

EVALUATION DE LA NORME IPTC-NAA

Dans le cadre de la construction d'une base de données iconographique
du projet Usine 3D
- Centre Alexandre Koyré/CRHST -
UMR 8560 du CNRS/EHESS/MNHN/CSI.

Julie Sastrada - CAK/ CRHST
Mardi 12 Avril 2005

I. Présentation de la norme IPTC-NAA	Pages 2-7
II. Comparaison de la Dublin Core et IPTC-NAA.....	Pages 8-9
III. Champs IPTC non admis dans la DC.....	Pages 10-11
IV. Logiciels et organismes.....	Pages 12
V. Remarques.....	Pages 13-15
VI. Tableau récapitulatif des avantages et inconvénients de l'IPTC-NAA	Page 16
VII. Bibliographie	Page 17

I. Présentation de la norme IPTC-NAA (ou IIM)

L' [IPTC](#) (International Press and Telecommunications Council) est une organisation internationale créée en 1965 pour développer et promouvoir des standards d'échange de données à destination de la presse. En association avec la [NAA](#) (Newspaper Association of America), l'*IPTC* a défini un modèle global de données appelé *IPTC-NAA Information Interchange Model* (la version 4 d'octobre 1997 est connue sous le nom *IIMv4*). Ce modèle a servi de base à la société *Adobe* pour définir dans son logiciel *Photoshop* les informations textuelles associées à une image. C'est ce sous-ensemble qui est communément appelé *informations* (ou *métadonnées* ou *champs* ou *légendes*) *IPTC*. Ces informations sont constituées de 33 métadonnées de type interne, c'est-à-dire stockées à l'intérieur des fichiers *Photoshop*, *JPEG*, ou *TIFF*, et permettent de décrire les images à l'aide de champs (*Titre*, *Auteur*, *Mots-clés*, *Catégorie*, *Copyright*, etc.).

La norme *IPTC-NAA* complète est constituée de différents sous-ensembles, chacun décrivant un aspect particulier de l'image. Chaque champ est encodé sous la forme « x :yz », x correspondant à l'ensemble qu'il identifie, et yz au numéro du champ dans le sous-ensemble identifié.

Il existe ainsi **9 sous-ensembles** :

1. « Enveloppe Record » = informations relatives à l'expédition de l'objet

Numéro	Nom	Description
1:00	Model Version	Version de l'IIM
1:05	Destination	Destination du document
1:20	File Format	Format du fichier
1:22	File Format Version	Version du format du fichier
1:30	Service Identifier	Identifiant du créateur ou du producteur
1:40	Envelope Number	Numéro du fichier
1:50	Product I.D.	Destinataire du document
1:60	Envelope Priority	Urgence de l'envoi
1:70	Date Sent	Date d'envoi
1:80	Time Sent	Heure d'envoi
1:90	Coded Character Set	Ensemble de caractères utilisés pour l'indexation
1:100	UNO	Identifiant unique du fichier
1:120	ARM Identifier	Ensemble de champs <i>IPTC</i> utilisés
1:122	ARM Version	Ensemble de champs <i>IPTC</i> utilisés

