Introduction au langage XML pour l'analyse de texte

une série d'ateliers

Les mardis de 10 h 30 à 12 h (FR)
Les mardis de 13 h 30 à 15 h (EN)

Avec Alice Defours et
Constance Crompton



Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**
3. Liens et chemins
4. **Encodage : Niveau 5**
5. Bonus : Réconciliation et Fichiers d’Autorité
6. Etude de cas : le projet LGLC



Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**





NIVEAU 4 - CONTENU (ENTITÉS ET MÉTADONNÉES)

Nuance et représentation des informations

C'est quoi? Encodage des informations thématiques, des entités du monde réel, personnes, lieux ,dates, etc.

<**persName**> : Noms de personnes cités dans le contenu textuel (ex: personnages, auteurs, artistes)
<**placeName**> : Lieux mentionnés dans le texte (ex: villes, pays, régions, etc.)
<**orgName**> : Noms d'organisations ou d'entreprises (ex: musées, galeries, institutions, etc.)

Ces trois éléments sont extrêmement permissifs et acceptent énormément d'attributs afin de raffiner l'information encodée. Allez consulter les TEI Guidelines pour avoir une idée complète des possibilités et des exemples précis de leur usage.

<**date**> : Dates importantes citées dans le texte (elles peuvent être détaillées, voir Slide Niveau 1)
<**title**> : Titres d'œuvres ou d'autres références citées dans le texte (peintures, sculptures, romans)
@type : Type de titre (ex. : "roman", "peinture")

<**term**> : Termes spécialisés (ex: styles artistiques, mouvements, mots compliqués à expliquer, etc.).



NIVEAU 4 - CONTENU (ENTITÉS ET MÉTADONNÉES)

C'est quoi? Encodage des informations thématiques, des entités du monde réel, personnes, lieux ,dates, etc. Nuance et représentation des informations

Exercice pratique de niveau 4 :

- Encodez les **entités** de votre document !
- Utilisez la fiche précédente pour trouver des éléments pertinents à votre document
  - Ou bien, faites des recherches sur les TEI Guidelines pour trouver les éléments qui correspondent le mieux aux besoin de vos structure physique.
- Référez-vous à la vidéo de démonstration !

Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**
3. Liens et chemins



3. Liens et chemins

Un chemin = Une adresse

Il correspond à la localisation d'un fichier dossier dans un espace de stockage.

On distingue deux types de chemins :

1. Les chemins **absolus** (adresse complète)
2. Les chemins **relatifs** (adresse en fonction de l'endroit où on se trouve déjà)

3. Liens et chemins

Un chemin = Une adresse

Il correspond à la localisation d'un fichier dossier dans un espace de stockage.

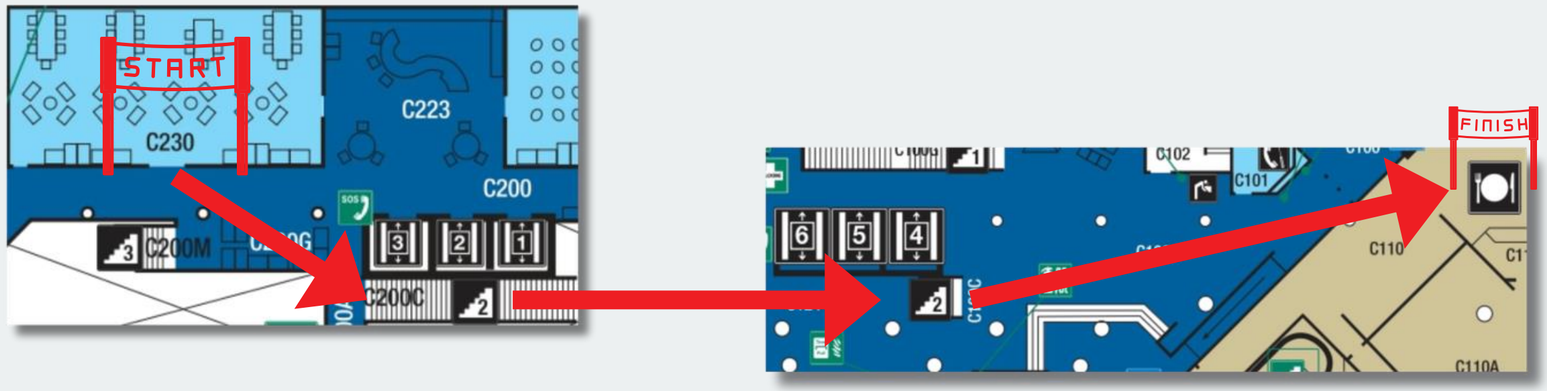
On distingue deux types de chemins :

1. Les chemins **absolus** (adresse complète)
2. Les chemins **relatifs** (adresse en fonction de l'endroit où on se trouve déjà)



Exemples, à notre échelle de l'université d'Ottawa :
Notre chemin absolu est l'adresse postale (permanente) du bâtiment CRX

Un chemin relatif serait l'adresse "de la salle où je suis "CRX 230", jusqu'au "café du RdC" où je veux aller dans le même bâtiment" si j'effectuais un déplacement.



3. Liens et chemins

a. Chemins absolus

Le départ du chemin absolu est la racine du système de fichiers, ou l'adresse complète d'une ressource sur le web (URL), exemple :

Sur Windows :
C:\Users\Nom\Documents\Projet\TEI.xml

Sur MacOS :
/Home/Nom/Projet/TEI.xml

Sur le web :
https://lglc.ca/images/lglc_reduced.jpg
https://www.wikidata.org/wiki/Q450923

En TEI, c'est utile quand on veut lier précisément une ressource
(autre fichier, ou lien permanent sur le web)

3. Liens et chemins

b. Chemins relatifs

Le chemin relatif dépend du dossier dans lequel on est. Il exprime la localisation d'un fichier précis par rapport à un autre dossier/fichier.

Les symboles importants à reconnaître sont :

1. Un point "." qui indique le dossier actuel,
2. Deux point ".." qui indique le dossier **parent** (donc un niveau au-dessus)

Exemple :

- Si votre fichier TEI est dans **/Projet/Texte/TEI.xml**,
- et qu'une image que vous voulez intégrer au document TEI est dans **/Projet/Images/photo.jpg**,
- alors, depuis TEI.xml, le chemin relatif vers l'image est "**../Images/photo.jpg"**

Dans notre TEI.xml, si on veut lier la photo.jpg avec un élément du contenu, alors l'encodage nécessaire et son chemin relatif seront :

<graphic url="../Images/photo.jpg"/>

".." remonte de /TEI.xml vers /Projet
puis "/Images/photo.jpg" descend vers le bon dossier

Pourquoi c’est important ?

Quand on fait de l’encodage en TEI, on a souvent des éléments qui pointent vers des ressources externes, comme des :

1. Images de manuscrits,
2. Fac-similés,
3. Fichiers audio,
4. Feuilles de transformation en XSLT,
5. D’autres fichiers XML liés.

Si les liens sont mal formés, votre document ne pourra pas trouver et lier les ressources au contenu.

Une astuce pour savoir si un chemin est absolu/relatif

1. **Où est le fichier actuel** ?

2. Où est la resource à lier ?
   a. Dans mon ordinateur (relatif) ?
   b. Dans le même fichier TEI.xml que je suis en train d'encoder ?
   c. Sur le web (absolu) ?

3. Si absolu
   a. Le chemin est **l'URL** dans cet exemple

4. Si relatif
   a. Où est la ressource à atteindre ?
      a. Dans mon ordinateur ?
         i. Faut-il monter au dossier parent avec ".." ?
         ii. Ou rester dans le même dossier avec "." ?
         iii. Dans l'un ou l'autre, le chemin sera "." ou ".."/Dossier/Fichier.extension
      b. Dans le même fichier TEI.xml ?
         i. Le chemin sera l'appel de l'attribut **correspondant** avec un
            <element corresp="#venus"**>**
            ou **<element corresp=**"#statue"**>**

Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**
3. Liens et chemins
4. **Encodage : Niveau 5**





NIVEAU 5 – DETAILLER LE CONTENU PAR DES ATTRIBUTS

C'est quoi? Toutes les informations connues des chercheurs qui souhaitent expliquer à l'ordinateur des nuances d'informations précises

@xml:id : Identifiants uniques pour les éléments (utile pour lier des éléments entre eux, comme des dialogues <said>, ou des <note> avec des <persName> ou des <placeName>)

@n : Numérotation (ex: chapitres,)

@rend : Mise en forme (ex: titres en italique)

@type : Catégories d'éléments (ex: type de division, type de note, types de concept)

@who : Indiquer l'auteur d'une citation ou d'un discours, d'un dialogue, etc.

