

Manuel de description des textes pour la Base de Français Médiéval

Céline Guillot (cguillot@ens-lsh.fr)
Alexei Lavrentiev (Alexei.Lavrentev@ens-lsh.fr)

CNRS / ENS-LSH, UMR 5191 ICAR



VERSION 2.3

Ce document de travail est élaboré dans le cadre de la normalisation des descripteurs des textes de la BFM, de la mise au point des entêtes TEI pour ces textes et des recherches sur la typologie textuelle. Il sert de référence à l'équipe de la BFM et dans le cadre des échanges de textes que nous pratiquons avec nos partenaires.

Ce document est publié librement sur le web à destination de la communauté scientifique dans le cadre de la licence Creative Commons « Paternité-Pas d'Utilisation Commerciale-Partage des Conditions Initiales à l'Identique 2.0 France ». En accord avec cette licence, si vous utilisez ce document dans vos travaux, vous êtes prié de mentionner sa référence (projet BFM, titre, auteurs).



Table de mise à jour du document

- 10 janvier 2007, v 1.0, Alexei Lavrentiev (AL), création du document
- 10 février 2007, v 1.1, AL & Céline Guillot (CG), réglages
- 21 mars 2007, v 1.2, SH, mise en place de la table de mise à jour du document
- 7 novembre 2008, v 2.0, AL, mise à jour de l'ensemble du document compte-tenu des évolutions du projet
- 11 février 2009, v. 2.1, AL, validation des modifications
- 25 mars 2009, v. 2.2., CG & AL, révision de l'ensemble du document
- 16 juin 2009, CG, quelques modifications
- 21 décembre 2009, v. 2.3., AL, validation des modifications, mise à jour du schéma des relations

SOMMAIRE

Table de mise à jour du document	1
Principes généraux	2
1. Définition des descripteurs.....	2
Fonctions des descripteurs	3

Fonctionnement du système des descripteurs	3
Base de données BFM.MDB.....	4
1. Structure générale.....	4
2. Description générale de l'œuvre (table Oeuvre)	5
2.1. Descripteurs actualisés	5
2.2. Descripteurs non actualisés	7
3. Caractéristiques typologiques	8
3.1. Descripteur DOMAINE	8
3.2. Descripteur GENRE.....	9
3.3. Autres descripteurs typologiques actualisés.....	10
3.4. Descripteurs typologiques non actualisés	11
4. Description de l'édition de référence	11
5. Description du texte numérisé.....	13
6. Description des révisions du texte numérisé	14
7. Références des personnes et d'organisations responsables d'édition, de datation et de numérisation des textes	15
6. Description de manuscrits	15
Références	15
Annexes	16
1. Relations dans la base BFM.MDB.....	16
2. Domaines et genres actuellement représentés dans la BFM	17

Principes généraux

1. Définition des descripteurs

Un descripteur correspond à une dimension de description d'une unité textuelle (et non, pour l'instant du moins, d'une partie d'une unité textuelle).

La définition de l'unité textuelle sur laquelle porte le descripteur repose sur la distinction de trois objets différents :

- l'œuvre (ou l'acte) ;
- l'édition ;
- le manuscrit.

L'unité textuelle correspond à une œuvre représentée par une édition (ou éventuellement par un manuscrit) et qui est numérisée afin d'être intégrée à la base textuelle.

Chaque œuvre et chaque édition sont dotées d'un identifiant unique (sigle DEAF entier pour une édition ; sigle DEAF moins l'initiale de l'éditeur pour une œuvre, quand cela est possible).

Un descripteur d'une unité textuelle correspond donc à une propriété que cette unité hérite soit de l'œuvre (auteur, date de composition, genre, etc.), soit du manuscrit, soit de l'édition (éditeur scientifique, etc.).

L'ensemble des descripteurs rassemble les métadonnées disponibles et/ou utiles pour l'interprétation et l'exploitation des textes.

Les descripteurs sont définis *a priori* sur les textes.

A chaque descripteur correspond une liste de valeurs. Ces valeurs peuvent être libres ou contrôlées par une liste d'autorité.

Ces données sont stockées et gérées dans une base ACCESS (bfm.mdb), qui n'est pas directement interrogeable par les utilisateurs mais qui alimente le site de la BFM et les différents corpus accessibles par Weblex.

Seules certaines des informations qui se trouvent dans la base ACCESS sont actuellement accessibles aux utilisateurs de la BFM.

