and Future works



- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions and Future works



- Experiments
- We use a Passage Retrieval system (IR-n system).
 - The system only shows the most relevant passage of each document instead of whole document.
- The relevant passages presented to the user have been automatically translated into Spanish by means of Systran.
- Moreover, the results are compared with Z-Prise system, which is based on the reading of the whole document.



- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions

■ Passage Retrieval Systems

- Use short fragments of documents instead of whole documents to evaluate the relevance or similarity
- These fragments are called passages
- Each document is divided into passages before calculating the relevance

■ Passage Retrieval Advantages

- Consider proximity between keywords
- Consider high relevant fragments embedded in long low-relevant documents
- Can find and detect the relevant fragment of the document

IR system based in whole document

The death of General Custer

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las **death** que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información

General Custer que almacenar sobre ella.

Same
relevance

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la muerte que **Death** y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, **General Custer** que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

IR system based on passages

The death of General Custer

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las death que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos <u>simples, de tal forma</u> que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.
Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos <u>simples, de tal forma</u> que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.
Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos <u>simples, de tal forma</u> que nos facilitará enormemente el tratamiento de la información General Custer que almacenar sobre ella.

More relevant



Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos <u>simples, de tal forma</u> que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.
Una estructura de datos es una colección de datos que se puede definir en base a la muerte que death y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos General Custer <u>simples,</u> que nos facilitará <u>enormemente el</u> tratamiento de la información que tengamos que almacenar sobre ella.
Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos <u>simples, de tal forma</u> que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

IR system based in whole document

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará

enormemente el tratamiento de la información

que almacenar sobre ella.

