

# OpenMT: Open Source Machine Translation Using Hybrid Methods

- Kepa Sarasola, Iñaki Alegria  
**Ixa Group. Euskal Herriko Unibertsitatea**
- Núria Castell, Lluís Màrquez  
**TALP, Universitat Politècnica de Catalunya**
- Nerea Areta, Xabier Saralegi  
**Elhuyar Fundazioa**

TIN2006-15307-C03

<http://ixa.si.ehu.es/openmt>

# Goals

- Defining open source architectures for machine translation.
- Combining different MT paradigms into a hybrid system.
  - Rule-Based MT (RBMT),
  - Statistical MT (SMT) and
  - Example-Based MT (EBMT).
- Integration of syntactic and semantic processing in machine translation.
- Evaluation.
  - A general open-source evaluation tool
  - Merging different criteria and using different languages and domains.

# Goals by subprojects

- Coordinator: IXA (EHU)
  - Combining different MT paradigms into a hybrid system
  - Enrich it with semantics captured by ML from EBMT system.
- TALP (UPC)
  - Statistical MT technology needed for the combined MT systems
    - Enrichment with syntactic-semantic information captured by ML
    - Linguistically-based transfer rules from EBMT
  - Evaluation work package, adaptation of IQMT for MT
- Elhuyar
  - Basic infrastructure for MT resources and tools
  - Evaluation
  - Exploitation and Dissemination

# Project Progress and Achievements

- **Basic tools and resources**
- Improving actual RBMT and SMT systems
- Combining EBMT and RBMT
- Hybrid architectures
- Advanced semantic processing for MT
- Evaluation and demonstration

# Basic tools and resources

- Web tool for managing and storing parallel corpora [51]
- Several new reference corpora
  - Collecting translation memories.
  - Domains: news, divulgative texts, manuals, environment and public administration.
- Elexbi: Extraction of specific domain bilingual terminology [2]
- Terminology extraction from comparable corpora [38, 39]
- Framework to represent tagged bitext [10,11]
- Semantic Role Labeling (ca, en & es) [41,34,33,42,40,32,44,45]

# Project Progress and Achievements

- Basic tools and resources
- Improving actual RBMT and SMT systems
- Combining EBMT and RBMT
- Hybrid architectures
- Advanced semantic processing for MT
- Evaluation and demonstration

# Improving the actual RBMT and SMT systems

- Improvements in modules [3,4]
  - Lexicon, parsing (Freeling), translation of prepositions, ...
- en-eu language pair included.
- Comparing RBMT and SMT systems [28,43]:
  - Automatic metrics: RBMT < SMT on the in-domain data.
  - HTER evaluation : RBMT > SMT for both corpora.
- Improving basic SMT systems [21,20,18]:
  - Combining shallow-syntactic translation models based on linguistic data views
  - Domain adaptation issue by porting an en-es phrase-based SMT system

# Project Progress and Achievements

- Basic tools and resources
- Improving actual RBMT and SMT systems
- **Combining EBMT and RBMT**
- Hybrid architectures
- Advanced semantic processing for MT
- Evaluation and demonstration

# Combining EBMT and RBMT

- Spanish-to-Basque **MultiEngine MT** [6,8]
- Specific domain (Labour agreements)

**Hierarchical strategy**, selecting the best output:

- first, EBMT translation patterns,
- second, SMT (if its confidence score > a fixed threshold)
- and, finally, RBMT.

**Important improvement in translation quality**

- 193.55% relative increase for BLEU (EBMT+SMT / SMT)
- 15.08% relative increase (EBMT+SMT / EBMT)
- Those improvements would be difficult to obtain by single-engine systems.

# Project Progress and Achievements

- Basic tools and resources
- Improving actual RBMT and SMT systems
- Combining EBMT and RBMT
- **Hybrid architectures**
- Advanced semantic processing for MT
- Evaluation and demonstration

# Hybrid architectures based on SMT

- Two experiments on **Statistical Post Editing (SPE)** [12]
  - **Using a morphological component** in both RBMT and SMT
  - Limited size of the available corpora.
- Results are coherent with previous literature
  - **Large improvements using the RBMT+SPE on a restricted domain.**
  - 200% improvement on BLEU scores for a RBMT+SPE approach, when comparing to raw translation output
  - The relative improvement with respect to the single SMT system is 40%

# Project Progress and Achievements

- Basic tools and resources
- Improving actual RBMT and SMT systems
- Combining EBMT and RBMT
- Hybrid architectures
- **Advanced semantic processing for MT**
- Evaluation and demonstration

# Advanced semantic processing for MT

- Inclusion of discriminative phrase selection models in SMT [23,24]
  - Local classifiers trained using WSD techniques instead of Maximum Likelihood.
  - Same techniques applied to English--Arabic, obtaining comparable improvements [13,14]
- Usage of external lexical-semantic resources to improve coverage and quality of SMT when applied across domains.
  - Enriching the original translation model with information derived from a multilingual LKB
  - Translation probabilities using a set of heuristics based on WordNet and local context.
  - Applied this to several translation tasks with significant improvement in translation quality.
- Inclusion of Semantic Role Labeling information (semantic dependencies) is left for the third year of research.