La base des descripteurs de la BFM a été développée progressivement à partir de 2002 et a évolué au cours de la progression du projet. Certains éléments de sa structure sont hérités des états antérieurs de son développement et ont été préservés pour assurer le bon fonctionnement des requêtes élaborées antérieurement.

Certains descripteurs (champs de la base de données) ne sont plus actualisés ni renseignés pour les nouveaux textes décrits. Ceux-ci sont présentés dans une section à part à la fin de la description de la table correspondante de la base de données.

Fonctions des descripteurs

On distingue au moins trois fonctions différentes :

- fonction documentaire : pouvoir retrouver un document répondant à un certain critère de description ; se faire une idée plus ou moins claire du contenu du document ; permettre les échanges de documents avec nos partenaires ;
- fonction méthodologique : à la fois organiser et enrichir le corpus en évaluant de quoi il est représentatif, dans un souci de diversification maximale ; pouvoir "partitionner" le corpus en sous-corpus selon plusieurs critères explicites et parfois combinés ;
- fonction scientifique : mener à partir du corpus des recherches synchroniques et diachroniques en langue (avec un niveau de généralité le plus grand possible) ; rendre compte de la variation des usages compte-tenu de l'ensemble des variables pertinentes ; décrire et modéliser le plus adéquatement possible les types de textes produits au Moyen Age compte-tenu des variables pertinentes.

Fonctionnement du système des descripteurs

Des trois fonctions principales assignées aux descripteurs découlent deux conséquences apparemment opposées :

- Pour un texte donné et pour certains descripteurs, on peut vouloir choisir plusieurs valeurs différentes ; par exemple, le *Roman de la Rose* peut être décrit comme un texte à la fois « littéraire » et « didactico-scientifique » (deux valeurs différentes du descripteur DOMAINE), car cela permet de retrouver ou d'inclure ce texte quand on travaille sur les textes littéraires ou didactiques (logique documentaire)
- En revanche, quand on utilise des outils statistiques, il est nécessaire de ne comptabiliser les occurrences d'un même texte qu'une seule fois (logique méthodologique).

Deux solutions sont envisageables selon le degré de « mixité » d'un texte donné :

- A. Il est possible de définir une valeur principale (ou par défaut). Une ou plusieurs autres valeurs seront considérées comme secondaires et ne seront pas prises en compte dans les analyses statistiques ;
- B. Il est impossible de sélectionner une valeur principale d'un descripteur sans fausser l'exactitude du classement. Dans ce cas, toutes les valeurs sont données comme

principales, et leur combinaison sera considérée comme une valeur indépendante. Lors des analyses statistiques contrastives les textes de ce type seront soit exclus, soit considérés comme une catégorie à part.

- On distingue les descripteurs qui sont obligatoirement renseignés, de ceux qui sont facultatifs ; dans certains cas, quand il est impossible d'attribuer une valeur précise à une variable, on peut ne pas la renseigner (par exemple, on ne connaît pas la région de production du ms). Les descripteurs obligatoires sont ceux qui sont *a priori* les plus pertinents pour l'exploitation des textes et qui peuvent donner lieu à la sélection de sous-corpus ciblés.
- On distingue les descripteurs « universaux », applicables normalement à n'importe quelle unité textuelle et les descripteurs conditionnels, dont la pertinence dépend de la valeur d'un autre descripteur (par exemple, le descripteur Thème est pertinent uniquement pour le domaine « didactico-scientifique »).

Base de données BFM.MDB

1. Structure générale

Les métadonnées des textes de la BFM sont stockées dans une base de données relationnelle. La structure des relations entre les principales tables de cette base est présentée dans l'Annexe 1.

Les tables de la base correspondent aux principales « facettes » de description d'un texte de la BFM :

