HAL
open science

La plate-forme "Jibiki" dans le projet LexALP

Gilles Sérasset

To cite this version:

Gilles Sérasset. La plate-forme "Jibiki" dans le projet LexALP. Normazione, armonizzazione e pianificazione linguistica; Normierung, Harmonisierung und Sprachplanung; Normalisation, harmonisation et planification linguistique, Feb 2008, Bolzano, Italie. pp.31-47. hal-00965438

HAL Id: hal-00965438

https://hal.science/hal-00965438

Submitted on 26 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La plate-forme « jibiki » dans le projet LexALP

Gilles Sérasset

Le but du projet LexALP[1] est d'harmoniser la terminologie en aménagement du territoire et développement durable au sein de la Convention Alpine[2] afin de permettre aux états membres de communiquer et coopérer efficacement dans les quatre langues officielles (allemand, français, italien et slovène). Dans ce but, LexALP utilise la plate-forme « jibiki » pour construire une banque de termes permettant l'analyse contrastive de la terminologie spécialisée utilisée dans six systèmes légaux nationaux, trois systèmes supra-nationaux et quatre langues. Dans cet article, nous montrons comment la plate-forme Jibiki est utilisée pour implanter la banque de termes du projet LexALP.

1. Introduction

Une protection globale et homogène des Alpes ne peut être effective que si les différents états impliqués s'entendent sur des politiques communes. C'est la raison de la rédaction et de la signature de l'accord international connu sous le nom de Convention Alpine et des Protocoles qui en sont issus. Malgré les efforts conjoints de tous les états membres, une série de difficultés, aussi bien légales que linguistiques/ terminologiques sont survenues durant cette rédaction. Les aspects suivants en sont les plus significatifs:

- Les textes de la Convention cadre et de ses neuf Protocoles ont dû être rédigés sous une forme qui permettait sa ratification dans tous les états membres. Il reflète donc un compromis politique. Cela entraîne l'usage de termes généraux[3]. La Convention cadre, en particulier est un document très général.

1 LexALP : *Legal Language Harmonisation System for Environment and Spatial Planning within the Multilingual Alps* (http://www.eurac.edu/lexalp). Le projet est co-financé par le programme INTERREG IIIb « Espace Alpin ».

2 La Convention Alpine est un accord-cadre international signé par tous les états de l'arc Alpin et par l'UE (http://www.conventionalpine.org).

3 Ex : En allemand le Protocole de Protection de la nature et entretien des paysages fait référence à *geschützte oder schützenswerte « Objekte »* au lieu de *Umweltgüter*.

- La Convention cadre et les Protocoles ont été principalement rédigés dans une langue, puis traduits vers les trois autres langues officielles[4] (sauf le cas du slovène qui n'est jamais langue source). Les traductions ont souvent été rédigées avec des contraintes de dates importantes[5].

Ainsi, les problèmes légaux et terminologiques sont entremêlés. Parmi les états signataires de la Convention, on trouve des pays ayant la même langue officielle (comme l'Allemagne, l'Autriche, une partie de la Suisse et la province italienne du Tyrol du Sud). Ils peuvent néanmoins utiliser le même terme pour représenter différents concepts légaux[6], définis dans leurs traditions légales. D'autre part, le même concept peut être exprimé de manière différente selon le système légal[7]. Des termes, qui peuvent apparaître superficiellement comme traduction les uns des autres, peuvent faire référence à des notions légales différentes[8]. Enfin, certains concepts qui sont bien définis dans un système légal national peuvent être absent dans un autre[9].

Afin de résoudre concrètement ces problèmes, différentes institutions ont mis en commun leurs compétences en traduction, terminologie, loi et informatique linguistique au sein du projet LexALP. Son objectif est d'harmoniser la terminologie utilisée dans et autour de la Convention Alpine et de ses Protocoles dans les quatre langues officielles. Pour atteindre cet objectif, la terminologie de six pays signataires (Allemagne, Autriche, France, Italie, Slovénie et Suisse[10]) et comparée à celle utilisée dans la Convention Cadre et les Protocoles. Ce travail contrastif sert de base pour un groupe d'experts en charge de l'harmonisation (la détermination et la validation) des traductions allemandes, françaises, italiennes et slovènes (correspondances « une à une ») pour les rédactions et traductions futures. Ceci favorisera une communication et une compréhension plus claires et cohérentes entre les états à un niveau supra-national[11]. De plus, grâce à la disponibilité de dé-

4 Ex : La langue originale du Protocole Transports est le français alors que le Protocole de Protection de la nature et entretien des paysages a été rédigé en allemand, puis traduit.

5 Ceci conduit à l'utilisation de traductions incorrectes comme *cura del paesaggio* utilisé en italien pour traduire le terme allemand *Landschaftspflege*, au lieu du terme utilisé en Italie *tutela del paesaggio*.

6 Ex : Dans la province germanophone de Bolzano (Italie), le *Landeshauptmann* est le président du conseil de province, avec une compétence bien plus limité qu'un *Landeshauptmann* autrichien, qui est le chef de l'un des états (*Bundesland*) de la fédération autrichienne.