# Project Progress and Achievements

- Basic tools and resources
- Improving actual RBMT and SMT systems
- Combining EBMT and RBMT
- Hybrid architectures
- Advanced semantic processing for MT
- **Evaluation and demonstration**

# Evaluation of MT systems

- **Framework for MT automatic evaluation: IQMT [19]**
  - Set of heterogeneous metrics at different linguistic levels.
  - IQMT manual published [17].
  - software freely available: <http://www.lsi.upc.edu/~nlp/IQMT/>
- **Meta-evaluation of metrics.**
  - Combination of metrics and the assessment of the quality of metric sets.
  - Metrics based on syntax and semantics more reliable than the widespread lexical-based metrics as BLEU [22].
  - Considering also semantic features at the discourse level [18,25]
  - Study the behaviour of a non-parametric metric combination scheme [27]

# Evaluation of MT systems

- The previous linguistic-based metrics and their combinations have been evaluated
  - ACL Workshop on Machine Translation 2007 [22]
  - ACL Workshop on Machine Translation 2008 [23]
  - NIST Metrics MATR workshop
- Other usages of the IQMT metric set explored:
  - Parameter adjustment for combinations with maximum descriptive power [30]
  - Viability of performing heterogeneous automatic MT error analyses [26]

# Results and Indicators

## IXA-EHU

### Results, relevance and production

- The goals have been achieved.
  - but RBMT for the en-eu pair postponed a few months.
  - Better results than initially expected in SMT and hybrid systems
- 15 publications and 2 demos.
  - Two papers in AMTA and one in MT-Summit, the main conferences in the MT area.
  - Publications in CICLING and in IJCNLP conference.

# Demo: MT from Spanish to Basque (4 systems)

OpenMT - Mozilla Firefox

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://ixa2.si.ehu.es/openmt-demo/itzuli

ca | en | es | gl | eu

## Testuen Itzulpena

Testu originala:

Dos adultos y dos niños mueren en un incendio en Burgos

Itzulpena:

2 heldu eta 2 haur sute batean hiltzen dira Burgosen

Matxin es-eu

- Matxin es-eu
- Matxin en-eu
- SMT es-eu
- RBMT-SMT es-eu
- EBMT SMT RBMT hierarkikoa es-eu

Itzuli

IXA Taldea

Eginda

# Demo: MT from English to Basque

The screenshot shows a Mozilla Firefox browser window titled "OpenMT - Mozilla Firefox". The address bar contains the URL "http://ixa2.si.ehu.es/openmt-demo/itzuli" and the search engine is set to Google. The page features a navigation menu with options: Eitxategia, Editatu, Ikusi, Historia, Laster-markak, Tresnak, Hiztegiak, and Laguntza. A large "OPENMT" logo is displayed at the top, with language links for ca, en, es, gl, and eu. The main content area is titled "Testuen Itzulpena" and contains a text input field with the English sentence "Citizens pay double for fares in railways." and a corresponding output field with the Basque translation "Herritarrek bikoitza ordaintzen dituzte salneurriak trenbideetan." Below the input field is a dropdown menu set to "Matxin en-eu" and an "Itzuli" button. The footer includes "IXA Taldea" and a status bar with the text "Eginda".

OpenMT - Mozilla Firefox

Eitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://ixa2.si.ehu.es/openmt-demo/itzuli openmt

ca | en | es | gl | eu

## Testuen Itzulpena

Testu originala:

Citizens pay double for fares in railways.

Itzulpena:

Herritarrek bikoitza ordaintzen dituzte salneurriak trenbideetan.

