

# MIATU Erraz

Eskulibua

2007ko apirila



ingeniaritza  
**ELEKA**  
linguistikoa

ELEKA Ingeniaritza Linguistikoa, S.L.

[www.eleka.net](http://www.eleka.net)

## Aurkibidea

1 Miatu-ren instalazioa.....	3
1.1 Oinarrizko liburutegiaren instalazioa.....	3
WINDOWS.....	3
LINUX.....	3
1.2 Miatu Java.....	3
1.3 Miatu .NET.....	3
1.4 Miatu PHP.....	4
1.5 Miatu Python.....	4
2 Miatu-ren dependentziak.....	4
2.1 Oinarrizko dependentziak.....	4
Windows.....	4
Linux.....	5
2.2 Miatu Java.....	5
Windows.....	5
Linux.....	5
2.3 Miatu .NET.....	5
Windows.....	5
Linux.....	5
2.4 Miatu PHP.....	5
Windows.....	6
Linux.....	6
2.5 Miatu Python.....	6
Windows.....	6
Linux.....	6
3 Miaturen erabilera.....	6
3.1 Miaturen APIa.....	6
3.2 Miaturen galdera lengoaia.....	9

## 1 Miatu-ren instalazioa

Miatu bi zati nagusitan banatzen da. Alde batetik, Miatu oinarritzko liburutegia, eta bestetik, liburutegi horretan oinarritzen diren goi mailako lengoaien interfazeak. Goi mailako lengoaien interfaze guztiek behar dute oinarritzko liburutegia instalatuta egotea, beraz, lehenengo eta behin oinarritzko liburutegia nola instalatu azalduko dugu labur-labur. Miaturen instalazioa oso erraza da ikusiko duzuen moduan. Dena dela, badaude aldez aurretik instalatu beharreko zenbait liburutegi. Hauei buruzko argibideak hurrengo atalean aurkituko dituzue.

### 1.1 Oinarritzko liburutegiaren instalazioa

#### WINDOWS

eleka\_miatu.dll fitxategia PATH ingurune aldagaia adierazitako direktorioen barruan egon behar da. Horretarako bi aukera daude:

- eleka\_miatu.dll C:\windows\ karpetan kopiatu (beste karpeta batzuk ere onargarriak izan daitezke).
- eleka\_miatu.dll liburutegia zure gustuko karpeta batean kopiatu eta PATH inguruneko aldagaian karpeta hori gehitu.

#### LINUX

libeleka\_miatu.so fitxategia estekatzaileak aurkitzeko moduko toki batean kopiatu behar da. Horretarako aukera desberdinak daude:

- libeleka\_miatu.so /usr/local/lib karpetan kopiatu (beste karpeta batzuk ere onargarriak izan daitezke).
- libeleka\_miatu.so zure gustuko karpeta batean kopiatu eta LD\_LIBRARY\_PATH inguruneko aldagaian balioan karpeta hori erantsi.

### 1.2 Miatu Java

Instalazio prozedura berbera da Linux eta Windowsentzat.

1. Lotura liburutegia (libmiatu.so linuxen eta miatu.dll windowsen) instalatu 1.1 atalean azaltzen den moduan.
2. miatu.jar fitxategia java inguruneak aurkitzeko moduko tokian jarri. Ingurunearen arabera modu desberdinean egiten da hori: konfigurazio fitxategi bidez, miatu.jar fitxategiaren bide osoa CLASSPATH inguruneko aldagaian gehituz, java komandoaren -cp parametroan miatu.jar gehituz, ...

### 1.3 Miatu .NET

Instalazio prozedura berbera da Linux eta Windowsentzat.

1. Lotura liburutegia (libeleka.so linuxen eta eleka.dll windowsen) instalatu 1.1 atalean azaltzen den moduan.
2. miatu.dll fitxategia .NET inguruneak aurkitzeko moduko tokian jarri.

Ingurunearen arabera modu desberdinean egiten da hori: Visual Studio proiektuetan “Agregar Referencia” funtzioa erabiltzen da .NET assembly bat proiektuan gehitzeko. Mono modu oinarritzkoena mcs komandoaren -r parametroan gehitzea da.

## 1.4 Miatu PHP

Miatu PHP, PHPko “extension” bat da. Instalatzeko modu errazena php.ini fitxategia egokitzea da:

1. Sistemako php.ini fitxategian *extension\_dir* aldagaiak esaten duen direktorioan kopiatu php\_miatu.so fitxategia.
2. Sistemako php.ini fitxategian *extension* = “*php\_miatu.so*” lerroa gehitu.

Miatuarekin batera banatzen den php.ini fitxategian agertzen da prozedura honen adibide bat.