7 Ex : L'Union Européenne utilise le terme *zona protetta*, alors que la loi italienne exprime le même concept par *area protetta* ou *area naturale protetta*.

8 Ex : Le concept slovène *kulturna krajina* n'est pas équivalent au terme italien *paesaggio culturale*. Il s'agit en fait d'un-terme plus large, qui recouvre les concepts italiens *paesaggio culturale* et *paesaggio rurale*.

9 Le terme allemand ou autrichien *Verbandsklage* est très difficile à traduire en italien, car le concept qu'il représente n'existe pas en tant que tel. Le terme *azione colletiva* utilisé récemment pour représenter une action en justice par un groupe de citoyens ayant subit un préjudice commun semble inapproprié car le concept allemand réfère plutôt à une action déposée par une association dans l'intérêt de ses membres.

10 Le Liechtenstein et Monaco sont aussi signataires, mais leurs systèmes légaux et terminologies ne sont pas abordées dans ce projet.

11 Le Secrétariat Permanent de la Convention Alpine est représenté dans ce groupe d'experts, ainsi que plusieurs ministères et institutions publiques des pays alpins. De plus, les experts légaux et les terminologues qui ont contribué à l'élaboration de la terminologie sont présents aux réunions en tant que conseillers (cf. Chiocchetti p. 22, note 4).

finitions claires et homogènes pour chaque concept, l'implantation et l'interprétation des Protocoles seront facilitées pour les experts nationaux.

Parmi les outils développés pour atteindre ces objectifs se trouvent un corpus de textes légaux et une banque de termes (Lyding et al. 2006, Sérasset et al. 2006). Le corpus a été développé en adaptant le système *bistro* (Streiter et al. 2006, Streiter et al. 2004). La banque de termes est basée sur la plate-forme jibiki (Mangeot et al. 2003, Sérasset 2004).

Cet article détaille la manière dont la plate-forme jibiki est utilisée pour prendre en charge la structure de la banque de termes. Cette plate-forme en source ouvert fournit des services d'accès et d'édition de dictionnaires divers en minimisant les développements spécifiques.

Après une brève vue d'ensemble de la plate-forme jibiki, nous décrivons les choix faits par le consortium LexALP pour la structure et l'organisation de la banque terminologique. Enfin, nous montrons comment cette structure est décrite en utilisant le langage de description des méta-données de jibiki.

2. Jibiki, plate-forme générique de développement de dictionnaires

2. 1. Présentation

La plate-forme jibiki permet le développement collaboratif de dictionnaires multilingues. Cette plate-forme est utilisée comme base pour le site Web du projet Papillon. Elle offre différents services aux utilisateurs :

- accès à différents dictionnaires grâce à un formulaire de recherche unique et simple,
- accès à des entrées particulières des dictionnaires à travers un formulaire de recherche avancée,
- création et édition d'entrée.

Ce qui rend jibiki tout à fait unique est le fait qu'elle fournit ces services indépendamment de la structure du dictionnaire, et sans imposer de restriction particulière sur cette structure. En d'autres termes, elle peut être employée pour donner accès à n'importe quel type de dictionnaire (monolingue, bilingue ou multilingue) et permettre son édition de manière collaborative. La principale restriction d'usage est dans sa licence d'utilisation : le dictionnaire résultant doit être librement accessible en ligne.

2. 2. Architecture de la plate-forme jibiki

La plate-forme jibiki permet de produire un site web consacré au développement collaboratif de dictionnaires multilingues. Elle est organisée en une architecture 3-tiers classique : une couche présentation (responsable de l'interface avec les utilisateurs), une couche traitement (qui fournit les services) et une couche données (responsable du stockage des données persistantes).

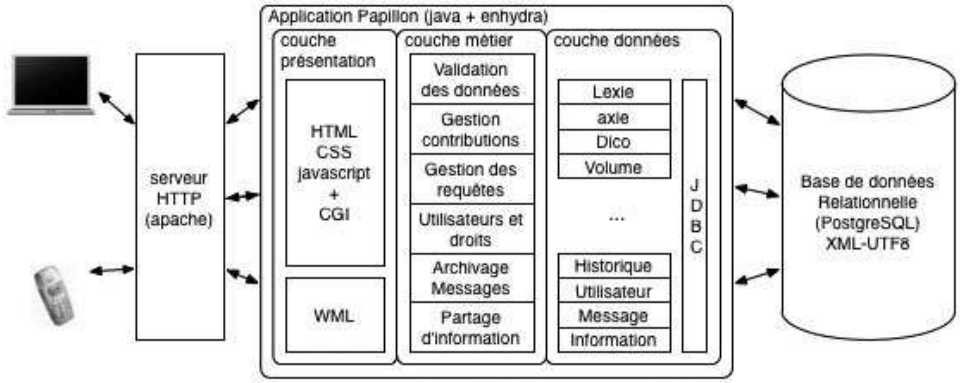

Figure 1 : architecture générale de la plate-forme jibiki

Pour utiliser jibiki, le concepteur d'un dictionnaire n'a pas à écrire de code spécifique en Java ni de pages Web dynamiques spécifiques. Les seules informations nécessaires à la plate-forme consistent en :

- une description des volumes du dictionnaire et leurs relations,
- une liaison entre la structure du dictionnaire et une structure simple d'un dictionnaire hypothétique (appelé CDM, *Common Dictionary Markup* (Mangeot 2001)),
- une définition de la structure XML de chaque volume envisagé (à l'aide d'un schéma XSD),
- une interface d'édition spécifique basée sur un formulaire standard XHTML (optionnelle, cette interface est générée automatiquement à partir du schéma XSD, mais peut être modifiée pour améliorer l'interaction avec l'utilisateur).

