

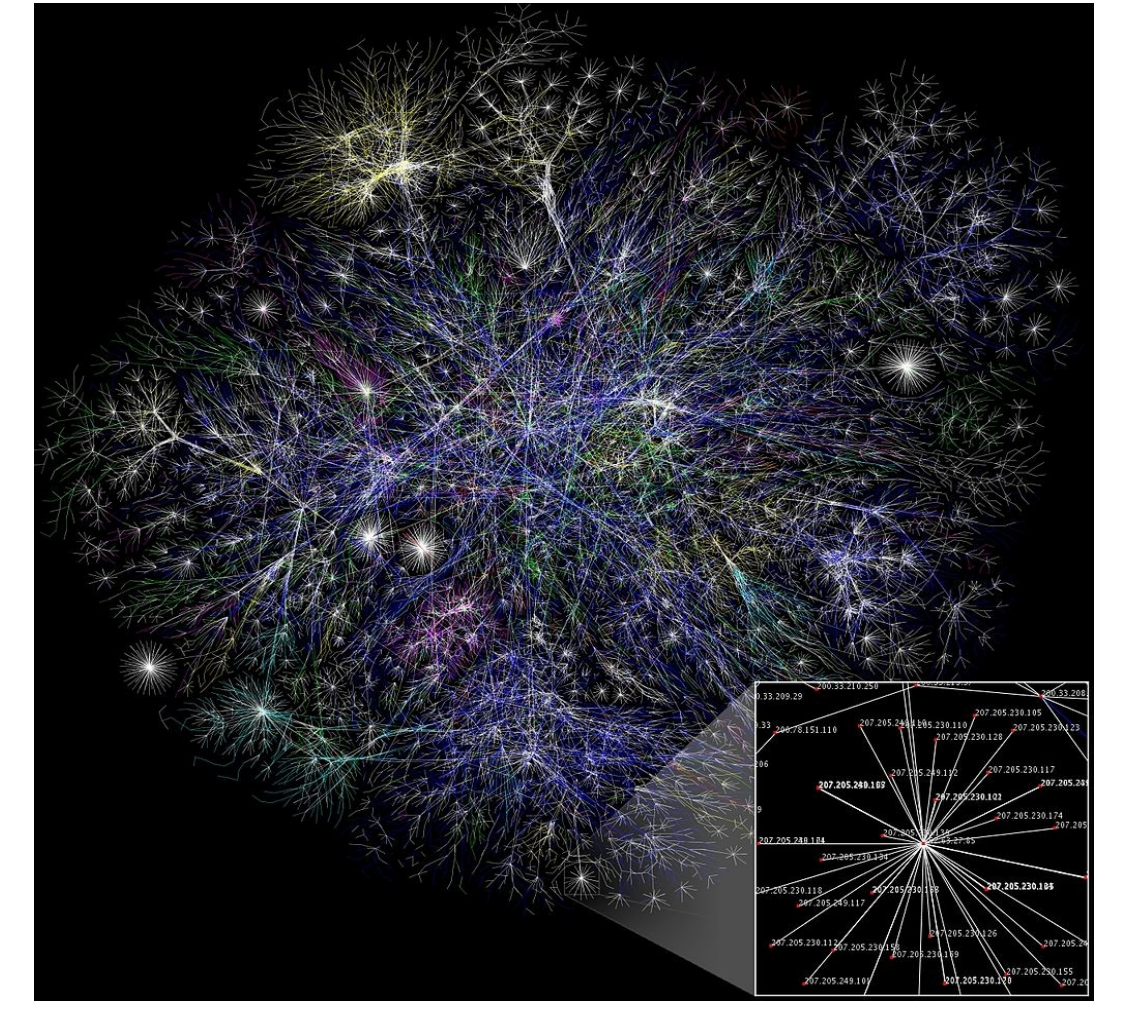
Format conteneur et métadonnées ufologiques