@ref : Liens vers des ressources externes (fichiers, bases d'autorité, DONNEES LIEES!)

@target : Un des nombreux attributs pouvant pointer vers un **@xml:id**. Connecte deux éléments entre eux.

@corresp : Peut aussi pointer vers un **@xml:id,** ex: **<note corresp="#para3">** Contenu**</note>** relie cette **<note>** avec le paragraphe 3 du même document (**<p xml:id="para3">**).

Exercice 1

En observant les liens et chemins ci-dessous, répondez à la question suivante :

Où puis-je m'attendre à trouver les informations liées aux éléments ci-dessous ?

```
<name corresp="#statue">la déesse de cuivre</name>

<p>Le
    <ref target="./venus.xml#ch7">chapitre 7</ref>
   de cet ouvrage en donne une description.
</p>

<p>Voir la publication du
     <date corresp="../contexts/timeline/1880-90.xml#note7"> 8 Novembre 1890</date>
     pour plus d’informations.
</p>

<p>Ce village est situé en
     <ref target="https://www.wikidata.org/wiki/Q142">France</ref>.
</p>

<p>
    <rs sameAs="https://viaf.org/fr/viaf/182225067">
        <persName>La Vénus d’Ille</persName>
    </rs> est sont nom.
</p>
```

Exercice 1

En observant les liens et chemins ci-dessous, répondez à la question suivante :

Où puis-je m'attendre à trouver les informations liées aux éléments ci-dessous ?

```
<name corresp="#statue">la déesse de cuivre</name>
```

MÊME DOCUMENT TEI

```
<p>Le
   <ref target="./venus.xml#ch7">chapitre 7</ref>
   de cet ouvrage en donne une description.
</p>
```

UN AUTRE FICHIER .XML DANS VOTRE ORDINATEUR

```
<p>Voir la publication du
    <date corresp="../contexts/timeline/1880-90.xml#note7"> 8 Novembre 1890</date>
    pour plus d'informations.
</p>
```

```
<p>Ce village est situé en
    <ref target="https://www.wikidata.org/wiki/Q142">France</ref>.
</p>
```

UNE RESSOURCE SUR LE WEB

```
<p>
   <rs sameAs="https://viaf.org/fr/viaf/182225067">
      <persName>La Vénus d'Ille</persName>
   </rs> est sont nom.
</p>
```



NIVEAU 5 – DETAILLER LE CONTENU PAR DES ATTRIBUTS

C'est quoi? Toutes les informations connues des chercheurs qui souhaitent expliquer à l'ordinateur des nuances d'informations précises

Exercice 2 de pratique de niveau 5 :

- Donnez des **attributs aux entités** de votre document !
- Utilisez la fiche précédente pour trouver des éléments pertinents à votre document
  - Ou bien, faites des recherches sur les TEI Guidelines pour trouver les éléments qui correspondent le mieux aux besoin de vos structure physique.
- Référez-vous à la vidéo de démonstration !

Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**
3. Liens et chemins
4. **Encodage : Niveau 5**
5. Bonus : Réconciliation et Fichiers d’Autorité



ÉTAPE BONUS : RÉCONCILIATION ET LIENS AVEC DES FICHIERS D'AUTORITÉ

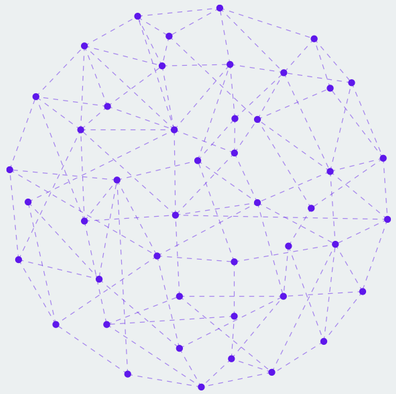
C'est quoi? Commencer à créer des fichiers connectés, intelligents, liés!

Lier les données à des fichiers d'autorité externes !

<**persName**> : Utiliser @ref pour pointer vers **VIAF** ou **Wikidata** pour les auteurs ou artistes
<**placeName**> : Utiliser @ref pour des liens vers **GeoNames**
<**orgName**> : Référencer les **sites** des institutions comme des musées dans un @ref !

C'est véritablement une étape qui ajoute une valeur en diamant à votre travail, car ça signifie que vous appliquez une partie des principes FAIR.

Principes FAIR

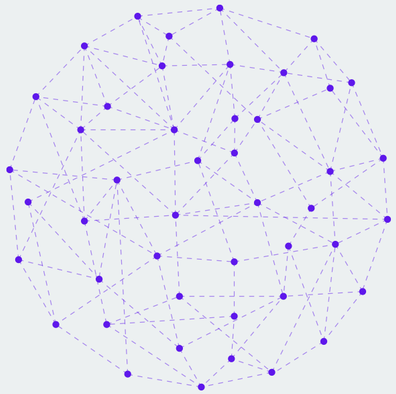


ÉTAPE BONUS : RÉCONCILIATION ET LIENS AVEC DES FICHIERS D'AUTORITÉ

C'est quoi? Commencer à créer des fichiers connectés, intelligents, liés!

Exercice pratique bonus :

- **Réconciliez les entités** de votre document,
- Utilisez la fiche précédente pour trouver des éléments pertinents à votre document
  - Ou bien, faites des recherches sur les TEI Guidelines pour trouver les éléments qui correspondent le mieux aux besoin de vos structure physique.
- Référez-vous à la vidéo de démonstration !



ÉTAPE BONUS : RÉCONCILIATION ET LIENS AVEC DES FICHIERS D'AUTORITÉ

C'est quoi? Commencer à créer des fichiers connectés, intelligents, liés!

Exercice pratique bonus :

- **Réconciliez les entités** de votre document,
- Utilisez la fiche précédente pour trouver des éléments pertinents à votre document
  - Ou bien, faites des recherches sur les TEI Guidelines pour trouver les éléments qui correspondent le mieux aux besoin de vos structure physique.
- Référez-vous à la vidéo de démonstration !

ET APRES ?

Atelier IV

7 avril 2026

1. Plan de l’atelier
2. **Encodage : Niveau 4**
3. Liens et chemins
4. **Encodage : Niveau 5**
5. Bonus : Réconciliation et Fichiers d’Autorité
6. Etude de cas : le projet LGLC









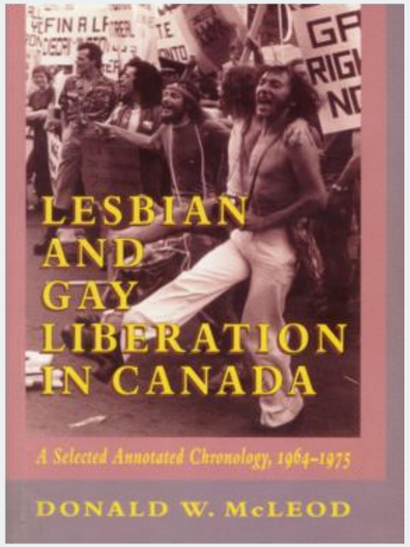

LABO DE DONNÉES EN SCIENCES HUMAINES

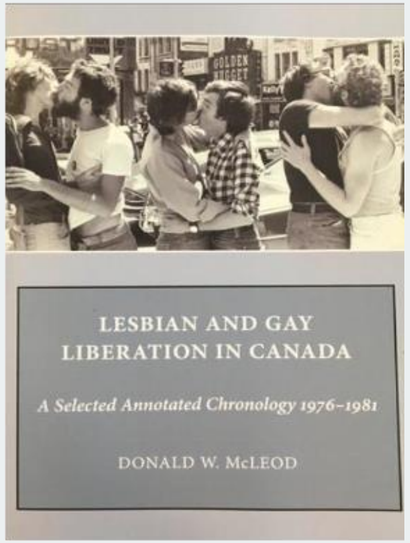
THE HUMANITIES DATA LAB











PIs Michelle Schwartz and Constance Crompton lglc.ca and prosopography.lglc.ca

Communauté
Représentation
Activisme

Environ 8000 gais et lesbiennes dans les rues de Montréal

Quelques politiciens parmi les manifestants

Un «témoignage de solidarité» de Boulerice, Lavallée et Forcillo

KATHLEEN LÉVESQUE
LE DEVOIR

Sous un soleil de plomb, environ 8000 gais et lesbiennes ont envahi hier après-midi les rues de Montréal dans une grande démonstration de fierté.

