

Statut du document	Structure et vocabulaire reconnu par la communauté scientifique
Titre	Text Encoding Initiative (TEI)
Mot clé	Domaines sémantiques
Source	Text Encoding Initiative http://www.tei-c.org/
Date de publication	
Nombre de pages	Quelques centaines
Langue	Version originale créée en anglais. Version française par François Role disponible à : http://www.oasis-open.org/cover/tei.html
Lien avec autres normes	L'une des productions les plus riches de la mouvance SGML développée par des scientifiques de diverses disciplines entre 1990 et 1993 principalement.
Situation actuelle	Référence reconnue, implantation sur outils d'accès public.
Description	<p>Un document comporte un <i>en-tête</i> servant à véhiculer des métadonnées sur le fichier physique principalement, une étiquette avec des renseignements devant servir à en faciliter l'ouverture. Ce qui est le plus réutilisable du TEI, c'est plutôt sa façon générique de décomposer un document en une séquence de trois grandes parties ainsi qu'en définissant une série d'éléments pouvant être insérés dans n'importe laquelle de ces parties. Les trois parties sont :</p> <ul style="list-style-type: none"> - Préliminaires : <i>Page-de-titre, Titre-du-document, Partie-de-titre, Auteur-du-document, Date-du-document, Date-d'impression.</i> Les principaux types de <i>Subdivision</i> sont : <i>Avant-propos, Préface, Dédicace, Sommaire, Remerciements, Table-des-matières, Frontispice.</i> - Corps : <i>Paragraphes, Subdivision</i> (niveaux 1 à 7) avec <i>Titre</i> et <i>Partie-finale</i>. Notons que l'introduction et la conclusion d'un document font partie de son Corps. - Post-liminaires : Les principaux types de <i>Subdivision</i> sont : <i>Annexe, Glossaire, Notes, Bibliographie, Index, Colophon.</i> <p>Les éléments « flottants » sont nombreux à être définis : <i>référence bibliographique, citation, note, liste, tableau, liens générique-spécifique, précédent-suivant, antérieur-postérieur, marques de transcription, d'éditeur, de présentation.</i> Beaucoup de sophistication a été introduite dans la TEI pour ceux qui auraient besoin d'une plus grande précision que pour les besoins administratifs courants. La réutilisation des définitions de la TEI est souhaitable à cause de la reconnaissance déjà acquise. Voici une liste alphabétique des noms en français d'éléments choisis de la TEI, suivis du nom de balise en anglais de la « TELLite » (ce qui signifie la Text Encoding Initiative, en version simplifiée). La liste est extraite de celle qui se trouve à la section 21.2 de l'URL mentionné ci-haut. Ne sont pas repris les éléments jugés savants, certains de ceux que le Chantier a définis avec le même sens, ceux qui sont jugés désuets. Quelques définitions ont été reformulées pour les clarifier. Une validation éventuelle de ces choix et révisions serait souhaitable. Pour ceux qui voudraient mieux comprendre la TEI, on peut lire l'article de Diane Beaudry dans la revue de l'EBSI à :</p>

Remarque

Lexique anglais-français Par ordre alphabétique du nom français suivi, entre crochets, du nom abrégé en anglais.

TEI simplifiée (éléments choisis)

abréviation <abbr>^{TEI}
contient une abréviation de n'importe quel type ; le développement peut être donné dans l'attribut **expan** ;

adresse <address>^{TEI}
contient une adresse postale ou d'un autre type, par exemple celle d'un éditeur, d'une organisation, ou d'un individu ;

ajout <add>^{TEI}
contient des lettres, des mots, ou des expressions insérées dans le texte par un auteur, scribe, annotateur, ou correcteur ;

argument<argument>^{TEI}
une liste formelle ou une description en texte libre des sujets traités par une subdivision d'un texte ;

auteur <author>^{TEI}
dans une référence bibliographique, contient le nom de l'auteur ou des auteurs, soit personnel soit corporatif, d'un ouvrage; constitue la mention de responsabilité principale pour toute unité bibliographique ;

auteur-du-document <docAuthor>^{TEI}
contient le nom de l'auteur du document, tel qu'il est indiqué sur la page de titre (souvent mais pas toujours contenu dans un élément <byline>) ;