## 1.5 Miatu Python

eleka.py eta \_eleka.py fitxategiak PYTHONPATH barruan sartu behar dira. Horretara bi aukera nagusi daude:

- eleka.py eta \_eleka.py zure gustuko karpeta batean kopiatu eta PYTHONPATH aldagaian karpeta hori gehitu.
- eleka.py eta \_eleka.py fitxategiak /usr/local/lib/python2.4/site-packages karpetan kopiatu (beste karpeta batzuk ere onargarriak izan daitezke).

## 2 Miatu-ren dependentziak

Atal honetan Miatu instalatu ahal izateko kontutan hartu behar diren xehetasunak eta ase behar diren dependentziak zeintzuk diren azaltzen da.

### 2.1 Oinarritzko dependentziak

Miatu C++en garatutako liburutegi bat da. Liburutegi honek hainbat oinarritzko dependentzia ditu. Oinarritzko dependentziak dira Miatuaren interfaze guztiek (java, php, .NET...) behar dituzten dependentziak. Interfaze hauetako bakoitzak bere dependentzia gehigarri propioak ditu.

#### WINDOWS

Miatuk glib liburutegia (bere dependentziekin), Berkely DB 4.4, zlib eta pcre liburutegia behar ditu. Liburutegi hauek jarraian datozen tokietan lor ditzakezu behar izanez gero.

- glib, bere dependentzia diren gnulibconv eta gettext, eta zlib:  
<http://www.gimp.org/~tml/gimp/win32/downloads.html>
- Berkeley DB:  
<http://www.oracle.com/technology/software/products/berkeley-db/db/index.html> (4.4 bertsioa gomendatzen da).
- pcre: <http://www.psyon.org/projects/pcre-win32/>

## **LINUX**

Miatu linux bertsioak dituen dependentziak Windows bertsioaren berdina dira. Miatu Debian Etch makina baten garatu eta probatu da. Etch-en beharrezko paketeak aptitude bidez eskuratzea posible da. Hau da instalatuta egon behar diren paketeen zerrenda:

- libglib2.0 eta bere dependentziak.
- libdb4.4: Berkeley DB 4.4 bertsioa.
- libpcre3.
- Zlib1g.

Beste edozein Linux banaketatan ere egongo dira paketa hauek. Zure banaketaren eskuliburua kontsultatu.

## **2.2 Miatu Java**

Miatu Java Miatu Javatik erabiltzea ahalbidetzen duen liburutegia da. Beharrezkoa da java instalatuta egotea erabiltzeko. Ez du oinarritzoez aparte dependentzia gehigarririk.

## **WINDOWS**

Sunen javaren 1.3 bertsioarekin dago estekatuta. Bertsio berriagoekin ere funtzionatuko du baina probatu gabe dago.

## **LINUX**

Sunen javaren 1.5 bertsioarekin dago estekatuta.

## **2.3 Miatu .NET**

Miatu .NET Miatu .NET plataformatik erabiltzea ahalbidetzen duen liburutegia da. Beharrezkoa da .NET instalatuta egotea erabiltzeko. Ez du oinarritzoez aparte dependentzia gehigarririk.

## **WINDOWS**

MS .NET Framework 1.1ekin dago estekatuta.

## **LINUX**

Mono 1.2rekin dago garatuta, Debian Etch-en datorren bertsioa.

## **2.4 Miatu PHP**

Miatu PHP Miatu PHPtik erabiltzea ahalbidetzen duen liburutegia da. Beharrezkoa da PHP instalatuta egotea erabiltzeko. Ez du oinarritzoez aparte dependentzia gehigarririk.

## WINDOWS

Windows plataformarako Miatu PHPren bertsiorik ez dago. Produktu honetan interesa izanez gero **ELEKA Ingeniaritza Linguistikoa, S.L.**rekin harremanetan jartzea gomendatzen dizugu ([www.eleka.net](http://www.eleka.net)).

## LINUX

PHP 4.3 bertsioarekin dago konpilatuta.

## 2.5 Miatu Python

Miatu Python Miatu Pythonetik erabiltzea ahalbidetzen duen liburutegia da. Beharrezkoa da Python instalatuta egotea erabiltzeko. Ez du oinarritzoez aparte dependentzia gehigarririk.

## WINDOWS

Windows plataformarako Miatu Pythonen bertsiorik ez dago. Produktu honetan interesa izanez gero **ELEKA Ingeniaritza Linguistikoa, S.L.**rekin harremanetan jartzea gomendatzen dizugu ([www.eleka.net](http://www.eleka.net)).

## LINUX

Python 2.4 bertsioarekin dago konpilatuta.

## 3 Miaturen erabilera

Miaturen erabilera aztertzeko gomendagarriena liburutegiarekin batera banatzen diren adibideak aztertzea da. Atal honetan adibide guzti horiek amankomunean dituzten ezaugarriak eta kokapen finkorik ez duten beste zenbait xehetasun zerrendatzen dira.