Jean-Marc Wattecamps

Comité Belge d'Etude des Phénomènes Spatiaux (COBEPS)

Problématique

Internet est une immense banque de données universelle. Il concrétise progressivement l'idée de Paul Otlet (1868-1944), co-créateur de la Classification Décimale Universelle, pour qui l'une des tâches essentielles des travailleurs intellectuels était de réaliser « une Encyclopédie universelle et perpétuelle » ayant pour collaborateurs « tous les savants de tous les temps et de tous les pays ». Cet outil fabuleux offre des perspectives de conservation de la mémoire certes, mais aussi d'échanges et de recherches collectives, notamment dans le cadre de l'ufologie. Internet offre l'avantage de non seulement permettre la diffusion de tableaux de synthèse, de graphiques mais aussi de textes de notifications originales (signalements), de verbatims, d'interviews, de photos, de vidéos ainsi que des résultats scientifiques. Le réseau constitue donc l'infrastructure mondiale qui met potentiellement des données à disposition de tous. Celles-ci sont fournies par ceux qui le souhaitent et en ont l'opportunité, mais pour le moment de façon non coordonnée. Nous proposons la création d'un format conteneur (*wrapper* ou *container* en anglais). Ces conteneurs seraient des fichiers zip ou rar. Ils contiendraient, suivant une nomenclature à établir, tous les éléments d'un dossier d'enquête : témoignages intégraux ou rapports d'audition, photos, vidéos, analyses, rapports d'enquêtes. Les contenus seraient aux formats standards comme par exemple pdf et jpeg. Ces fichiers conteneurs renfermeraient également des métadonnées à la façon d'un exif pour les fichiers de photographie numérique.



Contenu du fichier

Le fichier conteneur est destiné aux données ufologiques. Ces données, selon Hynek¹, sont les rapports ou les notifications. Il s'agit bien des déclarations « brutes », des croquis accompagnés de tous éléments de preuve tels que photos, films, enregistrements radar... Nous ajoutons cependant à ce matériel : tous documents susceptibles d'aider à l'analyse de la notification réalisée par les enquêteurs qui ont recueilli le témoignage et qui l'ont exploité (cartes, fiche explicative sommaire, compte rendu d'enquête simple, rapport d'enquête approfondi...). Il s'agit en quelque sorte **du dossier d'instruction d'un cas ufologique**.

Chaque fichier conteneur correspond à une notification, c'est-à-dire les témoignages d'une ou de plusieurs personnes qui étaient rassemblées au même endroit, qui ont pu interagir entre-elles, et qui ont fait l'observation d'un PAN. En voici la structure :

1. TEMOIGNAGES

• TEXTES => format TXT (ASCII)

Il s'agit de textes bruts, non mis en forme, permettant une lecture par la plupart des programmes informatiques, dont des outils lexicométriques.

Les contenus reprennent les témoignages bruts sous forme de lettres, d'e-mails, de verbatims, de procès verbaux d'audition.

• IMAGES => format JPEG

C'est un format de compression mis au point pour la photographie qui permet de reproduire des nuances parfois très peu visibles.

Les contenus sont les schémas, dessins et cartes établis par le témoin. Ces documents sont de préférence scannés à plat ou photographiés.

• AUDIOS => format MP3 (ISO MPEG - 3 couches - Audio)

Il s'agit d'un codage compressif de fichier sonore qui détruit une partie du signal. La qualité requise demande un codage minimum de 192kb/sec.

Les contenus concernent les enregistrements de l'audition des témoins.

• VIDEOS => format MPEG4

Il s'agit d'un format conteneur vidéo qui reprend différents éléments : images, sons, sous-titres... La qualité requise demande au minimum la définition correspondant à une image TV PAL/SECAM ou NTSC.

Les contenus concernent les enregistrements de l'audition des témoins.

