
El Corrector: un projecte obert

Martí Quixal, Judith Domingo, Francesc Benavent



Barcelona
Media

Jornades de Programari Lliure,
Barcelona, 1 al 4 de Juliol del 2009

Sumari

1.Context i història

2.Resultats

3.Objectius de futur

Sumari

1.Context i història

2.Resultats

3.Objectius de futur

Context

- La Generalitat de Catalunya obre a concurs:
 - Un projecte per a la creació d'un corrector ortogràfic i gramatical
 - de codi obert
 - normatiu
 - integrat en una sèrie d'aplicacions ofimàtiques
 - Accés a les especificacions del concurs
 - http://www10.gencat.net/dursi/pdf/se/plec_se_ctti_51_05_ca.pdf

Història

- Duració prevista inicialment: 11 mesos (de febrer a desembre del 2006)
- Tancament real: novembre del 2008
- Raons:
 - Ambició del projecte
 - Manca de previsió en aspectes de llicències
 - Inexperiència en el desenvolupament “industrial” i en la tradició del codi obert

Història (II)

- Objectius que van caure pel camí
 - Integració en MS Outlook Express
 - Integració en WordPad
 - Alguns dels tipus d'errors que es proposa tractar en el Bloc B (que ja s'intuïa que serien complicats i que no es posaven com a condició necessària per a la compleció del projecte)

Història (III)

- Finançament rebut (aprox. 200000 EUR, de 240000 EUR previstos)
- Cost real del projecte

	Personal	Maquinari	Altres	Total
Any 2006	157900	4000	3000	164900
Any 2007	98530	1200	0	99730
Any 2008	50000	1200	0	51200
Totals	306430	6400	3000	315830

Sumari

1.Context i història

2.Resultats

3.Objectius de futur

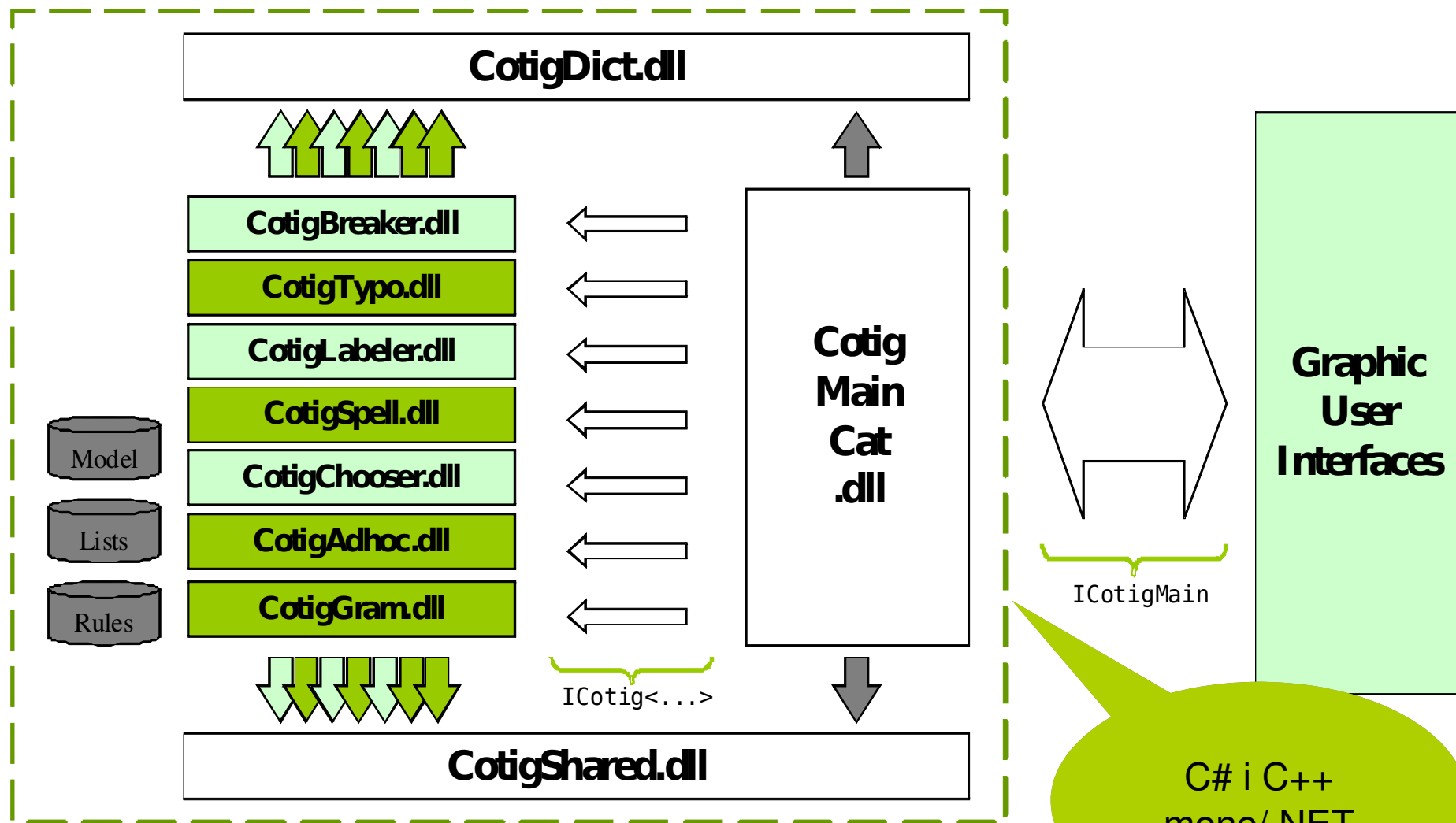
Productes resultants (I)

