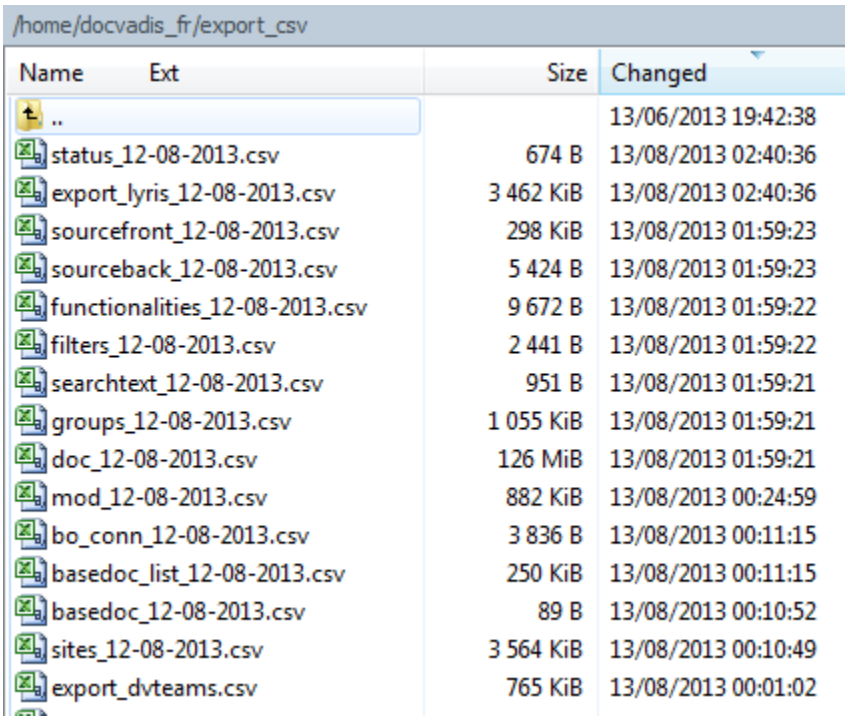
# CSV Exports


- Introduction
- Old csv export files
    - basedoc.csv
    - bo_conn.csv
    - export_lyris.csv
    - mod.csv
    - site.csv
    - export_dvteams.csv
    - groups.csv
- New csv export files
    - Functionalities tracking
    - Docvadis search filters tracking
    - Doc.csv
    - List of documents of the Documents library
    - Use of the "search engine"(keywords)
    - Physician website visit source Tracking
    - Physician back office Link source Tracking
    - Export status
- Export Manager
    - General parameters
    - Export Order


## Introduction

The Docvadis application exports every night (around midnight, but start time and frequency can be configured) some data from the database and the JCR repository.
The data exported are stored in 14 CSV files which are uploaded on an MSD SFTP server, once the complete export is done.

Example of the export on the PROD fr Docvadis:




| Name | Ext | Size | Changed |
|---|---|---|---|
| .. | | | 13/06/2013 19:42:38 |
| status_12-08-2013.csv | | 674 B | 13/08/2013 02:40:36 |
| export_lyris_12-08-2013.csv | | 3 462 KiB | 13/08/2013 02:40:36 |
| sourcefront_12-08-2013.csv | | 298 KiB | 13/08/2013 01:59:23 |
| sourceback_12-08-2013.csv | | 5 424 B | 13/08/2013 01:59:23 |
| functionalities_12-08-2013.csv | | 9 672 B | 13/08/2013 01:59:22 |
| filters_12-08-2013.csv | | 2 441 B | 13/08/2013 01:59:22 |
| searchtext_12-08-2013.csv | | 951 B | 13/08/2013 01:59:21 |
| groups_12-08-2013.csv | | 1 055 KiB | 13/08/2013 01:59:21 |
| doc_12-08-2013.csv | | 126 MiB | 13/08/2013 01:59:21 |
| mod_12-08-2013.csv | | 882 KiB | 13/08/2013 00:24:59 |
| bo_conn_12-08-2013.csv | | 3 836 B | 13/08/2013 00:11:15 |
| basedoc_list_12-08-2013.csv | | 250 KiB | 13/08/2013 00:11:15 |
| basedoc_12-08-2013.csv | | 89 B | 13/08/2013 00:10:52 |
| sites_12-08-2013.csv | | 3 564 KiB | 13/08/2013 00:10:49 |
| export_dvteams.csv | | 765 KiB | 13/08/2013 00:01:02 |