Matxin en-eu Itzuli

IXA Taldea

Eginda

# Results and Indicators

## IXA-EHU

### Person training

- Three students participate in OpenMT-Ixa.
  - One PhD thesis presented in 2007 by Aingeru Mayor [35]
  - Other PhD thesis to be presented in 2009 by Gorka Labaka
  - Gorka Labaka won the Albaycin evaluation award for MT integrated in the JTH2008 [29]
- Doctors from IXA are PhD advisors
  - 3 participants from Elhuyar: Antton Gurrutxaga (DEA), Xabier Saralegi and Igor Leturia. To be finished during 2010.
- Material from our project used in
  - HAP official master on LNP ([ixa.si.ehu.es/master](http://ixa.si.ehu.es/master))
  - conferences and dissemination courses (UEU, [www.ueu.org](http://www.ueu.org))

# Results and Indicators

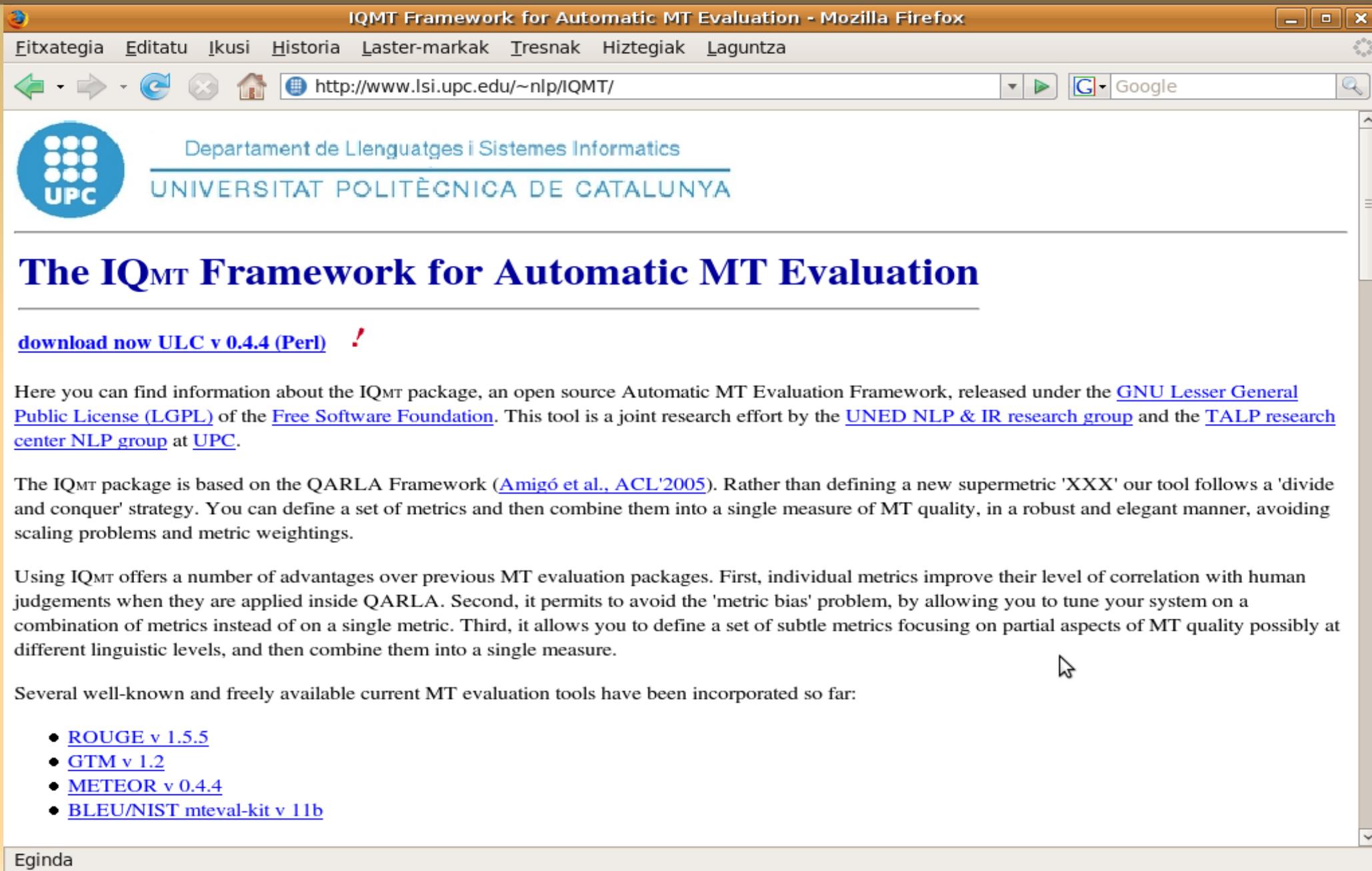
## TALP-UPC

### Results, relevance and production

- The goals have been achieved.
- 34 publications
  - 2 publications in Journal of Artificial Intelligence Research (JAIR). The impact factor of the journal is 1.107 (SCI/SSCI).
  - One publication in the ACL, the most prestigious in the NLP area
  - A 50-page chapter in a MIT Press book on Machine Learning for SMT (NIPS series)
  - Publications at the COLING-ACL and the IJCNLP conferences
- 2 demos:
  - JointParser joint analysis of syntactic and semantic dependencies.  
<http://www.lsi.upc.edu/~xlluis/jointparser/jointparser.php>
  - The IQMT Framework for Automatic MT Evaluation  
<http://www.lsi.upc.edu/~nlp/IQMT/>