1. Description générale d'une œuvre (table primaire « Œuvre »)
2. Caractéristiques typologiques (table « Typologie »)
 - cette table est en relation **1-1** avec Œuvre, elle en est séparée pour regrouper les champs typologiques
3. Description de l'édition de référence (table « Edition-source »)
 - un enregistrement correspond à un volume physique d'une édition d'une œuvre donnée ;
 - cette table est en relation ∞ -**1** avec Œuvre, car une même œuvre peut être représentée par plusieurs éditions et/ou éditée en plusieurs volumes.
4. Description de l'état de numérisation d'un texte (table « Numerisation »)
 - cette table est en relation ∞ -**1** avec Edition-source : il est possible qu'une édition soit représentée par des échantillons (cas des textes au programme de l'agrégation par exemple ou numérisation dans le cadre de certains projets)
5. Historique des modifications du texte (table « Corrections »)
 - cette table est en relation ∞ -**1** avec Numerisation, car chaque fichier subit plusieurs modifications au cours de son « cycle de vie ».
6. Transcriptions de sources primaires (manuscripts et incunables)
 - plusieurs tables décrites dans un document spécifique
 - à terme toutes les informations relatives aux sources primaires (date et histoire des manuscrits, dialecte du scribe, etc.) devront être gérées par ce module. Pour l'instant, les informations concernant les sources primaires des textes ne faisant pas partie du corpus « BFM –Manuscripts » sont gérées par la table « Edition – source ».

2. Description générale de l'œuvre (table Oeuvre)

2.1. Descripteurs actualisés

Référence

- identifiant de l'œuvre dans la base
 - ❖ les identifiants hérités des états antérieurs de la base seront remplacés progressivement par les sigles DEAF (sans l'initiale de l'éditeur)

Auteur (ou rédacteur)

- nom de l'auteur conventionnel (source d'autorité – Base bibliographique du DEAF, sinon le *Dictionnaire des lettres françaises*)
- « anonyme », si inconnu

L'identité de l'auteur/rédacteur pourra selon les cas correspondre à une personne ou à une institution (dans le cas des chartes par exemple), et on distinguera par ailleurs au moins deux rôles possibles (auteur/scribe). Des informations relatives à l'âge et à la catégorie sociale seront données dans les rares cas où elles sont disponibles et ces informations ne concerneront que les personnes.

- ❖ Dans l'avenir, une table spéciale sera créée pour gérer les informations supplémentaires relatives aux auteurs (l'âge, la catégorie sociale, etc.)

AuteurRef

- identifiant de l'auteur
- format : Initiale du prénom + Nom, par exemple *CTroyes*, *PCommynes* (sans accent, par exemple Beroul et non Béroul, comme cela figure dans le *Dictionnaire des Lettres françaises*)

Titre

- titre de l'œuvre conventionnel (source d'autorité – *Dictionnaire des lettres françaises* ; remarquer en particulier l'absence d'article au début du titre de l'œuvre (exemple : *Roman de la Rose* et non *Le roman de la Rose*).

Un groupe de champs est utilisé pour la localisation temporelle de la composition de l'œuvre :

Période

- « ancien » (date de composition formelle avant 1300) ou « moyen » (à partir de 1300)

Siècle

- « 12 », « 13 », etc.

Sous-siècle

- composition au début (de xx00 à xx33), au milieu (de xx34 à xx66) ou à la fin (de xx67 à xx99) d'un siècle

- format : xx_1, xx_2 ou xx_3 (par exemple 13_2 pour le milieu du 13^e s.).

Demi-siècle

- composition dans la 1^{ère} ou dans la 2^{nde} moitié d'un siècle
- format : xx_1 ou xx_2 (par exemple 13_2 pour la 2^{ème} moitié du 13^e s.).

Date de composition libre

- datation du texte en format libre, telle que donnée par la base bibliographique du DEAF, l'éditeur scientifique ou encore la littérature grise
- des règles sont utilisées pour passer de cette « prose libre » à un format utilisable dans les calculs (cf. ci-dessous)

Date début

- la date de composition la plus ancienne donnée par l'éditeur/la littérature grise ? ou calculée selon les règles établies.

Date fin

- la date de composition la plus tardive donnée par l'éditeur/la littérature grise ? ou calculée selon les règles établies.
 - ❖ on met 31/12 dans les champs « jour et mois ».

Date de composition formalisée

- date au format jj/mm/aaaa
 - ❖ on met 1/1 pour le jour et pour le mois, à moins qu'on ne connaisse la date précise ;
- moyenne calculée entre la « date début » et la « date fin ».