2. « Application Record » = description documentaire de l'image

Numéro	Nom	Description
2:00	Record Version	Version de la norme utilisée
2:03	Object Type Reference	Type de document
2:04	Object Attribute Reference	Type d'informations
2:05	Object Name	Nom de l'objet
2:07	Edit Status	Etat du traitement
2:08	Editorial Update	Code indiquant la présence d'autres langues
2:10	Urgency	Urgence du traitement
2:12	Subject Reference	Code indiquant les sujets traités
2:15	Category	Catégorie
2:20	Supplemental Category	Sous-catégorie
2:22	Fixture Identifier	Sujet traité fréquemment
2:25	Keywords	Mots-clés
2:26	Content Location Code	Code des pays où le document peut être publié
2:27	Content Location Name	Nom des pays où le document peut être publié
2:30	Release Date	Date de début de droit
2:35	Release Time	Heure de début de droit
2:37	Expiration Date	Date de fin de droit
2:38	Expiration Time	Heure de fin de droit
2:40	Special Instructions	Instructions spéciales
2:42	Action Advised	Conditions d'utilisation
2:45	Reference Service	Numéro du service envoyant le fichier
2:47	Reference Date	Date d'envoi du fichier
2:50	Reference Number	Numéro d'envoi du fichier
2:55	Date Created	Date de création
2:60	Time Created	Heure de création
2:62	Digital Creation Date	Date de création enregistrée par l'appareil
2:63	Digital Creation Time	Heure de création enregistrée par l'appareil
2:65	Originating Program	Programme utilisé pour créer l'image
2:70	Program Version	Version de ce programme
2:75	Object Cycle	Moment de création (a : matin, p : après-midi...)
2:80	By-line	Nom du créateur
2:85	By-line Title	Titre du créateur (photographe...)
2:90	City	Ville de prise de vue
2:92	Sublocation	Précision sur le lieu de prise de vue (rue...)
2:95	Province/State	Province ou Etat de prise de vue
2:100	Country/Primary Location Code	Code du pays de prise de vue
2:101	Country/Primary Location Name	Nom du pays de prise de vue
2:103	Original Transmission Reference	Référence de la transmission d'origine du fichier

2:105	Headline	Sous-titre du document
2:110	Credit	Crédit
Numéro	Nom	Description
2:115	Source	Propriétaire de l'image (agence...)
2:116	Copyright Notice	Copyright
2:118	Contact	Personne à contacter pour plus d'informations
2:120	Caption/Abstract	Légende
2:122	Writer/Editor	Auteur de la légende
2:125	Rasterized Caption	Prévisualisation d'une partie de l'image
2:130	Image Type	Contenu du fichier
2:131	Image Orientation	Orientaion de l'image (verticale/horizontale)
2:135	Language Identifier	Langue d'indexation
2:150	Audio Type	Champs concernant uniquement les documents sonores
2:151	Audio Sampling Rate	
2:152	Audio Sampling Resolution	
2:153	Audio Duration	
2:154	Audio Outcut	
2:200	ObjectData Preview File Format	Format de prévisualisation du document
2:201	ObjectData Preview File Format Version	Version du format de prévisualisation
2:202	ObjectData Preview Data	Format de prévisualisation des données

3. « Digital Newsphoto Parameter Record » = caractéristiques techniques de l'image

Numéro	Nom	Description
3:00	Record Version	Version de la norme utilisée
3:10	Picture Number	Numéro de l'image
3:20	Pixels Per Line	Nombre de pixels par ligne
3:30	Number of Lines	Nombre de lignes
3:40	Pixel Size In Scanning Direction	Nombre de pixels scannés
3:50	Pixel Size Perpendicular To Scanning Direction	Nombre de pixels perpendiculaires scannés
3:55	Supplement Type	Contenu supplémentaire de l'image
3:60	Colour Representation	Nombre de couleurs
3:64	Interchange Colour Space	Espace colorimétrique
3:65	Colour Sequence	Nombre de couleurs par octets
3:66	ICC Input Colour Profile	Profil ICC
3:70	Colour Calibration Matrix Table	Version du standard ICC
3:80	Lookup Table	Dimension du document (2D, 3D...)
3:84	Number of Index Entries	Nombre de palettes de couleurs utilisées
3:85	Colour Palette	Palettes de couleurs utilisées
3:86	Number of bits per Sample	Nombre de bits par pixels utilisés comme entrées dans la palette de couleur
3:90	Sampling Structure	Relation entre les pixels
3:100	Scanning Direction	Direction du scanner
3:102	Image Rotation	Rotation appliquée à l'image
3:110	Data Compression Method	Méthode de compression des données
3:120	Quantisation Method	Méthode de quantification
3:125	End Points	Nombre d'octets par composants
3:130	Excursion Tolerance	Tolérance d'agrandissement
3:135	Bits Per Component	Nombre de bits par composant
3:140	Maximum Density Range	Intervalle maximum de densité
3:145	Gamma Compensated Value	Valeur du gamma