# IQMT

## Framework for Automatic MT Evaluation



The screenshot shows a Mozilla Firefox browser window with the title "IQMT Framework for Automatic MT Evaluation - Mozilla Firefox". The address bar contains the URL "http://www.lsi.upc.edu/~nlp/IQMT/". The page header includes the UPC logo and the text "Departament de Llenguatges i Sistemes Informatics" and "UNIVERSITAT POLITÈCNICA DE CATALUNYA". The main heading is "The IQ<sub>MT</sub> Framework for Automatic MT Evaluation". Below the heading, there is a link "download now ULC v 0.4.4 (Perl) !" with a red exclamation mark. The text describes the IQ<sub>MT</sub> package as an open source Automatic MT Evaluation Framework, released under the GNU Lesser General Public License (LGPL) of the Free Software Foundation. It is a joint research effort by the UNED NLP & IR research group and the TALP research center NLP group at UPC. The text explains that the IQ<sub>MT</sub> package is based on the QARLA Framework (Amigó et al., ACL'2005) and follows a 'divide and conquer' strategy. It offers advantages over previous MT evaluation packages, such as improved correlation with human judgements, avoidance of the 'metric bias' problem, and the ability to define subtle metrics focusing on partial aspects of MT quality. Finally, it lists several well-known and freely available current MT evaluation tools that have been incorporated: ROUGE v 1.5.5, GTM v 1.2, METEOR v 0.4.4, and BLEU/NIST mteval-kit v 11b.

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://www.lsi.upc.edu/~nlp/IQMT/ Google

 Departament de Llenguatges i Sistemes Informatics  
UNIVERSITAT POLITÈCNICA DE CATALUNYA

## The IQ<sub>MT</sub> Framework for Automatic MT Evaluation

[download now ULC v 0.4.4 \(Perl\) !](#)

Here you can find information about the IQ<sub>MT</sub> package, an open source Automatic MT Evaluation Framework, released under the [GNU Lesser General Public License \(LGPL\)](#) of the [Free Software Foundation](#). This tool is a joint research effort by the [UNED NLP & IR research group](#) and the [TALP research center NLP group](#) at [UPC](#).

The IQ<sub>MT</sub> package is based on the QARLA Framework ([Amigó et al., ACL'2005](#)). Rather than defining a new supermetric 'XXX' our tool follows a 'divide and conquer' strategy. You can define a set of metrics and then combine them into a single measure of MT quality, in a robust and elegant manner, avoiding scaling problems and metric weightings.

Using IQ<sub>MT</sub> offers a number of advantages over previous MT evaluation packages. First, individual metrics improve their level of correlation with human judgements when they are applied inside QARLA. Second, it permits to avoid the 'metric bias' problem, by allowing you to tune your system on a combination of metrics instead of on a single metric. Third, it allows you to define a set of subtle metrics focusing on partial aspects of MT quality possibly at different linguistic levels, and then combine them into a single measure.

Several well-known and freely available current MT evaluation tools have been incorporated so far:

- [ROUGE v 1.5.5](#)
- [GTM v 1.2](#)
- [METEOR v 0.4.4](#)
- [BLEU/NIST mteval-kit v 11b](#)

Eginda

# JointParser: Joint syntactic and semantic analysis

jointparser results - Mozilla Firefox

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://www.lsi.upc.edu/~xlluis/jointparser/jointparser.php jointparser

## jointparser results

Sentence: Tomorrow, and tomorrow, and tomorrow creeps in this petty pace from day to day to the last syllable of recorded time.

