

## Unu traduttore automaticu pro sa limba sarda

**Su progetu, nàschidu dae sa collaboratzione intre sos Paìsos Catalanos e sa Sardigna, at retzidu su finantziamentu de Google**



The screenshot shows the Apertium website interface. At the top, there's a navigation bar with the Apertium logo and the text "Una prataforma libera a còdighe abertu pro sa tradutzione automàtica". Below this, there's a search bar and a language selector set to "sardu". The main content area features a translation tool with two input fields. The left field contains the text "In linea il primo traduttore automatico italiano-sardo!". The right field contains the text "In linia su primu traduttore automaticu italianu-sardu!". There are buttons for "Borta" and "Borta unu documentu". At the bottom, there are links for "Informatziones", "Iscàrriga", "Documentatzione", and "Cuntatade-nos". A small box at the bottom right says "Ais agatadu una faddina? Agiuade-nos a megiorare Apertium!".

Su primu traduttore automaticu pro sa limba sarda est nàschidu dae sa collaboratzione intre su grupu de investigatzione Tradumàtica ([www.tradumatica.net](http://www.tradumatica.net)), de s'Universitat Autònoma de Barcelona, e Prompsit ([www.prompsit.com](http://www.prompsit.com)), azienda valentziana chi dae annos meda gestit sa prataforma libera de tradutzione Apertium. Su finantziamentu pro s'isvilupu de su traduttore est arribadu dae Google, tràmite su programma Google Summer of Code.

Sa prataforma, creada in s'universidade de Alacant in su 2004 pro sa croba linguistica castillianu-catalanu, oe includet pagu prus o mancu baranta limbas comente a su brètone, a s'otzitanu, a su bascu e a àteras. Est unu traduttore basadu subra de regulas, duncas resurtat un'aina perfeta pro sas limbas in fase de istandardizatzione.

Pro s'elaboratzione de su traduttore, si sunt creados corpus e ditzionàrios leende comente riferimentu sa Wikipedia sarda e de unos cantos giornales. Su traduttore est presentadu cun 25.000 paràulas, una cobertura de tradutzione a de belle su 90% e unu tassu de faddina pro cada paràula de su 10%. In sas versiones imbenientes de su traduttore, s'ant a currègere sas faddinas, s'ant a creare regulas noas e s'ant a agiunghere àteras cumbinatzziones linguisticas. Sa pàgina printzipale de Apertium (bortada in sardu gràtzias a sa collaboratzione de su grupu de utentes Sardware) includet unu ligàmene in manera chi sos utentes potzant sinalare sas faddinas e cussigiare progetos noos.

[www.apertium.org](http://www.apertium.org)

Gianfranco Fronteddu  
Hèctor Alòs i Font  
Adrià Martín-Mor  
Francis Tyers

## Un traductor automàtic per a la llengua sarda

El projecte, nascut de la col·laboració entre els Països Catalans i Sardenya, ha rebut el finançament de Google



El primer traductor automàtic per a la llengua sarda ha nascut gràcies a la col·laboració entre el grup de recerca Tradumàtica ([www.tradumatica.net](http://www.tradumatica.net)), de la Universitat Autònoma de Barcelona, i Prompsit ([www.prompsit.com](http://www.prompsit.com)), empresa alacantina que de fa temps gestiona la plataforma lliure de traducció Apertium. El desenvolupament del traductor ha estat finançat per Google, mitjançant el programa Google Summer of Code.

La plataforma, creada a la universitat d'Alacant l'any 2004 per a la combinació lingüística castellà-català, avui inclou al voltant de quaranta llengües, entre les quals el bretó, l'occità i el basc. Es tracta d'un traductor basat en regles, per la qual cosa és un instrument particularment adequat per a les llengües en fase d'estandardització.

Per a l'elaboració del traductor, s'han creat corpus i diccionaris prenent com a referència la Viquipèdia en llengua sarda i diversos diaris. El traductor es presenta amb 25.000 paraules, una cobertura de traducció al voltant del 90% i una taxa d'error per paraula del 10%.

En futures versions del traductor, es corregiran errades, es crearan noves regles i s'afegiran noves combinacions lingüístiques. La pàgina principal d'Apertium (traduïda en sard gràcies a la col·laboració del grup d'usuaris Sardware) inclou un enllaç perquè els usuaris puguin comunicar possibles errors i proposar nous projectes.

