

Le document électronique dans le cadre juridique : une précieuse occasion pour la compréhension du droit¹

Marina Pietrangelo (1) and Maria Angela Biasiotti (2)

(1) Institut de Théorie et des Techniques de l'Information Juridique , Italie

(2) Conseil National de Recherches (CNR) - Institut de Théorie et des Techniques de l'Information Juridique (ITTIG), Italie

Résumé : Cet essai représente une contribution au débat scientifique sur l'utilisation de documents électroniques dans le contexte juridique, notamment en ce qui concerne les termes de connaissance du droit. En particulier, une description synthétique du cadre juridique général sera fournie pour ce qui en est de l'utilisation des documents électroniques pour la diffusion et la publication d'informations juridiques. Un examen plus approfondi portera sur les outils informatiques et juridiques employés de nos jours pour rédiger, publier et répertorier des documents électroniques au sein de systèmes juridiques avancés pour la connaissance du droit. Enfin, un système juridique fondé sur l'utilisation avancée de documents électroniques sera fourni en exemple, pour terminer dans la dernière partie avec des propositions pour une approche intégrée à la gestion des documents électronique dans le domaine juridique.

Mots-clés : systèmes avancés d'information juridique, accès au droit, structuration sémantique du domaine juridique

1 Introduction

Cet essai représente une contribution au débat scientifique sur l'utilisation de documents électroniques dans le contexte juridique, notamment en ce qui concerne les termes de connaissance du droit.

L'on considère, en effet, que la connaissance des informations juridiques provenant de différentes sources (règles, jurisprudence, doctrine) soit une valeur fondamentale de tout système démocratique, dans la mesure où il reflète le principe de certitude juridique. Une des conditions matérielles essentielles pour la certitude juridique - entendu ici comme la prévisibilité des conséquences juridiques d'une action - est justement la possibilité pour un agent de connaître les règles (ou la jurisprudence pertinente dans une

¹ Ce document est le résultat d'une réflexion commune des auteurs. Toutefois, l'introduction et la première partie doit être attribué à Marina Pietrangelo, les deuxième et troisième parties à Maria Angela Biasiotti, tandis que les conclusions ont été élaborées conjointement par les auteurs.

juridiction de *droit commun*) en vertu de laquelle une action peut être qualifiée par la loi.

Grâce à l'utilisation de technologies informatiques, au cours des dernières années, de nouvelles formes et méthodes ont été conçues et mises en œuvre pour la divulgation des lois et des informations juridiques qui s'y rapportent (débat parlementaire, arrêts des Cours, décisions des tribunaux de grande instance au niveau national, européen et international, doctrine), ce qui a permis un accès plus conscient au patrimoine juridique public.

Dans ce contexte, le rôle du document électronique, et notamment des documents électroniques rédigés conformément à certains standards (spécifiquement exposés dans la deuxième partie de cette contribution) a été sans doute décisif dans la construction de systèmes juridiques avancés pour la connaissance du droit. Il s'agit de recherches développées dans le cadre de l'informatique juridique, une discipline dans laquelle ont récemment été développées de nombreuses applications spécialisées de façon à soutenir l'évolution imprévisible du droit dans la société contemporaine.

Dans la première partie de ce travail, une description synthétique du cadre juridique général sera fournie pour ce qui est de l'utilisation des documents électroniques pour la diffusion et la publication d'informations juridiques, en référant en particulier de l'état de l'art au niveau communautaire.

La deuxième partie portera sur un examen plus approfondi des outils informatiques et juridiques employés de nos jours pour rédiger, publier et répertorier des documents électroniques au sein de systèmes juridiques avancés pour la connaissance du droit.

Dans la troisième partie de ce document sera présenté l'exemple d'un système juridique fondé sur l'utilisation avancée de documents électroniques, conçu et établi sous forme de prototype au sein d'un projet du Conseil National des Recherches italien.

Enfin, dans la dernière partie, nous allons faire des propositions pour une approche intégrée de la gestion électronique des documents dans le domaine juridique.

2 Première partie. Le document électronique pour la compréhension du droit.

Dans de nombreux pays de l'Union européenne ainsi que d'autres pays du monde, les expériences d'informatisation publique commencées au début des années quatre-vingt-dix ont porté à la prédisposition de systèmes informatifs aux caractéristiques techniques sophistiquées, qui permettent actuellement aux citoyens de récupérer facilement les textes juridiques publiés sur l'Internet par une entité publique. Ceci permet de réaliser, dans la plupart des cas, la publication électronique des normes juridiques aux fins de divulgation, ainsi que les arrêts des principales Cours nationales et supranationales.

En ce qui concerne notamment les normes, leur consultation s'effectue par un accès unique aux ressources publiées volontairement sur les sites institutionnels des différentes administrations publiques. Il s'agit généralement d'un méta-moteur de recherche dans lequel les documents contenant des informations juridiques sont mis à disposition par chaque entité gouvernementale sur leur respectif *site Web*. Dans de nombreux cas, les documents électroniques relatifs aux différents textes législatifs ont été rédigés et publiés de façon à permettre l'apport de modifications et amendements de manière semi-automatique, ainsi que de permettre à l'utilisateur de

consulter le contenu d'un texte législatif au cours de toute sa période de validité (la soi-disant "multi-validité").

Il s'agit, en fait, d'une modalité avancée pour la consultation des normes qui permettrait de faciliter la connaissance du droit, en particulier dans les systèmes de *droit civil* qui sont généralement composés d'un grand nombre de normes. Certaines des premières expériences à cet égard sont le projet italien «Norme in Rete» (désormais remplacé par le projet «Normattiva», consultable à l'adresse www.normattiva.it²) et le portail français «Legifrance» (www.legifrance.gouv.fr).

