

MANUAL DE INSTALACIÓN Y MONTAJE DE BASES DE DATOS NOSQL (COUCHDB Y ELASTICSEARCH) DE LA DATA PARA LOS APLICATIVOS DEL CLÚSTER DE SOLUCIONES NUVEM CNCFLORA BRASIL

VERSION 1.0 – 01-12-2016

Sede Principal: Calle 28A #15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



Elaborado Por: Jaime Alberto Gutiérrez Mejía

ÍNDICE

ÍNDICE.....	2
1. OBJETO DEL DOCUMENTO	3
2. PARTICIPANTES	3
3. OBJETIVOS.....	3
4. PREPARACIÓN DE LOS DATOS PARA EL PROCESO MIGRATORIO.....	4
5. EJECUCIÓN DEL CLIENTE DE GENERACIÓN DE ARCHIVOS DE CARGUE MASIVO EN COUCHDB Y ELASTICSEARCH	9
6. REALIZACIÓN DEL CARGUE AL AMBIENTE DE BASES DE DATOS NO RELACIONALES.....	23



1. OBJETO DEL DOCUMENTO

Este documento tiene como objeto ilustrar cómo debe realizarse el proceso de cargue de bases de datos noSQL (COUCHDB y ELASTICSEARCH) del conjunto de aplicaciones del clúster NUVEM CNCFLORES BRASIL, con el fin de poder registrar de forma automatizada para el core de soluciones de backoffice de la solución, los datos de las correspondientes bases de categorías (familias) de registros de especies biológicas.

2. PARTICIPANTES

En este proceso tiene principal participación el Administrador Técnico del Sistema, el jefe de gestión de aplicaciones del Instituto Alexander von Humboldt y los participantes del proceso de despliegue y paso a producción.

3. OBJETIVOS

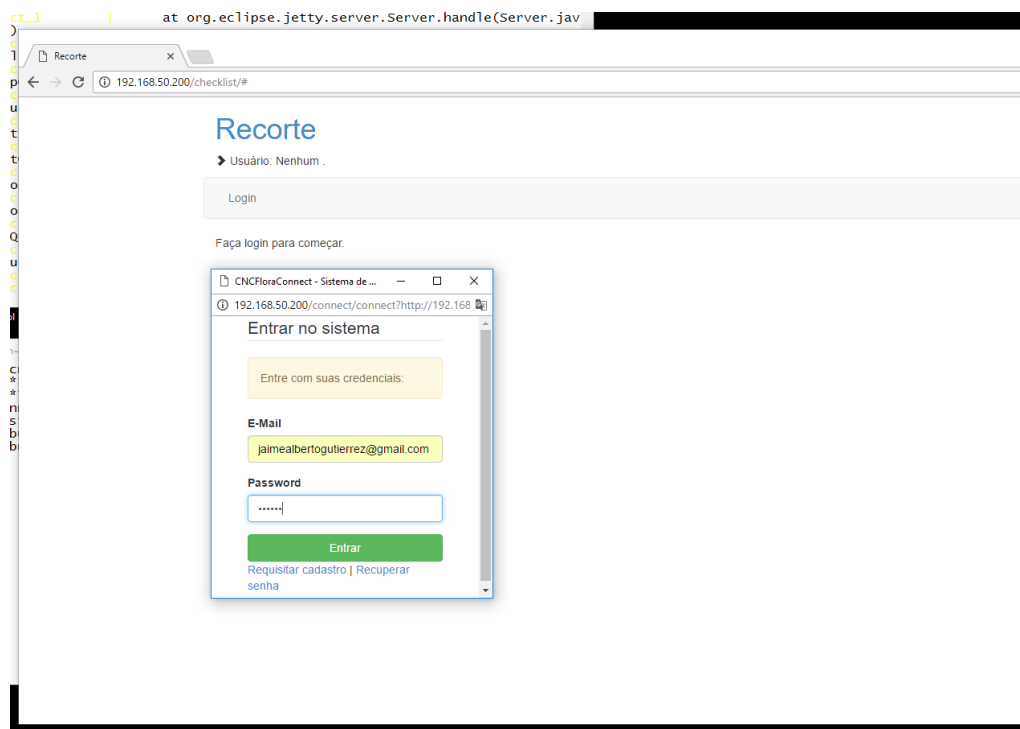
- a) Realizar la preparación de los datos para el cargue al Sistema
- b) Ejecutar el cliente de construcción de los archivos de generación de cargue masiva de registros para COUCHDB y la indexación de los mismos con Elasticsearch.
- c) Desarrollar el proceso de cargue de los datos desde la consola del Sistema operativo del ambiente de ejecución
- d) Validar la data cargada en los aplicativos del clúster NUVEM



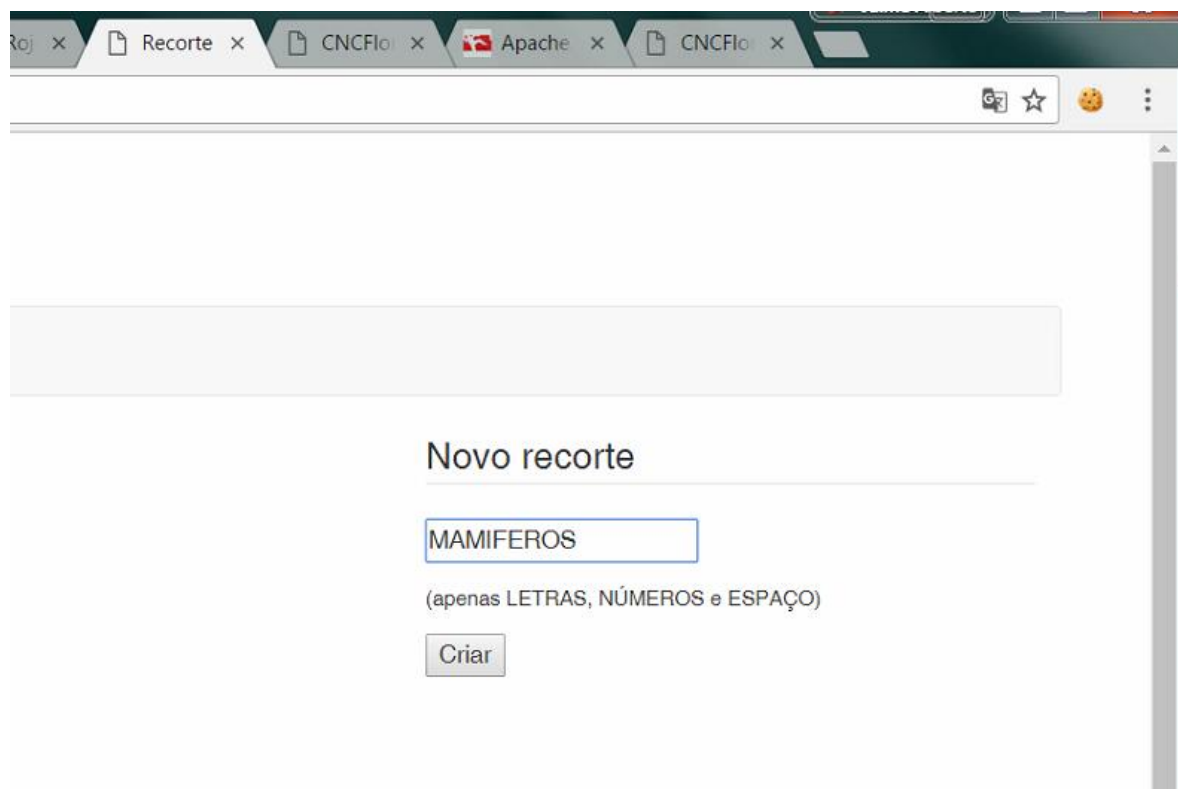
4. PREPARACIÓN DE LOS DATOS PARA EL PROCESO MIGRATORIO

Con el objeto de desarrollar el proceso de migración de datos a la solución de NUVEM CNCFLORA, el usuario deberá preparar la data de las diferentes bases de datos, previa creación de las bases concretas (denominadas en el ambiente de producción RECORTES). Para ello, se deberá ingresar al Sistema por alguna de las dos siguientes vías:

- a) **VÍA ADMINISTRACIÓN CHECKLIST:** Para construir las bases requeridas, el usuario deberá ingresar por el aplicativo Web de LISTA DE CHEQUEO (Checklist) <http://192.168.11.67/checklist>, autenticarse con el usuario previamente creado desde el aplicativo de gestión de SSO (ENTRYPOINT: <http://192.168.11.67/entrypoint/>), tal como se muestra a continuación:



Una vez el usuario se autentique se puede proceder a crear la base, escribiendo el nombre de la misma en el formulario Nuevo Recorte:

A screenshot of a web browser window showing a form titled "Novo recorte". The browser's address bar shows "CNCFlo" and the page title is "Recorte". The form has a text input field containing "MAMIFEROS" and a button labeled "Criar". Below the input field, there is a note in Portuguese: "(apenas LETRAS, NÚMEROS e ESPAÇO)".

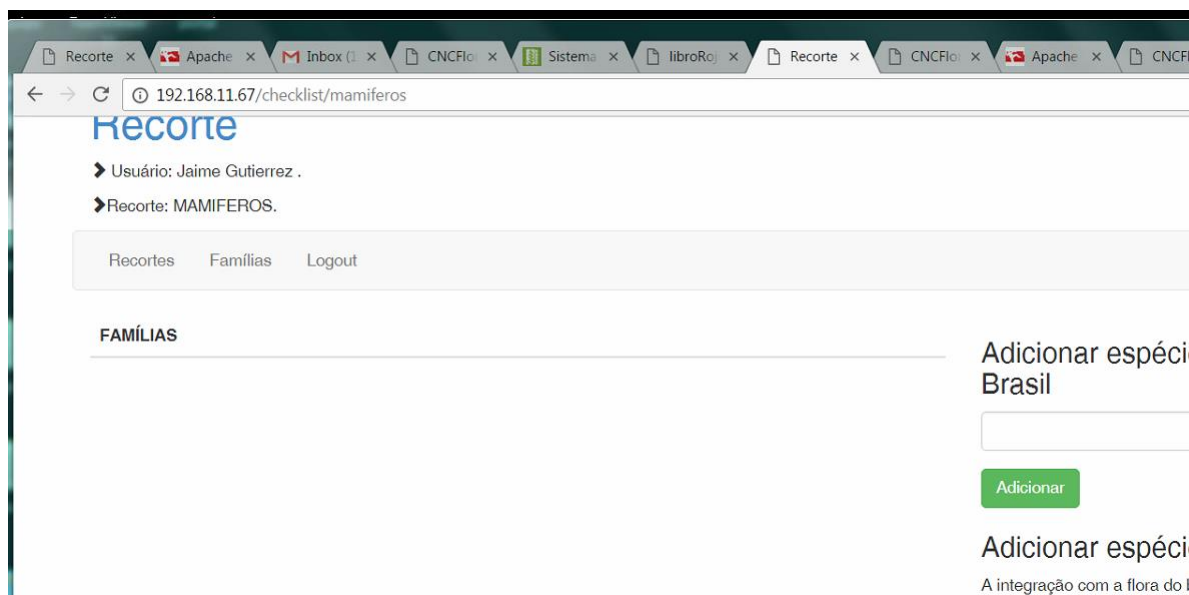
Novo recorte

(apenas LETRAS, NÚMEROS e ESPAÇO)



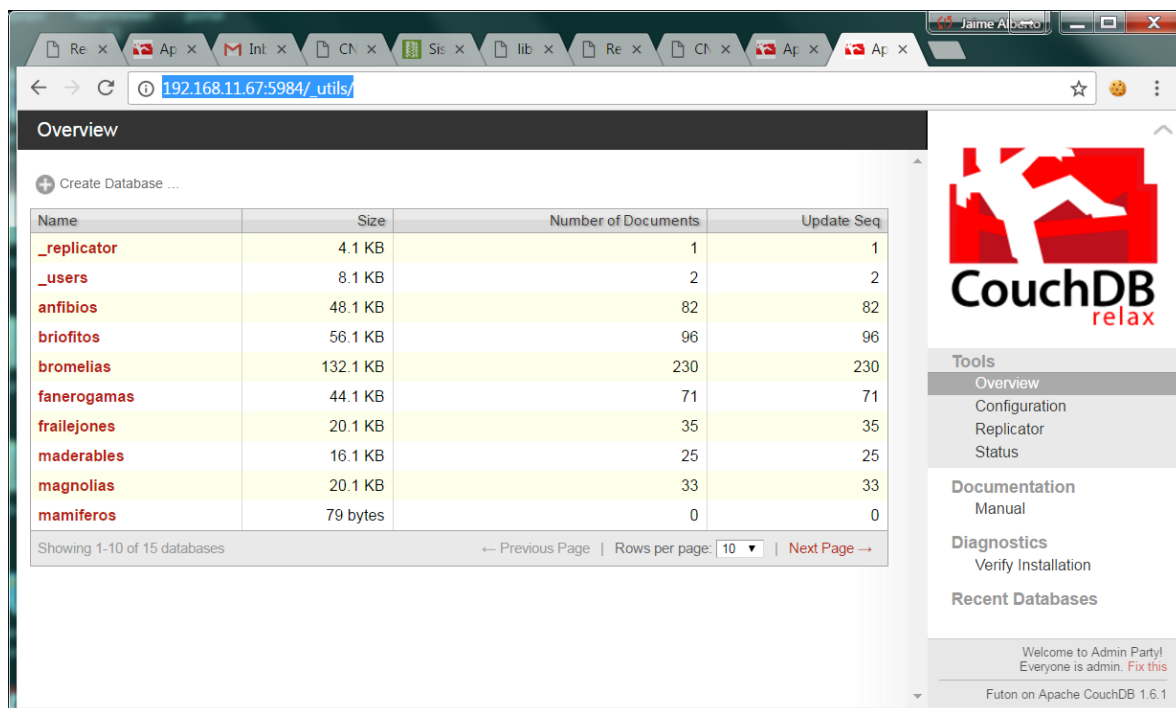
Recortes

- ANFIBIOS
- BRIOFITOS
- BROMELIAS
- FANEROGAMAS
- FRAILEJONES
- MADERABLES
- MAGNOLIAS
- MAMIFEROS
- MIRISITICACEAE
- ORQUIDEAS
- PALMAS



- b) **VÍA ADMINISTRACIÓN FUTON:** Si fuera a realizarse el proceso directamente desde el gestor de base de datos no relacional de COUCHDB, se deberá realizar inicio de sesión en la consola del administrador de la base (FUTON: <http://192.168.11.67:5984/ utils/>), para poder realizar la creación de las bases de agrupación de especies (familias).