brouillage <unclear>^{TEI}
contient un mot, une expression, ou un passage qui ne peut pas être transcrit avec certitude car illisible ou inaudible dans la source ;

catégorie <category>^{TEI}
contient une catégorie descriptive individuelle, éventuellement contenue dans une catégorie de niveau supérieur, le tout dans le cadre d'une taxonomie définie par l'utilisateur ;

cellule <cell>^{TEI}
contient une cellule d'un tableau ;

citation <cit>^{TEI}
citation empruntée à un autre document, accompagnée d'une référence bibliographique indiquant son origine ;

citation <q>^{TEI}

contient une citation ou quelque chose qui s'apparente à une citation ;

classification <classDecl>^{TEI}

contient un ou plusieurs taxonomies définissant tout code de classification employé ailleurs dans le texte ;

code <code>^{TEI}

contient un court fragment de code dans un langage formel donné (souvent un langage de programmation) ;

codification <classCode>^{TEI}

contient le code de classification employé pour ce texte dans un système de classification normalisé et identifié par l'attribut **scheme** ;

corps <body>^{TEI}

contient le corps entier d'un texte unitaire unique, à l'exclusion de toute pièce liminaire ou annexe ;

correction <corr>^{TEI}

contient la forme correcte d'un passage apparemment erroné dans la copie ;

date-du-document <docDate>^{TEI}

contient la date du document, telle qu'elle est donnée (habituellement) sur la page de titre ;

description-de-catégorie <catDesc>^{TEI}

décrit une certaine catégorie à l'intérieur d'une taxonomie ou typologie de texte, sous forme d'une courte description en texte libre ;

éditeur <publisher>^{TEI}

fournit le nom de l'organisation responsable de la publication ou de la diffusion d'une unité bibliographique ;

élément <item>^{TEI}

contient un seul composant de liste ;

épigraphe <epigraph>^{TEI}

contient une citation, anonyme ou attribuée, figurant au début d'une section ou d'un chapitre, ou dans une page de titre ;

figure <figure>^{TEI}

marque l'emplacement où un graphique doit être inséré dans un document; les attributs peuvent servir pour indiquer une entité SGML contenant l'image elle-même (dans une quelconque notation non-SGML) ; les paragraphes à l'intérieur de l'élément <figure> peuvent servir pour transcrire les légendes ;

formule <formula>^{TEI}

contient une formule chimique ou mathématique, éventuellement présentée dans une quelconque notation non-SGML. L'attribut **notation** est employé pour nommer la notation non-SGML utilisée pour transcrire la formule ;

glossaire <gloss>^{TEI}

marque un mot ou une expression qui fournit une glose ou une définition pour un autre mot ou une autre expression ;

groupe <group>^{TEI}
contient plusieurs textes unitaires ou groupes de textes ;

hiatus <gap>^{TEI}
indique un endroit où du texte a été omis dans une transcription, soit pour des raisons éditoriales décrites dans l'en-tête de la TEI et ayant trait à la méthode d'échantillonnage, soit parce que le texte est illisible ou inaudible ;

identifiant <ident>^{TEI}
contient un identifiant d'un certain type, par exemple, un variable de nom ou le nom d'un élément ou attribut SGML ;

identifiant-générique <gj>^{TEI}
contient un type d'identifiant spécial : un identifiant générique SGML, ou nom d'élément ;

lieu-de-publication <pubPlace>^{TEI}
contient le nom du lieu où une unité bibliographique a été publiée ;

liste <list>^{TEI}
contient toute séquence d'éléments organisés sous forme de liste, celle-ci pouvant être soit numérotée, soit identifiée par des puces ou tirets ;

mots-clés <keywords>^{TEI}
regroupe une liste de mots ou d'expressions servant à identifier le sujet ou la nature d'un texte ; si ces mots clés proviennent d'un vocabulaire contrôlé, ce dernier eut être identifié au moyen de l'attribut **scheme** ;

nom <name>^{TEI}
contient un nom propre ou une expression substantive ; les attributs peuvent indiquer son type, donner une forme normalisée, ou l'associer à une chose ou à un individu spécifique, au moyen d'identifiants uniques ;