Dans d'autres juridictions, en particulier celles de *droit commun*, cependant, la législation primaire et secondaire est accessible librement *en ligne*, mais la consultation ne peut être effectuée qu'après avoir accédé à une base de données centralisée spécialement conçue pour archiver les documents électroniques relatifs aux singles actes normatifs, ainsi que toute modification ou amendement qui leur ont été apportés par la suite. Il s'agit de documents pour la plupart publiés en format PDF, et qui sont donc statiques, non modifiable dans leur contenu, sinon par une substitution intégrale du document électronique original. Ainsi est le cas du Royaume-Uni avec le «UK Statute Law Database» (www.statutelaw.gov.uk) et de l'Irlande, avec l'«*Electronic Irish Statute Book database*» (<http://www.irishstatutebook.ie>).

A côté de ces services, qui visent à la divulgation maximale des normes juridiques, s'ajoutent aussi les procédures de numérisation des publications officielles, qui dans la plupart des Etats membres sont également publiées sur les sites *web* institutionnels, sur lesquels la consultation et l'extraction des textes est le plus souvent gratuite (comme dans le cas de la Bulgarie, de Chypre, du Danemark, de l'Allemagne, de la Grèce, de la Lettonie, de la Lituanie, du Luxembourg, de Malte, des Pays-Bas, de la Pologne, du Portugal, de Monaco, de la République Tchèque, de l'Espagne, de la Suède, et de l'Hongrie; pour l'Italie - malheureusement - la gratuité n'est que temporaire: www.gazzettaufficiale.it). Il est entendu, évidemment, que cette représentation du droit en format numérique est à titre exclusivement indicatif.²

Le droit communautaire est de même actuellement accessible que par édition imprimée, étant donné que la publication des documents communautaires en format numérique dans le système EUR-Lex ne constitue qu'un outil de diffusion. L'objectif, cependant, semble être un passage définitif à la publication numérique: "It is certain that the future of European Union law is eLaw – eLaw meaning law which is electronic, efficient, ergonomic and European law. The Official Journal of the Union in 2016 is probably an authentic electronic journal, with some paper copies distributed to the Member States"³.

Si un recours aux documents électroniques a permis de réaliser des systèmes informatiques pour la publication de normes, tout d'abord à titre uniquement d'information, et successivement à valeur légale, l'emploi de ces documents pourrait permettre de concevoir des systèmes dans lesquels la publication des normes serait accompagnée d'autres informations ou de méta-informations

² Pour une reconstruction complète du cadre communautaire en ce qui concerne la publication du droit sur le Web, voir *Access to legislation in Europe. Guide to legal gazettes and other official information sources in the European Union and European Free Trade Association*, Publications Office of the European Union, 2009.

³ K. Rissanen, *Access to EU law and eLaw – visions and challenges*, in *Séance académique « 25 années de droit européen en ligne »*, Publications Office of the European Union, 2006.

juridiques, qui conduiraient à l'établissement de systèmes juridiques plus raffinés et spécialisés.

Ces systèmes permettraient à l'utilisateur de repérer, dans un domaine d'intérêt sélectionné, tout type d'information juridique qui lui est associé: un arrêt, le commentaire d'une norme ou d'un arrêt, la législation connexe, ainsi que toute information contextuelle, pas nécessairement juridique, mais qui pourrait toutefois se révéler utile pour comprendre le cadre juridique d'intérêt.

Dans la société contemporaine, en effet, différentes catégories d'utilisateurs, plus ou moins spécialisés (les professionnels, l'administration publique, les opérateurs du droit, les citoyens) doivent pouvoir accéder chaque jour à une multitude de documents de contenu juridique ; c'est à eux que sont principalement adressés les systèmes mentionnés ci-dessus, étant donné qu'ils visent à faciliter l'accès à l'information à l'aide d'outils de recherche et de navigation spécialisés. La facilité d'accès, la capacité de récupérer rapidement tout les renseignements nécessaires, la gestion et la facilité d'utilisation représentent donc des paramètres clés.

3 Deuxième Partie. Outils informatiques et juridiques pour la compréhension de l'information juridique.

Actuellement, un utilisateur dispose de différents types d'outils ou de canaux de recherche pour repérer une particulière information⁴:

1. Recherche basée sur les détails d'une mesure
2. Recherche basée sur les références normatives
3. Recherche à l'aide de certains outils de soutien pour la recherche
4. Recherche textuelle ou au travers d'un mot clé

En fait, peut-être par souci de simplicité, que l'utilisateur soit avec ou sans expérience en matière juridique, et bien qu'il connaisse certains détails de la mesure qu'il recherche, il utilise généralement une recherche par mots-clés ou une recherche plus souvent définie en tant que "recherche textuelle". Cette option oblige l'utilisateur à préciser dans la requête les *mots* exacts qu'il considère pourraient se trouver dans le document, avec le résultat que si le choix de l'utilisateur (qui doit être particulièrement prudent à ne pas utiliser des mots trop généraux ou trop spécifiques) ne correspond pas au choix effectué, par exemple, par le juge ou les législateurs, l'utilisateur se trouvera incapable de trouver l'information qu'il recherche, bien que cette information existe. Si, au contraire, il a sélectionné le mot correct, l'utilisateur va certainement faire face à un large nombre de documents provenant de domaines différents et contenant le même mot, et devra donc effectuer lui-même une sélection pour obtenir l'information qu'il recherche. L'emploi du terme 'loyer' au lieu du terme technique «location» peut comporter que si le législateur n'a jamais utilisé le premier mais s'est limité au terme technique, l'utilisateur ne trouvera pas l'information qu'il recherche. Le domaine juridique est donc un domaine où la recherche rencontre des limitations que nous pouvons caractériser comme naturelles, étant donné qu'elles inhérent à la nature même de la matière qui exige que le langage commun utilisé au cours d'une requête soit successivement traduit en jargon technique de façon à ce que la recherche effectuée par des modèles de *Information Retrieval* soit efficace.