The screenshot shows the CouchDB Futon interface. The main table lists 15 databases. The first 10 are visible, and the rest are on the next page. The table has columns for Name, Size, Number of Documents, and Update Seq.

Name	Size	Number of Documents	Update Seq
_replicator	4.1 KB	1	1
_users	8.1 KB	2	2
anfibios	48.1 KB	82	82
briofitos	56.1 KB	96	96
bromelias	132.1 KB	230	230
fanerogamas	44.1 KB	71	71
frailejones	20.1 KB	35	35
maderables	16.1 KB	25	25
magnolias	20.1 KB	33	33
mamiferos	79 bytes	0	0

Showing 1-10 of 15 databases

← Previous Page | Rows per page: 10 | Next Page →

Tools: Overview, Configuration, Replicator, Status

Documentation: Manual

Diagnostics: Verify Installation

Recent Databases

Welcome to Admin Party!
Everyone is admin. [Fix this](#)

Futon on Apache CouchDB 1.6.1

Proceder a iniciar sesión en la parte inferior con las credenciales creadas para el administrador. En el caso de este proyecto:

Username: admin

Password: ncc&123x

Una vez se haya iniciado sesión en el administrador de FUTON, se procede a escoger la opción CREATE DATABASE y se ingresa el nombre del nuevo recorte o base para CouchdDB.



The screenshot shows the Apache CouchDB Futon interface. A red box highlights the 'Create Database ...' button in the top left. Another red box highlights the 'Create New Database' dialog box, which is open. The dialog box contains the following text:

Create New Database

Please enter the name of the database. Note that only lowercase characters (a-z), digits (0-9), or any of the characters `_`, `.`, `(`, `)`, `+`, `-`, and `/` are allowed.

Database Name:

Buttons: Create, Cancel

Name	Size	Number of Documents
_replicator	4.1 KB	1
_users	8.1 KB	2
anfibioscolombia	80.1 KB	83
brifitos	56.1 KB	96
bromelias	132.1 KB	230
fanerogamas	44.1 KB	71
frailejones	20.1 KB	35
maderables	16.1 KB	25
magnolias	20.1 KB	33
miristicaceae	16.1 KB	17

Showing 1-10 of 14 databases

Automáticamente la base quedará creada y lista para la inserción de nuevos registros de especies biológicas.

The screenshot shows the Apache CouchDB Futon interface. The 'mamiferos' database is now listed in the table. A blue box highlights the bottom of the table, showing the 'mamiferos' entry.

Name	Size	Number of Documents	Update Seq
_replicator	4.1 KB	1	1
_users	8.1 KB	2	2
anfibioscolombia	80.1 KB	83	88
brifitos	56.1 KB	96	96
bromelias	132.1 KB	230	230
fanerogamas	44.1 KB	71	71
frailejones	20.1 KB	35	35
maderables	16.1 KB	25	25
magnolias	20.1 KB	33	33
mamiferos	79 bytes	0	0

Showing 1-10 of 15 databases

Navigation: Previous Page | Rows per page: 10 | Next Page



5. EJECUCIÓN DEL CLIENTE DE GENERACIÓN DE ARCHIVOS DE CARGUE MASIVO EN COUCHDB Y ELASTICSEARCH

Para el proceso de cargue masivo de registros a las bases de datos del Sistema, se debe preparar la creación de los registros, empleando una plantilla de cargue masiva, la cual incluye los siguientes campos:

Campo	Descripción
id	Identificador único de la Especie
nombreCientifico	Nombre científico nominal de la Especie
kingdom	Familia taxonómica a la que se asocia la especie
phylum	Filum de agrupación de la Especie
clase	Clase de categorización
orden	Orden de distribución
family	Familia de la especie
genus	Genus taxonómico
specificEpithet	Epíteto de clasificación organizacional
rango	Rango de definición (especie, subespecie o derivación)
autor	Autor científico de investigación de la especie
nombreComun	Nombre común o genérico de la especie

Al realizar la organización, el documento debe quedar así:



reptiles - Microsoft Excel

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA COMPLEMENTOS TEAM

Calibri 11 A+ Fuente Alineación Número Estilos Estilos de celdas Insertar Eliminar Formato Celdas

Portapapeles Pegar Fuente Alineación Número Estilos Estilos de celdas Insertar Eliminar Formato Celdas