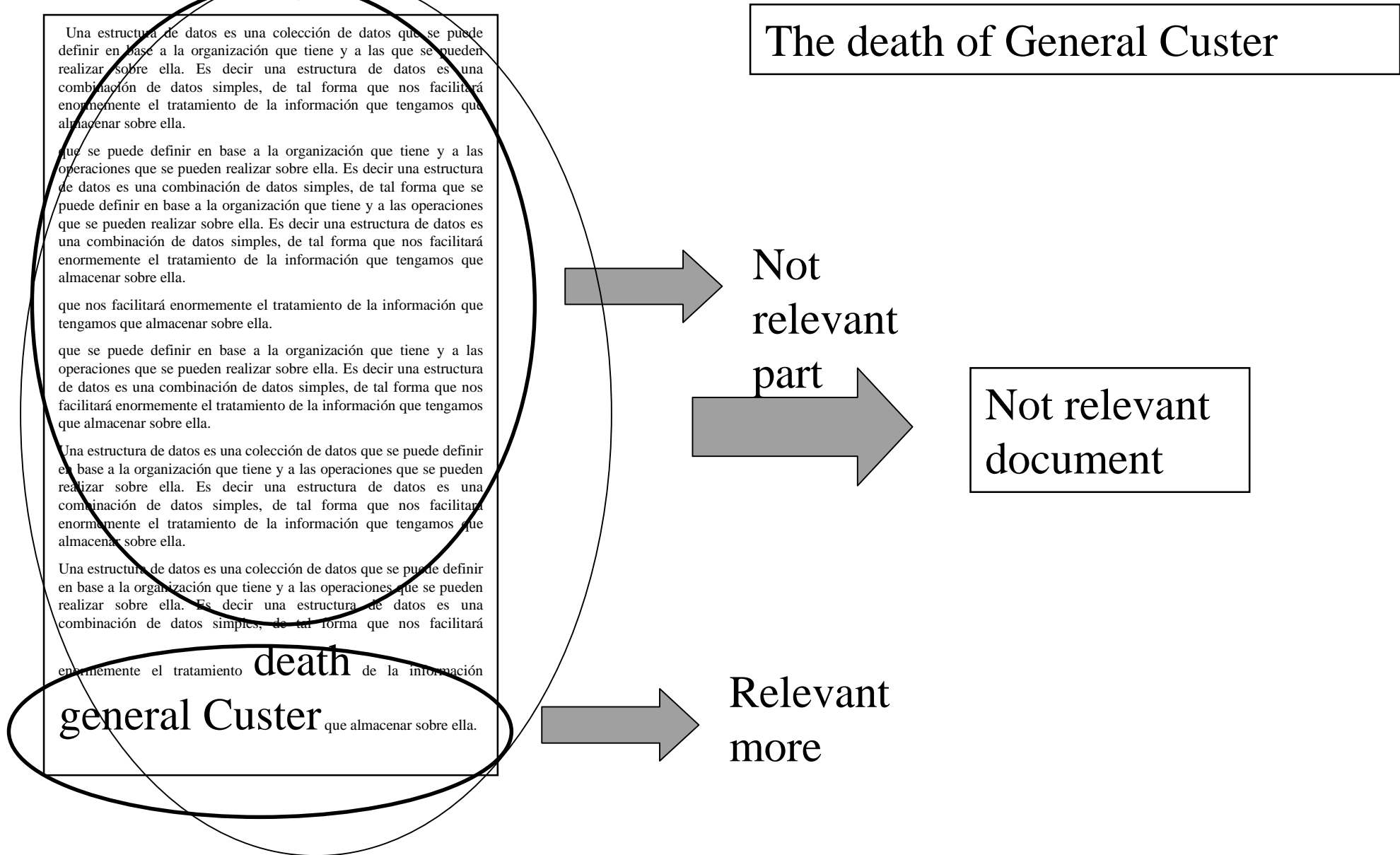
The death of General Custer

Not
relevant
part

Not relevant
document

Relevant
more

death
general Custer



IR system based on passages

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

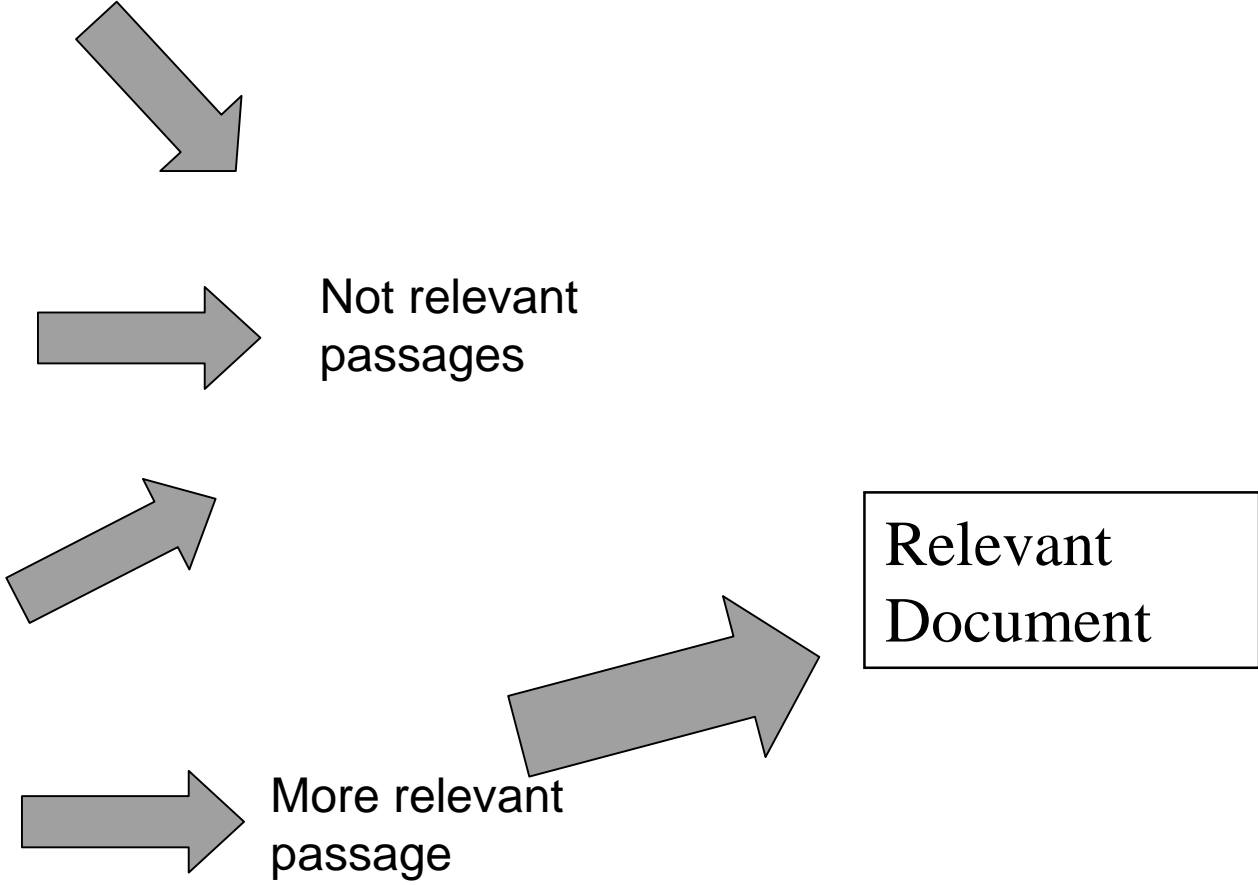
que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

enormemente el tratamiento **death** de la información **general Custer** que almacenar sobre ella.