Le mot assume donc une valeur fondamentale dans la recherche d'information juridique et en particulier au sein de la recherche spécialisée, démontrant par conséquent le besoin pressant de réduire la naturelle ambiguïté du langage. Disposer

⁴ G.Pascuzzi, *Cercare il diritto. Come reperire la legislazione, la giurisprudenza e la dottrina consultando libri e periodici specializzazione*, Bologna, Zanichelli, 1998.

d'outils capables de mieux gérer les informations électroniques en général et les informations juridiques en particulier est devenu un besoin de plus en plus évident. Ainsi, une telle facilité pour les utilisateurs peut dériver de l'adoption de systèmes pour la gestion et la recherche d'informations juridiques : des outils et des ressources sémantiques d'une part, des techniques de références croisées d'autre part, de façon à guider l'utilisateur en orientant la recherche vers un meilleur résultat. Les premiers sont des outils qui, au travers d'un système de métadonnées, ajoutent aux documents des informations de type sémantique relatives à leur signification ou à leur contenu. Grâce à ce procédé, se rendent souvent explicites, et donc reconnaissables et capable d'être professées par une machines, des informations qui étaient auparavant implicites au document, bien que extrêmement significatives et représentatives. Les secondes sont des techniques de connexion et de référencement qui relient différents types de données conformément à une logique de domaine, créant ainsi un ensemble d'informations complètes et intégrées (par exemple, une loi peut être accompagnée par des références aux décisions des tribunaux qui s'y rattachent, ainsi que les contributions doctrinales sur le même thème, etc ..).

4 Outils d'indexation sémantique

Les outils sémantiques (métadonnées sémantiques) sont des ressources qui s'appliquent au delà des limites de la recherche syntaxique, c'est à dire la recherche qui se base sur l'identification des textes d'une base de données qui contiennent les termes recherchés⁵

Les plus communs outils d'indexation sémantique sont les répertoires des taxonomies, les schémas de classification, les thésaurus, les ontologies⁷. En particulier:

- Les répertoires des taxonomies sont des listes de termes qui permettent d'identifier les questions pertinentes à un certain domaine disciplinaire;
- Les schémas de classification comportent une structuration hiérarchique des termes qui y figurent: les termes de portée plus restreinte sont subordonnés aux termes de plus grande envergure: la matière indiquée par un terme général englobe toute les matières indiquées par les termes plus précis;
- Un thésaurus est une liste structurée d'expressions qui vise à représenter sans ambiguïté, dans un système de documentation, les concepts contenus dans un ou plusieurs documents. Un thésaurus comporte (i) des descripteurs, à savoir des mots ou des expressions qui désignent de manière non ambiguë les concepts constitutifs du domaine couvert par le thésaurus, (ii) des non-descripteurs, à savoir des mots ou des expressions qui désignent, dans un langage naturel, un concept identique ou équivalent, ou même un concept considéré, dans la langue du thésaurus, comme étant équivalent aux concepts représentés par les descripteurs, (iii) des relations sémantiques, à savoir des relations relatives à la signification des termes, soit parmi les différents descripteurs que entre des descripteurs et des non-descripteurs. Les relations représentées dans un thésaurus sont des relations d'équivalence (entre descripteurs et non-descripteurs), ou des relations hiérarchiques et associatives (entre les descripteurs);
- Une ontologie définit les termes employés pour décrire et représenter un domaine de connaissance. Une ontologie peut être utilisés par des personnes, des

⁵ R.Baeza-Yates, B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley,1999; P. Bertolotti, *Semantic Web: analisi e utilizzo di strumenti per la sua realizzazione*, Université de Turin, 2002.

applications et des bases de données qui ont besoin d'échanger des informations d'un certain domaine de connaissance (c'est-à-dire un secteur spécifique de la connaissance, comme la médecine, la musique, la gestion d'une entité publique, etc.) Une ontologie comporte les définitions des concepts de base du domaine et les relations qu'ils entretiennent, en utilisant un langage compréhensible et utilisable par un ordinateur.⁶ Chaque ontologie encode la connaissance d'un domaine spécifique mais est également capable d'encoder la connaissance de différents domaines, c'est pour cela que l'on peut dire que les ontologies rendent la connaissance réutilisable. À l'aide d'une ontologie, tout objet peut être regroupé dans une certaine catégorie de concepts, dans laquelle chaque concept est défini en termes de propriétés distinctives, chaque élément appartenant à une catégorie partage avec les autres éléments une certaine propriété (nécessaire), de telle sorte que, s'il en était dépourvu, ce concept ne pourrait pas faire partie de cette classe ; il y a donc des propriétés prototypiques (la dimension spatio-temporelle, la masse sont des propriétés communes à tout objet physique) et d'autres propriétés qui servent à distinguer les sous-catégories.

