

**Initiatives 2003**  
**Normes et Standards pour l'apprentissage en ligne**  
**Journée Internationale de réflexion**

19 mars 2003. Château de Versailles-France

**Normalisation et EAD face aux contraintes  
linguistiques et culturelles de l'I18n dans le contexte  
arabophone**

**Mokhtar BEN HENDA**  
ISD, Université la Manouba-TUNISIE  
Mokhtar.Benhenda@isd.rnu.tn

**Rachid ZGHIBI**  
Université Paris 8-FRANCE  
rachid\_zghibi@yahoo.fr

# 1 - Objectifs :

- ❑ Introduire la langue arabe comme langue partenaire des langues latines dans les projets d'internationalisation et de normalisation menés par des instances régionales et internationale : LOM et RLM.
- ❑ LOM : Certaines particularités propres à la langue et à la culture arabes ne sont pas prises en compte.
- ❑ Réviser les catégories du LOM ayant un rapport avec les aspects linguistiques et culturels de la langue arabe dans une optique d'internationalisation (i18n) et de contribution au RLM.
- ❑ Une précision des définitions des catégories, des éléments de données, des valeurs et des types de données touchant des caractéristiques **linguistiques** et **culturelles**.

## 2. Les considérations linguistiques :

Envergure multilingue du LOM (LOM V1.0) :

(cf. 1.1 : Domaine d'application ;  
1.2 : Objectif, p.5),

= Multilinguisme "Souple" (langues latines)

Multilinguisme "Lourd" :

- deux ou plusieurs familles de langues dans la même ressource
- deux différences majeures :
  - ✓ le rendu visuel de la forme graphique des caractères
  - ✓ la directionnalité opposée de l'acte graphique

Partiellement résolu par ISO 8859-6, UNICODE et ISO 10646

## Partiellement ?

- ❑ ne permet pas d'avoir un environnement exclusivement arabe (tentative inachevée en 1985 pour la réalisation d'une informatique exclusivement arabe basée sur la norme ASMO 662)
- ❑ problèmes multilingues toujours concrets au niveau :
  - du traitement interne des données (tri et organisation des données)
  - des interfaces homme-machine (bidirectionnalité des bris et des incises, unidirectionnalité G-D des chiffres)
- ❑ ne répond pas aux exigences les plus subtiles de la langue et de la culture traitées ci-après.

### 3 - Les considérations culturelles :

- ❑ Des variations culturelles au niveau des usages et des habitudes des utilisateurs dans leurs rapports avec la langue et la culture :
  - des interférences linguistiques et culturelles externes agissant sur :
    - ✓ la démarche constitutive de la doctrine terminologique (traduction, translittération, emprunts)
    - ✓ les règles et les standards de l'usage de la langue (structures phonologiques, morphologiques et syntaxiques).
  
- ❑ Le LOM touche à ces polémiques linguistiques et culturelle :
  - Quand il s'agit de rapports entre entités arabophone et latine
  - Quand il s'agit de rapports dans l'entité arabophone elle-même.
  
- ❑ Les autres catégories du LOM les plus sensibles sont celles qui concernent les données numériques et dates :
  - chiffres (arabe/hindou),
  - dates (hégire/grégorien)

## 4 - Commentaires par catégorie

### a. LES DATES :

#### *LOM :*

- ISO 8601 : 2000,
- Calendrier grégorien pour les dates postérieures au 15 octobre 1528
- Calendrier julien pour les dates antérieures au 05 octobre 1528.

#### *Contexte arabe :*

- Majorité = calendrier grégorien (le Masihi : le calendrier «messianique» ou chrétien,
- Calendrier lunaire hégirien reste obligatoire pour déterminer les dates religieuses.
- Calendrier lunaire est officiellement utilisé dans tous les domaines par des pays comme l'Arabie Saoudite.
- Nom des mois : 3 systèmes de noms (voir tableau suivant) :