- Motor de correcció en C# i C++:
 - Reutilitzable en aplicacions .NET/Mono
- Recursos lingüístics
 - Diccionaris (general, d'errors, de noms propis, etc.)
 - Gramàtiques de detecció d'errors
 - Corpus d'errors (artificials)
 - Models lingüístics estadístics per a la desambiguació morfosintàctica

Productes resultants (II)

- Aplicacions i connectors
- Documentació
- Codi (línies i línies!)

Motor de correcció



Connectors i aplicacions

	Windows 2000	Windows XP	Linux (Fedora)	Mac
Editor propi	√	√	√	√
Office 2000	√	√		
Office XP	√	√		
Office 2003	√	√		
OpenOffice 2	√	√	√	
OpenOffice 3	√	√	√ (beta)	
FireFox	√	√	√	√
Thunderbird	√	√	√	√
SeaMonkey	√	√	√	√
Emacs (obsolet)			√	

Una web (www.elcorrector.cat)

- Corrector en línia (basat en TinyMCE)

El Corrector en línia

Per veure les propostes de correcció fes clic al botó dret sobre les paraules marcades en color vermell o verd. Per triar-ne una, fes un clic amb el botó esquerre. També pots fer servir el teclat. En acabar utilitzar el botó *Recorreix* o *Fet* (per esborrar les marques de correcció).

Opcions de correcció:

- Ortografia
- Gramàtica

Opcions dialectals:

- Català central
- Català nord-occidental
- Balear
- Valencià

Una avaria en **un trens** entre Sants i Catalunya causa **retars** de 30 minuts

- Error ortogràfic.
- retard
- retirs
- retards
- retis
- retir
- re tars

Recorreix

Fet

Recursos per a desenvolupar

- Documentació

- *Descripció general de l'arquitectura de processament*
- *Especificacions d'errors lingüístics tractats*
- *Manual d'usuari del connector de MS Office*
- *Manual d'usuari de l'editor bàsic*

- Repositori SVN del codi font



<https://projectes.lafarga.cat/projects/corrector/>

Sumari

1.Context i història

2.Resultats

3.Objectius de futur

De cara al futur

- Concentrar esforços (reduint l'abast)
- Connectors per processador de textos:

- OpenOffice Writer
- MS Word
- Editor bàsic (dubtem...)