Les outils d'indexation sémantique mentionnés⁷ sont des ressources qui se composent d'un certain nombre de termes / concepts / descripteurs / éléments dont le but est celui de décrire un certain domaine auquel ils se réfèrent, aussi bien à un niveau général que d'un point de vue plus spécifique. Ce qui les différencie est, d'une part, le degré de dépendance du langage, et d'autre part, le type de relation qui existe entre les différents concepts et la formalité des contraintes. Il s'agit de ressources caractérisées par une croissante expressivité des liaisons qui s'interposent entre les entités qui les constituent: en partant par les schémas de classification plus basiques dans lesquels les éléments ne sont liés que par une relation hiérarchique, l'on passe aux thésaurus, dans lesquels, en plus d'une relation hiérarchique, les termes sont également liés par une relation d'équivalence ou une relation associative, pour arriver ensuite aux ontologies dans lesquelles les relations sont une expression des liens sémantiques entre les différents concepts. Ces dernières comportent différents niveaux de complexité et exigent des techniques de construction différentes. Généralement, elles peuvent être réalisées, d'une part, sur la base de la connaissance du domaine auquel elles se réfèrent, et d'autre part, sur l'analyse du contenu concret des documents recueillis dans le système documentaire. Dans le cadre d'un projet visant à élaborer une base de données, il est possible de mettre à la disposition des utilisateurs différents outils d'indexation sémantique à des temps différents, dans un processus de raffinement progressif, en expérimentant éventuellement avec des versions multilingues (en ce qui concerne la documentation en une autre langue que l'italien).

Utiliser une ressource plutôt qu'une autre dans la recherche d'une information juridique peut avoir un impact considérable sur le succès de l'opération. Les relations qui existent parmi les concepts d'une ontologie sont l'expression des liens sémantiques qui existent parmi les concepts d'un domaine spécifique.

Les récents développements des technologies de l'information en matière juridique suggèrent de nos jours un outil qui s'inspire aux modèles ontologiques, capable d'offrir à l'utilisateur la possibilité de rechercher un document à partir d'un mot clé

⁶ Sur ce point, voir, entre autres: D. Tiscornia, *Il diritto nei modelli dell'intelligenza artificiale*, Bologna, Clueb, 1996; B. Smith, *Ontologia e Sistemi informativi*, Networks, 2006; Hirst G., *Ontology and the Lexicon*, in S. Staab and R. Studer (eds), *Handbook on ontologies*, Springer, Berlin-Heidelberg, 2004.

⁷ D.A. Cruse, *Lexical Semantics*, Cambridge University Press, 1986.

significatif, qui ne doit pas pour autant faire nécessairement partie du contenu textuel du document.

L'analyse linguistique constitue la saisie de données pour les modules d'extraction, d'acquisition et de structuration des connaissances. La connaissance qui en résulte représente une ressource pour l'utilisateur final, qui lui permet de remplir et d'élargir les répertoires de vocabulaire linguistique et terminologique utilisés pour l'analyse des documents. Ainsi se réalise un cycle vertueux entre les outils. Les ressources linguistiques, lexicales et textuelles permettent de construire, développer, d'exploiter, et d'évaluer les modèles, les algorithmes, les composants, des systèmes qui sont, à leur tour, des instruments nécessaires à alimenter et à développer de manière dynamique ces ressources.

4.1 La référence croisée

En outre des possibilités offertes par l'emploi de métadonnées sémantiques, il ya aussi les possibilités offertes par les techniques de référence croisée. Cette expression se réfère à tous les cas où le texte d'une loi renvoie à une autre loi (références externes) ou à une autre section de la même loi (référence interne). Bien que les références internes se basent généralement sur une logique purement textuelle plutôt qu'une logique de domaine, les références externes comportent des liens à d'autres documents qui se basent souvent une logique de domaine. Cette technique se prête donc à être utilisée de façons différentes à selon du résultat que l'on souhaite obtenir: si l'on veut, par exemple, permettre à l'utilisateur d'acquérir une image complète de la manière dont une institution juridique est régie par la loi, interprétée par les juges, conçue et comprise par la doctrine, différentes liaisons seront créés parmi les différentes sources qui concerne cette institution spécifique.

5 Un projet de système informatique juridique avancé: l'Observatoire sur les règles du droit agro-alimentaire (ORAAL)

Au sein des questions abordées dans cet essai est d'un intérêt particulier, en tant que *best practice*, la création d'un observatoire sur les règles de l'agriculture et l'alimentation⁸ Un cas unique en Italie pour sa taille et sa modalité d'action, survenu à l'initiative de l'Université de Pise et du Département d'identité culturelle du CNR, l'ORAAL a pour objectif de coordonner, développer et diffuser les connaissances dans le domaine du droit agro-alimentaire, ainsi que de promouvoir la diffusion et la formation sur les questions de sécurité et de qualité des aliments, de santé alimentaire et de protection des identités.

Dans une industrie caractérisée par une multitude de normes qui régissent la production, la commercialisation et la distribution alimentaire, l'Observatoire a pour objectif de créer une base de données et d'établir un réseau à disposition des différents acteurs qui interviennent dans le cycle économique de la production alimentaire, des citoyens consommateurs, ainsi que des institutions et des praticiens du droit, à fin de rendre leurs actions plus efficaces et mieux coordonnées. L'utilisateur est donc fourni avec un aperçu global des connaissances, ce qui comporte une globalité des sources et une globalité des contextes reliés entre eux dans une logique de domaine (des arrêts qui offrent une interprétation

⁸ ORAAL surgit à l'initiative de l'Université de Pise et du Département d'identité culturelle du CNR. Il profite de la collaboration de deux instituts du CNR, celui de Théorie et Techniques de l'Information Juridique (Ittig) - situé à Florence - et celui de Linguistique Computationnelle (ILC) - situé à Pise - qui s'occupent de l'analyse de l'information et du langage juridique.

spécifique des lois, des universitaires qui écrivent des notes sur le même sujet, la Cour de justice qui propose des idées particulières sur le même sujet, etc.). L'information devient donc un nœud duquel peuvent se dénouer de nombreux parcours connectés entre eux de différentes manières.