Des pancartes affichaient des slogans revendicateurs. Ici, un «vive la différence» et là, «le droit d'aimer sans peur et sans reproche». Mais globalement, l'ambiance était à la fête. Avant même que le défilé ne s'enclenche, des manifestants ont dansé sans plus de cérémonie, à l'angle de l'Avenue du Parc et de la rue Mont-Royal, chacun y allant de son déhanchement lascif ou plus sobrement, d'un tapement de pied.

«On voulait se montrer fièrement. C'est l'espoir de la communauté qui s'exprime ici», explique Puelo Deir, l'un des porte-parole de Divers Cité organisme en charge de l'événement. Depuis deux ans, le traditionnel défilé des gais et lesbiennes se déroule à l'extérieur des limites du village gai de Montréal. «La grande majorité des gais et lesbiennes de fréquentent pas le village, affirme Roger Leclerc, président du Centre communautaire des gais et lesbiennes de Montréal. Il fallait donc sortir du ghetto.»

Pour marquer le départ du défilé, les milliers de participants ont utilisé des sifflets dans une joyeuse cacophonie. Les chars allégoriques, colorés et bruyants, souvent chargés de jeunes gens qui dansent, alternaient avec des marcheurs.

Spectacle

Quelques politiciens se sont d'ailleurs mêlés à la foule. On pouvait y reconnaître le député péquiste André Boulerice, André Lavallée, membre du comité exécutif de la Ville de Montréal ainsi que Sammy Forcillo, conseiller municipal de l'opposition pour qui il s'agissait d'un «témoignage de solidarité».

Soixante orgnismes d'aide et de soutien ainsi que des groupes politiques ou de divertissements ont participé au défilé qui se rendait jusqu'au parc Lafontaine. Des spectacles y étaient prévus jusqu'à 20h.

Ne manquant aucune occasion de faire valoir leurs principes, des membres du Mouvement raëlien se sont faufilés dans la manifestation d'hier.

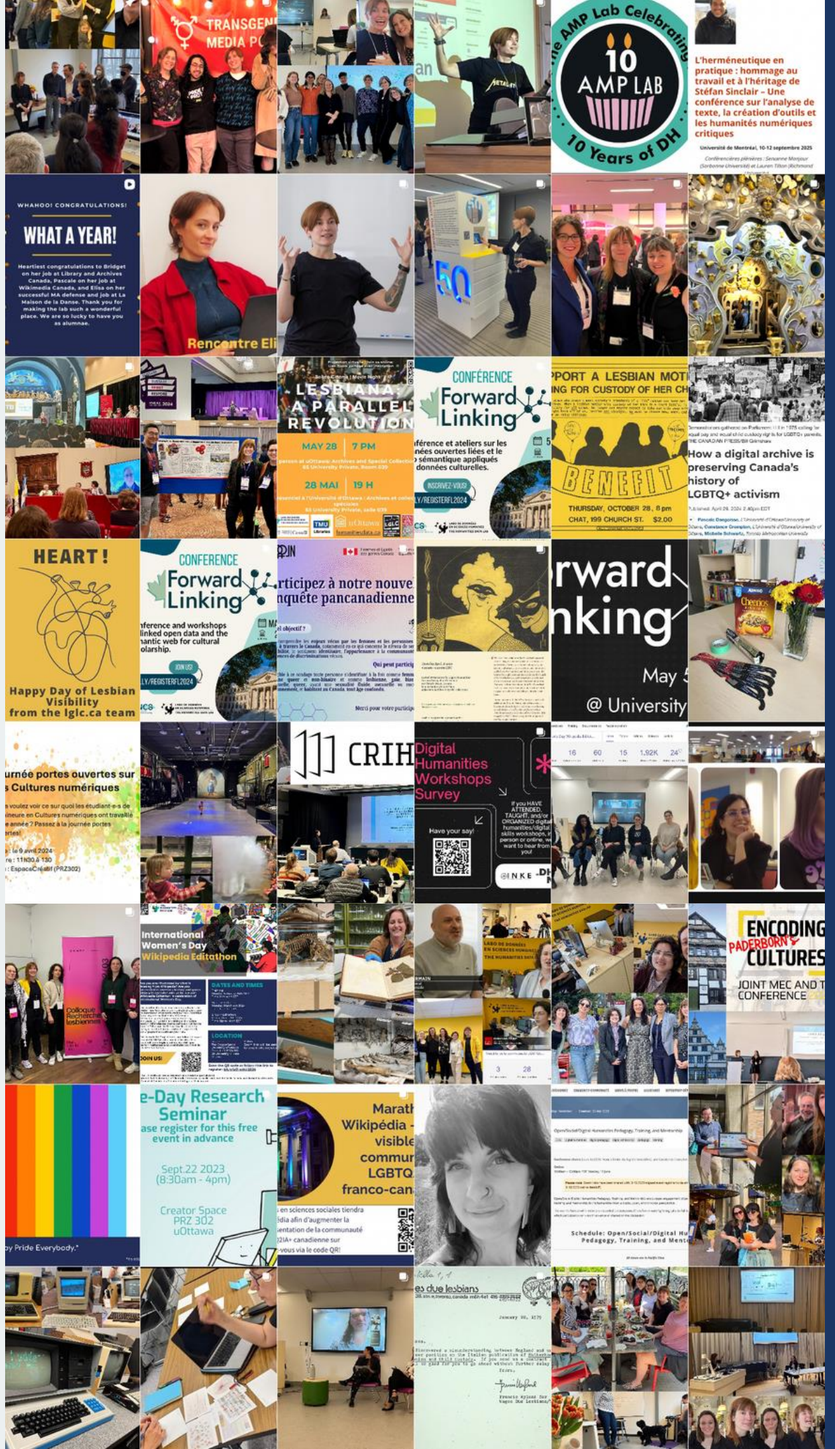
«Nous nous portons à la défense de toutes les minorités, de toutes les souffrances humaines», affirme Marie-Marcelle Godbout, transexuelle et guide humaniste au Comité international des diversités sexuelles chez les raëliens.

Au fur et à mesure que la joyeuse équipée sillonnait les rues, les curieux se sont agglutinés le long des trottoirs. «Moi et ma femme, on est là par curiosité. On veut les voir pour se faire une idée. Mais vous savez, les gens sont libres de faire ce qu'ils veulent», a commenté Raymond Beaucaire.

De ce point de vue, les éléments les plus extravagants de la communauté gaie, étaient très visibles. Ainsi, des hommes déambulaient vêtus de cuir, les fesses nues.Plus simplement, d'autres avaient troqué le pantalon pour le jupe écossaise. Et enfin, des travestis, perchés sur des talons hauts et fortement maquillés, roulaient des hanches dans des robes à paillettes.

«Je trouve ça dommage. Ça nuit à notre image», souligne Sylvie Marceau qui s'affichait comme lesbienne. «Je suis fier de cette minorité, affirme quant à lui Roger Leclerc du Centre commaunitaire des gais et lesbiennes de Montréal. Nous sommes une communauté qui aime s'amuser et on le montre. On veut d'ailleurs en venir à faire de cette manifestation, un grand festival.»





Homosexualité

Un référendum qui fera parler bien du monde

par Patrick POLGE

NEW YORK (AFP) — Le 6 juin prochain, les résidents du comté de Dade, en Floride, se prononceront par référendum pour ou contre le maintien de mesures discriminatoires en matière d'emploi et de logement à l'égard des homosexuels. Et ce vote, le premier du genre, a pris une importance nationale au terme d'une campagne très vive, les partenaires des deux camps s'acharnant à mélanger le sexe, la religion et les droits civiques. Et l'on crie aux "sorcières de Salem" ou à la décadence de la société.

L'affaire, anodine au départ, a été suscitée par un personnage très controversé, Anita Bryant, chanteuse de charme à succès, ancienne reine de beauté et promotrice du jus d'orange de Floride, selon eux, du référendum de lundi.