3. Structure et organisation de la banque terminologique

3. 1. Présentation

L'objectif du projet LexALP est d'harmoniser la terminologie utilisée dans et autour de la Convention Alpine et de ses protocoles dans les quatre langues officielles, l'allemand, le français, l'italien et le slovène. Dans ce but, la terminologie légale des états signataires est comparée à la terminologie spécialisée de la Convention Alpine. Les termes de la Convention Alpine sont aussi comparés à la terminologie européenne et aux termes définis et utilisés dans les lois internationales. Les entrées terminologiques sont élaborées et saisies dans la banque de termes LexALP pour chaque système légal. Le point de départ est toujours la terminologie de la Convention Alpine : le sens de chaque concept de la Convention est analysé en regard de ses équivalents dans les systèmes légaux nationaux et internationaux de référence, afin de mettre en lumière les différences conceptuelles ou terminologiques. Ces différences sont explicitées dans des notes spécifiques portées par le terme de la Convention Alpine.

Comme le projet traite de termes juridiques, les problèmes contrastifs sont la norme. Aussi, nous avons décidé d'adopter une organisation lexicale (macrostructure) basée sur la notion d'acceptions, telles qu'elles sont définies dans le projet de dictionnaire multilingue « Papillon » (Sérasset 1994a, Sérasset 1994b). Dans cette organisation, chaque entrée représente un concept approprié à une loi nationale ou à une convention internationale considérée. Ces entrées sont ensuite reliées au travers d'une « acception » représentant leur équivalence fonctionnelle. Lors de la construction de ces entrées, on commence toujours par les termes de la Convention Alpine.

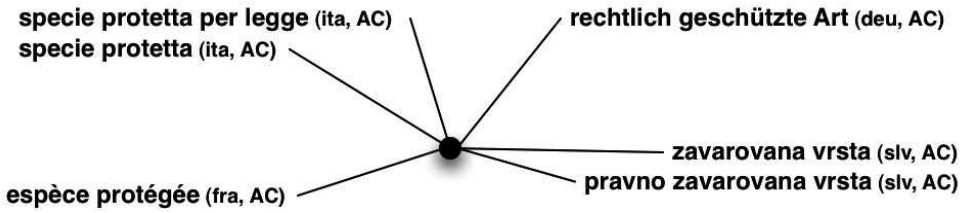

Figure 2 : un concept de la Convention Alpine dans les quatre langues

L'exemple donné dans la Figure 2 montre un concept défini dans la Convention Alpine. Ce concept a la même définition dans les quatre langues de la Convention Alpine, mais il est exprimé par différentes dénominations.

Tous les termes appartenant à différents systèmes légaux et considérés comme représentant le même concept sont reliés manuellement au même noeud central (nommé *axie*). La Figure 3 illustre le cas où des termes de systèmes légaux nationaux et supra-nationaux

différents dénotent de concepts fonctionnellement équivalents (s'il appartiennent à des systèmes différents) ou pleinement équivalents (s'ils appartiennent au même système). Les différences conceptuelles sont explicitées dans des notes présentes dans les entrées terminologiques respectives[12]. Ainsi, chaque terme est décrit de manière indépendante (car ils ne partagent ni définitions, ni contextes d'usage).

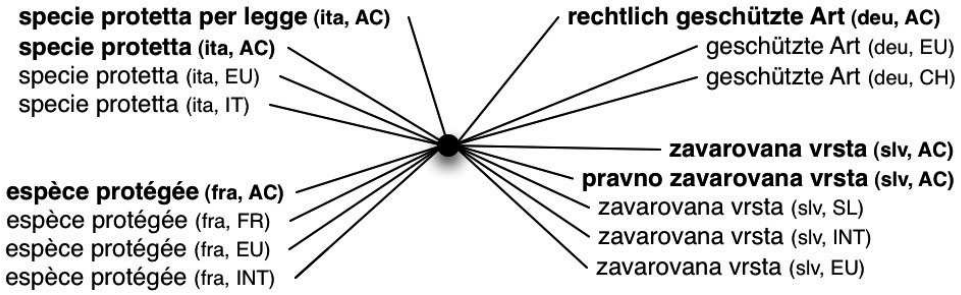

Figure 3 : un ensemble de termes fonctionnellement ou totalement équivalents

Lorsque des termes de différentes langues/ systèmes légaux ne partagent pas la même signification, par exemple dans les cas d'homonymie, hyperonymie ou d'équivalence partielle, etc., chaque terme est lié à une *axie* qui lui est propre. Ces axies peuvent à leur tour être liées à d'autres *axies* (par exemple à l'*axie* des termes de la Convention Alpine) afin de représenter l'absence d'équivalence complète ou fonctionnelle entre systèmes légaux dans des cas spécifiques. La Figure 4 illustre un tel cas où les termes allemands ont été considérés comme partiellement équivalents à ceux de la Convention Alpine.