Règles de formalisation de la date (quelques exemples)

Date libre	Date début	Date fin	Date formelle
XIII ^e	01/01/1200	01/01/1299	01/01/1250
fin XII ^e	01/01/1167	01/01/1199	01/01/1183
fin XIV ^{ème} ou début XV ^{ème}	01/01/1390	01/01/1410	01/01/1400
à partir de 1180	01/01/1180	01/01/1199	01/01/1190
seconde moitié du X ^{ème}	01/01/950	01/01/999	01/01/975
tout début du XV ^e	01/01/1400	01/01/1410	01/01/1405

Ces calculs peuvent être corrigés si l'on dispose de données supplémentaires. Par exemple, le *Livre des dix commandements* de Jean Gerson est daté du « début du 15^e », donc normalement de 1/1/1400 à 1/1/1433, mais on sait que l'auteur est mort en 1429, donc la date la plus tardive sera précisée.

Un groupe de champs est utilisé pour la caractérisation géographique de la langue de l'œuvre :

Langue originale

- langue originale s'il s'agit d'une traduction
- le champ est laissé vide pour les œuvres originales de français médiéval

Dialecte

- dialecte de l'auteur, s'il est identifiable
- Source d'autorité : Base bibliographique du DEAF

Actuellement, les étiquettes dialectales suivantes sont utilisées dans la BFM :

- lorrain
- franco-picard
- picard
- anglo-normand
- normand
- amiénois
- cambrésien
- arrageois
- parisien
- vendéen

Les champs suivants appartiennent à la caractérisation typologique, mais ils sont pour l'instant placés dans la table Œuvre (pour des raisons historiques).

Forme

- prose, vers, glose ou mixte (si un texte contient de la prose et du vers)

Type de vers

- octosyllabe, décasyllabe, alexandrin... (textes en vers uniquement)

Structure interne

- divisions du texte (chapitres, chansons...)
 - ❖ il faudrait distinguer les divisions de l'auteur de celle de l'éditeur
- hiérarchie des divisions (1, 2 ou 3 niveaux)
 - ❖ ce champ n'est pas renseigné à ce jour, les modalités de son application doivent encore être définies

2.2. Descripteurs non actualisés

Copiste et région

- remarques de l'éditeur scientifique/la littérature grise ? en « prose libre »
- champ à supprimer à terme, ces informations doivent figurer dans des tables spécifiques.

Langue et dialecte

- remarques de l'éditeur scientifique/ la littérature grise ? en « prose libre »

Factuality

- Le rapport de l'œuvre avec les faits réels
- 3 valeurs possibles (suggérées par la TEI) : fiction, fact et mixed

Date & Histoire du ms.

- remarques de l'éditeur scientifique sur son manuscrit de base ; des informations plus structurées sont données dans les tables de description de manuscrits
- champ à supprimer à terme, car ces informations sont prises en charge par des tables spécifiques

Dialecte scribe

- dialecte du scribe ou du copiste, s'il est identifiable et différent de celui de l'auteur
- champ à supprimer à terme, car ces informations sont prises en charge par des tables spécifiques

Destinataire (cf. angl. *Addressee*)

- destinataire de l'œuvre (nom et/ou position sociale), s'il est possible de l'identifier = dédicataire ? commanditaire ?
- ce champ n'est pas actuellement renseigné pour la plupart des textes

Destinateur (cf. angl. *Addresser*)

- caractérisation sociolinguistique de l'auteur du texte (position sociale), âge
- ce champ n'est actuellement pas renseigné pour la plupart des textes

Public

- milieu social dans lequel on pense que le texte a été reçu
- ce champ n'est pas actuellement renseigné pour la plupart des textes

Résumé

- résumé de l'œuvre en quelques lignes
- ce champ n'est pas actuellement renseigné pour la plupart des textes

Traits d'écriture

- remarques de l'éditeur scientifique sur son manuscrit de base ; des informations plus structurées sont données dans les tables de description de manuscrits
- champ à supprimer, car ces informations sont à gérer par le module des sources primaires

Type de document

- œuvre (littéraire ou non) ou un document d'archive
- cette information est intégrée désormais au descripteur DOMAINE

3. Caractéristiques typologiques

3.1. Descripteur DOMAINE

Définition du DOMAINE : trait fonctionnel qui correspond à la destination principale du texte et au domaine d'activité auquel il se rattache :

- divertir → « littéraire »
- enseigner, instruire (ce qui concerne le savoir et sa transmission) → « didactico-scientifique »
- édifier (ce qui concerne le rituel et la diffusion du message chrétien) → « religieux »
- consigner/relater les événements du passé → « historique »
- influencer le comportement des personnes et des institutions au pouvoir → « politique »
- réguler la vie sociale → « juridique »

- régler des questions pratiques (actes de la pratique) → « source documentaire »

Liste des valeurs du descripteur DOMAINE

- Littéraire
- Didactico-scientifique¹
- Religieux
- Historique
- Politique²
- Juridique
- Source documentaire (acte de la pratique)

Utilisation du descripteur DOMAINE :

On sait qu'il est difficile, parfois impossible, de proposer un domaine unique pour un texte.

