Le “Document” à l’ère de la différenciation numérique 14e colloque international sur le document électronique Auteurs / Editors :

Mostafa Bellafkih, Joël Gardes, Mohamed Ramdani, Khaldoun Zreik

Edité par / Published by :

**e**uropia Productions 15, avenue de Ségur 75007 Paris, France

Tel +31 1 45 51 26 07

Fax +31 1 45 51 26 32

Email: [info@europia.fr](mailto:info@europia.fr) [http://www.europia.fr](http://www.europia.fr/) [http://www.europiaproductions.com](http://www.europiaproductions.com/)

ISBN13 : 979-10-90094-07-9

© 2012 **e**uropia Productions

Tous droits réservés. La reproduction de tout ou partie de cet ouvrage sur un support quel qu'il soit est formellement interdite sauf autorisation expresse de l’éditeur : Europia Productions.

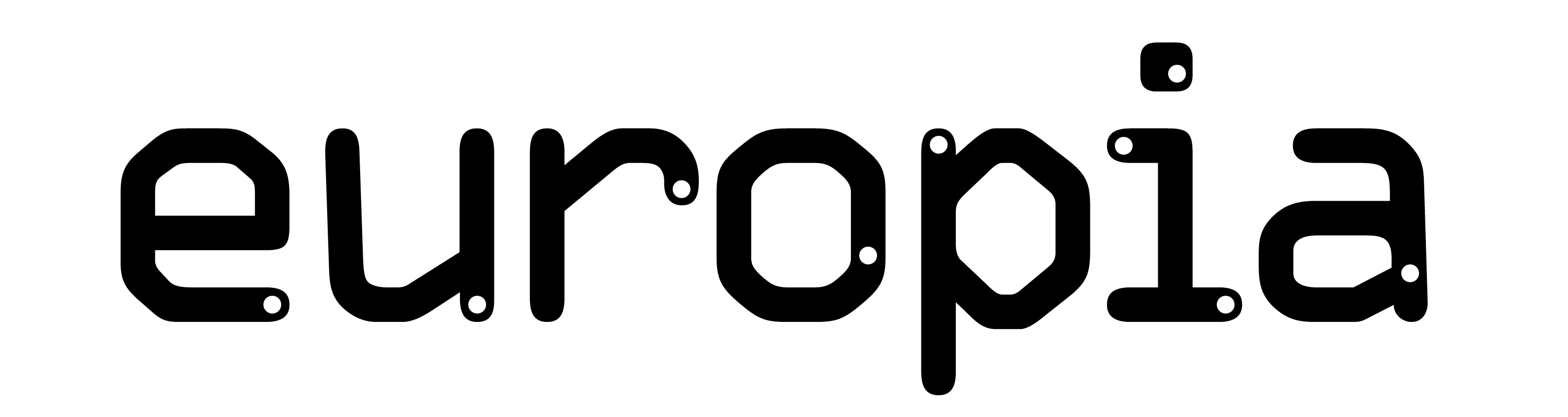
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher Europia Productions.

## Le “Document” à l’ère

**de la différenciation numérique**

**14e colloque international sur le document électronique**

Mostafa BELLAFKIH, Joël GARDES, Mohamed RAMDANI, Khaldoun ZREIK



**Présidents du colloque CiDE.14** Mostafa BELLAFKIH, INPT, Maroc Joël GARDES, Orange, France

Mohamed RAMDANI, Université Mohamadia, Maroc Khaldoun ZREIK, Université Paris 8, France

**Comité d’organisation local**

Mostafa BELLAFKIH, INPT, Maroc

Mohamed ERRAIS, INPT et Université Mohamedia, Maroc Ferdaous LAHMAR, CITU-Paragraphe, Université Paris 8, France Mohamed RAMDANI, Université Mohamedia, Maroc

Brahim RAOUYANE, INPT et Université Mohamedia, Maroc Karima TOUNSI, INPT, Rabat, Maroc

**Comité Permanent des colloques CiDE** Ghislaine AZEMARD, Université Paris8, France Mostafa BELLAFKIH, INPT, Maroc

Jean CAELEN, CLIPS-IMAG, Grenoble, France Jacques DUCLOY, DRRT Lorraine, France Patrice ENJALBERT, Université de Caen, France Mauro GAIO, Université de Pau, France

Joël GARDES, Orange, France Jean-Luc HAINAUT, Belgique

Maryvonne HOLZEM , Université de Rouen, France Madjid IHADJADENE, Université Paris 8, France Peter KING, Université de Manitoba, Canada Jacques LABICHE, Université de Rouen, France Abdelkarim MEZIANE, CERIST, Algérie

Mustapha MOJAHID, Université de Toulouse Le Mirail, France Ghassan MOURAD, Université Libanaise, Liban

Giovanni DE PAOLI, Université de Montréal, Canada Jean-Pierre RAYSZ, Jouve, France

Jean Marc ROBERT, Ecole Polytechniques – Université de Montréal, Canada Zaidi SAHNOUN, Université de Constantine, Algérie