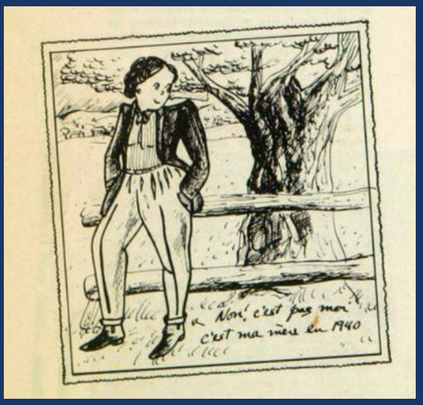
Mais l'homosexualité, révèle Anita Bryant, est un péché, "une bête sauvage de l'enfer" comme elle le proclame dans ses prières publiminatoires en matière d'emploi et de logement à l'égard des homosexuels. Et ce vote, le premier du genre, a pris une importance nationale au terme d'une campagne très vive, les partenaires des deux camps s'acharnant à mélanger le sexe, la religion et les droits civiques. Et l'on crie aux "sorcières de Salem" ou à la décadence de la société.

L'affaire, anodine au départ, a été suscitée par un personnage très controversé, Anita Bryant, chanteuse de charme à succès, ancienne reine de beauté et promotrice du jus d'orange de Floride, selon eux, du référendum de lundi.

Mais l'homosexualité, révèle Anita Bryant, est un péché, "une bête sauvage de l'enfer" comme elle le proclame dans ses prières publiminatoires, où elle demande à Dieu de lui donner la force de combattre et de vaincre.

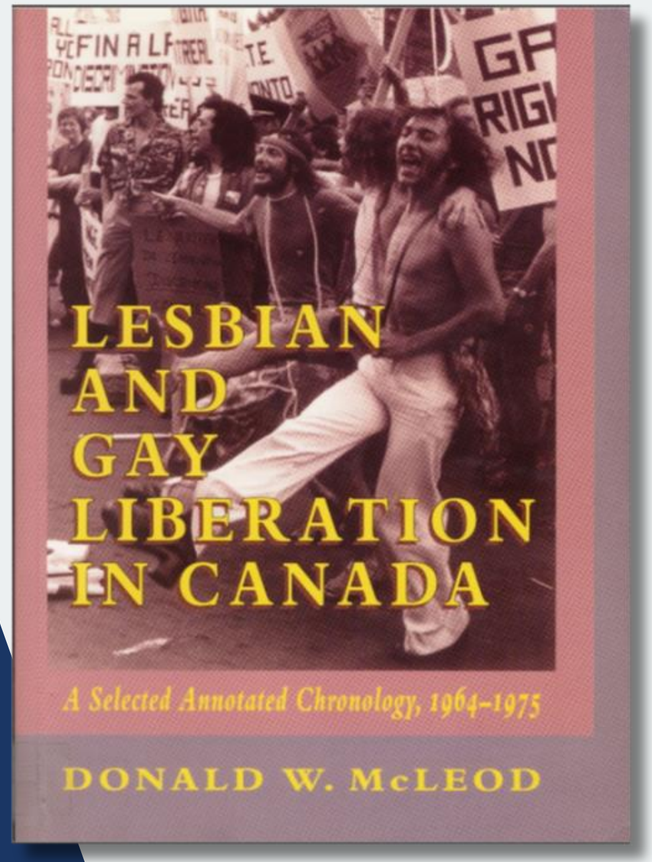
Evoquant Sodome et Gomorrhe, le Leviathan et le Pentateuque, Anita Bryant assure que la bible exècre les homosexuels, que ceux-ci sont de véritables dangers publics qui tentent de séduire les garçons impubères... et qu'ils n'ont qu'à changer, puisqu'avec un peu de bonne volonté, ils s'attacheraient aux charmes féminins et se marieraient pour contribuer utilement à la survie de l'espèce.

Avis divers





Architecture du Projet LGLC



June 18 * New York, N.Y.; Toronto * Jane Rule's novel *This Is Not for You* was published by McCall Publishing Company in New York and Doubleday of Canada in Toronto.

["A Bibliography," *Canadian Fiction Magazine*, no. 23 (1976), pp. 133, 136; Jane Rule, "Jane Rule: The Woman behind *Lesbian Images*" (interview), *Body Politic*, no. 21 (1975), p. 15.]

August * Montréal * Michel Tremblay's musical comedy *Demain matin, Montréal m'attend* was first performed at the Jardin des Étoiles de Terre des Hommes. It was directed by André Brassard and featured music by François Dompierre. *Demain matin, Montréal m'attend* was a musical exploration of the area of boul. Saint-Laurent (the Main), complete with its seedy bars, transvestites, gay men, and prostitutes. A second version of the play was performed in March 1972 at the Théâtre Maisonneuve de la Place des Arts; this version proved to be more popular, and was published in Montréal in 1972 by Leméac.

[Paul Lefebvre, "*Demain matin, Montréal m'attend*," in *Dictionnaire des oeuvres littéraires du Québec*, ed. Maurice Lemire et al., vol. 5 (1970–75) (Montréal: Fides, 1987), pp. 221–23.]

August * Vancouver * In a letter reprinted in the *Georgia Straight*, Huey P. Newton, Minister of Defense of the Black Panther Party, declared that blacks should overcome their "insecurities" about homosexuality and should try to form a working coalition with gay liberation groups.

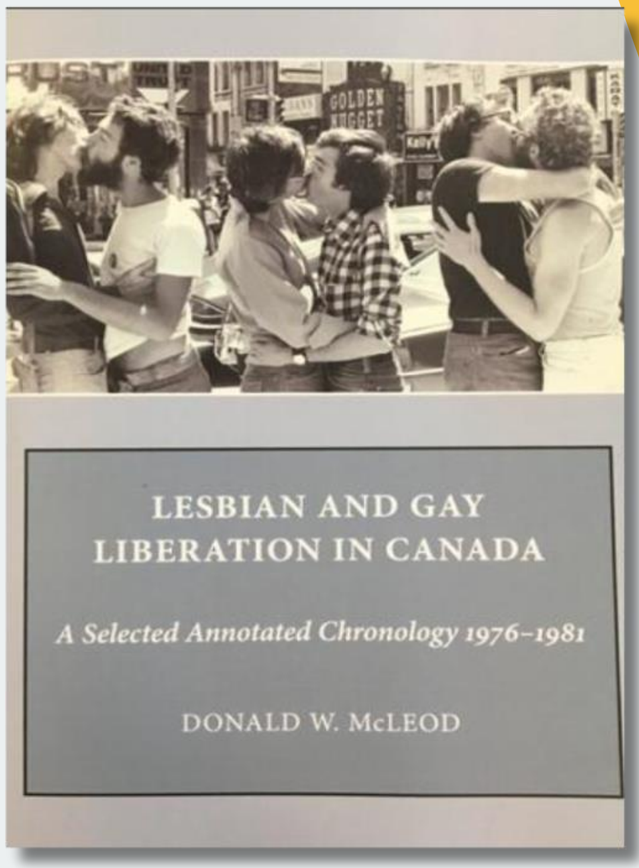
["A Letter from Huey P. Newton about the Women's Liberation and Gay Liberation Movements," *Georgia Straight*, 19–26 August 1970, p. 12.]

August * Vancouver * A group of gay men and lesbians formed the Gay Action Committee in an attempt to promote gay organizing and activities in Vancouver. The Committee's primary objectives were to form a gay social club as an alternative to the city's night clubs, to form various gay interest groups, and to establish a gay information centre and phone line. The Gay Action Committee was active for only a few months.

[Q.Q., "Page 69" (column), *Georgia Straight*, 19–26 August 1970, p. 21.]