The death of General Custer



IR system based in whole document

Relevant Document

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

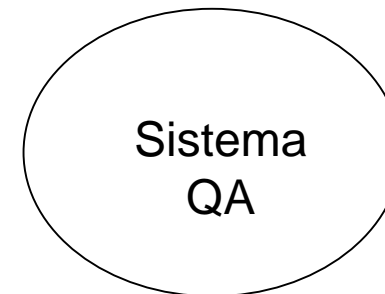
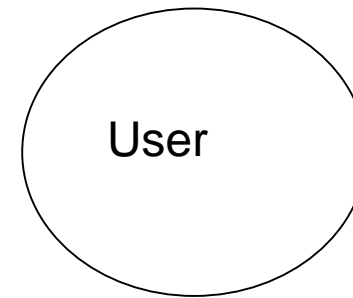
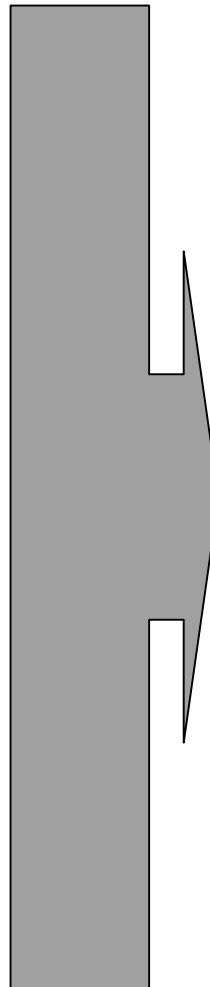
que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información general Custer que almacenar sobre ella.



IR system based on passages

RELEVANT PASSAGE

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

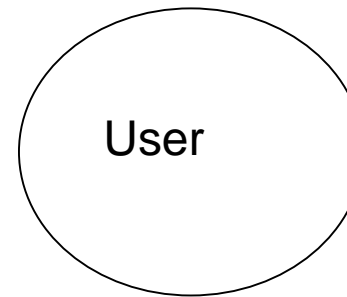
que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

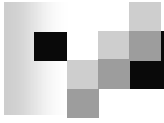
que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento muerte de la información general Custer que almacenar sobre ella.



But, Can a user determine if a document is relevant if only can see the more relevant passage ?



- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions and Future works

■ IR-n system

- Passage Retrieval system
- Its development began in year 2000
- It was presented at Clef'2001 conference (Llopis & Vicedo 2001) in monolingual task (Spanish) and bilingual task (English-Spanish)
- It was used in Question Answering task in Trec'2001 conference (Vicedo et al 2001), to select the more relevant passages from a collection, to make easier the Question Answering process

■ IR-n system

- A document is divided into passages formed by a n sentences.
- The number of sentences depends on collection type and query type.
- This number can be fixed after evaluating the query
- The system uses overlapping windows
- The system does not use normalization measures of document or passage size.

- A document is divided into passages formed by n sentences.
 1.
 - A sentence usually represents a “idea” in the document.
 - While the paragraphs can be used due to visual appearance.
 2.
 - A sentence is a complete unit that subsequently can be treated by another system
 - While a passage that begins in any word can lose the structure of the document.



- The system uses overlapping windows

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las death que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información General Custer que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las death que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información que tengamos que almacenar sobre ella.

Una estructura de datos es una colección de datos que se puede definir en base a la organización que tiene y a las operaciones que se pueden realizar sobre ella. Es decir una estructura de datos es una combinación de datos simples, de tal forma que nos facilitará enormemente el tratamiento de la información General Custer que almacenar sobre ella.