-Root-	Tomorrow	,	and	tomorrow	,	and	tomorrow	creeps	in	this	petty	pace	from	day	to	day	to	the	last	syllable
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
tomorrow	,	and	tomorrow	,	and	tomorrow	creep	in	this	petty	pace	from	day	to	day	to	the	last	syllable	
NN		CC	NN		CC	NN	VBZ	IN	DT	JJ	NN	IN	NN	TO	NN	TO	DT	JJ	NN	
Tomorrow	,	and	tomorrow	,	and	tomorrow	creeps	in	this	petty	pace	from	day	to	day	to	the	last	syllable	
tomorrow	,	and	tomorrow	,	and	tomorrow	creep	in	this	petty	pace	from	day	to	day	to	the	last	syllable	
NN		CC	NN		CC	NN	VBZ	IN	DT	JJ	NN	IN	NN	TO	NN	TO	DT	JJ	NN	

Syntactic dependencies in green.  
Semantic dependencies and nominal and verbal predicates in blue.  
Token properties in gray.

Eginda

# Results and Indicators

## TALP-UPC

### Person training

- Three students participate in OpenMT.
  - PhD thesis presented in 2008 by Jesús Gimenez [18]
  - Master Thesis presented in 2008 by Cristina España [13]
  - Master Thesis presented in 2008 by Xabier Lluís [31]
  - Stefan Bott achieved a post-doc grant from the Catalan Government.

# Results and Indicators

## Elhuyar

### Results, relevance and production

- The goals have been achieved.
  - Parallel corpora collected for several domains
  - Tool for managing and storing parallel corpora
  - Bilingual lexicon from the administrative--text domain was created making use of the terminology extractor Elexbi.
  - Works on terminology extraction from comparable corpora
- 3 publications
- Demo: Website repository of Translation Memories  
[http://ixa.si.ehu.es/openmt/demoak\\_html](http://ixa.si.ehu.es/openmt/demoak_html)

# Demo: TMR translation Memory Repository

Aplikazioak Lekuak Sistema Translation-memory Manager - Mozilla Firefox ots 19, og., 13:27 Kepa Sarasola

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://beam.to/tm-m/ openmt castellano - euskara - english

## TMR

OpenMT - TMR (Translation Memory Repository)

**Corpusak**

[TMR](#)

**Utilitateak**

[NireTMXa](#)

Aukeratutako dokumentuak: 0

### Consumer-en itzulpen memoria

[NireTMX-era gehitu](#)

<b>es =&gt; eu</b> 63124 segmentu	<b>Sorkuntzarako Tresna:</b> <b>Segmentu mota:</b> <b>O-TMF:</b> <b>Datu mota:</b> <b>Sorkuntza data:</b> <b>Sorkuntza IDa:</b>
--------------------------------------	--

#	Sorburu Hizkuntza: es	Xede Hizkuntza: eu
1	Menús escolares: estudiados los de 200 centros escolares	Eskolako menuak: 200 ikastetxetakoak aztertu dira
2	Pero este dato dista de ser homogéneo en las regiones.	Bizkalde, eskola publikoen menuetan pribatuenetan baino akats larri gutxiago antzematen da.
3	En Castellón, y Alicante y Burgos se sitúan en torno al 24% mientras que apenas suponen el 9% en Vizcaya, el 13% en Barcelona y el 15% en Alava y Navarra.	Arrainari dagokionez, berriz, ikastetxe pribatuen %23,5ak ez du menuetan arrainik eskaintzen, eta publikoetan %18 dago egoera horretan.
4	Y, a su vez, disminuir el consumo de frituras y platos precocinados, postres dulces y bollería.	Janari berri bat sartzean, errazio txikia ipini.
5	Recomendaciones para los menús escolares	Eskolako menuetarako gomendioak
	Elaborar platos ricos en proteínas vegetales,	Barazki-proteina, karbohidrato konplexu eta

Eginda

MEC\_2009\_OpenMT-... OPENMT\_defentsa\_0... Translation-memory ...

# Results and Indicators

## Elhuyar

### Person training

- Two participants finishing their Official Masters
  - Xabier Saralegi
  - Igor Leturia
- Another participant obtained the PhD research diploma (DEA) and is currently finishing the PhD thesis.
  - Antton Gurrutxaga

# Results and Indicators

## EHU + UPC + ELHUYAR

### International collaboration (1)

- Workshop organization 2008/02/14:  
Mixing Approaches to Machine Translation [5,48]
  - More than 50 persons came from Spain, Europe and Japan.
  - Invited talks were given by M. Federico, P. Koehn and A. Way.
  - They were very useful for the main aim of the project: discussing different methods for hybridization in MT.
  - Dissemination of the results has been carried out
- Network of Excellence on MT designed during the MATMT workshop 2008 in San Sebastian.
  - Coordinator: H. Sommers from DCU university.
  - Consortium: Ixa group the DCU (Dublin), UPC (Barcelona), Edinburgh, Aachen and Charles (Praga) universities, Xerox company (Grenoble) and Bruno Kressler (Trento) Foundation.