nombre <num>^{TEI}
contient un nombre, écrit sous une forme quelconque, avec une valeur normalisée dans l'attribut **valeur** ;

note <note>^{TEI}
contient une note ou une annotation, avec des attributs pour indiquer le type, l'emplacement et la source de la note ;

numéro-de-subdivision <div1\$>\$. . . \$<div7>^{TEI}
niveau de subdivision indiqué dans une hiérarchie descendante allant de 1 jusqu'à 7 ;

page-de-titre <titlePage>^{TEI}
contient la page de titre d'un texte, tel que figurant à l'intérieur des parties liminaires ou des annexes ;

paragraphe <p>^{TEI}
marque les paragraphes écrits en prose ;

partie-de-titre <titlePart>^{TEI}

contient une subdivision ou division du titre d'une oeuvre, comme indiqué sur une page de titre ; également employé pour les fragments « flottants » de la page de titre ne faisant pas partie du titre de document, comme par exemple le nom de l'auteur ;

partie-finale <trailer>^{TEI}

contient un bas de page ou un titre final paraissant à la fin d'une division de texte ;

partie-initiale <opener>^{TEI}

contient la date, la signature, la salutation et les autres expressions similaires susceptibles d'apparaître au début d'une division, surtout dans une lettre ;

point-d'ancrage <anchor>^{TEI}

précise un emplacement ou un endroit à l'intérieur d'un document afin de permettre de pointer sur cet emplacement ;

point-d'insertion-automatique <divGen>^{TEI}

indique l'emplacement où doit paraître une division textuelle générée automatiquement par une application de traitement texte ; l'attribut type précise s'il s'agit d'un index, d'une table des matières, ou de quelque chose d'autre ;

portée-finale <closer>^{TEI}

regroupe la date, la signature, la salutation et toute formule similaire pouvant apparaître en tant que groupe final à la fin d'une division, surtout dans le cas d'une lettre ;

post-liminaires <back>^{TEI}

contient tout appendice, etc., suivant la partie principale d'un texte ;

préliminaire <front>^{TEI}

contient tout texte liminaire (en-tête, page de titre, préfaces, dédicaces, etc.) se trouvant avant le début du texte proprement dit ;

présentation <rendition>^{TEI}

fournit l'information concernant la présentation prévue pour un ou plusieurs éléments ;

rangée <row>^{TEI}

contient une rangée d'un tableau ;

référence-de-catégorie <catRef>^{TEI}

spécifie une ou plusieurs catégories définies à l'intérieur d'une certaine taxonomie ou typologie de texte ;

segment <seg>^{TEI}

identifie un passage ou un segment de texte à l'intérieur d'un document afin d'en permettre le pointage; l'attribut **type** catégorise le segment ;

série <series>^{TEI}

contient une information concernant la série dans laquelle un livre ou une autre unité bibliographique a paru ;

subdivision <div>^{TEI}

contient une subdivision hiérarchique des parties d'une séquence délimitée que ce soit dans les préliminaires, le corps ou les post-liminaires d'un document ;

tableau <table>^{TEI}

contient du texte affiché sous forme tabulaire en rangées et en colonnes ;

taxonomie <taxonomy>^{TEI}

définit une typologie employée pour classer des textes soit implicitement, au moyen d'une citation bibliographique, soit explicitement au moyen d'une taxonomie structurée ;

terme <term>^{TEI}

contient un mot unique, plusieurs mots ou une désignation symbolique considérée comme un terme technique ;

titre <title>^{TEI}

contient le titre d'une oeuvre, que ce soit un article, livre, journal, ou une série, y compris tout sous-titre ou titre alternatif ;

titre-du-document <docTitle>^{TEI}

contient le titre d'un document, y compris tous ses constituants, tels qu'ils apparaissent sur une page de titre; doit être partagé en éléments <titlePart> ;

Rédacteur : Richard Parent
Organisation source : Secrétariat du Conseil du trésor
Date de publication : 21 août 2000
Raison d'être : Connaissance technologique
Programme gouvernemental : Inforoutes et ressources informationnelles
Nom du modificateur :
Date de dernière modification :
Note numéro : 44