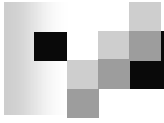
1st Passage. Sentences from 1 to 15

2nd Passage. Sentences from 2 to 16

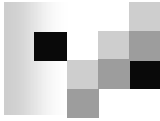
3rd Passage. Sentences from 3 to 17



- Experimentally we obtain better results overlapping the windows by sentence than when overlapping is not used or overlapping in more sentences is used.
- This increments the computational time, although not in great value because we consider that
 - 1st passage begins in the first sentence where a query term is found
 - Last passage finishes in the last sentence of the document where a query term is found

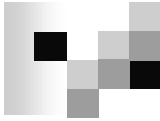


- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions and Future works



■ Objectives

- Try to show Correct Spanish Sentences
- Use a Passage Retrieval system to show results
or Can a user determine if a document is relevant,
if only can see the more relevant passage?
- Develop an interactive system via web

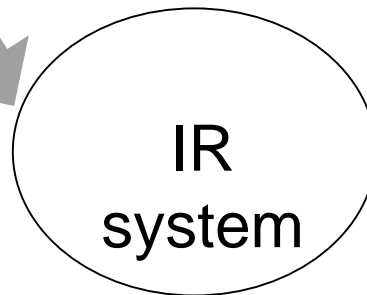


- We work in document selection task
- Available Document Collections
 - LA Times (English)
 - LA Times (Spanish Translated)
 - Efe (Spanish)
- Task
 - User asks in Spanish
 - System shows documents in Spanish
 - User must say if a document is relevant or not
 - The documents must be obtained from LA Times

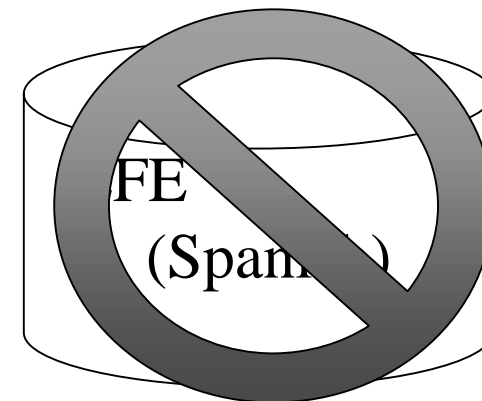


Answer in Spanish

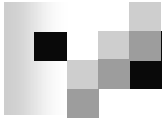
LA TIMES
(Spanish
Translated)



More relevant passage
in Spanish



Initial experiments
obtained bad results



- How can be shown English Documents automatically translated to Spanish without the user dies in the work?

IR-n system - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Atrás Búsqueda Favoritos Multimedia

Dirección http://gplsi.dlsi.ua.es/cgi-bin/llopis/buscar.cgi Ir Vinculos

PLSI IR-n system

Consulta Vista

LA112594-0043
GRUPO DE VIRUS
LA GENETICA
Encontraron es...
de Cocksackie pero también contra GAD.
Los ratones entonces desarrollaron la diabetes.
los ratones no-susceptibles, en contraste, desarrollaron anticuerpos solamente contra el viral la proteína y no desarrolló la diabetes.
Porque los genes que controlan susceptibilidad de la diabetes en ratones son similares a éstos en los seres humanos, Kaufman dicho, la observación sugieren que los virus de Cocksackie puede accionar la enfermedad en seres humanos también.
Los investigadores convinieron, sin embargo, que otros estudios serán necesarios a prueba realmente que los virus de Cocksackie causan la enfermedad.
Maclaren advirtió que los virus de Cocksackie puedan no ser la causa de todo el tipo 1 diabetes.
Cerca de uno en cada cinco diabéticos, él observó, no produce la inmunorespuesta a GAD, y la enfermedad en esos casos

Estoy bastante seguro

Relevante No relevante

The sentences are shown in different lines in the interface.

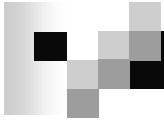
The following sentence after the title is the previous sentence to the most relevant passage



■ Results

Average Precision	IR-n	BASE (Z-prise)
Topic 1	0,4601	0,6371
Topic 2	0,8098	0,5925
Topic 3	0	0
Topic 4	0,7643	0,3748
Avg	0,5085	0,4011

	Average F_alpha
BASE (Z-Prise)	0.2166
IR-n	0.3248



- Introduction
- Passage Retrieval Systems
- IR-n system
- IR-n system at iClef-2002
- Conclusions

- The users find the automatic translations into Spanish quite unreadable most of the times (more than it was expected).
- It has been difficult to find users to carry out the experiments, which explains the reduction of the number of documents to study (only 25) for each query.

- We are quite happy with the obtained results, since the users have found a high percentage of relevant documents in not much time (an average between 8 and 9 minutes per query). This is because the piece of text that has to be read is only formed by nine sentences.

- We pretend to improve the automatic translations
- We have to improve the interactivity with the system adding to system Query expansion techniques (tested in Clef-2002)