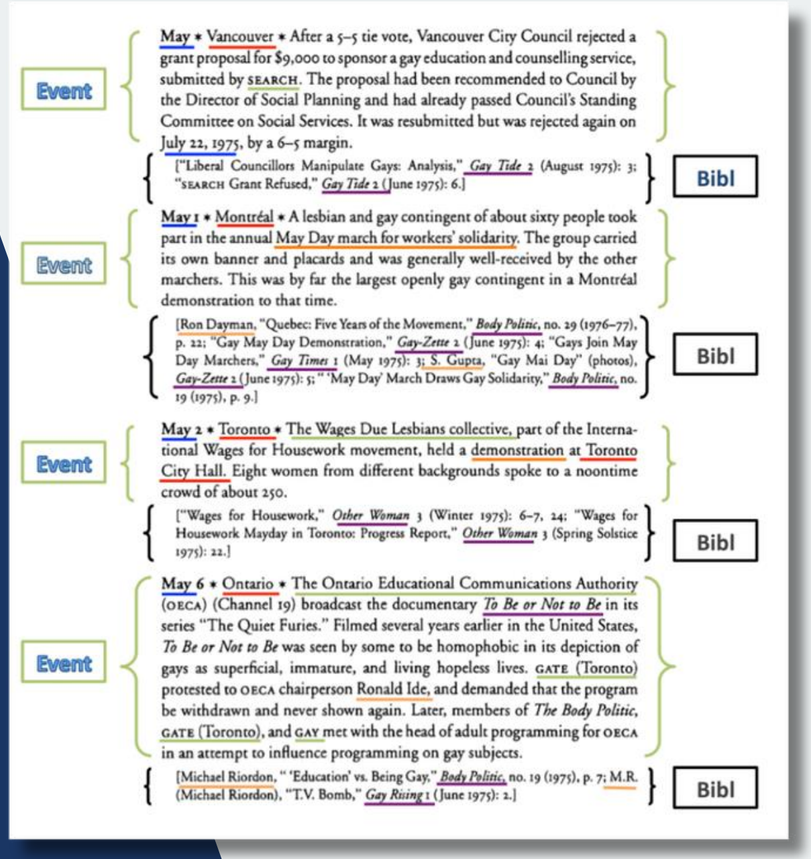
August 17 * Victoria, B.C. * Rev. Troy Perry, founder of the Metropolitan Community Church (MCC) in Los Angeles, spoke at the University of Victoria on "Metropolitan Community Church and the Homosexual Community in America." Perry's talk was at the invitation of the Victoria Youth Council and was part of their annual Youth Week program.

["Perry to Speak in Canadian Youth Week," *Advocate* (Los Angeles) 4 (2–15 September 1970): 24.]



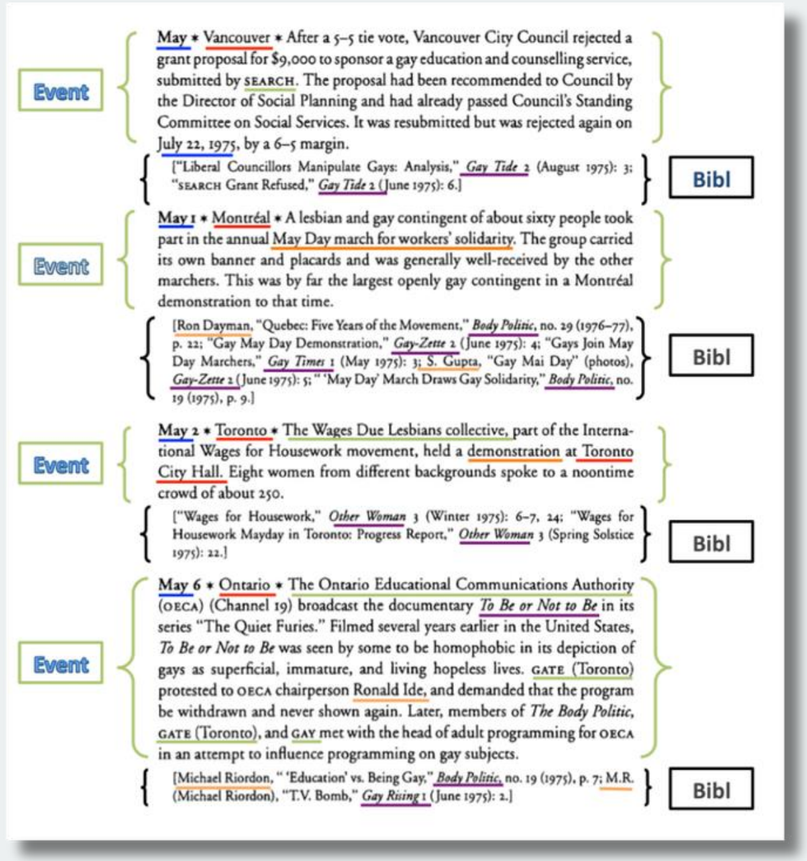
Informations recueillies en TEI

1. Noms
2. Pseudonymes
3. Dates
4. Lieux
5. Publications
6. Organisations
7. Manifestations
8. Législation



TEI text body listEvent event bibl seg

```
96979        <event xml:id="n81.425">
96980           <p>
96981              <date when="1981-12-29">December 29</date>
96982              <placeName corresp="places.xml#NY">New York</placeName>, N.Y. Professor <persName
                      sameAs="people.xml#MLYN">Michael Lynch</persName> of the <placeName
                      corresp="places.xml#UOFT">University of Toronto</placeName> was awarded the
96985              first Crompton-Noll Award in Gay Studies for his unpublished essay <title
                      level="m">"Here Is Adhesiveness: From Friendship to Homosexuality."</title> The
96987              award was sponsored by the <orgName corresp="organizations.xml#GCML">Gay Caucus
                      for the Modern Languages</orgName>, an allied organization of the <orgName
                      corresp="organizations.xml#MLA">MLA</orgName>, and was presented at the
96990              caucus's first annual awards banquet, held during the 1981 MLA Convention. </p>
96991           <bibl>
96992              <title level="a">"Lynch Receives Crompton-Noll Award,"</title>
96993              <title level="j" corresp="periodicals.xml#p411">Gay Studies Newsletter</title>
96994              (Toronto), 9, no. 1 (March 1982): <seg corresp="#n81.425.1">1–2</seg>.</bibl>
96995
96996        </event>
96997        <event xml:id="n81.426">
96998           <p>
96999              <date when="1981-12-31">December 31</date>
97000              <placeName corresp="places.xml#TOR">Toronto</placeName> A benefit dinner and dance
97001              was held for <title level="j" corresp="periodicals.xml#p226">Broadside</title> (a
97002              feminist newspaper), <orgName corresp="organizations.xml#WWP">Womynly Way
                      Productions</orgName>, and <orgName corresp="organizations.xml#MQ">Mama Quilla
                      II</orgName> at the <placeName corresp="places.xml#TDT">Toronto Dance
                      Theatre</placeName>. The event became controversial after a group of lesbian
97006              separatists took control and forced a woman-only policy on the celebration. Men
97007              were turned away at the door, including those who had already purchased tickets. </p>
97008           <bibl>
97009              <persName sameAs="people.xml#CBEA">Chris Bearchell</persName>, <title level="a"
97010                 >"The Closet and the Umbrella,"</title>
97011              <title level="j" corresp="periodicals.xml#p10">Body Politic</title>, no. 81
97012              (1982), <seg corresp="#n81.426.1">p. 7</seg>.</bibl>
97013
97014        </event>
97015        <event xml:id="n81.427">
97016           <p>(EN) <date from="1981-03" to="1981-05">Spring</date>
97017              <placeName corresp="places.xml#MON">Montreal</placeName>
97018              <persName sameAs="people.xml#FPEL">Francine Pelletier</persName> published an
97019              opinion piece in <title level="j" corresp="periodicals.xml#p592">La Vie en
                      Rose</title> titled <title level="a">"Une femme, deux femmes...beaucoup de
97021                 voyeurs"</title> reviewing an academic conference on lesbian eroticism.
97022              Pelletier criticized the event on two fronts. She said most attendees came as
97023              voyeurs rather than as people genuinely interested in women's eroticism, and that
```