The CSV export is managed by the `AllExportsCsvGenerator` class.
It can be triggered through an HTTP call to `http://bo.docvadis.fr/export/all.csv`

## Old csv export files

| | Duration | File size | Database | JCR |
|---|---|---|---|---|
| **sites** | ~15min | ~ 3.5 MB | Yes | Yes |
| bo_conn | <1min | ~ 4 kB | Yes | |
| basedoc | <1min | ~ 1 kB | | Yes |
| mod | ~16min | ~ 880 kB | Yes | Yes |
| doc (refactored) in new exports | ~3h | ~ 126 MB | Yes | Yes |
| export_lyris | ~2h15 | ~ 3.4 MB | Yes | Yes |
| groups | <1min | ~ 1 MB | Yes | |
| export_dvteams | <1min | | Yes | |

## basedoc.csv

The values are: NB_TOTAL;NB_CARD;NB_LINK;NB_VIDEO;NB_FILE;NB_CALCULATOR

- NB_TOTAL: number of BOMSD's documents
- The other values are the details of this NB_TOTAL.

## bo_conn.csv

The values are: USERID;NB_CONN;DURATION_CONN;MODIF

- USERID: physician's login
- NB_CONN: nb of connection in the last 24h
- DURATION_CONN: Average connection time

MODIF: true if the user has modified one of his websites

## export_lyris.csv

The values are:
LOGIN;DVSITEID;DVURLSITE;DVDESIGN;DVTYPESITE;DVSPE;DVSITECDV;DVSITEACTIF;DVSITECREA;ABONLDOC;DVCREASTATUS;
DVNBDOC;DVLASTVISIT

- LOGIN: physician's login
- DVURLSITE: dns of the physician's website
- DVSITEID: ID in Mysql database of the physician's website
- DVDESIGN: website layout id
- DVTYPESITE: type of the website
- DVSPE: specialty of the user
- DVSITECDV: has site online doc?
- DVSITEACTIF: is the site online?
- DVSITECREA: site creation date
- ABONLDOC: is user registered to the newsletter?
- DVCREASTATUS : website creation status (creation step)
- DVNBDOC : nb of tagged doc
- DVLASTVISIT : date of last login in bo

## mod.csv

The values are: URL_SITE;DVSITEID;NON_MOD;DETAIL;POS_MOD;COUNT;TOTAL_REFILL

- URL_SITE: dns of the physician's website
- DVSITEID: ID in Mysql database of the physician's website
- NON_MOD: Module type
- DETAIL: Title of the module (only free modules)
- POS_MOD: Position if the module
- COUNT: number of contact or refill emails
- TOTAL_REFILL: nb of prescription (in total) required by FO's users

## site.csv

The values are:
USERID;URL_SITE;DVSITEID;DATE_CREA;TYPE_SITE;SPE;ON_LINE;THEME;VISIT_CARD;ABO_NL;PERSO_THEME;PERSO_PHOTO;DIRECTORY;NB_MEMBER;NB_MEMBER_OK;CREATION

- USERID: Login of the site creator
- URL_SITE: dns of the physician's website
- DVSITEID: ID in Mysql database of the physician's website
- DATE_CREA: site creation date
- TYPE_SITE: type of the website
- SPE: specialty of the user
- ON_LINE: is the site online?
- THEME: website layout id
- VISIT_CARD: has site online doc?
- ABO_NL: is creator registered to the newsletter?
- PERSO_THEME: has the site personalized colors?
- PERSO_PHOTO: has the site a personalized banner picture?
- DIRECTORY: visibility of site on search sites
- NB_MEMBER: number of members for group website
- NB_MEMBER_OK: number of members with accepted invitation
- CREATION: is website creation process done?

## export_dvteams.csv

The values are: TVFCODE;DVSITEACTIF;LOGIN

- TVFCODE: user's tvfCode
- DVSITEACTIF: is website creation done?

LOGIN: physician's login

## groups.csv

The values are:
USERID;URL_SITE;DVSITEID;DEGREE;FIRSTNAME;MIDDLENAME;LASTNAME;DVSPE;EMAIL;INV_STATUS;ABO_NL;DVBOSS;DVADMIN;DVWRITER;DVPUBLISH

- USERID: user's login (univadis login or temp login)
- URL_SITE: dns of the physician's website
- DVSITEID: ID in Mysql database of the physician's website
- DEGREE: Title of the user
- FIRSTNAME;MIDDLENAME;LASTNAME: user's first name, middle name, last name
- DVSPE: user's specialty
- EMAIL: user's email
- INV_STATUS: Status of a member's invitation
- ABO_NL: is user registered to the newsletter?