4. }
 5. } ces deux sous-ensembles restent libres pour une éventuelle évolution de la norme
-

6. méthode de description utilisée pour traiter le document (*utilité ????*)

Numéro	Nom	Description
6:00	Abstract Relationship Record	Ensemble de champs IPTC utilisés

7. « Pre-objectdata Descriptor Record »

8. « Objectdata Record »
traitement

9. « Post-objectdata Descriptor Record »

} description de l'image à différentes étapes de son

Numéro	Nom	Description
7:10	Size Mode	Connaissance du poids du fichier
7:20	Max Subfile Size	Poids maximum des métadonnées
7:90	ObjectData Size Announced	Poids prévu du fichier sans les métadonnées
7:95	Maximum ObjectData Size	Poids maximum du fichier sans les métadonnées
8:10	Subfile	Poids des métadonnées
9:10	Confirmed ObjectData Size	Taille final du fichier

SCHEMA SIMPLIFIE DE L'IPTC-NAA

Principalement utilisée par les photothèques et les logiciels (photoshop....) : 33 métadonnées

DataSet	Nom	Traduction et commentaire
5	Object name	Nom de l'objet : 64 caractères
7	Edit status	Statut éditorial
10	Urgency	Priorité
15	Category	Catégorie (champ obsolète dans la version <i>IIMV4</i>) : 3 caractères
20	Supplemental category	Catégorie supplémentaire (idem dans la version <i>IIMV4</i>)
22	Fixture identifier	Identificateur
25	Keywords	Mots-clés : nombre illimité de mots clé ou phrases n'excèdent pas plus de 64 caractères.
30	Release Date	Date de disponibilité
35	Release Time	Heure de disponibilité (suit la norme ISO 860)
40	Special Instructions	Instructions spéciales : 256 caractères
45	Reference service	Services de référence
47	Reference Date	Date de référence
50	Reference Number	Numéro de référence
55	Date Created	Date de création de l'objet : 16 caractères
60	Time Created	Heure de création de l'objet
65	Originating Program	Programme ayant créé l'objet
70	Program version	Version du programme ayant créée l'objet
75	Object cycle	Cycle de l'objet
80	By-line	Créateur de l'objet (nom du rédacteur, du photographe etc.) : 64 caractères
85	By-line Title	Titre du ou des créateurs
90	City	Ville : 32 caractères
95	Province/State	Province/ Etat : 32 caractères
100	Country/Primary Location Code	Code du pays
101	Country/Primary Location Name	Libellé du pays : 64 caractères
103	Original Transmission Reference	Référence de la transmission
105	Headline	Titre : 256 caractères
110	Credit	Crédit : 32 caractères maximum
115	Source	Source : 32 caractères
116	Copyright Notice	Copyright : 128 caractères
118	Contact	Contact
120	Caption/Abstract	Description, résumé, commentaire : pas plus de 2000 caractères
122	Writer/Editor	Auteur de la description
130	Image Type	Type d'image

II. Dublin core : similitudes avec l'IPTC-NAA

IPTC NAA 33 champs personnalisables		DUBLIN CORE 16 champs avec 30 raffinements	
Nom	Traduction et commentaire	champs	Raffinement(s)
Object name	Nom de l'objet : 64 caractères		
Edit status	Statut éditorial		
Urgency	Priorité		
Category	Catégorie (champ obsolète dans la version <i>IIMV4</i>) : 3 caractères		
Supplemental category	Catégorie supplémentaire (idem dans la version <i>IIMV4</i>)		
Fixture identifier	Identificateur	identifier	
Keywords	Mots-clés : nombre illimité de mots clé ou phrases n'excèdent pas plus de 64 caractères.	Subject (mots clé)	
Release Date	Date de disponibilité	date	available
Release Time	Heure de disponibilité (suit la norme ISO 860)		
Special Instructions	Instructions spéciales : 256 caractères		
Reference service	Services de référence		
Reference Date	Date de référence		
Reference Number	Numéro de référence		
Date Created	Date de création de l'objet : 16 caractères	date	Created Valid Issued Modified Dateaccepted Datecopyrighted datesubmitted
Time Created	Heure de création de l'objet		
Originating Program	Programme ayant créé l'objet		
Program version	Version du programme ayant créée l'objet		
Object cycle	Cycle de l'objet		
By-line	Créateur de l'objet (nom du rédacteur, du photographe etc.) : 64 caractères	creator	
By-line Title	Titre du ou des créateurs		
City	Ville : 32 caractères		
Province/State	Province/ Etat : 32 caractères		
Country/Primary Location Code	Code du pays	coverage	spatial
Country/Primary Location Name	Libellé du pays : 64 caractères		
Original Transmission	Référence de la transmission		