2. PIECES A CONVICTION

• AUDIOS => format selon l'appareil ayant produit l'enregistrement si possible en accordant la préférence aux fichiers bruts de type RAW ou aux formats sans pertes FLAC, WAV. Pour des enregistrements amateurs, le plus courant reste le MP3. Une copie MP3 est conservée.

Il s'agit d'enregistrements par le témoin de sons produits ou en relation avec le PAN.

• PHOTOS => format selon l'appareil ayant produit l'enregistrement si possible en priorisant des fichiers bruts de type RAW. Pour des enregistrements amateurs, le plus courant reste le JPEG. Une copie JPEG réduite et de basse résolution est conservée pour la visualisation.

Il s'agit des photos du PAN réalisées par les témoins.

• VIDEOS => format selon l'appareil ayant produit l'enregistrement. Ces formats sont très variés. La vidéo exige de très grands flux d'images qui dépendent de la qualité de l'électronique de la caméra et de grandes capacités de stockage. Il n'y a pas réellement de fichiers bruts standardisés, il faut privilégier la haute définition et les faibles taux de compression.

Il s'agit de vidéos du PAN réalisées par les témoins.

• MESURES => il s'agit des enregistrements de capteurs divers. Les paramètres sont mesurés par intervalle de temps ou d'espace. L'idéal est de disposer du fichier de mesures en format texte brut. Donc des données ASCII.

Il s'agit de mesures réalisées par les témoins.

3. DOCUMENTS D'ENQUETE

• Tous les documents tels que les rapports des enquêteurs, les cartes géographiques, les photos de l'enquête sont intégrés dans un fichier PDF continu.

Ce fichier conteneur contiendra donc toujours ces 3 dossiers en racine (témoignages, pièces, documents) et des sous-dossiers correspondant à chaque point repris ci-dessus sous rubriques. Ces dossiers contiennent ou non des fichiers.

Métadonnées

Les métadonnées sont des informations structurées qui accompagnent le fichier et sont visibles dans les propriétés du fichier. Elles sont peu nombreuses dans la plupart des cas et plutôt contextuelles. Elles doivent permettre de se faire une idée de l'importance du cas avant d'ouvrir le contenu.

Voici les données proposées :

- date de notification
- date d'observation
- heure de l'observation
- date de clôture de l'enquête
- latitude de la position des témoins
- longitude de la position des témoins
- pays
- localité
- nom du responsable de l'enquête et organisation
- adresse web du fichier ou courrier électronique
- type d'enquête réalisée
- nombre de témoins
- classement GEIPAN²
- typologie Hynek
- indice d'étrangeté Poher³
- indice de crédibilité Poher

Conclusions

La réalisation de ces fichiers demanderait une certaine discipline de la part des fournisseurs de données ufologiques. Elle nécessiterait un supplément de travail conséquent. Voici ce qui pourrait justifier cet investissement :

- uniformité de la structure des données.
- maintien de la diversité des sources.
- facilité de partage et d'échange.
- permanence et conservation des données.
- démultiplication intégrale et facilitée de l'information.

Pour faciliter le travail des ufologues, il serait utile de développer des logiciels spécifiques qui pourraient, d'une part, réduire l'effort à fournir pour produire ce type de fichier et d'autre part, pour en visualiser les contenus et les exploiter dans des recherches et des analyses.

Ce projet nécessiterait la création d'un groupe de travail interdisciplinaire qui approfondirait cette norme et ces outils.

Références

1. J. A. Hynek, *Les objets volants non identifiés mythe ou réalité?*, J'ai lu, Paris, collection "L'Aventure mystérieuse", n° A327, 1975, 413 p.
2. <http://www.cnes-geipan.fr/>
3. C. Poher, *Etude statistique des rapports d'observation du phénomène OVNI*, 1976, http://www.cnes-geipan.fr/typo3conf/ext/dam_frontend/pushfile.php?docID=1627 (dernière consultation le 31/12/2011).