On propose de distinguer deux cas de figure :

- un texte relève plutôt d'un domaine mais a aussi des affinités avec un ou plusieurs autre(s) : on distingue un domaine principal et un ou plusieurs domaines secondaire(s) ; la conséquence est que si l'on veut établir des contrastes/comparaisons entre textes de plusieurs domaines, ce texte sera classé à l'intérieur de son domaine principal
- un texte ne peut vraiment pas être classé dans un seul domaine principal : on admet qu'il ait deux ou plusieurs domaines principaux ; la conséquence est que lors des opérations de contrastes, ce texte sera dans une catégorie à part ; pour cette raison cette solution doit être limitée aux cas les plus extrêmes.

Dans la base de données, les champs suivants correspondent à ce descripteur :

Domaine principal

- plusieurs domaines principaux peuvent être indiqués, séparés par une virgule, au cas où il est impossible de déterminer un seul domaine principal.

Domaine secondaire

- ce champ est rempli si un texte peut être placé à la fois dans plusieurs domaines
- plusieurs domaines peuvent être indiqués, séparés par une virgule

3.2. Descripteur GENRE

Trait qui correspond à certaines propriétés formelles internes au texte, mais qui sont difficiles à définir précisément ; au bout du compte, on s'accorde en général pour voir dans les genres textuels des catégories « intuitivement reconnues » et qui sont le produit d'une sorte de consensus à une époque donnée (Lee, Jauss pour le Moyen Age, etc.) ; catégorie dont les limites sont évidentes mais dont on peut en même temps difficilement se passer.

Il s'agit d'un descripteur dont la liste des valeurs est ouverte dans l'état actuel des connaissances.

Trois taxonomies au moins sont possibles :

- les GENRES tels qu'on pense que les médiévaux les voyaient ;

¹ En ce qui concerne la partition entre « scientifique » et « didactique » ou leur association à l'intérieur d'un seul DOMAINE, nous proposons dans un premier temps de réunir les deux et de voir dans une seconde phase si au vu de la masse des données il est pertinent ou non de scinder les deux valeurs.

² Ce domaine n'est pas d'actualité pour les textes antérieurs au XV^e siècle.

- les GENRES médiévaux tels qu'on les voit aujourd'hui (du point de vue de la tradition philologique) ;
- les GENRES tels qu'on les définit dans la société actuelle.

Pour l'instant on s'est surtout concentré sur la seconde.

Le descripteur GENRE est d'une certaine façon subordonné au descripteur DOMAINE, mais un même GENRE peut se rencontrer dans différents DOMAINES (par exemple le GENRE « dramatique » se rencontre dans les DOMAINES « religieux » et « littéraire »).

Le DOMAINE « source documentaire » comporte une série de valeurs de GENRE qui lui sont propres

Sources documentaires	Charte-lettre
	Registre ³
	Compte
	Censier-cadastre

Il existe une certaine affinité entre le descripteur GENRE et l'opposition entre textes narratifs, argumentatifs, expositifs et autres (lyriques, etc.). On proposera une mise en correspondance entre les genres et cet aspect de la typologie textuelle.

Dans la base de données, les champs suivants correspondent à ce descripteur :

Genre traditionnel

- genre du texte dans la tradition de la philologie médiévale (cf. *Dictionnaire des lettres françaises*)
- éventuellement le terme générique figurant dans le titre (*Chanson, Roman, Chantefable...*)

Genre moderne

- caractérisation de l'œuvre dans une taxonomie moderne (par exemple *Récits brefs* pour tous les recueils de petites histoires comme le *Roman de Renart*)

GenreRef

- genre retenu pour les partitions dans la BFM, une sorte de compromis entre le genre traditionnel et le genre moderne
- écrit sans espaces blancs et sans accents

3.3. Autres descripteurs typologiques actualisés

Thème

- thème (discipline) d'une œuvre didactico-scientifique

Commentaire-typologie

- remarques éventuelles sur le classement typologique

³

Registre de délibérations, de sentences, de plaidoiries etc.