Autosuma Rellenar Ordenar y filtrar Buscar y seleccionar

Inicio sesión

	A	B	C	D	E	F	G	H	I	J	K
1	effaf81a-29c1-102b-9a4a-00304854f821	Anolis ienderenae	Animalia	Chordata	Reptilia	Squamata	Dactyloidae	Anolis	ienderenae	species	Rueda y Hernández, 1988
2	effaf81a-29c1-102b-9a4a-00304854f822	Ptychoglossus danieli	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Ptychoglossus	danieli	species	Harris, 1994
3	effaf81a-29c1-102b-9a4a-00304854f823	Anadia pamplonensis	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Anadia	pamplonensis	species	Dunn, 1944
4	effaf81a-29c1-102b-9a4a-00304854f824	Anolis calimae	Animalia	Chordata	Reptilia	Squamata	Dactyloidae	Anolis	calimae	species	Ayala, Harris y Williams, 1983
5	effaf81a-29c1-102b-9a4a-00304854f825	Anolis ruizii	Animalia	Chordata	Reptilia	Squamata	Dactyloidae	Anolis	ruizii	species	Rueda y Williams, 1986
6	effaf81a-29c1-102b-9a4a-00304854f826	Lepidoblepharis miyatai	Animalia	Chordata	Reptilia	Squamata	Sphaerodactylidae	Lepidoblepharis	miyatai	species	Lamar, 1985
7	effaf81a-29c1-102b-9a4a-00304854f827	Riama columbiana	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Riama	columbiana	species	Andersson, 1914
8	effaf81a-29c1-102b-9a4a-00304854f828	Riama simotera	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Riama	simotera	species	O'Shaughnessy, 1879
9	effaf81a-29c1-102b-9a4a-00304854f829	Morunasaurus groi	Animalia	Chordata	Reptilia	Squamata	Holpocercidae	Morunasaurus	groi	species	Dunn, 1933
10	effaf81a-29c1-102b-9a4a-00304854f830	Lepidoblepharis williamsi	Animalia	Chordata	Reptilia	Squamata	Sphaerodactylidae	Lepidoblepharis	williamsi	species	Ayala y Serna, 1986
11	effaf81a-29c1-102b-9a4a-00304854f831	Anadia antioquiensis	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Anadia	antioquiensis	species	Arredondo, 2013
12	effaf81a-29c1-102b-9a4a-00304854f832	Aristelliger georgeensis	Animalia	Chordata	Reptilia	Squamata	Sphaerodactylidae	Aristelliger	georgeensis	species	Bocourt, 1873
13	effaf81a-29c1-102b-9a4a-00304854f833	Anadia bogotensis	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Anadia	bogotensis	species	Peters, 1862
14	effaf81a-29c1-102b-9a4a-00304854f834	Riama laevis	Animalia	Chordata	Reptilia	Squamata	Gymnophthalmidae	Riama	laevis	species	Boulenger, 1908
15	effaf81a-29c1-102b-9a4a-00304854f835	Enyaliodes oshaughnessyi	Animalia	Chordata	Reptilia	Squamata	Holpocercidae	Enyaliodes	oshaughnessyi	species	Boulenger, 1881
16	effaf81a-29c1-102b-9a4a-00304854f836	Morunasaurus annularis	Animalia	Chordata	Reptilia	Squamata	Holpocercidae	Morunasaurus	annularis	species	O'Shaughnessy, 1881
17	effaf81a-29c1-102b-9a4a-00304854f837	Dendrophidion boshelli	Animalia	Chordata	Reptilia	Squamata	Colubridae	Dendrophidion	boshelli	species	Dunn, 1944
18	effaf81a-29c1-102b-9a4a-00304854f838	Coniophanes andresensis	Animalia	Chordata	Reptilia	Squamata	Colubridae	Coniophanes	andresensis	species	Bailey, 1937
19	effaf81a-29c1-102b-9a4a-00304854f839	Synophis plectroverbralis	Animalia	Chordata	Reptilia	Squamata	Dipsadidae	Synophis	plectroverbralis	species	Shell y Grant, 2001
20	effaf81a-29c1-102b-9a4a-00304854f840	Micrurus medemi	Animalia	Chordata	Reptilia	Squamata	Elapidae	Micrurus	medemi	species	Roze, 1967
21	effaf81a-29c1-102b-9a4a-00304854f841	Atractus punctiventris	Animalia	Chordata	Reptilia	Squamata	Colubridae	Atractus	punctiventris	species	Amaral, 1933
22	effaf81a-29c1-102b-9a4a-00304854f842	Saphenophis seniderni	Animalia	Chordata	Reptilia	Squamata	Dipsadidae	Saphenophis	seniderni	species	Myers, 1973
23	effaf81a-29c1-102b-9a4a-00304854f843	Micrurus sanglensis	Animalia	Chordata	Reptilia	Squamata	Elapidae	Micrurus	sanglensis	species	Nicéforo María, 1942
24	effaf81a-29c1-102b-9a4a-00304854f844	Atractus orcesi	Animalia	Chordata	Reptilia	Squamata	Dipsadidae	Atractus	orcesi	species	Savage, 1955

reptiles

LISTO

La estructura del ID debe seguir la nomenclatura que se cita a continuación:

effaf81a-29c1-102b-9a4a-00304854f826

- Primer bloque: 8 dígitos hexadecimales
- Segundo bloque: 4 dígitos hexadecimales
- Tercer bloque: 4 dígitos hexadecimales
- Cuarto bloque: 4 dígitos hexadecimales
- Quinto bloque: 12 dígitos hexadecimales

Y cada ID de registro debe ser absolutamente diferente del siguiente, para mantener la integridad de entidad en este modelo.

Una vez el archivo se organice en formato CSV, el paso siguiente, es verificar que no existan los símbolos de tabulación (TAB), punto y coma (;), dado que este último es el separador de campos en el archivo de interface y finalmente las comillas dobles o simples, las cuales deben ser removidas de los registros.

Sede Principal: Calle28A#15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



Una vez el archivo esté correspondientemente formado, debe procederse a la ejecución de la aplicación cliente JSE desarrollada para la construcción de forma automatizada de los archivos tanto para COUCHDB como para ElasticSearch. Desde la consola de comandos debe ejecutarse

java -jar GeneradorCouch.jar

Prevía validación de la instalación del ambiente JDK de Oracle y del HotSpot correspondiente para la ejecución de aplicaciones

```
C:\WINDOWS\system32\cmd.exe

C:\Users\Jaimiko\Desktop>java -version
java version "1.8.0_73"
Java(TM) SE Runtime Environment (build 1.8.0_73-b02)
Java HotSpot(TM) 64-Bit Server VM (build 25.73-b02, mixed mode)

C:\Users\Jaimiko\Desktop>_
```

Si no se puede ejecutar el comando, debe descargarse desde la página oficial de Oracle en la sección DOWNLOADS/JSE TECHNOLOGY (<http://www.oracle.com/technetwork/java/javase/downloads/index.html>)



Oracle Technology Network > Java > Java SE > Downloads

Java SE Downloads

Java Platform (JDK) 8u121

NetBeans with JDK 8

Java Platform, Standard Edition

Java SE 8u121
Java SE 8u121 includes important security fixes. Oracle strongly recommends that all Java SE 8 users upgrade to this release.
[Learn more](#)

Important planned change for MD5-signed JARs
Starting with the April Critical Patch Update releases, planned for April 18 2017, all JRE versions will treat JARs signed with MD5 as unsigned. [Learn more and view testing instructions](#)
For more information on cryptographic algorithm support, please check the JRE and JDK Crypto Roadmap.

Installation Instructions

Release Notes

JDK

DOWNLOAD

Java Developer Day hands on workshops group and other events

- Java Magazine

JDK 8u121 checksum

Java SE Development Kit 8u121

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

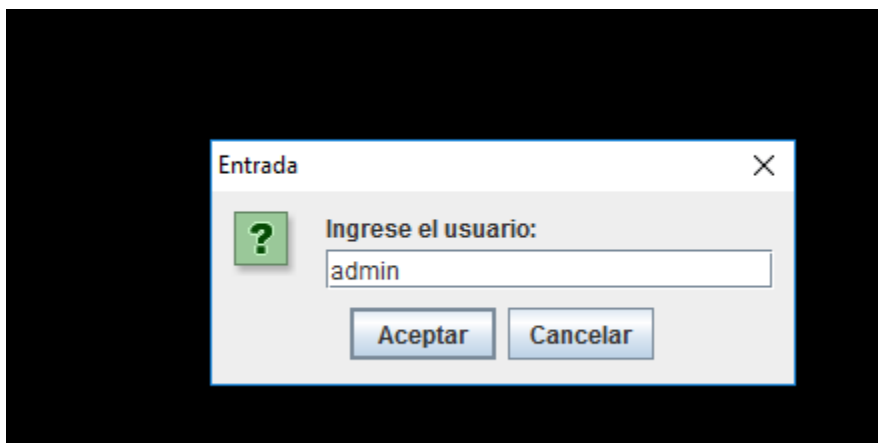
Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	77.86 MB	jdk-8u121-linux-arm32-vfp-hflt.tar.gz
Linux ARM 64 Hard Float ABI	74.83 MB	jdk-8u121-linux-arm64-vfp-hflt.tar.gz
Linux x86	162.41 MB	jdk-8u121-linux-i586.rpm
Linux x86	177.13 MB	jdk-8u121-linux-i586.tar.gz
Linux x64	159.96 MB	jdk-8u121-linux-x64.rpm
Linux x64	174.76 MB	jdk-8u121-linux-x64.tar.gz
Mac OS X	223.21 MB	jdk-8u121-macosx-x64.dmg
Solaris SPARC 64-bit	139.64 MB	jdk-8u121-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit	99.07 MB	jdk-8u121-solaris-sparcv9.tar.gz
Solaris x64	140.42 MB	jdk-8u121-solaris-x64.tar.Z
Solaris x64	96.9 MB	jdk-8u121-solaris-x64.tar.gz
Windows x86	189.36 MB	jdk-8u121-windows-i586.exe
Windows x64	195.51 MB	jdk-8u121-windows-x64.exe