# OPENMT

## MATMT workshop (2008/02/14)

MATMT 2008 - Mozilla Firefox

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://ixa2.si.ehu.es/matmt-2008/ matmt

**MATMT2008 workshop:**  
*"Mixing Approaches to Machine Translation"*

---

[Korta Research Center](#)  
[Avda. de Tolosa 72](#)  
**University of the Basque Country**  
[Donostia-San Sebastian](#), Thursday February 14th 2008

**CALL FOR PARTICIPATION**

---

**Contents** [hide]

- 1 Proceedings
- 2 Background
- 3 Workshop topics
- 4 Program
- 5 Registration
  - 5.1 Optional Dinner
- 6 Venue and Travel
- 7 Important Dates
- 8 Paper submission
- 9 Programme committee
- 10 Local organization
- 11 About OpenMT project

**Proceedings**

Iñaki Alegria, Lluís Márquez, Kepa Sarasola (ed.) 2008  
[Mixing Approaches to Machine Translation. MATMT2008. Proceedings](#)

Eginda

# Results and Indicators

## EHU + UPC + ELHUYAR

### International collaboration (2)

- STREP project was proposed in 2007
  - It was called EurOpenTrans (Large Scale Open Source Machine Translation for low density European Languages).
  - Partners included: DCU (Dublin), UPC(Barcelona), Edinburgh and Charles (Praga) universities, Alpinion (Ljubljana) and Translan (Dublin) companies and Elhuyar Foundation.
  - The project was evaluated with 9.0 points when the threshold was 10.0.
- Project proposal (TransBlog) in the European Work Programme 2009-2010, Challenge 2 call
  - Leader: Elhuyar. Partners: IXA-EHU.
  - The consortium is set up
  - Application of technologies developed in OpenMT in a framework for translating and increasing visibility of Blogs

# Results and Indicators

## EHU + UPC + ELHUYAR

### **International collaboration (3)**

- Andy Way's group at the Dublin City University.
  - Kepa Sarasola (IXA-EHU): Stay of 2 months in 2006.
  - Gorka Labaka (IXA-EHU): 2 months in 2007+ 2 months in 2008
  - Two papers generated [43,28]