Cette organisation permet de retrouver des relations indirectes entre systèmes légaux nationaux, au travers de leurs relations aux concepts de la Convention Alpine. Néanmoins ce type de relation d'est pas l'un des objectifs principaux du projet LexALP et demanderait plus d'analyses contrastives.

12 Par exemple, les termes *besonders geschützte Art* et *streng geschützte Art* présentés dans l'exemple développé contiennent une note expliquant qu'il n'y a pas, dans le droit allemand de notion générale d'espèces protégées, mais qu'une subdivision claire est faite entre deux sous catégories spécifiques. L'évolution de ces catégories est brièvement expliquée et permet de mieux comprendre que le terme allemand *geschützte Art* utilisé dans la Convention Alpine est parfaitement acceptable linguistiquement car il est plus général que les termes définis au noveau national, mais qu'il doit être clairement défini pour être compris d'un expert légal allemand.

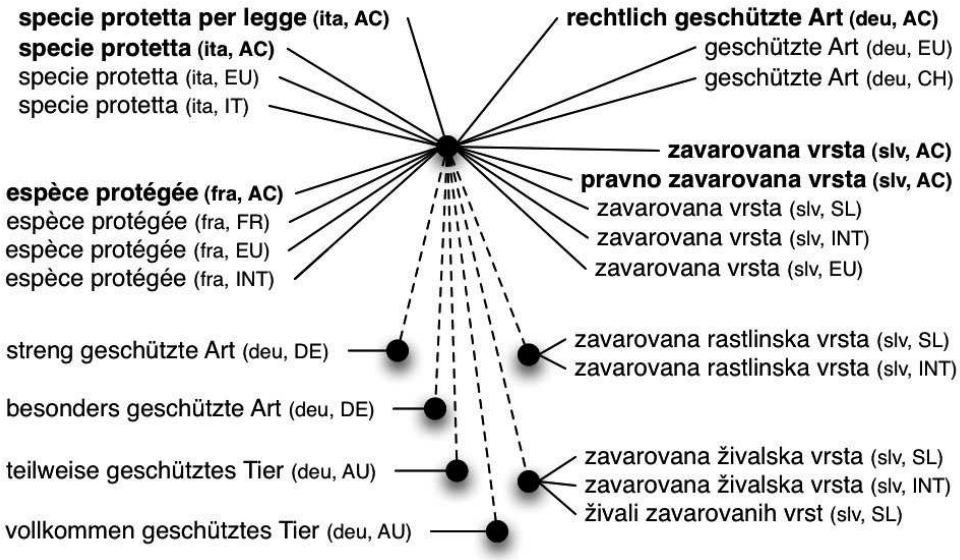

Figure 4 : un concept de la Convention Alpine, avec ses concepts reliés dans d'autres langues/ systèmes légaux

3. 2. Micro- et macro-structures

La banque terminologique du projet LexALP consiste en quatre volumes (pour l'allemand, le français, l'italien et le slovène) contenant les descriptions des termes. Les liens de traduction sont établis par un volume central d'acceptions interlingues (*axies*). Les Figures 2, 3 et 4 montrent des exemples de termes extraits des quatre volumes monolingues et d'axies extraites du volume central. Pour le projet LexALP, tous les volumes partagent la même microstructure. Cette structure est stockée sous forme XML.

La Figure 5 montre la structure xml du terme « espèce protégée », comme défini dans la Convention Alpine. L'entrée est associée à un identificateur unique employé pour établir des relations entre les entrées des volumes. Chaque entrée appartient à un (et seulement un) système légal. Le terme présenté en exemple appartient au système de la Convention Alpine. L'ensemble des systèmes légaux connus inclut naturellement les pays appartenant à l'espace alpin (Autriche, France, Allemagne, Italie, Liechtenstein, Slovénie, Suisse) mais également les traités ou conventions internationaux. L'entrée porte aussi des informations sur son statut (harmonisé ou rejeté). Le terme lui-même et sa catégorie sont également donnés, ainsi que le domaine général auquel le terme appartient, et quelques notes d'utilisation. Dans ces notes, l'attribut *geographical-code* tient compte de la discrimination entre les termes définis dans les lois nationales (ou fédérales) et les termes définis dans des lois régionales (certains des pays impliqués permettent plusieurs niveaux d'autorité