Maurice SZMURLO, Orange, France Loïc THOMAZO, Orange, France

Eric TRUPIN, Université de Rouen, France Christophe TURBOUT, Université de Caen, France

Jacques VIRBEL, Université de Toulouse Le Mirail, France Jean VIVIER, Université de Caen, France

Christine VANOIRBEEK, EPFL, Suisse Manuel ZACKLAD, CNAM-Paris, France

Khaldoun ZREIK, Université Paris8, France (Coordinateur)

### TABLE DES MATIERES

### Introduction 5

Mostafa BELLAFKIH, Joël GARDES, Mohamed RAMDANI, Khaldoun ZREIK

### Partie 1 - Indexation sémantique 9

« Un modèle sémantique pour l'indexation

de document arabes et anglais » 11

Taher ZAKI, Abdellatif ENNAJI, Stéphane NICOLAS, Driss MAMMAS

« Approche d’indexation automatique d’informations pédagogiques à partir de documents » 25

Boutheina SMINE, Rim FAIZ, Jean-Pierre DESCLES

« Indexation sémantique de documents textuels » 43

Fatiha BOUBEKEUR, Wassila AZZOUG, Sarah CHIOUT, Mohand BOUGHANEM

### Partie 2 - Document interactif 61

« Extension d’un algorithme Diff & Merge au Merge Interactif » 63

Xuan TRUONG VU, Pierre MORIZET-MAHOUDEAUX,

Joost GEURTS , Stéphane CROZAT

« La métaphore dans les relations intermédiatiques :

quelles remédiatisations interactives ? » 83

Pergia GKOUSKOU-GIANNAKOU

« LaSuli : un outil pour le travail intellectuel » 91

Aurélien BENEL, Jean-Pierre CAHIER, Matthieu TIXIER

[Partie 3 - Document participatif 107](#_TOC_250007)

« Analyse exploratoire d'un wiki académique :

le cas d'EFRARD 109

[Kahina BELGAID](#_TOC_250006)

[« Les références bibliographiques dans Wikipédia » 115](#_TOC_250005)

[Gilles SAHUT](#_TOC_250004)

« Enrichissement sémantique

du corpus iSPEDAL » 125

[Abd El Salam AL HAJJAR, Mohammad HAJJAR, Zeinab ABDEL NABI, Georges LEBBOS](#_TOC_250003)

[Partie 4 - Aspect cognitif du document numérique 133](#_TOC_250002)

[« Terminologie hypertexte :](#_TOC_250001)

[dynamique temporelle d'une taxonomie » 135](#_TOC_250000)

Nathalie PINEDE, David REYMOND,

Benoit LE BLANC, Véronique LESPINET-NAJIB

« Un modèle d'architecture de pages web

pour une accessibilité augmentée destinée aux non-voyants » 153

Mustapha MOJAHID, Bou Issa YOUSSEF, Bernard ORIOLA, Nadine VIGOUROUX

### Partie 5 - Pratique du document numérique

**dans l'univers de la recherche 171**

« Pratiques de lecture numérique et usages

des technologies de l’écrit chez le chercheur tunisien » 173

Abderrazak MKADMI, Besma BSIR

« Présentation de l'information

comme support d'aide à des processus cognitifs » 189

Mustapha MOJAHID, Nesrine NOUGHI, Philippe BOISSIERE

« Tendances lourdes et tensions

pour les filières du document numérique » 205

Ghislaine CHARTRON, François MOREAU

### Partie 6 - Edition hypertextuelle 219

« Ré-édition de Chrestien de Lihus dans l'hypertexte » 221

Thierry DAUNOIS

« Formalisation des processus d’éditique :

Proposition d’un guide d’assistance à la formalisation

de processus d’éditique à travers la transposition contextuelle

de la notion de veille vue comme un système cybernétique » 237

Sébastien BRUYERE, Vincent OECHSEL

« Accès aux collections de presse ancienne :

une étude exploratoire » 249

Céline PAGANELLI, Evelyne MOUNIER, Stéphanie POUCHOT

# Introduction

Le document à l’ère de la différenciation numérique

La tenue de CIDE à Rabat cette année revêt une valeur symbolique particulière : la première édition de notre colloque s’est tenue au même endroit à la fin du XXème siècle (1998) et, depuis, 13 ans ce sont écoulés au cours desquels le concept de document, sa perception ainsi que ses pratiques ont évolué, voire, subit des mutations.

Les bibliothèques numériques sont désormais une « réalité » en progression permanente et non plus des projets. Le succès des dispositifs collaboratifs de type Wiki confirme la nécessité et la volonté croissantes de partager des connaissances, même si la valeur éditoriale des contenus reste encore sujette à questions.

CIDE a vécu ces treize ans que l’on pourrait dorénavant qualifier les années du « big bang » des télécommunications mobiles et, en particulier, de terminaux que l’on considérait encore d’intelligents en 1998, quand on décrivait à quoi ils allaient ressembler et qui, aujourd’hui, s’appellent

« smartphones » ou « tablettes ».