# Results and Indicators

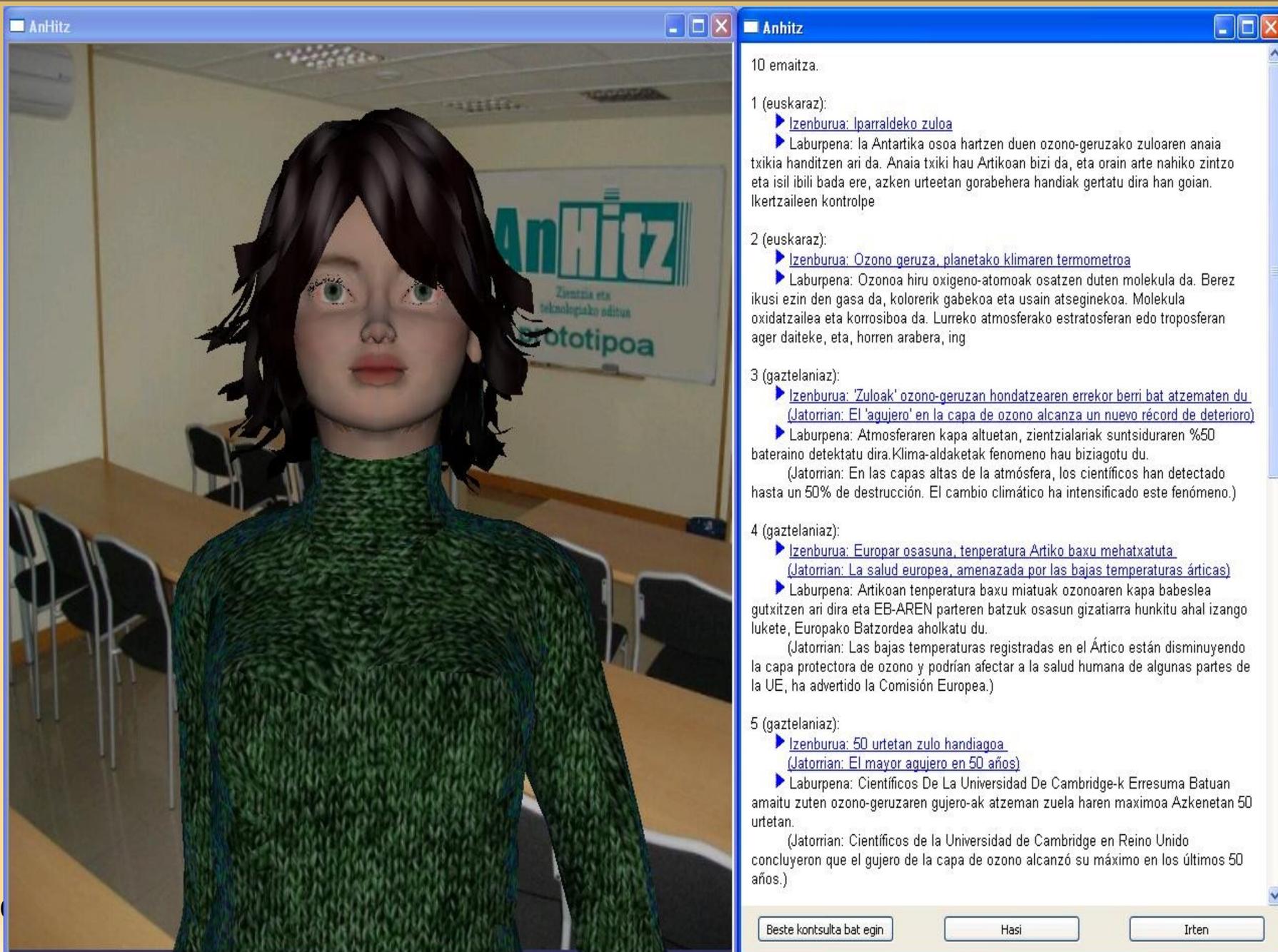
## EHU + UPC + ELHUYAR

### Technology transfer

- Most of the modules and technology are being transferred in the framework of two PROFIT projects to:
  - Eleka company ([www.eleka.net](http://www.eleka.net))
  - OpenTrad consortium ([www.opentrad.org](http://www.opentrad.org))
- Presented to a public call for a major contract in es-eu MT Basque Government 2008.
- Integrated in a CLQA system named Ihardetsi [53] which was included in the QA-CLEF evaluation 2008.
- MT technology encapsulated in AnHitz project [52], promoted by the Basque Government.

# ANHITZ:

## Speech CLIR and QA with avatar and MT



10 emaitza.

1 (euskaraz):

- ▶ [Izenburua: Iparraldeko zuloa](#)
- ▶ Laburpena: la Antartika osoa hartzen duen ozono-geruzako zuloaren anaia txikia handitzen ari da. Anaia txiki hau Artikoan bizi da, eta orain arte nahiko zintzo eta isil ibili bada ere, azken urteetan gorabehera handiak gertatu dira han goian. Ikertzaileen kontrolpe

2 (euskaraz):

- ▶ [Izenburua: Ozono geruza, planetako klimaren termometroa](#)
- ▶ Laburpena: Ozonoa hiru oxigeno-atomoak osatzen duten molekula da. Berez ikusi ezin den gasa da, kolorerik gabekoa eta usain atseginekkoa. Molekula oxidatzailea eta korrosiboa da. Lurreko atmosferako estratosferan edo troposferan ager daiteke, eta, horren arabera, ing

3 (gaztelaniaz):

- ▶ [Izenburua: 'Zuloak' ozono-geruzan hondatzearen errekor berri bat atzematen du](#)  
(Jatorrian: El 'agujero' en la capa de ozono alcanza un nuevo récord de deterioro)
- ▶ Laburpena: Atmosferaren kapa altuetan, zientzialariak suntsiduraren %50 bateraino detektatu dira. Klima-aldaketak fenomeno hau biziagotu du.  
(Jatorrian: En las capas altas de la atmósfera, los científicos han detectado hasta un 50% de destrucción. El cambio climático ha intensificado este fenómeno.)

4 (gaztelaniaz):

- ▶ [Izenburua: Europar osasuna, temperatura Artiko baxu mehatxatuta](#)  
(Jatorrian: La salud europea, amenazada por las bajas temperaturas árticas)
- ▶ Laburpena: Artikoan tenperatura baxu miatuak ozonoaren kapa babeslea gutxitzen ari dira eta EB-AREN parteren batzuk osasun gizatiarra hunkitu ahal izango lukete, Europako Batzardea aholkatu du.  
(Jatorrian: Las bajas temperaturas registradas en el Ártico están disminuyendo la capa protectora de ozono y podrían afectar a la salud humana de algunas partes de la UE, ha advertido la Comisión Europea.)