[www.apertium.org](http://www.apertium.org)

Gianfranco Fronteddu  
Hèctor Alòs i Font  
Adrià Martín-Mor  
Francis Tyers

## Un traduttore automatico per la lingua sarda

### Il progetto, nato dalla collaborazione tra i Paesi Catalani e la Sardegna, ha ricevuto il finanziamento di Google



The screenshot shows the Apertium website interface. At the top, there's a navigation bar with the Apertium logo and the text "Una prataforma libera a còdighe abertu pro sa traduzione automàtica". Below this, there's a search bar and a language selector set to "sardu". The main content area features a translation interface with two input boxes. The left box contains the Italian text "In linea il primo traduttore automatico italiano-sardo!". The right box contains the Sardinian translation "In llinia su primu traduttore automàticu italianu-sardu!". There are buttons for "Borta" (Copy) and "Borta unu documentu" (Copy document). At the bottom, there are links for "Informatziones", "Iscàrriga", "Documentatzione", and "Cuntatade-nos", along with a feedback link: "Ais agatadu una faddina? Agiuade-nos a megiorare Apertium!".

Il primo traduttore automatico per la lingua sarda è nato dalla collaborazione tra il gruppo di ricerca Tradumàtica ([www.tradumatica.net](http://www.tradumatica.net)), dell'Universitat Autònoma de Barcelona, e Prompsit ([www.prompsit.com](http://www.prompsit.com)), azienda valenziana che da anni gestisce la piattaforma libera de traduzione Apertium. Il finanziamento per lo sviluppo del traduttore è arrivato da Google, tramite il programma Google Summer of Code.

La piattaforma, creata nell'università di Alacant nel 2004 per la coppia linguistica castigliano-catalano, oggi include circa quaranta lingue, tra le quali il bretone, l'occitano e il basco. Si tratta di un traduttore basato su regole, dunque uno strumento particolarmente adatto alle lingue in fase di standardizzazione.

Per l'elaborazione del traduttore, sono stati creati corpora e dizionari prendendo come riferimento la Wikipedia in lingua sarda e vari giornali. Il traduttore viene presentato con 25.000 parole, una copertura di traduzione del 90% circa e un tasso di errore per parola del 10%.

Il traduttore potrà essere migliorato nelle future versioni, nelle quali si procederà alla correzione di errori, la creazione di nuove regole e l'aggiunta di altre combinazioni linguistiche. La pagina principale di Apertium (tradotta in sardo grazie alla collaborazione del gruppo di utenti Sardware) include inoltre un link che permette agli utenti di segnalare eventuali errori e proporre nuovi progetti.

[www.apertium.org](http://www.apertium.org)

Gianfranco Fronteddu  
Hèctor Alòs i Font  
Adrià Martín-Mor  
Francis Tyers

## A machine translation system for Sardinian

The project, a collaboration between the Catalan Countries and Sardinia, has received financing from the Google Summer of Code



The screenshot shows the Apertium website interface. At the top, there is a navigation bar with the Apertium logo and the tagline "Una prataforma libera a còdighe abertu pro sa tradutzione automàtica". Below this, there is a search bar and a language selector set to "sardu". The main content area features a translation tool with two input fields. The left field contains the text "In linea il primo traduttore automatico italiano-sardo!" and the right field contains "In linia su primu traduttore automàticu italianu-sardu!". There are buttons for "Borta" and "Borta unu documentu". At the bottom, there are links for "Informatziones", "Iscàrriga", "Documentatzione", and "Cuntatade-nos", along with a call to action: "Ais agatadu una faddina? Agiuade-nos a megiorare Apertium!".

The first machine translation system for Sardinian has been created thanks to the collaboration between the research group Tradumàtica ([www.tradumatica.net](http://www.tradumatica.net)) at the Universitat Autònoma de Barcelona and Prompsit Language Engineering ([www.prompsit.com](http://www.prompsit.com)), a company which participates in the management of the free/open-source software platform Apertium. The development of the system has been financed in the Google Summer of Code, a programme created by Google to introduce students to free/open-source software.

The Apertium platform, which was created at the Universitat d'Alacant in 2004 for translating between Spanish and Catalan, now supports more than 40 language pairs, including Breton, Occitan and Basque. The platform is rule-based, and as such is particularly suited to languages which are being standardised.

In order to create the system, a corpus of texts and dictionaries were created using sources like the Wikipedia in Sardinian and newspapers. The system currently contains a bilingual dictionary of 25,000 head words, a lexical coverage of around 90% and a word-error rate of 10%. In future versions of the translator, errors will be fixed and new rules created, and we will add new language pairs including Sardinian. The main page of the Apertium project (now available in Sardinian thanks to the efforts of the Sardware user group) includes a link where users can report possible errors and propose new projects.

[www.apertium.org](http://www.apertium.org)

Gianfranco Fronteddu  
Hèctor Alòs i Font  
Adrià Martín-Mor  
Francis Tyers