10/10/2005

Reference			
Headline	Titre : 256 caractères	titre	Alternative (abréviation/traduction)
Credit	Crédit : 32 caractères maximum	Publisher Contributor ?	
Source	Source : 32 caractères	source	
Copyright Notice	Copyright : 128 caractères	rights	
Contact	Contact	Audience	Mediator ?
Caption/Abstract	Description, résumé, commentaire : pas plus de 2000 caractères	description	Table of contents (table des matières) Abstract (résumé)
Writer/Editor	Auteur de la description		
Image Type	Type d'image	Type (nature/genre du contenu)	

III. Champs IPTC non admis dans la Dublin Core

Ainsi, les champs de la Dublin Core n'apparaissant pas dans l'IPTC-NAA standard sont les suivants :

- format (extend – taille – et medium – support physique -)
- langage
- relations (13 raffinements)
- coverage (spatial et temporal)
- audience (mediator et education level)

Certains d'entre eux sont nécessaires à la description bibliographique de l'image. En outre, certains de ces champs non présents dans l'IPTC-NAA standards sont disponibles dans la version complète de l'IPTC, tels que le format et le langage à titre d'exemple.

Il s'avère à l'inverse que de nombreux champs de l'IPTC-NAA ne sont pas présents dans la Dublin Core. Ces éléments manquants sont pourtant d'une précision non négligeable concernant les objets iconographiques à décrire. A titre d'exemple, la Dublin core, en dépit de ses raffinements, ne dispose pas des champs suivants :

- Catégorie
- Sous catégorie
- Auteur de la légende
- Orientation de l'image (notamment dans l'objectif d'un usage scientifique pour les chercheurs)

Néanmoins, certains champs contenus dans l'IPTC ne sont pas indispensables, tels que :

- heure de création
- code du pays de prise de vue...

C'est pour cette raison qu'il existe des champs personnalisés non standards :

Le standard IPTC/IIM décrit 33 types de métadonnées (*Titre, Auteur, Mots-clés*, etc.). Chaque type est défini par un nom anglais et par un numéro interne nommé *DataSet* compris entre 1 et 254: *Mot-clé = Keyword = 25, Titre = Headline = 105*, etc. Il est donc possible de définir des informations IIM dont les numéros *DataSet* sont effectivement compris entre 1 et 254 mais différents des numéros utilisés par le standard IPTC/IIM: ce sont les informations personnalisées non standards.

Ces informations personnalisées doivent être utilisées avec précaution car elles risquent d'être perdues lorsque l'image est traitée par un logiciel qui ne les supporte pas.

Champs communs entre DC et IPTC-NAA

- mots clé
- date de disponibilité
- identificateur (?)
- créateur
- titre
- crédit
- source
- droits

10/10/2005

- description
- type (nature/genre du contenu)
- coverage spatial

Champs nécessaires dans l'IPTC

- catégorie
- numéro de référence (?)
- titre du créateur (photographe)
- ville
- état ou pays

Champs nécessaires dans la DC

- dates et raffinements
- format (taille, support physique)
- Portée du document (coverage) spatiale et/ou temporelle