3.4. Descripteurs typologiques non actualisés

Bossuat Niveau1

- classement de l'œuvre dans le Bossuat

Bossuat Niveau2

- éventuellement, une sous-rubrique dans le classement du Bossuat

4. Description de l'édition de référence

Nom

- identifiant de l'édition dans la base
- il est identique à « référence » sauf dans deux cas :
 - ❖ s'il y a plusieurs volumes : dans ce cas on ajoute le numéro du volume à la référence de l'œuvre (par ex. *commyn1*)
 - ❖ si plusieurs éditions d'un texte sont dans la base : dans ce cas les initiales de l'éditeur scientifique sont ajoutées au moment où une nouvelle édition est intégrée dans la base (par ex. *qgraal_cm*)
- les anciens identifiants seront progressivement remplacés par des sigles DEAF

Sigle DEAF

- identifiant dans la base bibliographique du *Dictionnaire étymologique de l'ancien français* (<http://www.deaf-page.de/>).

Numéro du volume

- un chiffre

Nombre de volumes

- un chiffre

Titre complet

- titre de la publication (qu'on peut citer dans une bibliographie)
 - ❖ anciennement, ce champ était utilisé pour indiquer le texte de la page de garde.

Editeur scientifique

- initiale du prénom et nom de l'éditeur scientifique (par ex. *L. Foulet*) ;
- s'il y a plusieurs éditeurs, les noms sont séparés par des points-virgules.

Maison d'édition

- Plusieurs éditeurs commerciaux peuvent être indiqués, séparés par des points-virgules

Lieu d'édition

- Les noms de villes en écriture latine ne sont pas francisés (*London* et pas *Londres*)
- Plusieurs villes peuvent être indiquées, séparées par des points-virgules

Collection

- Nom de la collection

- Nom de la revue précédé de la mention [*Revue*]

Numéro dans la collection

- chiffre
- volume et/ou numéro pour les revues

Date de l'édition

- année de la première parution de l'édition source de la numérisation

Année de réimpression

- année d'impression de l'exemplaire de référence (si différente de l'année de l'édition)

Editions précédentes

- année d'une édition antérieure du même texte par le même éditeur (différente de l'édition numérisée)
- éventuellement, toute autre information utile sur les éditions précédentes

ISBN

- s'il y a lieu

Type de cote

- bibliothèque dans laquelle se trouve l'exemplaire de référence

Cote de bibliothèque**Photocopie**

- présence (oui ou non) d'une photocopie de l'édition au centre de doc.

Caractérisation édition

- commentaire sur la qualité de l'édition (données de la base bibliographique du DEAF, de la base du NCA, de comptes-rendus, de recherches propres)

Ms de base

- identification du manuscrit de base de l'édition

Ms région

- région de la rédaction du manuscrit de base de l'édition (d'après la base bibliographique du DEAF)

Ms date libre

- la date de la rédaction du manuscrit de base de l'édition (d'après la base bibliographique du DEAF)

Ms date début

- la date de la rédaction du manuscrit de base la plus ancienne calculée selon les règles établies (*cf.* les règles pour la date de composition, tableau « Oeuvre »).

Ms date fin

- la date de la rédaction du manuscrit de base la plus tardive calculée selon les règles établies (*cf.* les règles pour la date de composition, tableau « Oeuvre »).

Ms date formelle

- la date de la rédaction du manuscrit de base formalisée
- moyenne calculée entre la « Ms date début » et la « Ms date fin ».

Commentaire

- tout autre commentaire concernant l'édition de référence

5. Description du texte numérisé**Nom du fichier**

- le nom du fichier sans extension
- normalement identique au « nom » de l'édition, sauf le cas où une édition est représentée par des échantillons répartis dans plusieurs fichiers

Enrichissement

- information sur l'enrichissement linguistique (étiquetage morphologique ou autre) du texte

Droits

- information sur les droits relatifs à l'édition de référence et à la version numérisée

Pages saisies

- Ex. 1-99

Nombre de mots

- nombre de mots, d'après l'outil « Dimensions » de Weblex

Origine du texte

- personne ou institution et éventuellement le projet à l'origine de la numérisation du texte

BFM0, BFM1, BFM2, etc.