TEI text body listEvent event bibl seg

```
96979        <event xml:id="n81.425">
96980           <p>
96981              <date when="1981-12-29">December 29</date>
96982              <placeName corresp="places.xml#NY">New York</placeName>, N.Y. Professor <persName
96983                 sameAs="people.xml#MLYN">Michael Lynch</persName> of the <placeName
96984                 corresp="places.xml#UOFT">University of Toronto</placeName> was awarded the
96985              first Crompton-Noll Award in Gay Studies for his unpublished essay <title
96986                 level="m">"Here Is Adhesiveness: From Friendship to Homosexuality."</title> The
96987              award was sponsored by the <orgName corresp="organizations.xml#GCML">Gay Caucus
96988                 for the Modern Languages</orgName>, an allied organization of the <orgName
96989                 corresp="organizations.xml#MLA">MLA</orgName>, and was presented at the
96990              caucus's first annual awards banquet, held during the 1981 MLA Convention. </p>
96991           <bibl>
96992              <title level="a">"Lynch Receives Crompton-Noll Award,"</title>
96993              <title level="j" corresp="periodicals.xml#p411">Gay Studies Newsletter</title>
96994              (Toronto), 9, no. 1 (March 1982): <seg corresp="#n81.425.1">1–2</seg>.</bibl>
96995
96996        </event>
96997        <event xml:id="n81.426">
96998           <p>
96999              <date when="1981-12-31">December 31</date>
97000              <placeName corresp="places.xml#TOR">Toronto</placeName> A benefit dinner and dance
97001              was held for <title level="j" corresp="periodicals.xml#p226">Broadside</title> (a
97002              feminist newspaper), <orgName corresp="organizations.xml#WWP">Womynly Way
97003                 Productions</orgName>, and <orgName corresp="organizations.xml#MQ">Mama Quilla
97004                 II</orgName> at the <placeName corresp="places.xml#TDT">Toronto Dance
97005                 Theatre</placeName>. The event became controversial after a group of lesbian
97006              separatists took control and forced a woman-only policy on the celebration. Men
97007              were turned away at the door, including those who had already purchased tickets. </p>
97008           <bibl>
97009              <persName sameAs="people.xml#CBEA">Chris Bearchell</persName>, <title level="a"
97010                 >"The Closet and the Umbrella,"</title>
97011              <title level="j" corresp="periodicals.xml#p10">Body Politic</title>, no. 81
97012              (1982), <seg corresp="#n81.426.1">p. 7</seg>.</bibl>
97013
97014        </event>
97015        <event xml:id="n81.427">
97016           <p>(EN) <date from="1981-03" to="1981-05">Spring</date>
97017              <placeName corresp="places.xml#MON">Montreal</placeName>
97018              <persName sameAs="people.xml#FPEL">Francine Pelletier</persName> published an
97019              opinion piece in <title level="j" corresp="periodicals.xml#p592">La Vie en
97020                 Rose</title> titled <title level="a">"Une femme, deux femmes...beaucoup de
97021                 voyeurs"</title> reviewing an academic conference on lesbian eroticism.
97022              Pelletier criticized the event on two fronts. She said most attendees came as
97023              voyeurs rather than as people genuinely interested in women's eroticism, and that
```

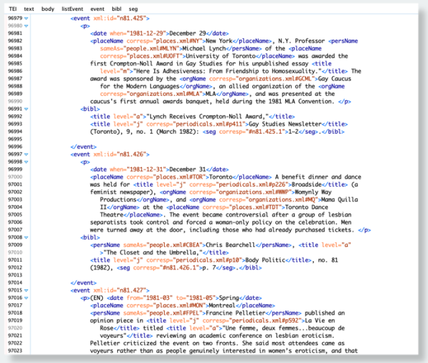
L'ANALYSE DU DOCUMENT

DEVIENT L'ENCODAGE EN TEI

QUI RÉPOND À L'OBJECTIF DU PROJET

1

L’encodage en TEI est effectué 1 fois



2

Puis transformé plusieurs fois (à l’aide du XSLT)

```
    <xsl:template match="primaryOccupation[text()]">
        <xsl:element name="{'meta'}" use-attribute-sets="occupationPrimary classDesc content_Q"/>
    </xsl:template>

    <xsl:template match="firstAppearance">
        <xsl:element name="{'meta'}" use-attribute-sets="firstAppearance classDesc content_F"/>
    </xsl:template>

    <xsl:template match="birthDate[text()]">
        <xsl:element name="{'meta'}" use-attribute-sets="birthDate classDate content_H"/>
    </xsl:template>

    <xsl:template match="deathDate[text()]">
        <xsl:element name="{'meta'}" use-attribute-sets="deathDate classDate content_K"/>
    </xsl:template>

    <!-- Calling <body>-->

    <xsl:template match="row">
        <xsl:for-each select="persName">
            <h1>Record for <xsl:value-of select="."/></h1>

            <p>Born in <xsl:value-of select="../birthYear"/> in <xsl:value-of select="../birthPlace"/><xsl:comment>item-1-
            <xsl:element name="{'img'}" use-attribute-sets="src alt"/>
        </xsl:for-each>
    </xsl:template>

    <!--Attribute-Sets for <head>-->

            <!-- Names -->

    <xsl:attribute-set name="persName">
        <xsl:attribute name="name">persName</xsl:attribute>
    </xsl:attribute-set>

    <xsl:attribute-set name="name">
        <xsl:attribute name="name">field</xsl:attribute>
    </xsl:attribute-set>

    <xsl:attribute-set name="gender">
        <xsl:attribute name="name">gender</xsl:attribute>
    </xsl:attribute-set>
```

1

L’encodage en TEI est effectué 1 fois



2

Puis transformé plusieurs fois (à l’aide du XSLT)



3

Transformé en quoi ?

1. **Statistiques**

Analyse statistique des personnes mentionnées dans LGLC :

On compte 2 596 noms répertoriés, dont 1 184 dans les comptes rendus d’événements et 1 402 dans les notices bibliographiques. Parmi ces noms, 1 076 sont uniques. 145 personnes apparaissent à la fois dans les notices bibliographiques et les comptes rendus d’événements.

Les auteurs les plus cités dans les notices bibliographiques :

1. Gary Kinsman (56) – *The Regulation of Desire: Sexuality in Canada*
2. Ron Dayman (43) – *Body Politic*, CLGA
3. Margaret Fulford (31) – T*he Canadian Women’s Movement, 1960-1990*
4. Ken Popert (27) – *Body Politic*
5. Ian Young (27) – *UTHA, Catalyst, Body Politic*

1

L’encodage en TEI est effectué 1 fois



2

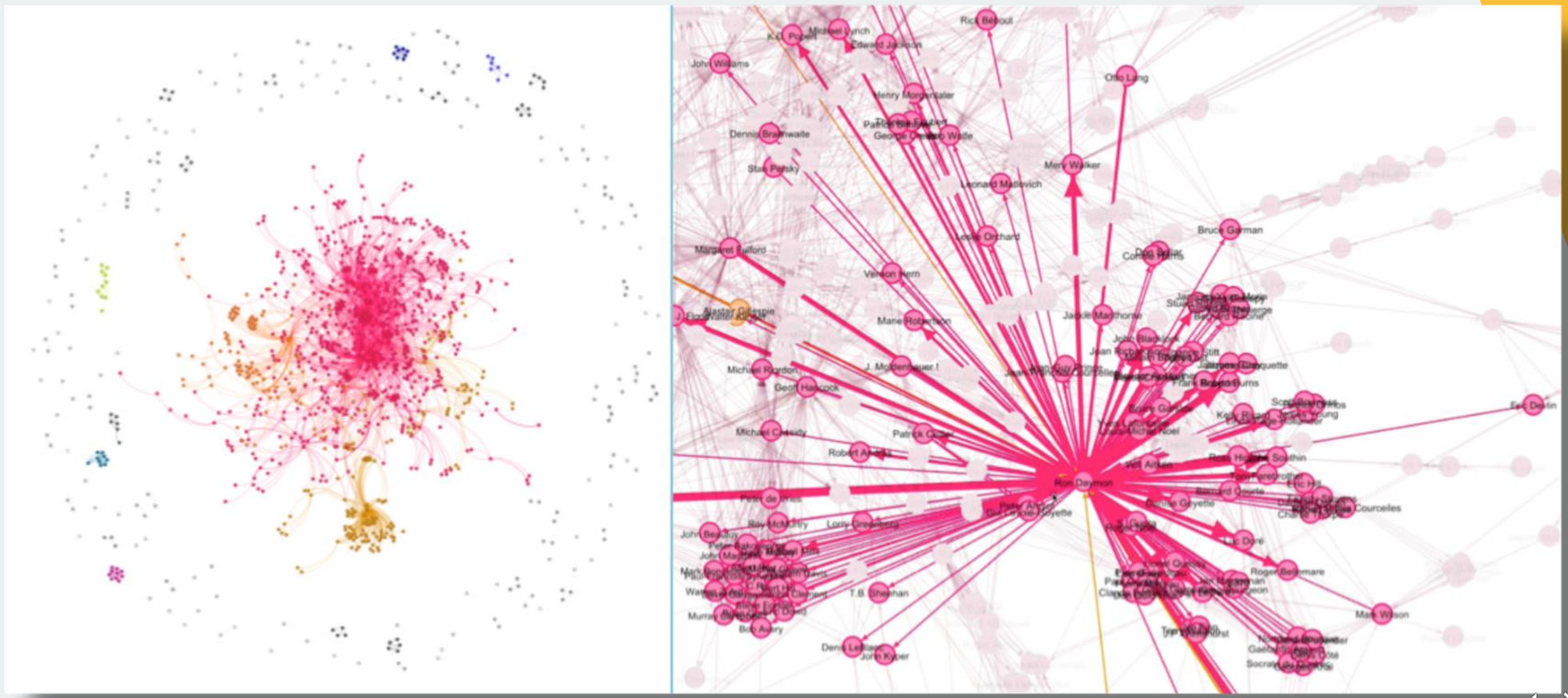
Puis transformé plusieurs fois (à l’aide du XSLT)



3

Transformé en quoi ?