IV. Logiciels et organismes

4.1. De nombreux logiciels sont capables de lire la norme IPTC-NAA et de l'encoder dans les images

- Soft Experience Kalimages et Kalimages Basic (gratuit), Win: base de données IPTC
- Adobe PhotoShop , File Browser - Win/Mac
- Armantis ArmadilloPHOTO/ArmadilloMEDIA , Win
- CodedColor PhotoStudio , Win
- FotoStation , Win/Mac
- Photools IMatch , Win
- PhotoThumb , Win
- Breeze Browser , Win
- Canto Cumulus , Win/Mac
- Cerious ThumbsPlus , Win
- Image Info Toolkit , Win/Mac
- SCC MediaGrid , Win/Mac
- Camera Bits Photo Mechanic , Win/Mac
- Graphic Converter , Mac
- ImageWalker , Windows, shareware
- iView MediaPro , Win/Mac
- J2S J2SEditor , Mac
- GrabbyWindow MetaPro , Mac
- Extensis PortFolio , Win/Mac
- EXIFUtils , Win/Mac/Linux
- Image Rodeo , Mac OSX
- Informator Pro , Mac
- Touchup , une applet Java par Peter Jones
- MaPiVi , Unix et Win, écrit en Perl/Tk par Martin Herrmann, licence **GPL**
- exiv2 , licence **GPL**
- IrfanView (avec son plugin IPTC), Win, **gratuit**
- XnView (avec son plugin IPTC), Win, **gratuit**
- Exifer , par Friedemann Schmidt, Win, **gratuit**
- PixVue , par Eamonn Coleman, Win, extension de l'Explorateur, traduit en français par *Soft Experience*, **gratuit**
- PhotoThumb IPTCExt , Win, extension de l'Explorateur, **gratuit**
- Rodeo Info , Mac OSX, affiche les informations IPTC, **gratuit**

4.2. Organismes utilisant la norme

- l'Agence France Presse (AFP)
- l'Associated Press (AP)
- Reuters

L'IPTC a défini un modèle global de données appelé *[IPTC-NAA Information Interchange Model](#)* (la version 4 date d'octobre 1997; elle est connue sous le nom *IIMV4*. La révision 4.1 date de Juillet 1999).

- Les noms et structures des champs IPTC/IIM dans les programmes d'édition ne sont pas identiques. *Headline* (n° 105) correspond par exemple au *Titre* dans le domaine de l'écrit. Or, à la suite des choix réalisés par Adobe dans Photoshop, *Object Name* (n° 5) est appelé *Titre (Title)* par la majorité des éditeurs...
Voir [IPTC - IIMv4 Fields mapped to Imaging Programs](#) qui compare la nomenclature des champs IPTC dans les logiciels suivants: Image Info Toolkit, Photoshop 7.01, Photoshop 6 et IrfanView.
Voir aussi le tableau de [Correspondance entre les informations IPTC/IIM et les champs disponibles de certains éditeurs IPTC](#)
- *Date & Time Created* (n° 55 et 60), c'est-à-dire la Date et l'Heure de création de l'image ne correspondent pas aux besoins des agences magazines qui travaillent par reportages souvent étalés sur plusieurs jours.
- *Release Date & Time* (n° 30 et 35), c'est-à-dire la Date et l'Heure de disponibilité de l'image, n'ont guère d'intérêt que dans un contexte "presse quotidienne" où une image peut être sous embargo de publication avant une certaine date. Par contre, d'autres professions auraient besoin d'autres informations concernant l'image mais non prévues dans la nomenclature *IPTC-NAA*.
- *By-line* (n° 80) correspond à *Auteur*, terme adopté majoritairement dans le domaine de l'écrit ou par d'autres professions de l'image.
Rappelons que le *Dublin Core* a retenu *Creator* (v. [ci-dessus](#)).
- Certains champs sont absents dans certains éditeurs: *Edit Status* (n° 7), *Contact* (n° 118)
- *Caption/Abstract* (n° 120) correspond à *Description*, *Résumé*, *Commentaire*, etc.
- *Category* (n° 15) et *Supplemental Category* (n° 20) sont considérés comme obsolètes dans l'*IIMV4* et sont pourtant encore largement utilisés.
- Il est possible de définir des champs personnalisés de type IPTC sur certains logiciels

10/10/2005

[IPTC Core](#) version 1.0 est le standard qui succède à l'*Information Interchange Model* (IIM version 4.1) de l'IPTC. Il est donc destiné à remplacer la technologie des "informations [IPTC](#) classiques" (dérivées de l'IIM) et s'appuie sur le framework [XMP](#) défini par Adobe en 2001. Il a été annoncé officiellement en Mars 2005.