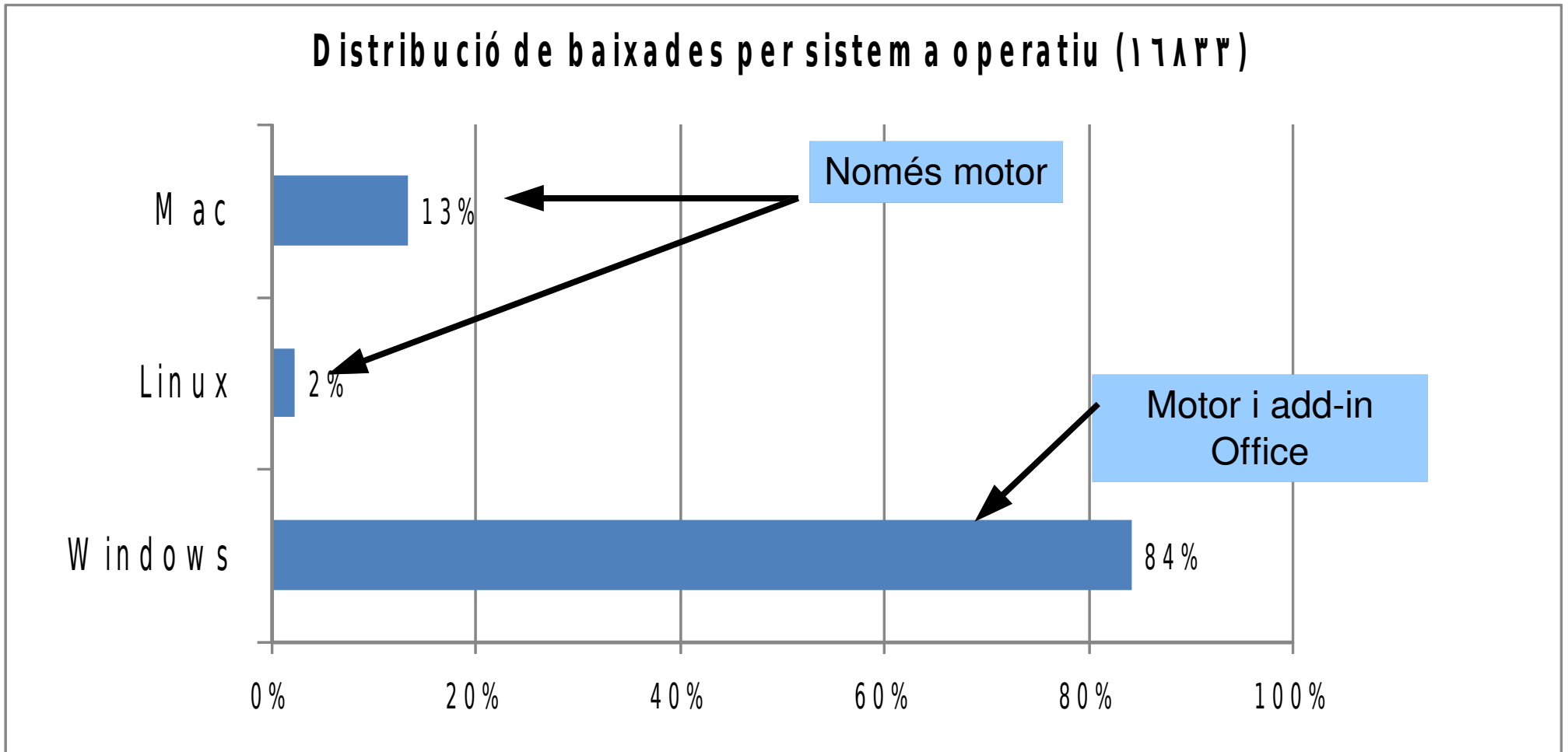
- Mantenir multiplataforma:
en Linux, Mac i Windows



De què disposem?