- état d'intégration (*oui* ou *non* pour chaque champ)

Statut

- état de saisie du texte : intégral ou partiel

Mise à jour

- date de la dernière modification (contenu ou balisage)

Comment-num

- tout commentaire relatif à la numérisation du texte

Des champs supplémentaires sont ajoutés en cas de création d'un corpus spécial destiné à un projet de recherche particulier. Le champ porte alors le nom du corpus et les valeurs possibles sont *oui* ou *non*.

6. Description des révisions du texte numérisé

Corrections ID

- numéro automatique servant de clé unique pour cette table

Etape

- étape de la correction dans le « cycle de vie » d'un texte :
- les valeurs possibles sont : "0 Desc";"1 Scan";"2 OCR";"3 Nett1";"4 R1";"5 Nett2";"6 R2";"7 ValFin";"9 Autre"

Nom

- identifiant du fichier (relié à Nom dans la table Numerisation)

Fonction

- rôle de la personne effectuant une révision (par ex., « Relecteur », « Encodeur », « Numeriseur », etc.)

RespID

- identifiant du rôle d'un correcteur dans le cycle de vie d'un texte (par ex. « proofreader1 », « encoder2 », etc.)
- cette information permet de relier le nom d'une personne à un type de responsabilité indiqué dans les balises d'interventions éditoriales

Personne

- le prénom et le nom de la personne

Date

- mois et année de l'opération de révision
- ce champ n'est plus utilisé

Date formelle

- date précise de l'opération de révision (aaaa-mm-jj)
- date de fin pour les opérations terminées de plusieurs jours
- date de début pour les opérations en cours

Nombre d'heures

- le nombre d'heures que l'opération a duré (chiffre avec une décimale éventuelle)

Corrections

- description de la nature de l'opération de révision

Statut

- statut de l'opération
- les valeurs possibles : "terminé";"en cours";"à faire"

Remarques

- toute remarque supplémentaire concernant une opération de révision

7. Références des personnes et organisations responsables des éditions, de la datation et de la numérisation des textes

[à compléter]

8. Description de manuscrits

Les champs correspondent aux éléments de description de manuscrit recommandés par la TEI.

Voir de *Manuel d'encodage BFM Manuscrits* pour plus de détail

[à compléter]

Références

Bossuat, R. (1954 -). *Manuel bibliographique de la littérature française du Moyen âge*, Paris, Librairie d'Argences.

Jauss, H. R. et al. (éd.) (1972-). *Grundriß der romanischen Literaturen des Mittelalters*, Heidelberg, C. Winter.

2. Domaines et genres actuellement représentés dans la BFM

(état du 7 novembre 2008)

Domaine	Genre	Nombre de textes	Nombre de mots
didactique	[indéterminé]	5	151190
	bestiaire	1	13788
	comput	1	14746
	exemples	1	94125
	lapidaire	2	0
	manuel	5	134420
	proverbes	3	0
	recettes	2	0
	roman	8	252356
	traite	7	321159
	Total	35	981784
historique	[indéterminé]	3	84467
	chronique	22	1047646
	epistolaire	1	0
	histoire	5	176367
	memoires	9	282848
	Total	40	1591328
juridique	[indéterminé]	3	4175
	ceremonial	1	0
	charte	10	292437
	chartes	1	0
	coutumier	1	0
	divers	1	0
	journal	2	132775
	jugements	1	0
	legislatif	1	0
	registre	2	335805
	serment	1	117
	traite	1	142507
	Total	25	907816
litteraire	[indéterminé]	12	377083
	biographie	1	111012
	comique	1	2467
	dramatique	5	28420
	epique	9	118496
	fabliau	14	24636
	lyrique	17	190620

	nouvelle	5	346973
	rbrefs	23	114506
	roman	50	2535732
	rvoyage	1	252890
	Total	138	4102835
politique	proposition	1	1351
	Total	1	1351
religieux	[indéterminé]	10	9080
	cantique	1	0
	commentaire	1	0
	dialogue	1	0
	dramatique	3	2749
	elucidaire	1	0
	epistolaire	3	0
	hagiographie	5	71197
	hagiographique	2	0
	lyrique	4	222185
	miracle	5	0
	miracles	1	0
	poeme	1	26111
	proverbes	1	0
	psautier	4	46094
	rbrefs	1	9453
	regle	1	0
	rpieux	1	0
	sermon	9	6663
	tradbib	1	90554
	traite	1	0
		Total	57