Concomitamment ces terminaux ont redonné, et ceci peu paraître de prime abord paradoxal de l’écrire, une place de choix au geste. Les interfaces sont largement tactiles, non seulement pour saisir de l’information, mais aussi, pour commander le système de l’on tient au creux de sa main et pour manipuler les données affichées à l’écran.

N’y a-t’il pas ici une meilleure illustration d’une des thématiques récurrente de CIDE, à savoir, le concept de document dynamique et interactif ? En y réfléchissant, ce concept n’est rien d’autre que l’interface utilisateur de nos terminaux mobiles ainsi que des nouvelles générations d’ordinateurs personnels qui tendent à mettre à mal progressivement le concept de clavier et de souris, en offrant également la possibilité de saisir et de manipuler directement des objets graphiques ou symbolique représentant de l’information.

Bref, tout ceci avait été pressenti comme « toile de fond » du document numérique lors des différentes éditions de CIDE. Mais un aspect de plus a surgi du fait de cette banalisation de l’information numérique et de l’énorme puissance de calcul et de stockage des terminaux : l’utilisateur destinataire de l’information désire s’approprier totalement le contenu restitué et accepte de moins en moins de se plier aux contraintes de l’émetteur.

Qu’est-ce que cela signifie dans notre concept de document ?

La personnalisation de la présentation de l’information devient une problématique centrale, en ce sens que le contenu restitué doit être non seulement adapté à la nature du terminal sur lequel le document s’affiche, mais aussi, aux capacités de lecture et de manipulation du contenu de l’utilisateur en fonction de sa situation de déficience permanente

**Le “Document” à l’ère de la différenciation numérique**

(handicap) ou temporaire (communication « bruitée »). Tout ce travail d’adaptation fait que l’on ne communique plus l’information au travers de messages et de signaux préétabli, mais au travers d’une médiation qui se traduit par autant d’instanciation de la présentation de l’information que de situations et de contextes d’usages.

Il devient indispensable de préserver des notions d’authenticité et d’intangibilité de l’information, puisque la présentation de celle-ci n’est plus fixée a priori lors de la composition du document, à l’instar de la chose imprimée ; mais est reconstruite dès que l’utilisateur y accède par l’intermédiaire de ses propres moyens techniques.

L’extension « naturelle » de cette personnalisation est, toujours, la multimodalité. Ici, on touche aux mécanismes intimes de l’interaction homme machine en posant clairement la question de l’efficacité du

« canal de communication » qu’est le document numérique ou l’interface homme machine et du maintien de la conformité de l’information transmise par ce canal. Prenons simplement la vocalisation d’un extrait de livre. La première condition est de disposer d’un texte cohérent et conforme à la source. La deuxième condition est d’offrir à l’utilisateur une sorte de « balisage » du document audio lui permettant de se repérer et ce balisage remet en avant le récit du texte. Ce dernier, s’il est monocorde, égare l’utilisateur, alors que s’il contient une prosodie bien adaptée, permet à notre utilisateur de s’approprier la « lecture » audio du document.

Au départ : une information à transmettre en la codant. A l’arrivée, une information restituée selon de profonds critères de personnalisation tout en préservant conformité et authenticité. Ainsi, ce n’est pas l’information elle-même qui se différencie en fonction des contextes de situation et d’usage du numérique, mais sa présentation : une même information revêt désormais une large variété de présentations dont nous devons tenir compte. Cette variété dépend du codage que l’on adopte. Historiquement, ce codage, pour le document, correspondait à tout l’art de l’éditeur associé à celui de l’imprimeur. L’éditeur garantissait la conformité à l’œuvre et l’imprimeur fixait ces marques de conformité sur un support intangible. Avec le numérique, le contenu se réédite autant de fois que l’on y accède : que deviennent les marques de conformité et d’authenticité ? Par quels mécanismes peut-on préserver la conformité ? Quelle est « l’autorité » permettant de garantir l’authenticité ?

L’utilisateur « ouvrant » un document, ouvre en fait, un fichier contenant des instructions codées. Le document est devenu une sorte de « kit d’assemblage » de l’information, mais est-on certain que le « mode d’emploi », on devrait dire les « modes d’emplois » de montage de ce kit a été également fourni ?

Ainsi peut-on résumer le « cas » du document numérique. Il a conservé sa vocation historique de communication de connaissances (document venant de doceo, docere : communiquer pour enseigner). Mais

**Introduction**

désormais, le document, devenu multi[media, modal, lingue, culturel, …] (pardon pour cette pseudo écriture), n’est plus qu’une instance dépendant du contexte de la communication. C’est la raison pour laquelle nous avons proposé de parler d’ère de la différenciation numérique pour cette édition de CIDE. L’information se comporte comme une cellule souche en biologie, évolue en autant de présentations fonctionnelles qu’il y a d’usages et de contextes. Ceci ouvre la voie à de vastes champs de recherche qui, on peut le parier, nous mènent aux frontières des théories de l’information actuelles en nous faisant découvrir les limites de tous nos modèles de documents. En effet, la présentation de l’information devient relativisée aux contextes d’utilisation et de personnalisation du contenu.