Dans ce contexte, la recherche peut donc fournir des services complets, intégrés et facilement accessibles et disponibles à tous indépendamment de leur origine. L'outil informatique combinée à l'expertise du secteur alimentaires (dans ce cas, bien que cela s'applique aussi à d'autres domaines de compétence) permet de traiter de manière complexe, uniforme et harmonisée un contexte tel que celui du droit agro-alimentaire qui semble être inconstant, incohérent et internement déconnecté. Une fenêtre privilégiée sur les règles et les dynamiques qui régissent la discipline de ce secteur, ainsi qu'un instrument de surveillance continue par rapport aux changements qui interviennent dans l'industrie alimentaire dans d'autres contextes. Cela favorise la fertilisation croisée de la réglementation provenant de différents contextes.

Tout matériel pertinent en matière de droit agro-alimentaire est recueilli dans sa version intégrale et officielle, lorsque celle-ci est disponible. La documentation juridique qui provient de sources différentes est donc rassemblée et organisée selon les méthodes et techniques propre des bases de données. En particulier, l'objectif est celui de créer un système d'information à travers duquel il est possible d'effectuer une recherche soit par matière que par source, en laissant à l'utilisateur la possibilité de combiner deux options différentes : restreindre la recherche à seulement l'une ou l'autre des possibilités ou l'étendre à toutes les deux. Le traitement des documents mis à disposition des utilisateurs prévoit l'établissement de systèmes de repérage des informations qui reposent sur une organisation uniforme des connaissances contenues dans les documents et sur l'utilisation d'outils de langage universel, en s'appuyant sur une analyse sémantique des textes qui en augmente la capacité de traitement et de recherche.

Il s'agit d'une méthodologie de travail organisée et de solutions technologiques accolées, qui conduisent à un choix d'outils spécialisés pour l'interopérabilité, avec une attention particulière pour ce qui concerne les standards. Tout ceci compte tenu de la croissance globale et pour le plus partagée des connaissances et des technologies, et en vue d'une représentation des connaissances juridiques, comme précédemment exposé, capable de rendre extrêmement efficace les systèmes de recherche d'information (pour ce qui en est de l'organisation et la récupération d'information et de la représentation sémantique des documents) utilisés soit par les spécialistes, que par les simples citoyens et entreprises.

À ces fins, trois éléments peuvent se regarder essentiel: (i) le recueil et l'organisation de textes intégraux, (ii) la fourniture et l'application de schémas de métadonnées appropriées, (iii) l'utilisation d'outils pour l'indexation sémantique.

5.1 Le recueil et l'organisation de textes intégraux

Le premier objectif du système informatif est celui de rendre disponible aux utilisateurs, dans la mesure du possible, tous les textes intégraux pertinents à un document spécifique.

Afin d'obtenir un aperçu utile et approprié des normes relatives à l'objet en question, les textes normatifs doivent être identifiés, reconstruits dans leur cycle de vie (amendements, abrogations, etc.), interprétés et mis en relation entre eux. Une opération qui n'est pas sans critiques en raison de la multiplicité des documents concernés (par nombre et par type) et en vue de la division des responsabilités entre le gouvernement central et les entités périphériques (sans qu'il n'y ait eu

nécessairement un transfert direct de compétences du niveau central au niveau local) au bénéfice des citoyens et des entreprises, ainsi que des opérateurs de l'industrie et du droit en général. Les documents normatifs seront donc successivement marqués en langage XML en conformité avec les normes établies par le projet «Norme in Rete» (NIR) et incorporées dans les Circulaires AIPA du 6 Novembre 2001, n AIPA/CR/35 "Assignement de noms uniformes pour les documents juridiques » et du 22 avril 2002 N ° AIPA/CR/40 "Format pour la représentation électronique des mesures législatives par le langage de balisage XML ». Telle représentation permettra de profiter des fonctionnalités avancées pour la consultation, la recherche et le marquage des documents au plus haut niveau de détail, ainsi que de la gestion des schémas de classification spécifiques au secteur agro-alimentaire, qui bénéficient d'une façon plus détaillée de marquer les textes normatifs et qui permettent ainsi de produire un patrimoine à partager avec la communauté de professionnels du secteur, les références automatiques ... et le mode d'affichage en modalité multi-valide (à savoir la version du texte en vigueur à tout moment de modification).

6 Métadonnées

Afin de systématiser l'information recueillie, ORAAL a adopté un schéma de métadonnées qui peut être associé aux textes intégraux de manière à pouvoir représenter soit les éléments formels (auteur du document, titre, année et numéro d'identification. ...) que les contenus (par exemple au moyen de sommaires ou de notes et commentaires). D'une part, il a fallu tenir compte des meilleures expériences dans le contexte scientifique de référence (l'ensemble de métadonnées développées dans le cadre de l'initiative de Dublin Core), d'autre part, l'expérience acquise par les institutions participantes a été réutilisée pour ce qui en est de la préparation des bases de données en matière d'informations juridiques.

Notamment, les documents législatifs les plus importants sont accompagnés par des notes de coordination et des commentaires pour la plupart orientées aux utilisateurs non-spécialistes. Ces notes contiennent généralement une brève synthèse de la loi, les détails relatifs à toute modification de la législation précédente, ainsi qu'une référence aux normes corrélées les plus importantes (en absence de références formelles).