législative). Ensuite, une définition (avec sa source) et un contexte peuvent être associé au terme. La définition et le contexte doivent être extraits de textes légaux identifiés pour un domaine donné. Enfin, une note et une note d'harmonisation peuvent apparaître si cela s'avère nécessaire pour préciser les différences contrastives observées ou pour aider au travail d'harmonisation lui-même.

```
<entry id="fra.espèces_protégées.1038.e" lang="fra"
        legalSystem="AC" status="UNKNOWN">
    <term>espèce protégée</term>
    <grammar>n.f.</grammar>
    <phraseological-unit>false</phraseological-unit>
    <domain>1.4</domain>
    <domain>4.4</domain>
    <usage frequency="common" geographical-code="AC" technical="false"/>
    <definition>
      <text>Espèce menacée ou vulnérable qui est protégée de
                l'extinction par des mesures préventives.</text>
      <source date="2006/08/31" sourceType="Url">
          http://glossary.eea.europa.eu/EEAGlossary/P/protected_species
      </source>
      <source date="2006/09/15" sourceType="Author">Randier (trad)</source>
    </definition>
    <context>
      <text>[L'] [é]tat des inventaires des espèces animales sauvages
      [comprend l'élaboration de] [l]istes des espèces protégées [.]</text>
      <source documentId="3491828" >Prot. PNEP, ann. I, art. 1.1.2.
      </source>
    </context>
    <note>...</note>
    <harmonisation-note>...</harmonisation-note>
</entry>
```

Figure 5 : structure XML du terme espèce protégée

```
<axie  id="axi..1432634.e">
    <termref idref="ita.specie_protette_per_legge.1037.e" lang="ita"/>
    <termref idref="fra.espèces_protégées.1038.e" lang="fra"/>
    <termref idref="deu.rechtlich_geschützte_Arten.1039.e" lang="deu"   />
    <termref idref="slv.pravno_zavarovana_vrsta.1040.e" lang="slv"/>
    <termref idref="ita.specie_protetta.1200075.e" lang="ita"/>
    <termref idref="slv.zavarovana_vrsta.1256879.e" lang="slv"/>
    ...
    <axieref idref="axi..2184588.e"/>
    <axieref idref="axi..2184615.e"/>
    ...
</axie>
```

Figure 6 : structure XML d'une acception interlingue (axie)

La Figure 6 montre la structure XML de l'acception interlingue illustrée par la Figure 3. Chaque acception interlingue peut être reliée à plusieurs entrées grâce à l'élément *termref* et à d'autres acceptions interlingues grâce à l'élément *axieref*.

4. La banque de termes LexALP

4. 1. Présentation

La construction d'une telle banque terminologique peut uniquement être envisagée comme un travail collaboratif entre des lexicographes, des traducteurs et des juristes de tous les pays impliqués. Par conséquent, le consortium LexALP utilise un système d'information centralisé pour recueillir toutes les données textuelles et lexicographiques. Ce système d'information est organisé en deux parties. La première est consacrée à la gestion du corpus. Celle-ci permet aux utilisateurs de télécharger les textes légaux qui serviront à amorcer le travail lexicographique : extraction de termes candidats, de définitions, etc. La seconde partie est consacrée au travail lexicographique proprement dit. Elle a été développée avec la plate-forme jibiki (décrite dans la section 2). Dans la suite, nous présentons les fonctionnalités du système d'information LexALP et les méta-données associées ayant permis son implémentation par jibiki.

4. 2. Recherche de termes

Le premier service consiste à rechercher dans le dictionnaire en cours de développement. Cela consiste en deux interfaces de recherche (recherche simple et avancée, illustrées par les Figures 7 et 8) et une interface de présentation du résultat.

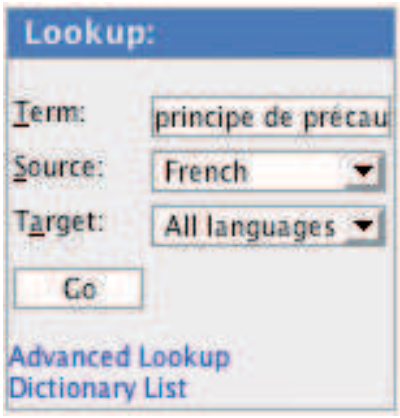

Figure 7 : interface de recherche simple

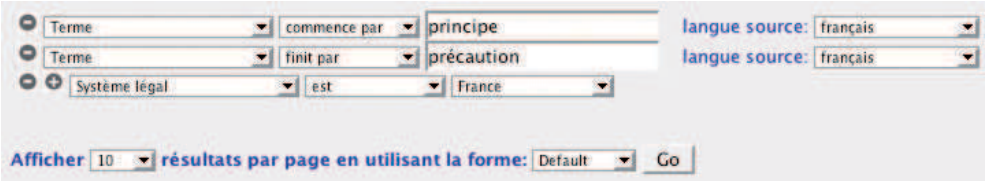

Figure 8 : interface de recherche avancée

Pour effectuer une recherche, l'utilisateur indique un mot-clé relatif à l'entrée souhaitée (mot, catégorie, définition, etc.) et une langue dans laquelle la recherche doit être faite. La recherche peut alors être réalisée si et seulement si :

- le système sait dans quel volume la recherche doit être effectuée,
- le système sait où, dans l'entrée, le mot-clé doit être trouvé,
- le système peut produire une présentation pour les structures XML retrouvées.