Les noms arabes		Les noms syriaques		Les noms latins			
1	محرم	1	كانون الثاني	1	يناير	جانفي	
2	صفر	2	شباط	2	فبراير	فيفري	
3	ربيع الأول	3	آذار	3	مارس		
4	ربيع الثاني	4	نيسان	4	أبريل	أفريل	
5	جمادى الأولى	5	أيار	5	مايو	ماي	مايو
6	جمادى الثانية	6	حزيران	6	يونيو	جون	يونيه
7	رجب	7	تموز	7	يوليو	جويلية	يوليوز
8	شعبان	8	آب	8	أغسطس	أوت	غشت
9	رمضان	9	أيلول	9	سبتمبر	سبتمبر	
10	شوال	10	تشرين الأول	10	أكتوبر	أكتوبر	
11	نور القعدة	11	تشرين الثاني	11	نوفمبر		
12	نور الحجة	12	كانون الأول	12	ديسمبر	ديسمبر	

## Les catégories concernées par cette remarque sont :

- **Catégorie 2 : Cycle de vie, élément de données 2.3.2 : La date de contribution ;**
- **Catégorie 3 : Méta-métadonnées, élément de données 3.2.3 : La date de contribution ;**
- **Catégorie 8 : Commentaire, élément de données 8.2 : Date.**

## **b. LES DURÉES :**

### ***LOM :***

ISO 8601 : 2000 :

- P [yY][m M][d D][T[h H][nM][S[.s]S]]

Exemple :

- P1Y3MT2H30M = une année et 3 mois et 2 heures et 30 minutes.

### ***Contexte arabe :***

Problèmes avec l'usage des unités exprimées en caractères latins.

Nécessité de traduire ces unités en arabe.

## **Les catégories concernées par cette remarque sont :**

- Catégorie 4 : Technique, élément de données 4.7 : Durée de l'utilisation continue de l'objet pédagogique ;
- Catégorie 5 : Pédagogique, élément de données 5.9 : Temps d'apprentissage type.

## c. LA VOCALISATION

Écriture arabe : écriture consonantique

La prononciation : 3 voyelles brèves et 7 signes orthographiques.

La vocalisation est indispensable pour :

- l'apprentissage (langue et écriture)
- certains domaines d'enseignement spécialisé (théologique, philologique, etc.).

LOM : standard voué à l'enseignement et à l'apprentissage à distance

➔ la mention objet pédagogique « vocalisé » ou « non vocalisé » est nécessaire pour :

- identifier l'objet pédagogique
- identifier l'utilisateur final.

### **Exemple1 :**

un objet pédagogique destiné à des apprenants (i.e. 5.5 : Pédagogique.Rôle présumé de l'utilisateur final ("LOMv1.0", "Apprenant")), utilisé dans le contexte d'une école (i.e. 5.6 : Pédagogique.Contexte ("LOMv1.0", "École")) doit forcément être vocalisé.

### **Exemple 2 :**

un objet pédagogique de type expérience (i.e. 5.2 : Pédagogie.Type de ressource pédagogique ("LOMv1.0", "Expérience")) utilisé dans le contexte d'un enseignement supérieur (i.e. 5.6 : Pédagogique.Contexte ("LOMv1.0", "Enseignement supérieur")) ne doit pas forcément être vocalisé.

### **Proposition :**

Mentions « vocalisé » ou « non vocalisé » indiquées au niveau de la catégorie 1.Général comme complément de l'élément de données 1.3 Langue.

### **Exemple :**

"ar, vocalisée"

"ar, non vocalisée"

## **d. LES CODES DES LANGUES ET DES PAYS**

### ***LOM :***

- norme ISO 639 : 1998 code de la langue
- norme ISO 3166-1 1997 code de pays.

➤ Codes représentés en utilisant des lettres latines

### ***Contexte arabe :***

- Prévoir des équivalents arabes à ces codes pour le besoin du LOM.

### **Les catégories concernées par cette remarque sont :**

- Catégorie 1 : Général, élément de données 1.3 : Langue ;
- Catégorie 3 : Méta-Métadonnées, élément de données 3.4 : Langue ;
- Catégorie 5 : Pédagogique, élément de données 5.11 : Langue.

## e. LES CHIFFRES

Deux types de graphies pour représenter les chiffres :

- Les chiffres arabes : Maghreb (Tunisie, Algérie, Maroc et Mauritanie)
- Les chiffres hindous : A partir de la Libye vers l'Orient du monde arabe

0	1	2	3	4	5	6	7	8	9	10
٠	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠

Correspondance entre chiffres arabes et chiffres hindous

### *LOM :*

- Seuls les chiffres arabes sont utilisés.

### *Recommandation :*

- Prévoir la possibilité d'utiliser la graphie hindoue dans le cadre d'un LOM arabisé.