IPTC Core est issu du groupe de travail [IPTC For XMP \(IPTC4XMP\)](#) constitué par l' [IPTC](#) , [IDEAlliance](#) , et la société Adobe.

Les avantages de la technologie XMP sur les informations IPTC "classiques" sont nombreux: pas de limitation de taille des champs, pas de problèmes d'accents (codage Unicode), possibilité de légendes multilingues, extensibilité et personnalisation des métadonnées.

IPTC Core définit un ensemble de métadonnées exprimées en XMP et destiné à faciliter la transition des informations [IPTC](#) "classiques" vers le nouveau standard XMP.

IPTC Core reprend la plupart des informations de la version précédente (IIM v4.1) comme les champs *Keywords*, *By-line*, *Headline*, *Caption*, etc. Certains champs sont abandonnés: *Urgency*, *Category*, *Supplemental Categories*. Enfin, de nouveaux champs font leur apparition: *Intellectual Genre*, *Rights Usage Terms*, *IPTC Scene*, etc.

VI. Récapitulatif des avantages et inconvénients de la norme IPTC-NAA :

AVANTAGES	INCONVENIENTS
<ul style="list-style-type: none"> • informations et document indissociables • données très complètes prenant en compte différents aspects • facilité de lecture et possibilité de recherche multicritère 	<ul style="list-style-type: none"> • Le jeu d'informations IPTC/IIM est fermé et non extensible de manière standard. Il n'est donc pas possible de représenter de façon simple des informations importantes comme la notion de <i>Reportage</i>, non prévue dans le jeu IPTC/IIM. Il est possible néanmoins de définir des informations IIM personnalisées et non standards (voir ci-dessous) mais avec le risque de perdre ces informations lorsque l'image est traitée par un logiciel qui ne les supporte pas. Kalimages permet de gérer 16 champs IIM personnalisés. • La description multilingue des images est également difficile à l'aide du standard IPTC/IIM. • Enfin, les légendes rédigées dans une écriture non latine et les légendes qui possèdent des caractères accentués peuvent aussi poser des problèmes (voir ci-dessous). <p>De nouvelles techniques comme XMP permettent de s'affranchir de ces inconvénients</p>

Ainsi, la norme IPTC-NAA est spécialisée dans le traitement des images, contrairement à la Dublin Core qui reste très générale.

En revanche, cette dernière dispose de raffinements permettant une précision convenable dans la description des éléments iconographiques.

L'IPTC-NAA, elle, est certainement trop spécialisée dans les détails techniques liés à la photographie, créée initialement pour et par les agences de presse.

Ainsi, dans le cadre de notre projet, les avantages de l'une ne combrent pas les inconvénients de l'autre. Il est donc à envisager un mix entre les deux normes : l'IPTC-NAA pour la base de données documentaire, et la Dublin Core pour l'interface du site web, indexable par les plus gros moteurs de recherche tels que Google.

La correspondance entre les différents champs est donc à étudier précisément, en définissant les champs à choisir pour la description physique et intellectuelle de l'icône.

Par ailleurs, il est à se demander s'il est préférable d'utiliser la norme IPTC-NAA ou sa nouvelle version : l'IPTC Core (ou XMP) ?

<http://peccatte.karefil.com/Software/Metadata.htm#IPTC>

http://www.controlledvocabulary.com/imagedatabases/iptc_naa.html

http://www.controlledvocabulary.com/imagedatabases/iptc_naa.html#standards

<http://peccatte.karefil.com/Kalimages/FR/KalimagesFAQ-FR.html>

http://www.openweb.eu.org/articles/dublin_core/

<http://www.iptc.org/IIM/>

<http://peccatte.karefil.com/software/IPTCPropositionsPixpalace.pdf>