Java SE Development Kit 8u121 Demos and Samples Downloads

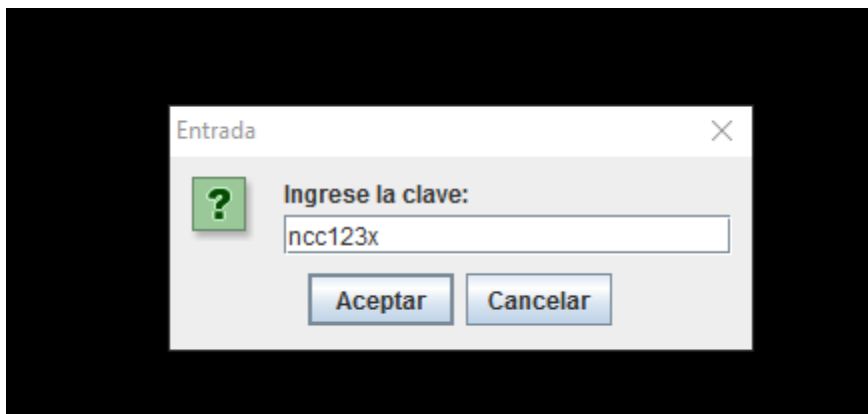
Sede Principal: Calle28A#15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



- 1) Lo primero que hará la aplicación será solicitar el usuario de COUCHDB que está configurado en la consola de FUTON:

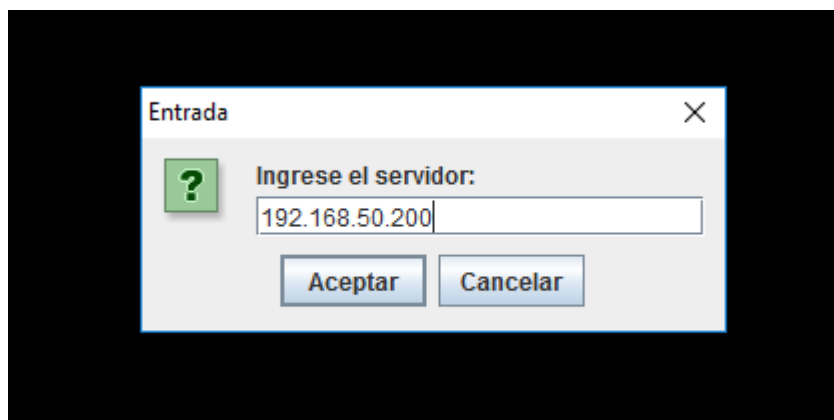


- 2) Posteriormente se ingresa la contraseña asignada a dicho administrador

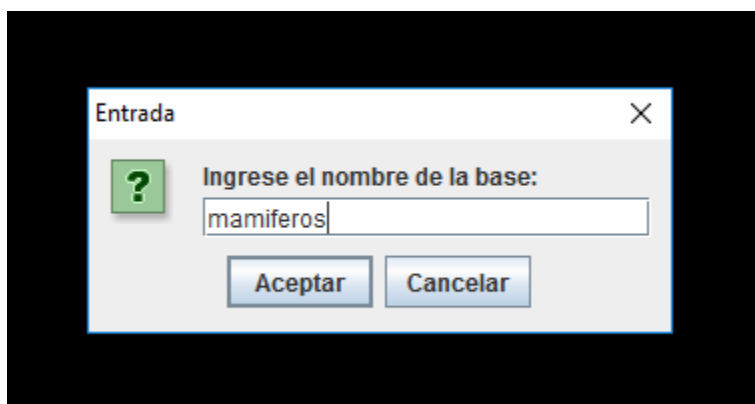


- 3) Acto seguido debe ingresarse la dirección IP o nombre DNS del servidor que aloja el servicio de CouchDB (No debe ingresarse el puerto, ya que este por defecto está configurado en el 5498).



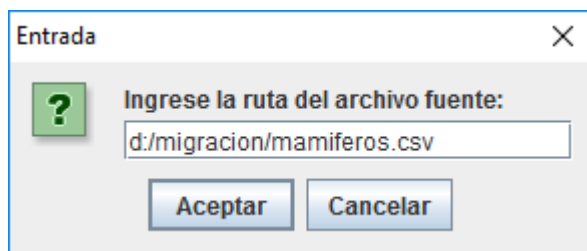


- 4) El aplicativo solicitará el nombre de la base (Esta se corresponde al nombre exacto (en letras minúsculas del índice o relatorio creado en COUCHDB) o a través del aplicativo de CHECKLIST.

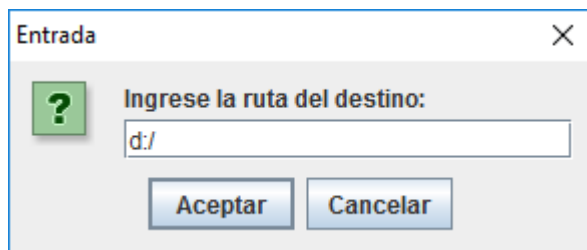


- 5) El aplicativo solicitará la ruta completa (Linux o Windows) de la ubicación del archivo fuente en formato CSV que contiene el listado completo de todas las especies. Se deberá digitar, como en el ejemplo que se ilustra a continuación:



A screenshot of a Windows-style dialog box titled "Entrada" with a close button (X) in the top right corner. Inside the dialog, there is a green square icon with a white question mark. To the right of the icon, the text "Ingrese la ruta del archivo fuente:" is displayed. Below this text is a text input field containing the path "d:/migracion/mamiferos.csv". At the bottom of the dialog, there are two buttons: "Aceptar" and "Cancelar".

- 6) Posteriormente, digitar el path o ruta de destino en donde va a quedar el archivo destino tanto en COUCHDB como en Elasticsearch.

A screenshot of a Windows-style dialog box titled "Entrada" with a close button (X) in the top right corner. Inside the dialog, there is a green square icon with a white question mark. To the right of the icon, the text "Ingrese la ruta del destino:" is displayed. Below this text is a text input field containing the path "d:/". At the bottom of the dialog, there are two buttons: "Aceptar" and "Cancelar".