1. Statistiques
2. **Visualisations**



1

L’encodage en TEI est effectué 1 fois



2

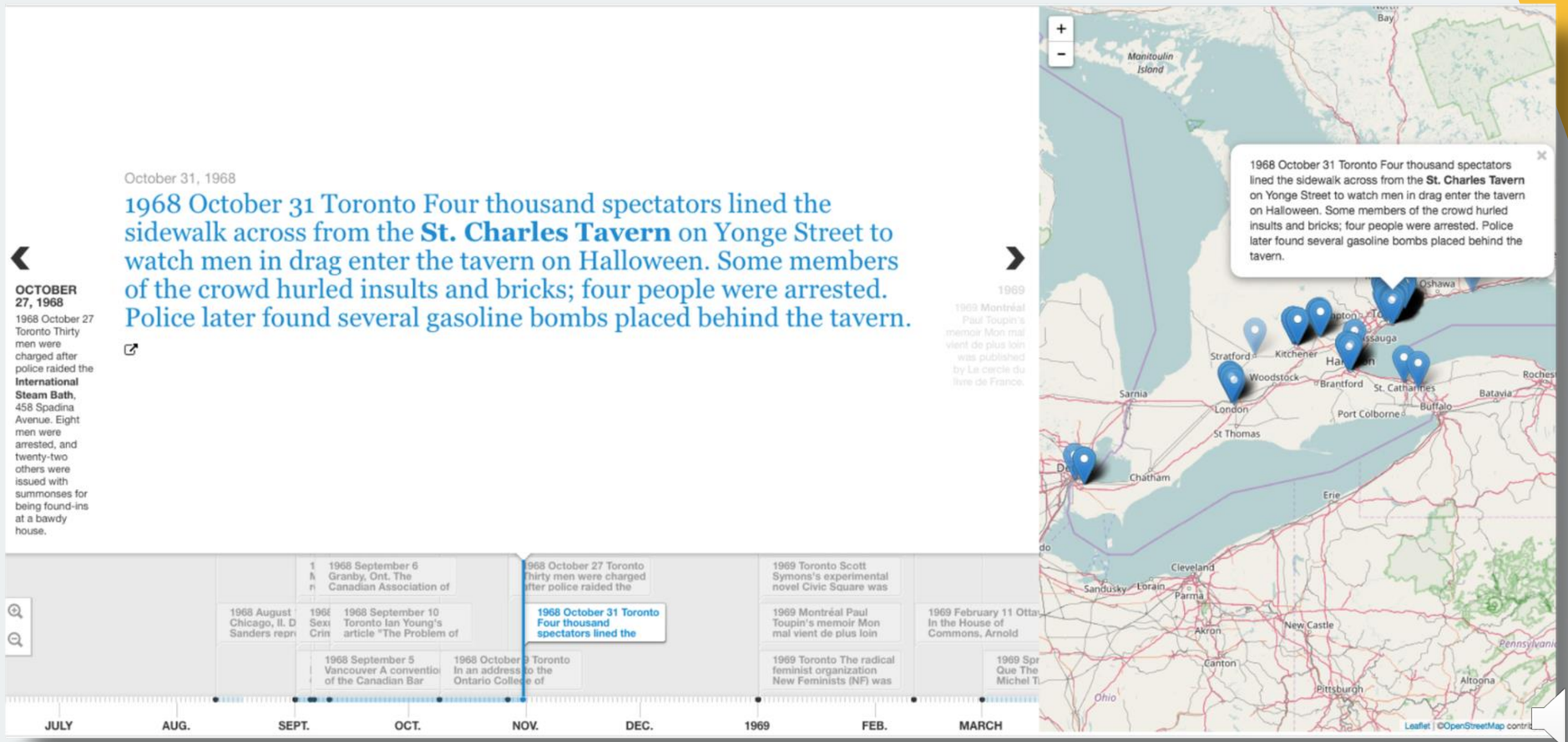
Puis transformé plusieurs fois (à l’aide du XSLT)



3

Transformé en quoi ?

1. Statistiques
2. **Visualisations**
3. **Cartes**



1

L’encodage en TEI est effectué 1 fois



2

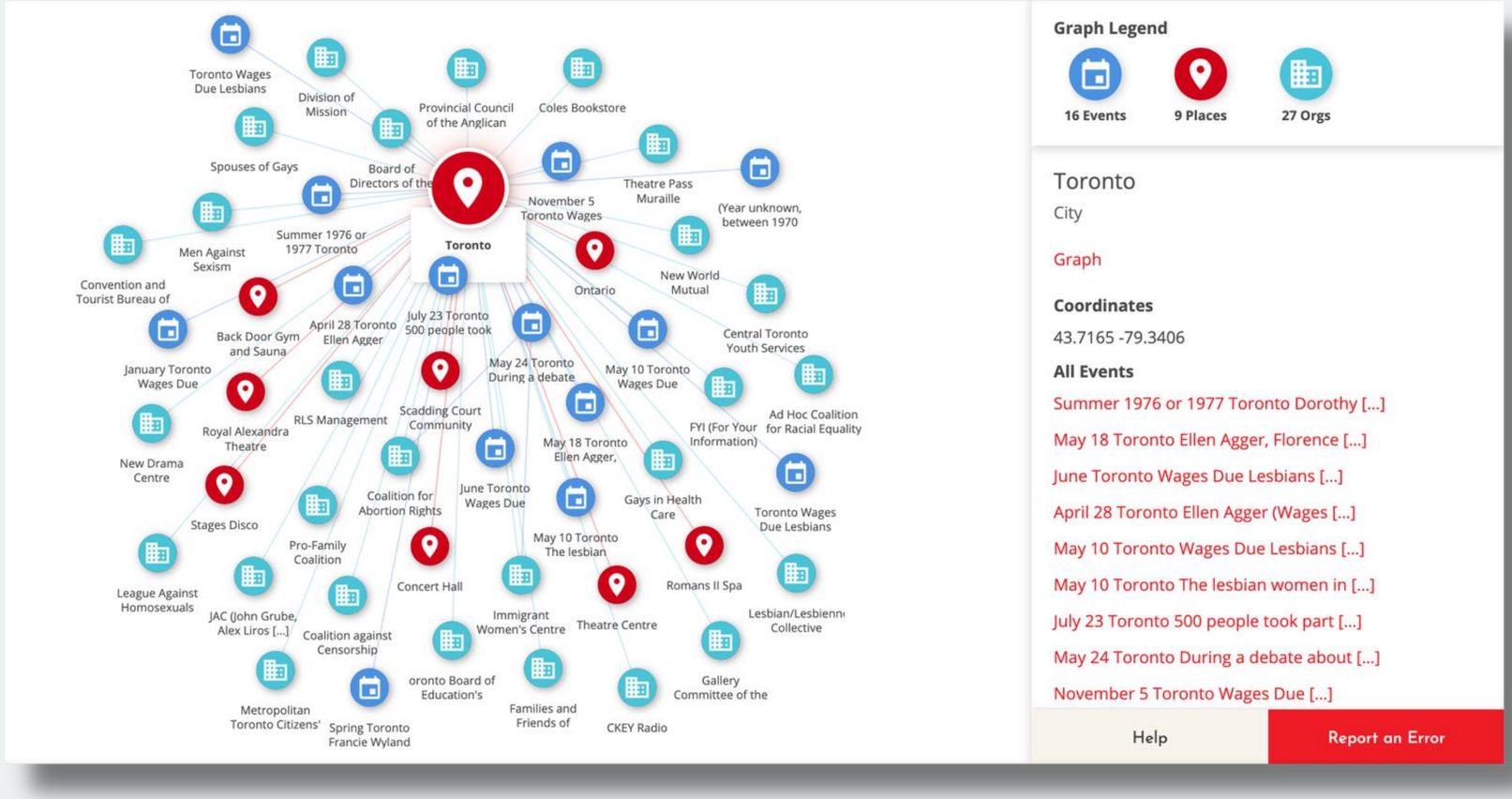
Puis transformé plusieurs fois (à l’aide du XSLT)



3

Transformé en quoi ?