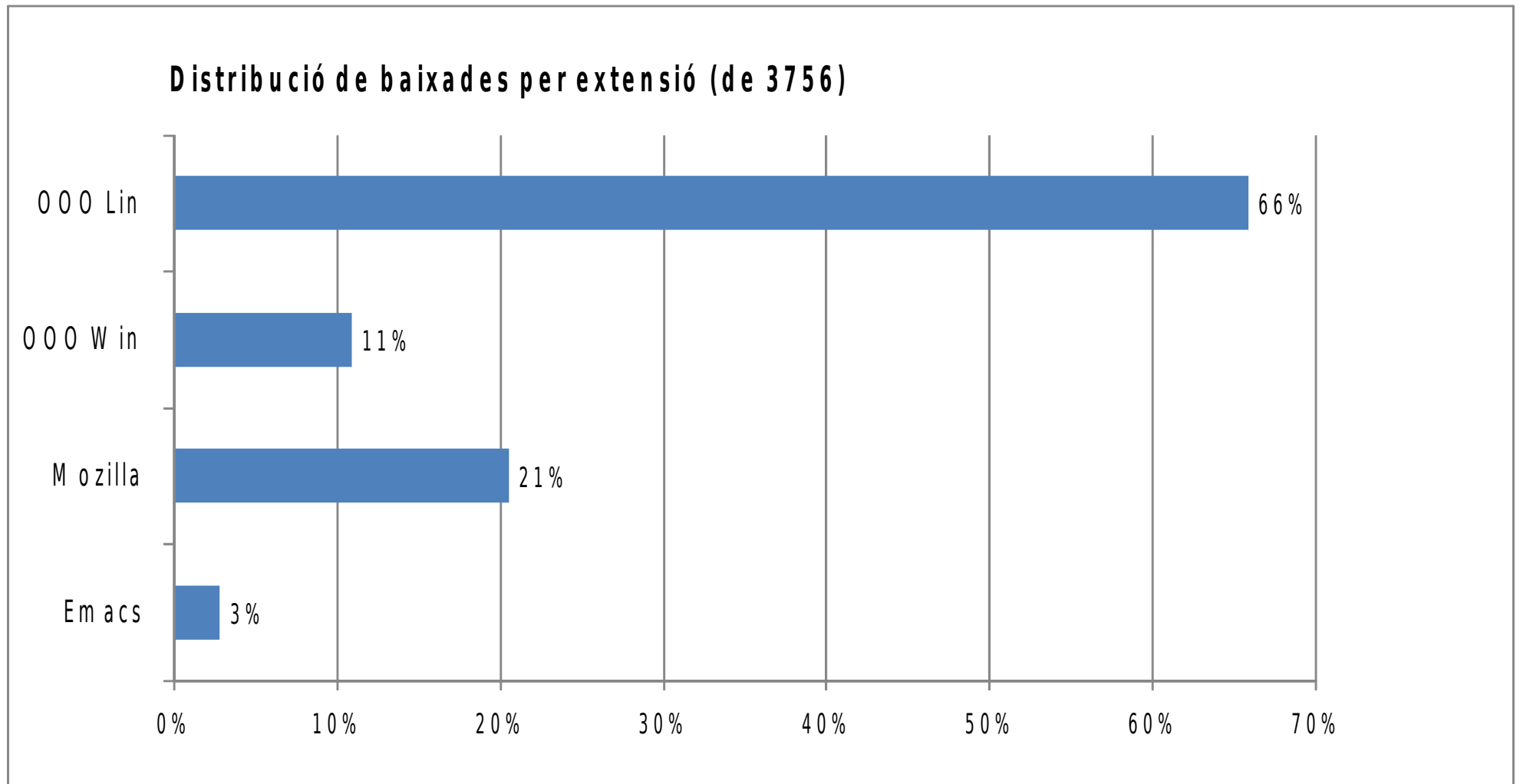
- Web
- Documentació
- Respositori svn
- Fòrum (en construcció)

Amb qui comptem? (I)



*Fins a finals d'abril

Amb qui comptem? (II)



Amb qui més voldríem comptar?

Demandes dels usuaris

- Sistemes operatius
 - Ubuntu
 - Windows Vista
- Aplicacions
 - Connectors per MS Office 2007 (Windows) i MS Office 2008 (Mac)

Què ens caldria?

- Ajuda/Consell
 - Volem aprendre del que ja saben les persones que fa temps que són a la comunitat de codi lliure.
 - Com coordinar-nos
 - Criteris i tècniques per incorporar pedaços i extensions
 - Desenvolupar segons peticions d'usuaris o finançament *ad hoc*

Qui tenim per treballar?

Equip de desenvolupadors	
Lingüistes	1/10 de jornada
Programador	1/10 de jornada
Voluntaris/Encàrrecs	Hores perdudes

Qui ens caldria?

- Persones amb ganes de col·laborar

Saps programar?

- C#, C++, o Java i experiència en Win, Lin i Mac

- Ets un filòleg, lingüista o lletraferit sense por als ordinadors?

- Coneixes bé la norma? Ets corrector?

- Escrius en català i t'importa escriure'l bé?

- Vols fer d'usuari beta?

Conclusions

Si t'interessa col·laborar en el projecte

www.elcorrector.cat

elcorrectordecatala@gmail.com

Un editor bàsic

- Edició i correcció bàsica, però *multiplataforma*