Dans le contexte d'un accès multilingue aux ressources normatives, telles que celui des ressources communautaires, les potentialités offertes par ces deux différents types de ressources deviennent beaucoup plus évidentes. Surtout dans un domaine spécifique comme celui du vocabulaire juridique, un accès lexico-conceptuel ne suffit pas à lui seul à garantir à l'utilisateur la fiabilité d'un système de recherche. Dans le contexte européen, en effet, la traduction des concepts en différentes langues a été réalisée *re ipsa*, c'est à dire que ce sont les concepts qui ont été transposés en différentes versions linguistiques à servir de pivot - en fait les deux types d'accès ont nécessairement besoin de cohabiter et de fonctionner, bien que d'une manière diamétralement opposée par rapport à une recherche monolingue: c'est l'accès ontologique qui permet de contextualiser la requête de l'utilisateur en transformant une notion juridique, que l'utilisateur ne connaît que dans sa propre langue, en un terme qui identifie le même concept dans une autre langue pour lui permettre d'effectuer une recherche des ressources juridiques étrangères en utilisant les expressions lexicales appropriées pour identifier ce concept spécifique.

7 Conclusions

À la lumière des expériences analysées dans le présent document et en vue des instruments informatiques réalisés jusqu'à présent dans le domaine de l'information juridique, il apparaît évident que l'énorme quantité de données juridiques contenues dans les documents électroniques en ligne exige d'une organisation appropriée pour pouvoir être effectivement consultés et reconnus par les citoyens. À ces fins, la tendance la plus désirable est celle de créer des outils qui facilitent une approche simplifiée à la complexité juridique et qui rendent les systèmes de recherche juridique, qui devront être utilisés soit par des utilisateurs spécialistes que par des citoyens ordinaires, beaucoup plus efficaces.

En effet, autant plus grande est la complexité du domaine, autant plus puissant devront être les instruments informatiques pour la récupération et pour la représentation sémantique des documents juridiques, de façon à rendre la recherche du système plus efficace. Des systèmes informatifs de ce genre doivent par conséquent garantir un accès intégré, complet et global au domaine juridique d'intérêt (normes, jurisprudence et doctrine) en fournissant à l'utilisateur non seulement les informations qu'il recherche, mais aussi tout ce qui résulte implicitement contenu dans le document de référence, ainsi que tout ce qui lui est logiquement associé mais que l'utilisateur ignore. Les instruments d'information juridique doivent donc être au service du droit fondamental du citoyen à la connaissance et à la compréhension des normes qui régissent ses actions.

La difficulté principale à l'utilisation du document électronique dans un contexte juridique est sans doute celle de créer des ressources normalisées qui soit reconnues par une certaine communauté et utilisées de manière uniforme afin d'assurer un désirable niveau d'interopérabilité.

8 Bibliographie

- 1 R. Baeza-Yates, B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley, 1999
- 2 Benjamins, Casanovas, Breuker, Gangemi (eds.), *Law and the Semantic Web*, Berlin, Springer, 2004
- 3 C. Ciampi, *Legal Information on the Web. The NIR Portal for the Citizen*, in *Proceedings of LEFIS Workshop on "Legal Aspects of E-Government"* (Vilnius, 20th September 2003), pp. 40 ss.
- 4 D.A. Cruse, *Lexical Semantics*, Cambridge University Press, 1986
- 5 European Union, *Access to legislation in Europe. Guide to legal gazettes and other official information sources in the European Union and European Free Trade Association*, Publications Office of the European Union, 2009
- 6 G. Hirst, *Ontology and the Lexicon*, in S. Staab and R. Studer (eds), *Handbook on ontologies*, Springer, Berlin-Heidelberg 2004, pp. 210 ss.
- 7 K. Rissanen, *Access to EU law and eLaw – visions and challenges*, in *Séance académique «25 années de droit européen en ligne»*, Publications Office of the European Union, 2006.
- 8 B. Smith B., *Ontologia e Sistemi informativi*, Networks, 2006
- 9 D. Tiscornia, *Il diritto nei modelli dell'intelligenza artificiale*, Bologna, Clueb, 1996

Lire numérique : nouveaux dispositifs - nouvelles pratiques ?

Animateur. Alexandra Saemmer, Paragraphe-Université Paris 8

Participants. Joël Gardes, Orange Labs / INSA Lyon-LIRIS ; Claire Bélsile, Lire-CNRS ; Emmanuël Souchier et Etienne Candel, CELSA-Université Paris-Sorbonne ; Caroline Courbières, LERASS-Université Toulouse 3.

Sont abordées dans cette table ronde les possibilités ouvertes pour la lecture numérique par les dispositifs numériques (de l'écran d'ordinateur aux « liseuses »). Au-delà des questionnements suscités par les caractéristiques des dispositifs et des interrogations concernant les potentialités hypermédiatiques au sein des créations (pensons aux relations intersémiotiques entre texte, image, mouvement et son couplées à l'interaction), il nous semble d'une part important de poser la question de l'intérêt de la lecture : les nouveaux dispositifs vont-ils susciter de nouvelles motivations, de nouvelles émotions, bref, un « plaisir du texte » renouvelé ? D'autre part, nous examinons les imaginaires et les représentations visuelles du lire comme une activité sociale située. Comment les dispositifs sont-ils pris dans une « mise en scène » de la lecture comme pratique médiatique ?

Les attentes des lecteurs évoluent avec la confrontation aux dispositifs, leurs contraintes et compositions, et en fonction des contenus présentés sur ces dispositifs : nous observons par exemple l'émergence de gestes spécifiques qui ne conditionnent pas seulement l'accès à une création numérique, mais prennent sens en fonction des mots ou images manipulables. Les attentes sont également forgées par les apprentissages, les discours ambiants (promotionnels ou critiques) dans un contexte socio-culturel. Se croisent, s'entrelacent des pratiques de lecture encore et toujours apprises sur papier, des pratiques forgées par les écritures numériques (qu'elles soient informationnelles, commerciales ou artistiques, qu'elles s'exercent lors de la lecture de blogs, de profils ou d'états facebook, de twits ou de résultats fournis par un moteur de recherche), des pratiques de manipulation de la matière numérique issues du jeu vidéo... Est-ce que ces pratiques de lecture multiples s'entrecroisent pourtant toujours sans heurt ?