## **f. LA BIDIRECTIONNALITE :**

L'une des sources de difficultés entre les écritures arabe et latines :

- Gérer la bidirectionnalité dans un même document et sur la même ligne.

Quatre aspects de complexités :

- les incises,
- les bris,
- les caractères neutres (ponctuation),
- les chiffres.

Dans le LOM :

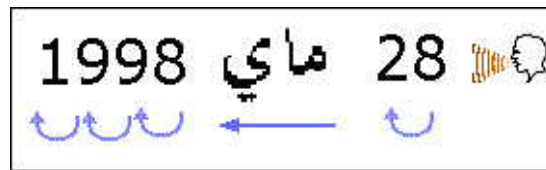
- La bidirectionnalité est source de polémiques au niveau des catégories Date, Durée et Chiffres.

# Cas de problèmes de la Bidirectionnalité :

## Date :

- ◆ En Mode Maître Arabe (MMA) à directionnalité de base D-G :  
une logique d'évolution du plus petit au plus grand :  
Date du jour, Nombre/nom du Mois et Année dans une suite D-G.

- ◆ La date du 28 Mai 1998 prononcée comme suit :  
« Huit et Vingt Mai, Huit et Quatre vingt Dix et Neuf Cent et Mille ».



Structure de la date arabe conforme à la norme ISO 8601

-> Conformité graphique (!) avec la norme ISO 8601 : 2000

- ◆ En Mode Maître Latin (MML), cette directionnalité D-G est n'est pas inversée.

*Ajouter une note d'application pour ne pas inverser systématiquement l'ordre de la date avec l'inversement de l'acte scriptural dans un LOM entièrement arabisé.*

## **Cette caractéristique correspond aux catégories suivantes :**

Tableau 1 : Structure du schéma élémentaire du LOM,

- Catégorie 2 : Cycle de vie, Elément de donnée 2.3.3 : Date.
- Catégorie 3 : Méta-Métadonnées, Elément de donnée 3.2.3. : Date.
- Catégorie 8 : Commentaire, Elément de données 8.2 : Date.

Tableau 3 : Date-Heure,

- catégorie 1 : DateHeure

## Durée :

- ❑ Les mêmes remarques que celles relatives à la date,
- ❑ Les mêmes considérations pour l'ordre séquentiel des unités de mesure du temps ,
- ❑ Les mêmes considérations pour la directionnalité définie par la norme ISO 8601.

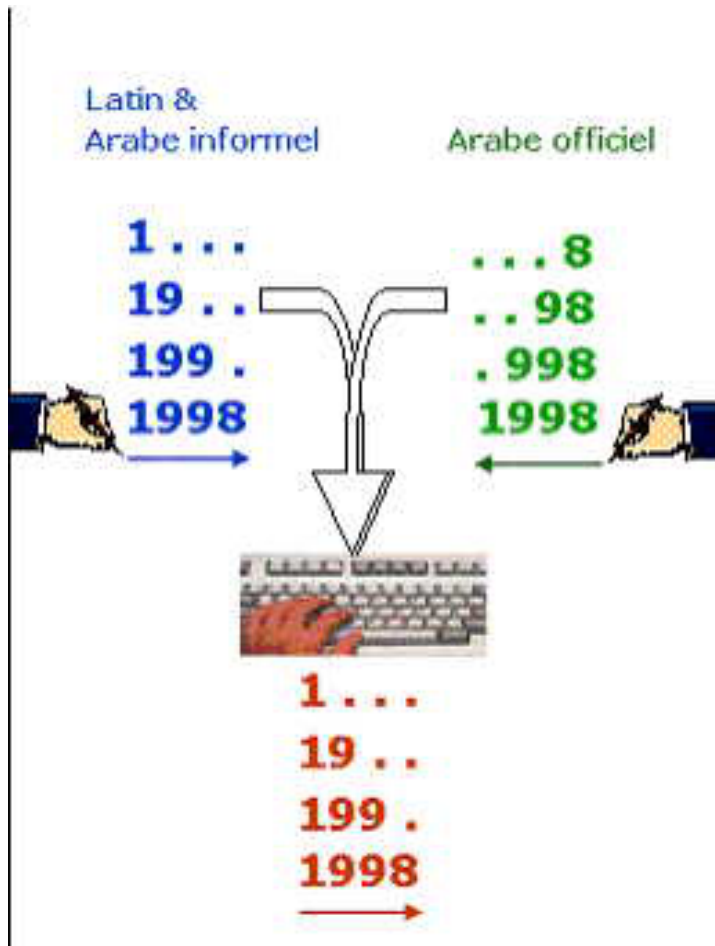
**Cette caractéristique correspond aux catégories suivantes :**

Tableau 4 : Durée,

- Catégorie 1 : Durée

## Chiffres :

En arabe littéraire, toute valeur numérique est traitée conformément à sa structure arithmétique interne ainsi qu'au mode de sa transcription graphique.