BVBOSS;DVADMIN;DVEDITOR;DVPUBLISH: is the user group responsible, website administrator, document editor, document publisher

# New csv export files

| | Duration | File size | Database | JCR |
|---|---|---|---|---|
| functionalities | <1min | ~ 10 kB | Yes | |
| filters | <1min | ~ 2.5 kB | Yes | |
| basedoc_list | <1min | ~ 250 kB | Yes | Yes |
| searchtext | <1min | ~ 1 kB | Yes | |
| sourcefront | <1min | ~ 300 kB | Yes | |
| sourceback | <1min | ~ 5 kB | Yes | |

# Functionalities tracking

**File name:** functionalities_DD-MM-YYYY.csv

**Description :**

Every day the file must provide, for each user, sites, documents and origin of the action, the counts of Docvadis V3 usage for the following functionalities: print, PDF, Email and Tag

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| USERID | String | Not null | User id |
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | | Url of the site, null for database users (optional) |
| DOC_ID | String | Not null | Id of the document |
| ORIGIN | String | Not null | Origin of the document |
| DOC_POOL | String | Not null | Place from where action on document is done (“Documents library” ,"My documents" or Document Viewing mode) |
| NB_EMAIL | Number | Not null | Count of sent email this day |
| NB_PRINT | Number | Not null | Count of printed document this day |
| NB_PDF | Number | Not null | Count of pdf extract this day. |
| NB_TAG | Number | Not null | Count of the tag functionality use this day |

**Frequency of delivery:** daily

**Granularity:** USERID, DVSITEID, DOC_ID, ORIGIN,DOC_POOL.

Every day the file shall contain only yesterday's data (**incremental delivery**).

**Data:**

The file must contain the actions done by website users and database users only.

Possible values for DOC_POOL field:

- "doclib" : for actions done through the "document database" page
- "mydoc": for actions done through "my documents" page
- “viewmode”: for actions done when the documents are in viewing mode.

ORIGIN: values of the field

- "dbdoc" for documents from MSD Document database.
- "physician" for a document created by a Physician.

# Docvadis search filters tracking

**File name:** filters_DD-MM-YYYY.csv

**Description :**

Every day the file must provide, for each user, sites and the origin of the search, the number of times that each filter has been used.

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| USERID | String | Not null | User id |

| | | | |
|---|---|---|---|
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | | Url of the site, null for database users (optional) |
| DOC_POOL | String | Not null | Tab from which search is done ("Document database" or "My documents") |
| NB_DCV_TAG_FILTER_USED | Number | Not null | Number of times that speciality filter has been used |
| NB_USER_TAG_FILTER_USED | Number | Not null | Number of times that filter has been used |
| NB_SPE_FILTER_USED | Number | Not null | Number of times that speciality filter has been used |
| NB_PATHO_FILTER_USED | Number | Not null | Number of times that filter has been used |
| NB_THEME_FILTER_USED | Number | Not null | Number of times that filter has been used |
| NB_SOURCES_FILTER_USED | Number | Not null | Number of times that sources filter has been used |
| NB_FORMAT_FILTER_USED | Number | Not null | Number of times that format filter has been used |
| NB_LANG_FILTER_USED | Number | Not null | Number of times that language filter has been used |
| NB_STAT_FILTER_USED | Number | Not null | Number of times that statut filter has been used |
| NB_SEARCH_FILTER_USED | Number | Not null | Number of times that search filter has been usedn |

**Frequency of delivery:** daily

**Granularity:** USERID, DVSITEID, DOC_POOL.

Every day the file shall contain only yesterday's data (**incremental delivery**).

**Data:**

The file must contain the filters actions done by website users and database users only.

Possible values for DOC_POOL field:

- "doclib" : for search done from the "document database" page
- "mydoc": for search done from "my documents" page

## Doc.csv

**File name:** doc_DD-MM-YYYY.csv

**Description :**

Every day the file must provide, for each user and attached site, the list of selected documents.