1. Statistiques
2. **Visualisations**
3. Cartes
4. **Base de données interactive** (1er site web)



3

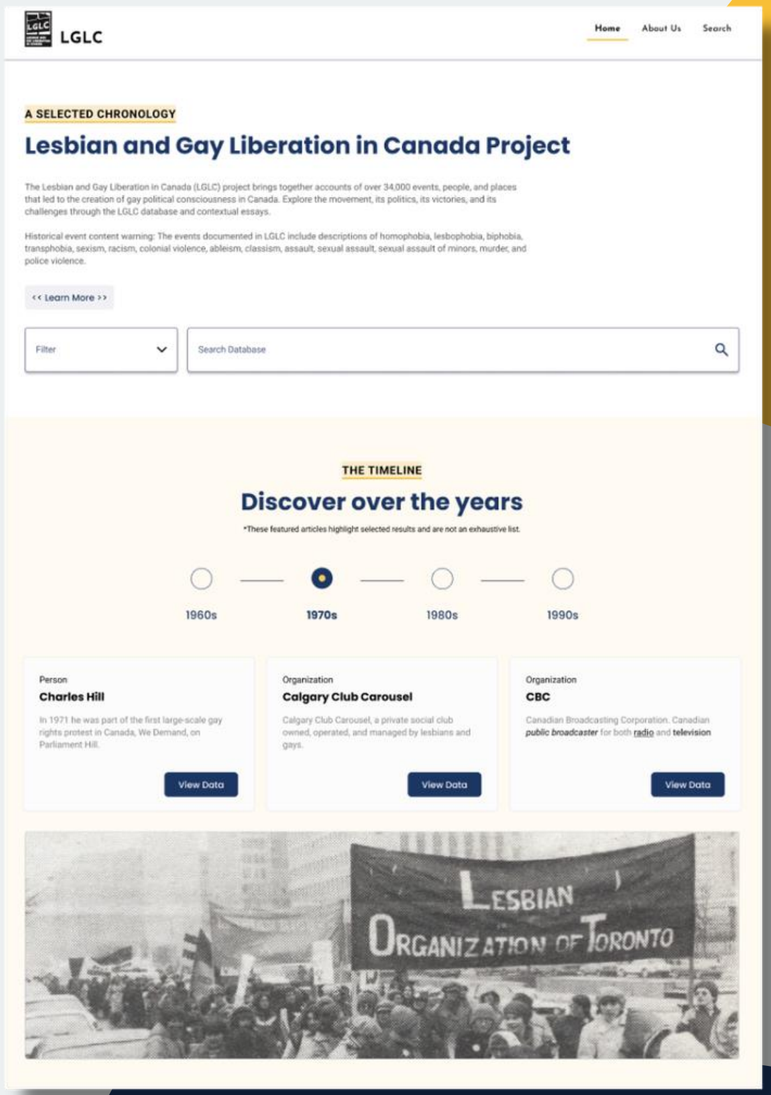
5. **Et bientôt en site web statique**

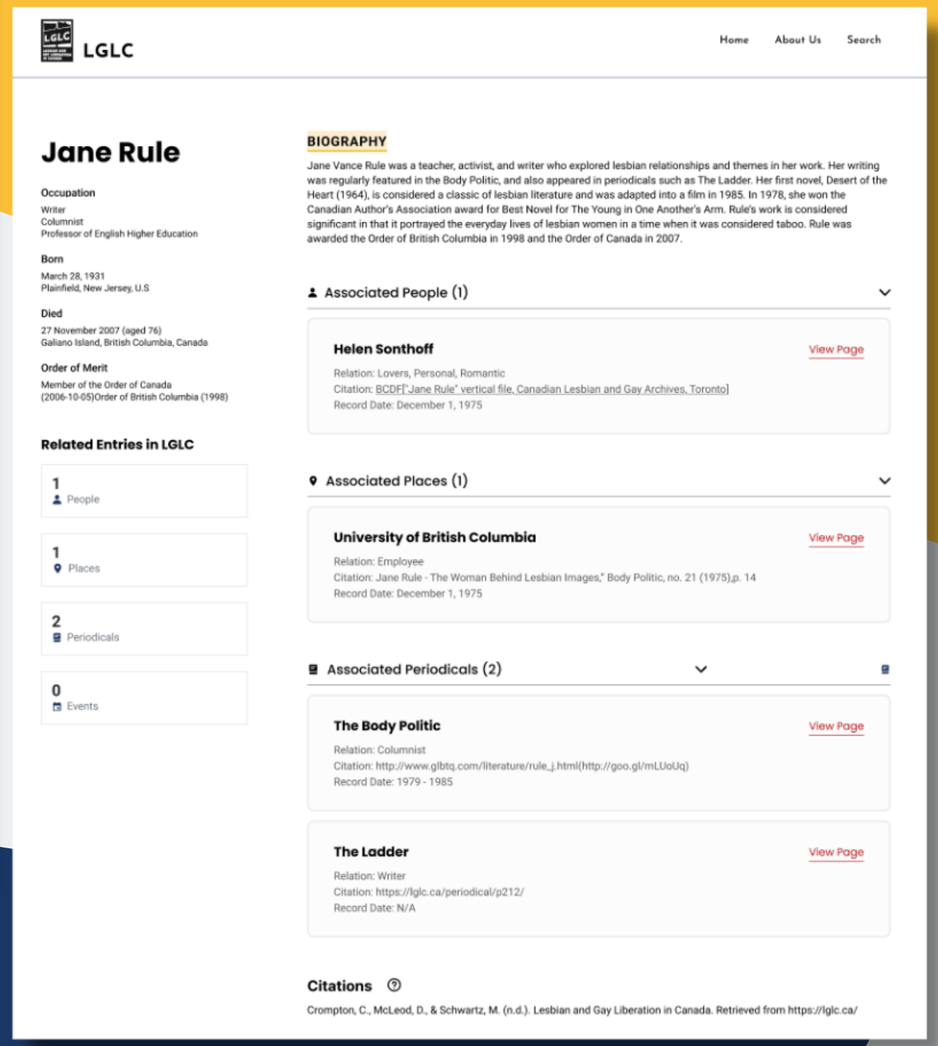
La conception web contemporaine s'appuie souvent sur

- des systèmes de gestion de contenu **automatiques**,
- des bases de données liées
- et de la **génération de pages à la demande**.

LGLC en fait partie
Mais ce format **n'est pas durable**

Pour assurer le **fonctionnement** des sites, et **l'accessibilité** des données indépendamment des systèmes dynamiques, on fait un **site web statique**





Le projet Endings

Construire des projets durables en humanités numériques

Merci

LABO DE DONNÉES EN SCIENCES HUMAINES
THE HUMANITIES DATA LAB

uOttawa
Bibliothèque
Library

Alice Defours (she/elle)

Part-time professor | Professeure à temps partiel
Senior RA at the Humanities Data Lab | Assistante de recherche Senior au Labo de Données en Sciences Humaines
PhD Candidate in Digital Humanities | Candidate au doctorat en Humanités Numériques

University of Ottawa | Université d'Ottawa
adefours@uottawa.ca

Introduction au langage XML pour l'analyse de texte

une série d'ateliers

Les mardis de 10 h 30 à 12 h (FR)
Les mardis de 13 h 30 à 15 h (EN)

Avec Alice Defours et
Constance Crompton