Miatu liburutegi moduan aurkezten da. Liburutegi horrek 2 klase eskaintzen ditu. Erabiltzen den lengoaiaren arabera klase hauen eta beraien funtzioen deien sintaxia aldatu egiten da baina funtzionamendua mantendu egiten da. Jarraian bi klase horien funtzio nagusiak azaltzen dira labur C++ sintaxia erabiliz.

### 3.1 Miaturen APIa

Jarraian datorren taulan MIATUren interfaze publikoa osatzen duten funtzioak deskribatzen dira erraz kontsultatzeko formatuan.

## MIATU Erraz. Eskuliburua

Izena	Sarrerak	Azalpena
Class Miatu	-	Miatu klasearen instantziak indizearen abstrakzioa dira. Nahi adina izan ditzakezu batera martxan, prozesu desberdinetatik ere bai eta Miatu arduratuko da beharrezkoa koordinazioaz.
Miatu* Miatu::createInstance(const char *dbdir)	dbdir: DB fitxategia (miatu.index) sortu behar duen direktorioaren izena. Existitzen den direktorio bat izan behar da. Bertan, indizeaz gain, sinkronizazio eta atzipen konkurrenterako beharrezkoak diren zenbait fitxategi sortuko dira.	Miatu klasearen instantzia berri bat. Direktorio horretan indize bat existitzen bada ireki egingo du eta bestela berri bat sortu.
MiatuErrorCode Miatu::index(const char* url, const char* content, const char* mimetype)	<ul style="list-style-type: none"> <li>● url: dokumentu honek izango duen izena (normalean url bat).</li> <li>● content: dokumentuaren edukia.</li> <li>● mimetype: edukiaren mime mota. "text/plain" edo "text/html" izan behar da. Content mota honekin bat etorri behar da. Baldintza hauek betetzen ez badira Miatu-ren funtzionamendua ez dago definituta. Beste dokumentu motak erabili ahal izateko <i>ELEKA Ingeniaritza Linguistikoarekin</i> kontsultatu.</li> </ul>	MEC_SUCCESS dokumentua ondo indexatu bada, errore kode bat bestela.
void Miatu::close()	-	Indizera erreferentzia askatzen du. Lengoia gehienetan ez dago deitu beharrik. Salbuespena Java da. Erreferentzia guztiak askatzen indize datu basea itxi egiten da.
ResultIterator* Miatu::lookup(const char *query)	query: ebatzi beharreko galdera.	Miatuk galdera sintaxi sinple baina ahaltsua eskaintzen du, kasu gehienetan nahikoa izango dena. Indizearen gainean bilaketa konplexuagoak egitea beharrezkoa den kasuetarako kontsultatu <i>ELEKA Ingeniaritza Linguistikoarekin</i> . <b>Sintaxiaren nondik norakoak eskuliburuko 3.2 atalean azaltzen dira.</b> Emaizta ResultIterator instantzia bat da edo NULL erroreren bat gertatu bada.

## MIATU Erraz. Eskuliburua

Izena	Sarrerak	Azalpena
<code>class ResultIterator</code>	-	Eraitza moduan URL multzo baten abstrakzioa den klasearen instantzia bat ematen du. Instantzia honek indizera erreferentzia bat gordetzen duenez, posible da Miatu::close deitu eta gero ere instantzia erabiltzea.
<code>const char * ResultIterator::next()</code>	-	Eraitza multzoko hurrengo osagaia ematen du edo NULL, osagai gehiago ez badaude.
<code>void ResultIterator::close()</code>	-	Indizera erreferentzia askatzen du. Lengoia gehienetan ez dago deitu beharrik. Salbuespena Java da. Erreferentzia guztiak askatzen indize datu basea itxi egiten da.



### 3.2 Miaturen galdera lengoaia

Miaturen galdera lengoaia sinplea baina ahaltsua da eta kasu gehienetan nahikoa izatea espero da. Dena dela bilaketa ahaltsuagoak egiteko premia izanez gero, jo *ELEKA Ingeniaritza Linguistikoara*.

Miaturen galdera lengoaia lookup funtzioaren deietan agertzen da. Karaktere alfabetikoz osatutako hitzak bakarrik onartzen ditu. Eraitza moduan galderan emandako hitz guztiak dituzten dokumentuen zerrenda ematen du, hau da, hitz guztien arteko AND bat egiten du.

Indizean bilaketa egin aurretik hitzak lematizatu egiten ditu. Honek ematen dizkio Miaturi bere ezaugarri bereziak, Euskararako ezinbestekoak direnak. Adibidez, bilaketa gako moduan “amaren” emanez gero, indizean “amaren” bilatzeaz gain “ama”, “amak”, “amek”, “amaren”, “amatik”, ... bilatuko ditu, baina ez “amonaren” edo “amigdala”, MIATUK dituen lematizazio funtzio aurreratuak ez dituen bilatzaile batek egingo lukeen moduan.