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| USERID | String | Not null | User id |
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | | url of the site (null for database user) |

| | | | |
|---|---|---|---|
| DOC_ID | String | Not null | Id of the document |
| LINK | String | | Link of the document (only for online documents) |
| TYPE | String | Not null | Type of the document |
| ORIGIN | String | Not null | Origin of the document |
| ON_LINE | Boolean | Not null | is document on line ? |
| MODIF | Boolean | | true if BOMSD doc with comment |

**Frequency of delivery:** daily

**Granularity:** USERID, DVSITEID, DOC_ID, ORIGIN.

Every day the file shall contain only yesterday's data (**incremental delivery**).

**Data:**

The content of this export is all the documents existing in 'My documents' (documents available for database users only and website users).

The values of the following are as defined below:

ON_LINE: true if the document is published and false if the document is only selected (currently, this field is set to false only if the document has been published before)

ORIGIN: values of the field

- "dbdoc" for documents from MSD Document database.
- "physician" for a document created by a Physician.

## List of documents of the Documents library

**File name:** basedoc_list_DD-MM-YYYY.csv

**Description :**

This file will contain the list of the documents of the documents database (currently doc.csv only contains published documents; basedoc.csv only contains the count of the document)

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| DOC_ID | String | Not null | Id of the document |
| TYPE | String | Not null | Type of the document |
| PATHO | String | | Pathologies of BOMSD documents |
| SPE | String | | Specialties of BOMSD documents |
| THEME | String | | Topics of BOMSD documents |
| SOURCE | String | | Source of BOMSD documents |
| IC | String | | International code of the document |
| ONLINE_DATE | Date | | Online Date |
| LAST_UPDATE | Date | Not null | Last update |

**Frequency of delivery:** daily

**Granularity:** DOC_ID

Every day the file shall contain only yesterday's data (**incremental delivery**).

**Data:**

Possible values for SOURCE field are:

- ‘Committee’: if the source of the document is a committee,
- <Partner name>: if the source of the document is a partner,
- Null: if no source has been defined.

## Use of the "search engine"(keywords)

**Nom du fichier:** searchtext_DD-MM-YYYY.csv

**Description :**

This file will contain all the searches done by the users in the document database, through the “keywords” area.

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| ID | Number | Not null | Technical id |
| USERID | String | Not null | User id |
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | | Url of the site, null for database users (optional) |
| SEARCH_BOX | Text | Not null | Text of the search |

**Frequency of delivery:** daily

**Granularity:** ID

Every day the file shall contain only yesterday's data (**incremental delivery**).

## Physician website visit source Tracking

**File name:** **sourcefront**_DD-MM-YYYY.csv

**Description :**

This file will contain for each physician website visited, the source of the visit (QRCode, …)

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | Not null | Url of the site |
| SOURCE | String | Not null | Link source (=’QRCode’, ...) the value « other » could be set if source value is not specified in the link |
| NB_VISIT | Number | | Number of visits associated to each source |

**Frequency of delivery:** daily

**Granularity:** DVSITEID, SOURCE

Every day the file shall contain only yesterday's data (**incremental delivery**).

## Physician back office Link source Tracking

**File name:** **sourceback**_DD-MM-YYYY.csv

**Description :**

This file will contain for each physician website which back office is visited, the source of the visit (newsletter, …)

**Structure:**

| Fields | Type | Constraint | Description |
|---|---|---|---|
| DVSITEID | String | Not null | Id of the site |
| URL_SITE | String | Not null | Url of the site |
| USERID | String | Not null | Id of the connected user |
| SOURCE | String | Not null | Link source (= 'other', 'newsletter', ...), the value « other » could be set in the case of source value is not specified in the link.<br><br>**Autologin link is needed to get source parameter in URL.** |
| NB_VISIT | Number | | Number of visits associated to each source |

**Frequency of delivery:** daily

**Granularity:** DVDSITEID, USERID, SOURCE

Every day the file shall contain only yesterday's data (**incremental delivery**).

## Export status

Anyware Services provides to create a text file which will allow knowing the export status associated to each file to be exported.