Opposition de l'acte graphique des chiffres entre les contextes arabe et latin

- ❑ En arabe littéraire, la transcription des chiffres est souvent conforme à l'acte graphique des caractères alphabétiques (D-G)
- ❑ Parfois, des fractions et des valeurs décimales (virgule flottante) dans un texte arabe officiel sont de Droite à Gauche.
  - $2/3 \rightarrow 3\backslash 2$
  - $25,10 \rightarrow 10,25$
- ❑ Ceci n'est pas très commun dans les systèmes d'information publics (médias, textes juridiques, documentation financière...).
- ❑ La littérature : dernier bastion des origines classiques :
  - directionnalité Droite-Gauche est respectée aux niveaux de la transcription graphique, de la structure orale et de la calculabilité arithmétique.

*Une remarque supplémentaire au niveau des catégories correspondantes du LOM permettrait d'unifier la forme de saisie des valeurs des éléments de données chiffrées de façon uniforme et éviter ainsi l'interprétation aléatoire dans un LOM entièrement arabisé.*

## **Cette caractéristique correspond aux catégories suivantes :**

Catégorie 2 : Cycle de vie, Elément de donnée 2.1 : Version.

Catégorie 4 : Technique, Elément de donnée 4.4.1.3 : Version minimale.

Catégorie 4 : Technique, Elément de donnée 4.4.1.4 : Version maximale.

Catégorie 5 : Pédagogique, Elément de donnée 5.7 : Tranche d'âge type.

## **g. LA TERMINOLOGIE :**

- ❑ Objectif du LOM : Un ensemble minimal de ressources terminologique homogènes et uniformes pour la description des éléments métadonnées.
- ❑ Deux grands problèmes terminologiques majeurs dans le monde arabe :

### **1 - L'actualisation de la langue arabe par rapports aux sciences nouvelles :**

- ❑ Industrialisation : enrichir le lexique conformément aux innovations technologiques.
- ❑ L'usage des sigles, des acronymes et des abréviations.
  - Aucun équivalent pour LOM, EAD, ISO, AUF
  - Des exceptions : UNESCO, ISBN, ISSN.

## 2 - La normalisation terminologique :

- ❑ Un grand vide de normalisation terminologique résolu par :
  - La Translittération
  - Les Emprunts dans des langues tierces.
  
- ❑ Les disciplines scientifiques sont les plus touchées par la pauvreté lexicale,
- ❑ Les sciences humaines sont par contre fertiles en synonymie.
- ❑ Absence d'une uniformisation des concepts (Synonymie),
  - Des variations d'usage terminologiques (interférences linguistiques i.e. Machrek-Maghreb)
- ❑ Besoin d'une démarche normative pour renforcer une doctrine terminologique.
- ❑ S'aligner peut être sur la Normalisation terminologique et Normalisation des techniques documentaires régies respectivement par la norme ISO/TC 37 et la norme ISO/TC 46.

## **Cette caractéristique correspond aux catégories suivantes :**

Tableau 1 : Structure du schéma élémentaire du LOM,

- Catégorie 1.5 : Général.Mot-clé
- Catégorie 9 : Classification, Tous les éléments de données

Tableau 5 : Vocabulaire,

- Élément de donnée 1 : Source.

# CONCLUSION

Objectif de ce constat :

- contribuer au futur RLM

Recommandation :

- Travailler en étroite collaboration avec deux autres projets d'internationalisation de normes :
  - Arabisation de Dublin Core (Initiative H.Hadouti)
  - Internationalisation des noms des domaines (DNS) sur Internet : AINC (Arabic Internet Names Consortium)

Initiative :

- Nous commençons déjà l'arabisation du texte du LOM pour le soumettre aux commentaires de la communauté arabophone..