- 7) El aplicativo realiza todo el procesamiento y genera la salida en la carpeta destino de los archivos de la forma NOMBRE_FUENTE_RegistrosCouchDBXXXXXXXXXXXXXXXXXXXXX.json y NOMBRE_FUENTE_RegistrosElasticXXXXXXXXXXXXXXXXXXXXX.json, donde NOMBRE_FUENTE representa el nombre del archivo fuente de origen y XXXXXXXXXXXXXXXXXXXXXXXX un serial autonumérico extraído de la fecha hora del Sistema.




```
C:\WINDOWS\system32\cmd.exe
orden: Anura
family: Hemiphractidae
genus: Gastrotheca
specificEpithet: ruizi
rango: species
autor: Duellman y Burrowes, 1986
nombreComun: Rana marsupial de Ruiz
*****
*****
Leyendo registros de la fuente
*****
*****
ffffaf81a-39c1-202b-8a4a-00304854f876;Hemiphractus johnsoni;Animalia;Chordata;Amphibia;Anura;Hemiphractidae;Hemiphractus;
johnsoni;species;Noble, 1917;Rana cornuda de Johnson
Registro actual de ID = fffaf81a-39c1-202b-8a4a-00304854f876
id: fffaf81a-39c1-202b-8a4a-00304854f876
nombreCientifico: Hemiphractus johnsoni
kingdom: Animalia
phylum: Chordata
clase: Amphibia
orden: Anura
family: Hemiphractidae
genus: Hemiphractus
specificEpithet: johnsoni
rango: species
autor: Noble, 1917
nombreComun: Rana cornuda de Johnson
*****
*****
```

- 8) Al final de la ejecución de la aplicación, se generarán los correspondientes formatos de cargue de consola CURL, con su notación de acuerdo a los parámetros de entrada:

```
-8a4a-00304854f882", "scientificName": "Bolitoglossa pandi - Brame y Wake, 1963", "scientificNameWithoutAuthorship": "Boli
roglossa pandi", "scientificNameAuthorship": "Brame y Wake, 1963", "family": "Plethodontidae", "taxonRank": "species", "taxo
nomicStatus": "accepted", "acceptedNameUsage": "Bolitoglossa pandi - Brame y Wake, 1963", "metadata": {"identifier": "ffffaf
81a-39c1-202b-8a4a-00304854f882", "type": "taxon", "created": 1479335354, "modified": 1479329478, "creator": "Jaime Alberto Gu
tierrez", "contributor": "Jaime Alberto Gutierrez", "contact": "jaimealbertogutierrez@gmail.com", "source": "User"}}
curl -X POST -d @mamiferos-RegistrosCouchDb1485204497115.json -H "Content-Type: application/json" http://admin:ncc123x@1
92.168.50.200:5984/mamiferos/_bulk/docs
ruta de salida del archivo de ELASTICSEARCH: d://mamiferos-RegistrosElasticSearch1485204497115.json
curl -XPOST 192.168.50.200:9200/mamiferos/_bulk --data-binary @mamiferos-RegistrosElasticSearch1485204497115.json
C:\Users\Jaimiko\Desktop>
```

Ejemplo:

Sede Principal: Calle28A#15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



Cargue para COUCHDB:

curl -X POST -d @mamiferos-RegistrosCouchDb1485204497115.json -H "Content-Type: application/json" http://admin:ncc123x@192.168.50.200:5984/mamiferos/_bulk_docs

Cargue del Índice para ELASTICSEARCH:

curl -XPOST 192.168.50.200:9200/mamiferos/_bulk --data-binary @mamiferos-RegistrosElasticSearch1485204497115.json

9) Los archivos resultantes en formato JSON (JAVASCRIPT NOTATION OBJECT), aparecerán en la carpeta de destino:

ArchivosMartillo	19/07/2014 10:15 a...	Archivo WinRAR Z...	1.054 KB
data	10/12/2016 4:53 a...	JSON File	19 KB
Desarrollo	28/04/2016 12:24 a...	Archivo WinRAR Z...	1.088.512 KB
eclipse-jee-kepler-SR2-win32	26/04/2014 2:28 p...	Archivo WinRAR Z...	256.362 KB
mamiferos-RegistrosCouchDb1485204497115	23/01/2017 4:17 p....	JSON File	59 KB
mamiferos-RegistrosElasticSearch1485204497115	23/01/2017 4:17 p....	JSON File	60 KB
RegistrosCouchDb1480354894461	28/11/2016 12:42 ...	JSON File	35 KB
RegistrosElasticSearch1480354894461	28/11/2016 12:42 ...	JSON File	36 KB

Archivo de cargue COUCHDB:



[illegible][illegible]

Archivo de cargue ELASTICSEARCH:




```

{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f801"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f801","scientificName":"Allobates ranoides - Boul"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f802"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f802","scientificName":"Atelopus eusebianus - Rive"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f803"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f803","scientificName":"Atelopus farci - Lynch, 1"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f804"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f804","scientificName":"Atelopus ignescens - Corn"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f805"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f805","scientificName":"Atelopus lozanoi - Osorno"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f806"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f806","scientificName":"Atelopus ebenoides - Rive"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f807"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f807","scientificName":"Atelopus minutulus - Ruíz"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f808"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f808","scientificName":"Atelopus muisca - Rueda-A"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f809"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f809","scientificName":"Atelopus pedimarmoratus -"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f810"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f810","scientificName":"Atelopus sernai - Ruíz-Ca"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f811"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f811","scientificName":"Atelopus subornatus - Wer"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f812"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f812","scientificName":"Osornophryne percrassa -"}
{"index":{"_index":"mamiferos","_type":"taxon","_id":"ffffaf81a-39c1-202b-8a4a-00304854f813"}}
{"_score":1.0,"taxonID":"ffffaf81a-39c1-202b-8a4a-00304854f813","scientificName":"Rhinella macrorrhina - Tru

```

```

"lenger, 1918","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f801","type":"taxon","created":1479
eusebianus - Rivero y Granados, 1993","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f802","type
"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f803","type":"taxon","created":1479335275,"modified":147932939
a, 1849","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f804","type":"taxon","created":1479335276
atus":"accepted","acceptedNameUsage":"Atelopus lozanoi - Osorno-Muñoz, Ardila-Robayo y Ruíz-Carranza, 2001",
3","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f806","type":"taxon","created":1479335278,"modi
","taxonomicStatus":"accepted","acceptedNameUsage":"Atelopus minutulus - Ruíz-Carranza, Hernandez-Camacho y
opus muisca - Rueda-Almonacid y Hoyos, 1992","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f808'
s - Rivero, 1963","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f809","type":"taxon","created":1
age":"Atelopus sernai - Ruíz-Carranza y Osorno Muñoz, 1994","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4
9","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f811","type":"taxon","created":1479335283,"modi
ccepted","acceptedNameUsage":"Osornophryne percrassa - Ruíz-Carranza y Hernández-Camacho, 1976","metadata":{"
1","metadata":{"identifier":"ffffaf81a-39c1-202b-8a4a-00304854f813","type":"taxon","created":1479335285,"modi

```



A continuación se presenta la estructura de cada entrada en el archivo de registro de alimentación de la base para COUCHDB:

```
{
  "_id": "ffffaf81a-39c1-202b-8a4a-00304854f801",
  "_rev": "1-49cdcc98dff1785f64a3d9fd02844601",
  "taxonID": "ffffaf81a-39c1-202b-8a4a-00304854f801",
  "scientificName": "Allobates ranoides - Boulenger, 1918",
  "scientificNameWithoutAuthorship": "Allobates ranoides",
  "scientificNameAuthorship": "Boulenger, 1918",
  "family": "Aromobatidae",
  "taxonRank": "species",
  "taxonomicStatus": "accepted",
  "acceptedNameUsage": "Allobates ranoides - Boulenger, 1918",
  "metadata": {
    "identifier": "ffffaf81a-39c1-202b-8a4a-00304854f801",
    "type": "taxon",
    "created": 1479335273,
    "modified": 1479329397,
    "creator": "Jaime Alberto Gutierrez",
    "contributor": "Jaime Alberto Gutierrez",
    "contact": "jaimealbertogutierrez@gmail.com",
    "source": "User"
  }
}
```