Cependant, comme la plate-forme jibiki est entièrement indépendante de la structure fondamentale du dictionnaire (ce qui la rend fortement adaptable), le résultat attendu ne peut être affiché que si des informations additionnelles sont fournies au système pour décrire le dictionnaire et ses volumes. Ces informations, les descripteurs, sont un ensemble de méta-données, écrites dans le langage DML (*Dictionary Metadata Language*) et associée à un espace de noms XML spécifique défini dans (Mangeot 2001).

```
<dictionary-metadata
  d:category="multilingual" d:fullname="LexALP multilingual Term Base"
  d:name="LexALP" d:owner="LexALP consortium" d:type="pivot">
    <languages>
        <source-language d:lang="deu"/>
        <source-language d:lang="fra"/>
        <target-language d:lang="deu"/>
        <target-language d:lang="fra"/>
        [...]</languages>
    <volumes>
        <volume-metadata-ref name="LexALP_fra" source-language="fra"
            xlink:href="LexALP_fra-metadata.xml"/>
        <volume-metadata-ref name="LexALP_deu" source-language="deu"
            xlink:href="LexALP_deu-metadata.xml"/>
            [...]
        <volume-metadata-ref name="LexALP_axi" source-language="axi"
            xlink:href="LexALP_axi-metadata.xml"/>
    </volumes>
</dictionary-metadata>
```

Figure 9 : extrait du descripteur du dictionnaire

La Figure 9 donne un extrait du descripteur du dictionnaire LexALP. La méta-donnée identifie d'abord le dictionnaire en lui donnant un nom et un type. Dans cet exemple,

le dictionnaire est un dictionnaire multilingue basé sur un pivot (DML peut également définir des dictionnaires monolingues et bilingues). Le descripteur définit également l'ensemble des langues sources et cibles. Enfin, le dictionnaire est défini comme un ensemble de volumes, chaque volume étant décrit dans un autre descripteur (fichier de méta-donnée). Le dictionnaire LexALP étant un dictionnaire pivot, il contient un volume pour une langue artificielle (*axi*), qui est le volume pivot.

```
<volume-metadata
  dbname="lexalpfra" dictname="LexALP" name="LexALP_fra"
  source-language="fra">
    <cdm-elements>
        <cdm-entry-id index="true" xpath="/volume/entry/@id"/>
        <cdm-headword index="true" xpath="/volume/entry/term/text()"/>
        <cdm-pos index="true" xpath="/volume/entry/grammar/text()"/>
        [...]
    </cdm-elements>
    <xmlschema-ref xlink:href="lexalp.xsd"/>
    <template-entry-ref xlink:href="lexalp_fra-template.xml"/>
    <template-interface-ref xlink:href="lexalp-interface.xhtml"/>
    <xsl-stylesheet name="LexALP" default="true" xlink:href="LexALP.xsl"/>
    <xsl-stylesheet name="short-list" xlink:href="short-list.xsl"/>
</volume-metadata>
```

Figure 10 : extrait d'un descripteur de volume

La Figure 10 montre un extrait de la description du volume français du dictionnaire LexALP. Après l'indication du nom du dictionnaire, le descripteur fournit un ensemble de *cdm-elements*. Ces éléments sont utilisés pour identifier les éléments standard d'un dictionnaire hypothétique[13]. Par exemple, le descripteur précise que la graphie d'une entrée, l'élément standard *cdm-headword*, est trouvée en appliquant un endroit spécifique de chaque entrée du dictionnaire (défini par un *xpath*[14]). Avec cet ensemble de méta-données, le système sait que :

- les requêtes sur le français devront être dirigées vers le volume *LexALP_fra*,
- les requêtes sur la graphie seront dirigées sur l'élément *term* de l'entrée.

Ainsi, le système peut facilement exécuter une requête et fournir les entrées XML désirées. La seule étape restante est de produire une présentation des entrées trouvées lisible par l'utilisateur. Ceci est réalisé par une feuille de style XSL[15]. Cette feuille de style est associée au dictionnaire ou au volume. Dans l'exemple donné, le concepteur du dictionnaire a fourni deux interfaces de présentation appelées LexALP (la présentation par défaut) et

13 L'ensemble des éléments connus du système, le *Common Dictionary Markup*, a été défini en étudiant des éléments fréquents dans les structures des principaux dictionnaires disponibles.

14 Un *xpath* est une expression dans une syntaxe standard qui permet de désigner une portion d'un document XML.

15 XSL est un langage permettant de spécifier une transformation d'un document XML en une autre forme (XML ou non).

short-list. La présentation par défaut est illustrée Figure 11. Ce mécanisme permet de définir des présentations différentes en fonction des sorties désirées, par exemple, au format XHTML (pour la visualisation en ligne) ou au format pdf (pour l'exportation et l'impression du dictionnaire).