A text file named "export_status_DD-MM-YYYY.csv" will be created in which the following information will be provided:

- *Export file name*: "groups"/"sites"/"mod"/"doc"/ ….
- *Export status* which possible values are :
  - "Generated",
  - "Partially generated" in case of the export process is killed because "**Export process timeout value"** has been reached**,** the current generated file is not completely generated,
  - "Not Generated" in case of the export file has been disabled in the settings or in case of the export process is killed because "**Export process timeout value"** has been reached.
- *Generation start time* (HH:MM:SS) : Null if "Export status" is "Not Generated"
- *Generation end time* (HH:MM:SS) : Null if "Export status" is "Not Generated"

Example:

Lines (samples) of this file could be:

groups;Generated;01:24:32;02:19:01

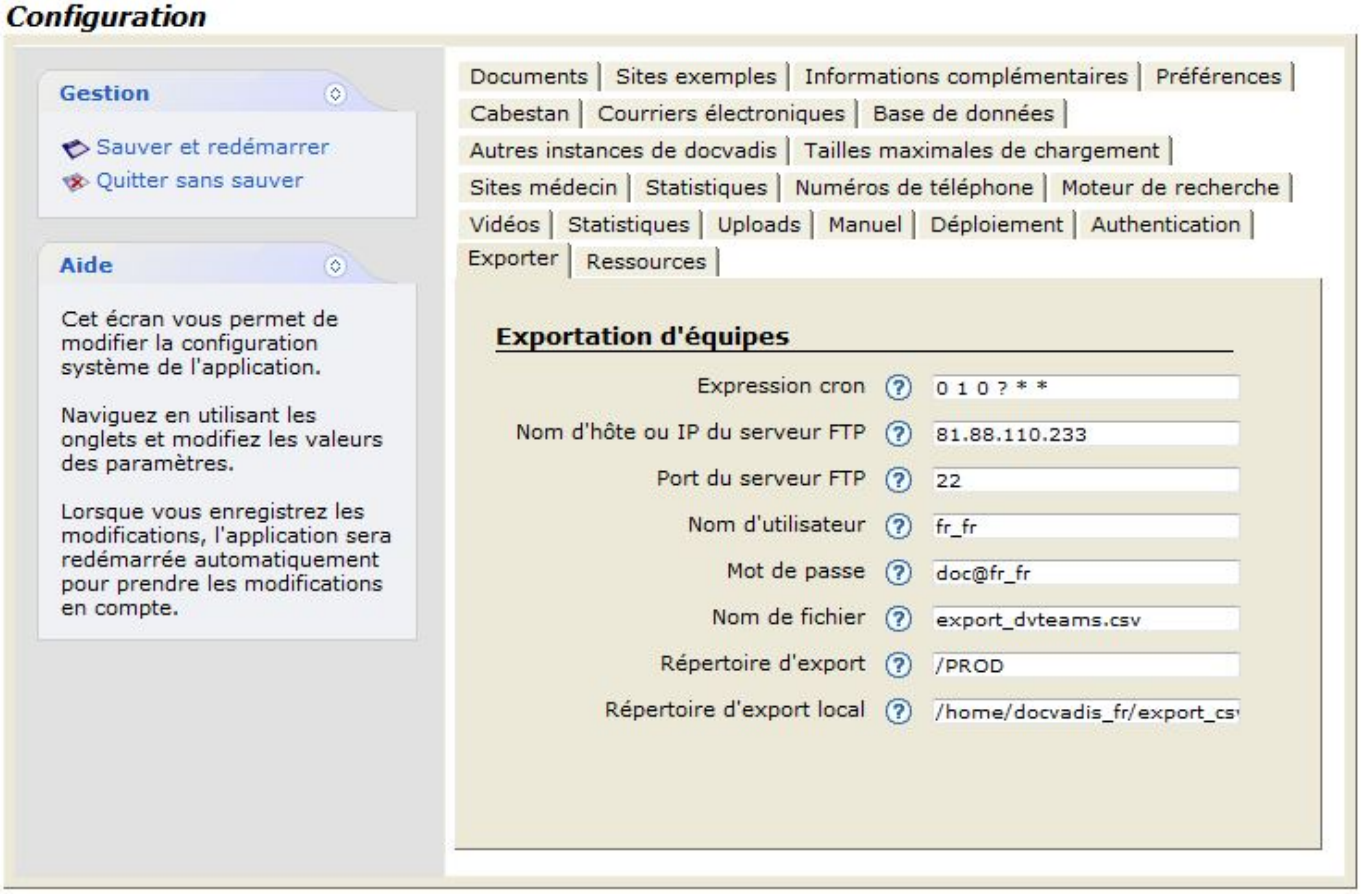
mod;Partially generated;06:35:24;07:00:00

sites;Not Generated;;