Dans des productions comme *Moving tales*¹ ou *Alice in Wonderland*² pour i-pad par exemple, qui présentent sur une même « page » du texte fixe et animé, de l'image fixe et manipulable, l'expérience d'une lecture textuelle intensive peut éventuellement disparaître au profit d'un jeu avec la matière des mots et des images, s'épuiser dans l'exploration (certes jouissive) des potentialités offertes du dispositif.

¹ <<http://www.moving-tales.com/>>

² <<http://itunes.apple.com/us/app/alice-for-the-ipad/id354537426?mt=8>>

Beaucoup d'études montrent déjà que le lecteur, lorsqu'il consulte les résultats hyperliés fournis par un moteur de recherche, n'épluche pas tous les résultats mot à mot ; à la recherche d'une information précise sur un portail, il parcourt les pages web affichées sans se concentrer sur tous les détails - même au cas où il a envie de se laisser surprendre par des informations dépassant ses attentes. R. Simanowski constate ainsi le web fournit certes un accès illimité et immédiat à l'information, induit la ramification et relativisation constante des données, et provoque la mise en place de nouvelles formes d'archivage et d'indexation ; une rhétorique « de l'arrivée et du départ », et une structuration de l'information « par petite bouchées » est pourtant également caractéristique du web³. Poussé par le désir de découvrir encore et encore, le lecteur ne prend peut-être plus suffisamment le temps de s'arrêter. Encore une fois, le but du lecteur ne semble plus être la compréhension, mais le mouvement lui-même. Nous posons donc entre autres la question de savoir si ce mouvement empêche de fait la pratique documentaire et/ou remet en cause le rapport au savoir.

R. Simanowski, pour sa part, parle d'une mort du lecteur au profit de l'interacteur⁴. Doit-on craindre effectivement que ces pratiques de « lecture post-alphabétique »⁵ empêchent la réflexion, induisent l'acceptation léthargique du statu quo ? Ou sont-elles au contraire positivement ludiques, insouciantes et irrespectueuses, comme le proclament certains artistes du numérique et certains critiques lorsqu'ils affirment : « Dans l'expérimentation qu'offre l'art numérique, rien à gagner sinon l'expérimentation pour elle-même, plus exactement la découverte intime du plaisir de l'expérimentation »⁶, ou : « *Playfulness is a defining quality of this new medium [...] No matter how competitive, the experience of reading in the electronic medium remains a game* »⁷ ?

Nous devons certainement élargir la définition de la « lecture » dès qu'elle s'exerce sur un dispositif numérique, et sortir de la comparaison encore trop largement convoquée avec le livre ou le journal papier ; nous devons néanmoins nous demander ce que le lecteur « gagne » dans ces nouvelles pratiques de la lecture numérique, et ce qu'il y perd peut-être.

Ces dernières années, c'est notamment la discussion autour des prodiges et vertiges procurés par l'hypertexte qui a occupé le devant de la scène. Investi d'espoirs parfois démesurés (comme support d'une interconnexion quasi neuronale des savoirs), l'hypertexte a ensuite été soupçonné d'induire des dissonances cognitives. D'un côté, un mot hyperlié reste en effet un signe linguistique ; d'un autre côté, il est affecté d'une signalétique qui rapproche sa fonction de celle d'un bouton. Qu'est-ce qui prime alors dans la perception et dans la compréhension de ce signe : l'incitation au déchiffrement ou la tentation d'une manipulation immédiate ? Avant de cliquer, lorsqu'il perçoit un texte parsemé de liens pour la première fois, le lecteur est rarement renseigné sur les contenus à découvrir. Certes, la relation entre les médias activables et activés par le geste de manipulation peut ensuite devenir signifiante, et permettre au lecteur d'approfondir ses connaissances, de faire de nouvelles découvertes, de satisfaire sa curiosité. Néanmoins, l'incitation permanente à l'interaction, qui repose d'une part

³ R. Simanowski, *Digitale Medien in der Erlebnisgesellschaft. Kultur – Kunst – Utopien*, Reinbek bei Hamburg, Rowohlt, 2008, p. 216.

⁴ *Ibid.*, p. 49.

⁵ *Ibid.*, p. 119.

⁶ J.-P. Balpe, « Quelques concepts de l'art numérique », <http://transitoireobs.free.fr/to/article.php3?id_article=42>.

⁷ J. D. Bolter, *Writing Space : The Computer, Hypertext, and the History of Writing*, Hillsdale, L. Erlbaum, 1991, p. 130.

sur les liens hypertextes internes, et d'autre part sur les barres interactives, les signets, la fenêtre de saisie du moteur de recherche, les icônes de la messagerie, du chat etc. entourant la fenêtre du navigateur, distraient en permanence le sujet de sa lecture d'un texte précis. Faut-il alors mettre en avant, comme le fait Paul Virilio, le risque majeur de l'interactivité ? Les dernières fonctionnalités du navigateur *Firefox*, qui permettent de lancer à partir de n'importe quel mot d'un texte une requête *Google* concernant ce mot (il suffit de sélectionner le mot et d'effectuer un clic droit), renforcent encore cette tendance à la « distraction » permanente : chaque mot paraît désormais virtuellement hyperlié. Par ailleurs, bien que beaucoup de critiques s'accordent pour affirmer que le support numérique demande un lecteur plus « actif », il faut regarder de près en quoi consiste exactement l'activité du lecteur, et ce qu'elle lui apporte : A partir du moment où il peut seulement cliquer sans produire lui-même du texte, à partir du moment où son geste ne sert qu'à faire fonctionner la machine ou à feuilleter les contenus proposés, cette activité n'est-elle pas en même temps incitante *et* frustrante ?