Figure 11 : le résultat d'une recherche, mis en forme grâce à la feuille de style fournie

4. 3. Édition de dictionnaires

Le deuxième service fourni par la plate-forme jibiki doit permettre aux lexicographes de développer en collaboration le dictionnaire envisagé (jibiki fédère, sur la même plate-forme la construction et la diffusion d'un dictionnaire). Comme vu précédemment, jibiki peut être employé pour éditer n'importe quel dictionnaire. Par conséquent, comme pour la recherche de terme, le système a besoin d'informations additionnelles afin de fonctionner :

- une définition de la structure XML des entrées à l'aide d'un schéma XSD[16],
- un modèle XML décrivant une structure d'entrée vide,
- un formulaire XHTML utilisé pour éditer la structure d'une entrée du dictionnaire (optionnel, ce formulaire est généré automatiquement à partir de la définition XSD, mais il doit être modifié pour améliorer l'interaction avec l'utilisateur lors de l'édition).

16 Le langage XSD permet de définir les éléments d'une structure XML et ses règles de bonne formation.

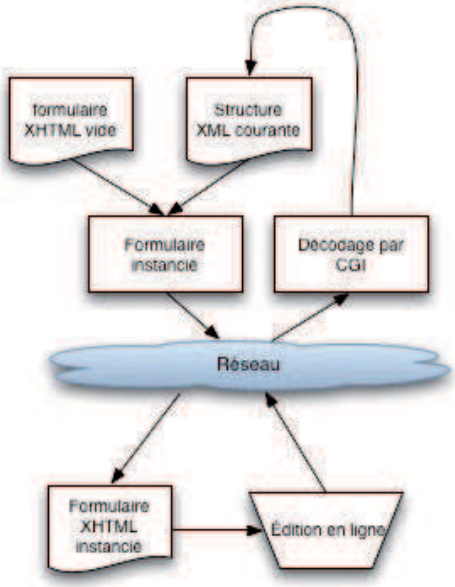

Figure 12 : flux de prise en charge de l'édition d'une entrée

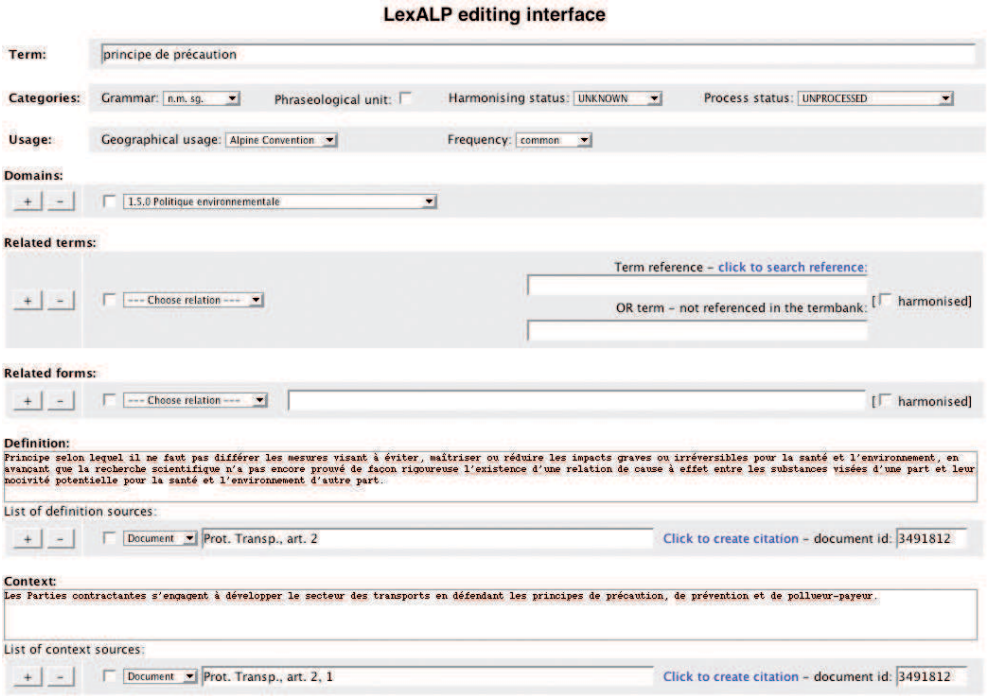

Figure 13 : interface d'édition d'un terme de LexALP

Quand ces informations sont connues, la plate-forme jibiki produit une page Web spécifique pour éditer l'entrée du dictionnaire. Comme représentée sur la Figure 12, la structure de l'entrée est projetée dans le formulaire XHTML vide. Ce formulaire instancié est une page Web standard, affichée par le navigateur client (la Figure 13 en donne un exemple). Après édition manuelle, le formulaire résultant est envoyé à la plate-forme jibiki qui décode les données et modifie la structure de l'entrée en conséquence. Le processus est réitéré aussi longtemps que nécessaire.

Après chaque mise à jour, la structure XML résultante est stockée dans la base de données de plate-forme. Cependant, elle n'est pas mise à la disposition des autres utilisateurs jusqu'à ce qu'elle soit marquée comme terminée par le contributeur. Si le contributeur quitte l'édition sans sauver l'entrée, il pourra la rechercher et la terminer plus tard.

À chaque étape de l'édition d'une entrée, l'état précédent est sauvé et le contributeur (ou l'administrateur du dictionnaire) peut passer en revue l'historique des changements et revenir à une version précédente de l'entrée si nécessaire.