Campo	Descripción
id	Identificador único de la Especie
_rev	Identificador del proceso de revisión
taxonID	Id del taxón asociado a la especie
scientificName	Nombre científico nominal de la Especie
scientificNameWithoutAuthorship	Nombre científico de la especie sin autor asociado
scientificNameAuthorship	Nombre científico dado por el investigador autor
family	Familia taxonómica a la que se asocia la especie
taxonRank	Rango del taxón

Sede Principal: Calle28A#15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



taxonomicStatus	Status taxonómico (aceptado o rechazado)
acceptedNameUsage	Nombre definido aceptado para la especie
identifier	Identificador general de metadatos de registro
type	Tipo (Especie, subespecie o derivación)
created	Fecha de creación del registro
modified	Fecha de modificación del registro
creator	Usuario responsable de la creación/modificación del registro
contributor	Persona responsable de la catalogación de la especie
contact	Correo electrónico o username de contacto del responsable del cargue
source	Fuente primaria de origen de la información de la especie

Dentro de la estructura de indización para cada registro, se presenta la forma del archivo de cargue masivo de Elastic, de la siguiente manera:

```
{
  "index":
    { "_index": "mamiferos",
      "_type": "taxon",
      "_id": "ffffaf81a-39c1-202b-8a4a-00304854f801" }
}

{
  "_score" : 1.0,
  "taxonID": "ffffaf81a-39c1-202b-8a4a-00304854f801",
  "scientificName": "Allobates ranoides - Boulenger, 1918",
  "scientificNameWithoutAuthorship": "Allobates ranoides",
  "scientificNameAuthorship": "Boulenger, 1918",
  "family": "Aromobatidae",
  "taxonRank": "species",
  "taxonomicStatus": "accepted",
  "acceptedNameUsage": "Allobates ranoides - Boulenger, 1918",
  "metadata":
    { "identifier": "ffffaf81a-39c1-202b-8a4a-00304854f801",
```

Sede Principal: Calle 28A #15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2



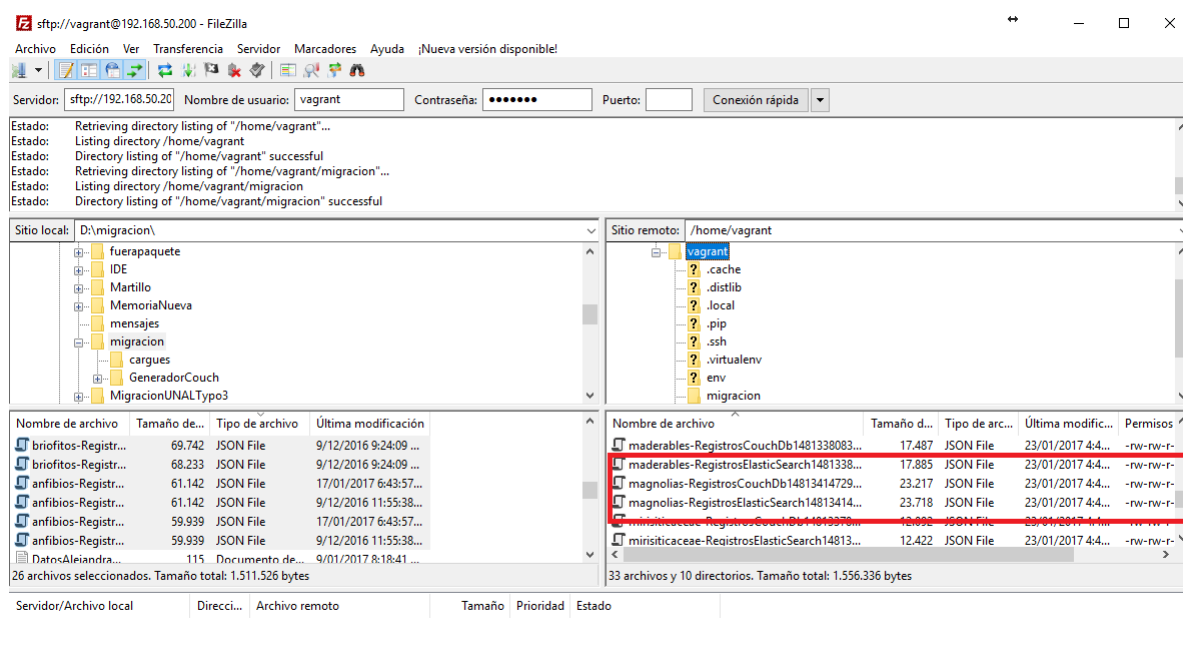
```
"type":"taxon",  
"created":1479335273,  
"modified":1479329397,  
"creator":"Jaime Alberto Gutierrez",  
"contributor":"Jaime Alberto Gutierrez",  
"contact":"jaimealbertogutierrez@gmail.com",  
"source":"User"},"rev":null  
}
```

Que cómo se observa, comparte los campos comunes con el archivo de COUCHDB.



6. REALIZACIÓN DEL CARGUE AL AMBIENTE DE BASES DE DATOS NO RELACIONALES

Una vez se hayan realizado los procesos de generación de cargue de todas las bases para las familias de especies, el administrador debe realizar el cargue de los archivos mediante su herramienta favorita de SSH y administración remota de archivos. Se recomienda usar FILEZILLA para ello.



Una vez los archivos se encuentren cargados, se debe ingresar por la consola y realizar el proceso de cargue con los comandos CURL generados por el aplicativo de transformación y migración explicado en el punto 6.



```
posh~git ~ nuvem [master]
vagrant@cncfloralocal:~$ ls mamif*
mamiferos-RegistrosCouchDb1485204497115.json
mamiferos-RegistrosElasticSearch1485204497115.json
vagrant@cncfloralocal:~$
```

Se debe ejecutar el primer comando para insertar los registros en la base de COUCHDB, así:

curl -X POST -d @mamiferos-RegistrosCouchDb1485204497115.json -H "Content-Type: application/json" http://admin:ncc123x@192.168.50.200:5984/mamiferos/_bulk_docs