5 (gaztelaniaz):

- ▶ [Izenburua: 50 urtetan zulo handiagoa](#)  
(Jatorrian: El mayor agujero en 50 años)
- ▶ Laburpena: Científicos De La Universidad De Cambridge-k Erresuma Batuan amaitu zuten ozono-geruzaren gujero-ak atzeman zuela haren maximoa Azkenetan 50 urtetan.  
(Jatorrian: Científicos de la Universidad de Cambridge en Reino Unido concluyeron que el agujero de la capa de ozono alcanzó su máximo en los últimos 50 años.)

Beste kontsulta bat egin      Hasi      Irten

# Results and Indicators

## EHU + UPC + ELHUYAR

### Project management

- Private web page of the project (wiki): <http://ixa2.si.ehu.es/openmt>
- Public website of the project: <http://ixa.si.ehu.es/openmt>
- Periodical meetings
  - (1) Donostia 19/20-XII-2006,
  - (2) Barcelona 9-X-2007,
  - (3) Donostia 14-II-2008,
  - (4) Barcelona: 2-X-2008.
- Technical meetings:
  - (1) Lexical selection, Donostia 19-IV-2007;
  - (2) Lexical selection, Barcelona 6-VII-2007,
  - (3) MT-hybridization, Donostia 14-II-2008.

# OPENMT public web site

The screenshot shows a Mozilla Firefox browser window titled "OpenMT project - Mozilla Firefox". The address bar contains the URL "http://ixa.si.ehu.es/openmt/demoak\_htn" and the search engine is set to Google with the search term "openmt". The website header features the "OPENMT" logo in large white letters on a dark background. A left-hand navigation menu lists various sections: Project, Goal, Participants, Publications, Intranet, Demos, Resources and Tools, State of the art '08, and References '08. The "Demos" section is currently selected and expanded, displaying a list of bullet points: "MT systems, en-eu and hybrids", "MATMT-2008 Workshop", "IQMT Framework for MT Evaluation", "JointParser", "Repository of TMs", "Anhitz", and "Ihardetsi". At the bottom left of the page, there is a language selection dropdown menu currently set to "English". The browser's status bar at the bottom shows the text "Eginda".

OpenMT project - Mozilla Firefox

Fitxategia Editatu Ikusi Historia Laster-markak Tresnak Hiztegiak Laguntza

http://ixa.si.ehu.es/openmt/demoak\_htn openmt

# OPENMT

**Project**

- Goal
- Participants
- Publications
- Intranet
- Demos**
- Resources and Tools
- State of the art '08
- References '08

English

- **MT systems, en-eu and hybrids**
- **MATMT-2008 Workshop**
- **IQMT Framework for MT Evaluation**
- **JointParser**
- **Repository of TMs**
- **Anhitz**
- **Ihardetsi**

Eginda

# OPENMT private web site

The screenshot shows a Mozilla Firefox browser window displaying the main page of the OpenMT project. The browser's address bar shows the URL `http://ixa2.si.ehu.es/openmt/index.php/Main_Page`. The page title is "Main Page - OpenMT".

The page content includes a navigation menu with options like "Main Page", "Community portal", "Current events", "Recent changes", "Random page", "Help", and "Donations". A search box is also present.

The main content area features a "Main Page" section with the title "OpenMT: Traducción automática en código libre mediante métodos híbridos". Below this, there is a list of project details:

- Código: TIN2006-15307-C03
- Inicio: 1-XII-2006
- Duración: 3 años
- Participantes:
  - EHU (IXA taldea) [↗](#)
  - UPC [↗](#)
  - Elhuyar Fundazioa [↗](#)

Below the list, there are sections for "Documentación inicial", "Web pública", and "PLANIFICACION". The "PLANIFICACION" section lists:

- **Planificación detallada y resultados 1. año**
- **Planificación detallada y resultados 2. año**

The "INFORMES ANUALES DE SEGUIMIENTO" section includes a sub-section "INFORMES EHU" with the following items:

- **Aceptación del presupuesto, enero de 2007** [↗](#)
- **Informe enviado en marzo de 2007** [↗](#) (simplemente aceptación)
- **Informe enviado en marzo de 2008** [↗](#)

The browser's status bar at the bottom shows "Eginda".

# OpenMT: Open Source Machine Translation Using Hybrid Methods

- Kepa Sarasola, Iñaki Alegria  
**Ixa Group. Euskal Herriko Unibertsitatea**
- Núria Castell, Lluís Màrquez  
**TALP, Universitat Politècnica de Catalunya**
- Nerea Areta, Xabier Saralegi  
**Elhuyar Fundazioa**

TIN2006-15307-C03

<http://ixa.si.ehu.es/openmt>