5. Conclusion

Dans cet article, nous donnons quelques détails sur le fonctionnement de la plate-forme Jibiki permettant la diffusion et l'édition en ligne d'un dictionnaire, quelle que soit sa structure (monolingue, bilingue – orienté ou non – multilingue – multi-bilingue ou à pivot).

Initialement développée pour l'édition du dictionnaire Papillon[17], la plate-forme jibiki s'est avérée utile pour le développement d'autres dictionnaires très différents. Elle est actuellement employée pour le développement du projet GDEF[18] (Grand Dictionnaire Estonien-Français). Cet article montre également l'utilisation de la plate-forme dans le projet LexALP pour le développement d'une banque terminologique de termes juridiques sur l'environnement et l'aménagement du territoire de l'espace alpin.

Adapter la plate-forme jibiki à un nouveau dictionnaire exige la définition de plusieurs méta-données, prenant la forme de plusieurs fichiers XML. Bien que non-triviales, ces définitions exigent peu de compétence informatique. Par conséquent cette adaptation peut être faite par des lexicographes expérimentés. De plus, quand la microstructure du dictionnaire doit évoluer, cette évolution n'exige pas de programmation. Par conséquent cette plate-forme donne aux lexicographes une grande liberté dans leurs décisions.

Un autre aspect positif concernant jibiki est qu'il intègre des services de diffusion et d'édition sur la même plate-forme. Ceci permet une collaboration plus étroite entre les lexicographes et les utilisateurs, et permet également la participation des utilisateurs mo-

17 http://www.papillon-dictionary.org

18 http://estfra.ee

tivés au processus d'édition. La plate-forme jibiki est utilisable gratuitement à condition que les données résultantes du dictionnaire soient librement disponibles en ligne.

Le projet LexALP nous a permis de redéfinir et de rendre plus efficace l'implantation actuelle de jibiki. Il a aussi permis de définir un ensemble de contraintes à prendre en compte pour une prochaine version. De plus, il a été l'occasion de tester l'utilisation d'une organisation interlingue par acceptions dans le cadre d'un projet de terminologie multilingue. Je tiens ici à remercier l'ensemble des terminologues et des traducteurs qui ont fait l'effort de travailler sur des structures différentes de celles dont ils avaient déjà la pratique et qui m'ont apporté leur point de vue sur cette organisation. Évidemment, tout ceci n'aurait pu être possible sans le soutien du programme INTERREG IIIB « Espace Alpin » qui a co-financé ce projet.

Références

Mangeot, Mathieu/ Sérasset, Gilles/ Lafourcade, Mathieu (2003) : Construction collaborative d'une base lexicale multilingue, le projet Papillon. *TAL*, 44:2, 151–176. http://jibiki.univ-savoie.fr/~mangeot/Publications/TAL03_MM-GS-ML.pdf

Mangeot, Mathieu (2001) : *Environnements centralisés et distribués pour lexicographes et lexicologues en contexte multilingue.* Thèse de doctorat, spécialité informatique. Université Joseph Fourier Grenoble I, septembre.

Lyding, Verena/ Chiocchetti, Elena/ Sérasset, Gilles/ Brunet-Manquat, Francis (2006) : The LexALP information system: Term bank and corpus for multilingual legal terminology consolidated. In COLING-ACL 2006 *Workshop on Multilingual Language Resources and Interoperability*. Sydney, Australie, 25-31. http://acl.ldc.upenn.edu/W/W06/W06-1004.pdf

Sérasset, Gilles (1994a) : An interlingual lexical organisation based on acceptions, from the parax mock-up to the nadia system. In *ICLA-94*. Penang, Malaisie, 21-33.

Sérasset, Gilles (1994b) : Interlingual lexical organisation for multilingual lexical databases in nadia. In Nagao, Makoto (Ed.) : *COLING-94*, vol. 1. Kyoto, Japon, 278-282.

Sérasset, Gilles (2004) : A generic collaborative platform for multilingual lexical database development. In Sérasset, Gilles (Ed.) : COLING 2004 *Workshop on Multilingual Linguistic Resources.* Genève, Suisse, 73-79.

Sérasset, Gilles/ Brunet-Manquat, Francis/ Chiocchetti, Elena (2006) : Multilingual legal terminology on the jibiki platform: The LexALP project. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, COLING-ACL 2006. Sydney, Australie, 937-944. http://www.mt-archive.info/Coling-ACL-2006-Serasset.pdf

Streiter, Oliver/ Voltmer, Leonhard/ Ties, Isabella/ Ralli, Natascia (2004) : BISTRO, the online platform for terminology management: structuring terminology without entry structures. In Temmerman, Rita & Knops, Uus (Ed.) : *The translation of domain specific languages and multilingual terminology*, n° 3, Linguistica Antverpiensia New Series. Hoger Instituut voor Vertalers en Tolken, Hogeschool Antwerpen.

Streiter, Oliver/ Voltmer, Leonhard/ Ties, Isabella/ Ralli, Natascia/ Lyding, Verena (2006) : BISTRO: Data Structure, Term Tools and Interface. In *Terminology Science and Research (TSR)* n° 16. Vienne : International Institute for Terminology Research (IITF).