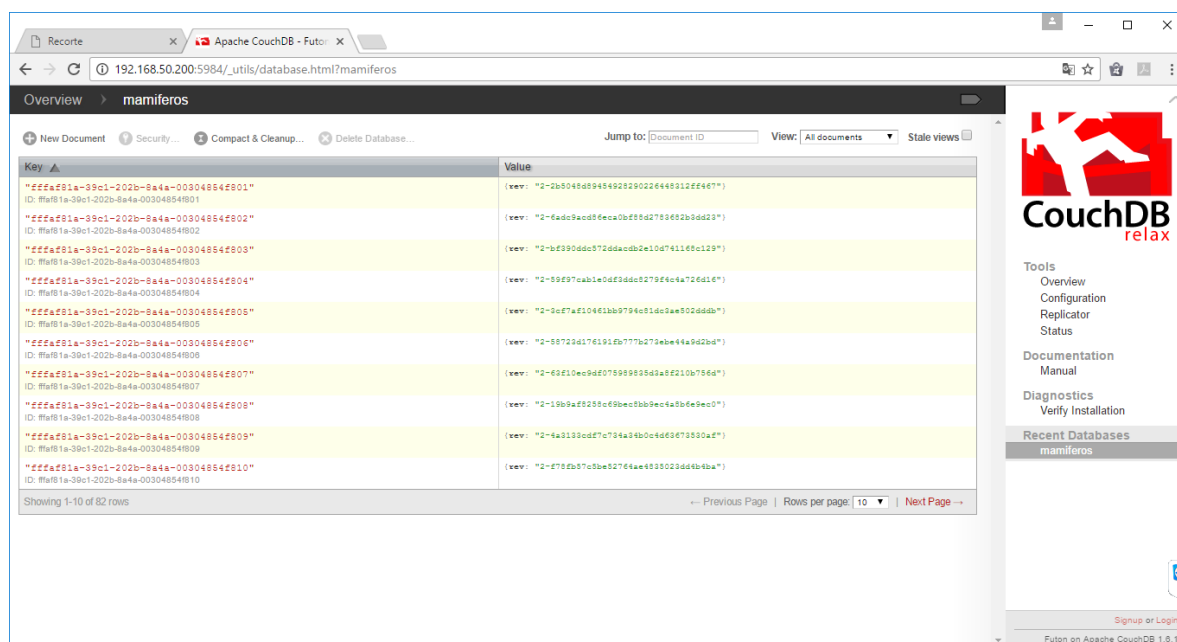
Se deberá generar en pantalla una salida de la siguiente manera:



```
posh~git ~ nuvem [master]
4f0c18b88bcec99eb57fd14530"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f865", "rev": "2-0ab57b808879a5566d2269f72aa3abe5"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f866", "rev": "2-45e581db394524563e95163f72b5f4d5"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f867", "rev": "2-52b64662be4d9e24318aadb73e996182"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f868", "rev": "2-9e4f1c92254e6754a76b17952d8fdb34"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f869", "rev": "2-2cab95e3558e056b5b25ef19daf3a44c"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f870", "rev": "2-0c18c76ce78ccb40473d7f6aa5052fdc"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f871", "rev": "2-ce7b2a5ba354b388d3b8ee0206a45f5d"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f872", "rev": "2-e44a301831462579c35fb0ac05d99059"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f873", "rev": "2-385f2a25fa57a3d7b485a6b94e268f76"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f874", "rev": "2-ee14281d93328da7921f3442e1aabe73"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f875", "rev": "2-3382d38b87ddb535611b7f4ec5951ffa"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f876", "rev": "2-13fbc4c9c5b66a129d02f360372919a6"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f877", "rev": "2-b1b66720c5f062cb217ad59df75a470c"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f878", "rev": "2-89efc36da5101c78ebd9c4619ef31325"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f879", "rev": "2-9db2ddfe8dcb25a033c624e1090b36ac"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f880", "rev": "2-8d22760eca9a0745c9976a7321520eb4"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f881", "rev": "2-38d1f12200804420b0c0198aa7bd5edc"}, {"ok": true, "id": "ffffaf81a-39c1-202b-8a4a-00304854f882", "rev": "2-21b0fc573628096e1bd51cb4790502c3"}]
vagrant@cncfloralocal:~$
```

Para verificar, ingresar por FUTON y validar los documentos cargados en la base:





Key	Value
"fffa81a-39c1-202b-8a4a-00304854f801"	{rev: "2-2b5048d945492329226449312f467"}
"fffa81a-39c1-202b-8a4a-00304854f802"	{rev: "2-6ad05c0d8feca0b88562785692b3dd23"}
"fffa81a-39c1-202b-8a4a-00304854f803"	{rev: "2-bf990ddc572ddacdb2e1d6741160c129"}
"fffa81a-39c1-202b-8a4a-00304854f804"	{rev: "2-58597cab1e0d3ddc3279f4c4a72d16"}
"fffa81a-39c1-202b-8a4a-00304854f805"	{rev: "2-3e07a51c4613b9794c81d02ae902dddb"}
"fffa81a-39c1-202b-8a4a-00304854f806"	{rev: "2-58722d2761915b777b2734be4a9d2bd"}
"fffa81a-39c1-202b-8a4a-00304854f807"	{rev: "2-632f10c9d5075959835d3a8f210b756d"}
"fffa81a-39c1-202b-8a4a-00304854f808"	{rev: "2-19b9a2f225c69b0cbb0ec6a3b0fe9ec0"}
"fffa81a-39c1-202b-8a4a-00304854f809"	{rev: "2-4a9130cdf7c734a34b0c4d63673530a2f"}
"fffa81a-39c1-202b-8a4a-00304854f810"	{rev: "2-2702b57c3be52764ae4535023dd4b4ba"}

No obstante, hasta que no se realice la configuración del índice, los registros no serán visibles desde la aplicación de CHECKLIST, tal como se ilustra a continuación:



Recorte

➤ Usuario: Jaime Gutierrez .

➤ Recorte: MAMIFEROS.

Recortes Familias Logout

FAMILIAS



A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Para realizar este proceso, con los registros de Elastic generados, debe procederse a ejecutar el comando que sigue:

```
posh~git ~ nuvem [master]
vagrant@cncfloralocal:~$ curl -XPOST 192.168.50.200:9200/mamiferos/_bulk --data-binary @mamiferos-RegistrosElasticSearch1485204497115.json
```

El cual debe generar el siguiente resultado luego del cargue:




```
posh~git ~ nuvem [master]
os", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f869", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f870", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f871", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f872", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f873", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f874", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f875", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f876", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f877", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f878", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f879", "status": 400, "error": "MappingException[failed to parse]; nested: JsonParseException[Invalid UTF-8 middle byte 0x7a\\n at [Source: [B@76c36b4e; line: 1, column: 121]]; }}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f880", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f881", "_version": 1, "status": 201}, {"index": {"_index": "mamiferos", "_type": "taxon", "_id": "ffffaf81a-39c1-202b-8a4a-00304854f882", "_version": 1, "status": 201}}]vagrant@cncfloralocal:~$
```

Lo cual, al validar en la aplicación de CHECKLIST, ya debe permitir visualizar los datos cargados:

RECORTE

➤ Usuario: Jaime Gutierrez .

➤ Recorte: MAMIFEROS.

Recortes Familias Logout

FAMILIAS

AROMBATIDAE / 1
BUFONIDAE / 10
CENTROLENIDAE / 11
CRAUGASTORIDAE / 22
DENDROBATIDAE / 8
HEMIPHRACTIDAE / 3
HYLIDAE / 2
PLETHODONTIDAE / 3

A
E
[
[
A
n
F
[
N





Instituto de Investigación de Recursos Biológicos Alexander von Humboldt

Elaboró:
Ingeniero
JAIME ALBERTO GUTIÉRREZ MEJÍA
CONSTRATISTA INSTITUTO ALEXANDER VON HUMBOLDT
DICIEMBRE 2016

Sede Principal: Calle 28A #15-09 Bogotá, D.C., Colombia | PBX: (57)(1) 3202767 | NIT 820000142-2