# Export Manager

## General parameters

The parameters used for CSV export files upload on MSD Univadis SFTP server are :

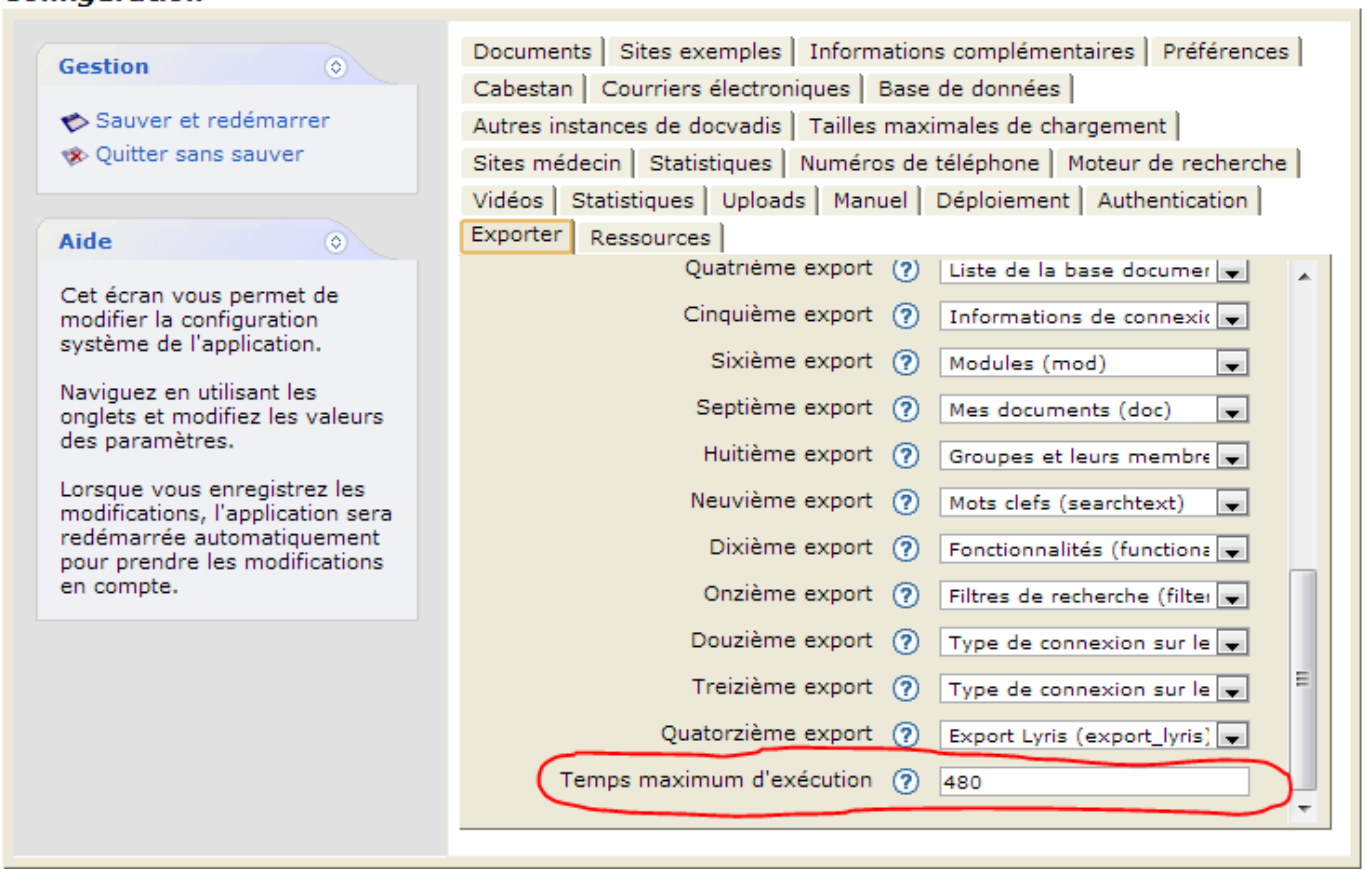



## Export Order

The export manager allows to choose in which order will be done the exports.

It is also possible to disable some exports.

The last item of the configuration tab allows to set a maximum duration process.

⚠ When the duration time set is reached, the export process will be killed when the current export being processed will be finished .

***Configuration***