Comme le font remarquer Y. Jeanneret et al., face à la machine, le lecteur est « placé dans une situation paradoxale de distanciation et d'engagement. La distance de l'homme à la machine est plus grande que celle de l'homme au livre, car le texte semble avoir disparu 'derrière' l'écran, laissant prise à l'espace du secret et du sacré. En revanche, l'engagement physique s'accroît, car le lecteur devient manipulateur et doit 'agir' la machine à des fins purement fonctionnelles »⁸.

En résulte un sentiment d'inconfort, de malaise relevé dans beaucoup d'études sur la lecture numérique⁹. Face au texte numérique, le lecteur semble par exemple à la fois jouir et souffrir d'une étrange impatience. Cette urgence est-elle due à la force de séduction des liens hypertexte (réels et virtuels), au désir de profiter immédiatement de la possibilité de non seulement lire le texte, mais de le manipuler - comme si la main ne pouvait plus rester immobile à partir du moment où elle a pris l'habitude des touches de la souris et du *touchpad* ? Effectivement, la main s'impatiente lorsqu'il y a trop de choses à faire, car elle ne sait pas où cliquer d'abord ; elle s'impatiente pourtant également lorsqu'il n'y a rien à faire. Y. Jeanneret et al.¹⁰ ont pu observer que le premier geste effectué par beaucoup de sujets devant un texte numérique, peu importe sa nature et son contenu, consiste à faire avec la souris un balayage global sur l'écran. Ce geste exploratoire s'explique par le désir de tester les potentialités de l'interface, et semble ainsi constituer une préparation à l'activité de consultation « sérieuse » des liens. D'autres raisons se rajoutent à cet argument principal pour expliquer l'impatience du lecteur devant le texte numérique. Même s'il ne contient ni hypertextes internes ni animations, le texte numérique est entouré d'un contexte interfacique manipulable qui envoie en permanence des signaux au lecteur. Le curseur lui-même indique la présence du lecteur dans le texte. Certains sujets essaient de ne pas perdre de fil de la lecture numérique en suivant les lignes du texte avec le curseur. Mais contrairement au

⁸ Y. Jeanneret, A. Béguin, D. Cotte, S. Labelle, V. Perrier, P. Quinton, E. Souchier, « Formes observables, représentations et appropriations du texte de réseau ». Lire, écrire, récrire. Objets, signes et pratiques des médias informatisés, éd. Emmanuel Souchier, Yves Jeanneret et Joëlle Le Marec, Paris, BPI Centre Pompidou, 2003 ; p. 98-99.

⁹ Voir par exemple l'enquête dirigée par C. Bélisle, mise en place en 2006 auprès d'étudiants universitaires pour étudier dans un premier temps leur usage des encyclopédies en ligne. Plus de six cent quarante étudiants ont répondu à un questionnaire en ligne sur le bureau virtuel des étudiants de l'université. Synthèse et analyses des résultats consultables à l'adresse : <<http://lire.ish-lyon.cnrs.fr/spip.php?rubrique88>>.

¹⁰ Y. Jeanneret et al., op. cit., p. 120.

doigt qui, lors d'une lecture papier, a toujours constitué un repère fixe pour l'écolier timoré, le curseur se transforme en butant sur les liens hypertexte, et incite au « divertissement ». En outre, comme le font remarquer Y. Jeanneret et al.¹¹, beaucoup d'indices avertissent le lecteur du temps qui passe : sur la plupart des ordinateurs, une horloge tourne en haut de l'écran, rappelant au lecteur que les secondes lui filent sous la main et sous les yeux.

Face à ces caractéristiques et difficultés, il nous paraît ainsi incontestable que ces nouvelles pratiques du « lire numérique » nécessitent de nouveaux apprentissages : savoir trouver l'information, savoir filtrer, savoir utiliser les outils, maîtriser de nouveaux gestes ; mais aussi mettre en perspective des contenus hyperliés sans se trouver découragé ou dispersé par le flux permanent de nouvelles informations ; savoir s'arrêter pour se concentrer sur un texte ou une vidéo, malgré les sollicitations permanentes envoyées par les différents cadres du dispositif. Quelles sont les compétences attendues en la matière notamment chez les jeunes, par exemple dans le cadre du « B2i – Brevet informatique et internet » (®) mis en place par l'Éducation nationale depuis quelques années afin de permettre aux futurs citoyens de faire « une utilisation raisonnée des TIC » ?

Dans cette table ronde, nous avons fait l'inventaire des caractéristiques principales de la lecture numérique. Nous avons abordé les potentialités induites par les nouveaux dispositifs dans leur contexte social, sans pour éviter les questionnements critiques. Car décidément, la lecture numérique constitue un défi : pour les créateurs et les éditeurs, qui sont amenés à inventer des formes et figures nouvelles de l'écrit numérique ; pour les lecteurs, dont les attentes et habitudes se retrouvent régulièrement remises en question.

Enfin, nous avons aussi posé la question de savoir si nous avons encore la bonne approche du document dans ce contexte actuel où l'on ne peut plus envisager de services sans multimodalité et sans multimédia ; ou si la distinction conceptuelle entre donnée et information ne se trouve pas au contraire réellement réaffirmée par le numérique. Quelles sont les nouvelles « contraintes » apportées par l'accessibilité aux contenus ?

¹¹ Ibid., p. 124.