

# Exploration des archives du Ceo-Mag

## Une difficile recherche d'informations

par André C.

La grosse collection des Ceo-Mag est une incroyable source de savoir sur L'Oric. Vu le nombre de pages accumulées au fil des années, il est bien difficile d'avoir une idée claire sur l'étendue de ses richesses. Non seulement, le mag est incollable sur le software et le hardware, mais on y trouve aussi des lectures divertissantes. Pas plus tard qu'hier, en cherchant tout autre chose, je suis tombé sur "Oric : une garantie 15 ans sans souci !", par Fabrice F. (Ceo-Mag n°105 page 5). Si vous avez un petit moment, je vous en recommande la lecture...

### Note préliminaire:

Les archives du Ceo-Mag couvrent toute la collection française du n°1 (1990) à aujourd'hui. En outre, pour les 80 premiers numéros, il existe une version anglaise.

Les originaux de tous les mags anglais, ainsi que des mags français du n° de 1 à 111 (c'était des maquettes montées avec colle et ciseaux) ont été perdus. Pour ces mags, nous ne possédons que de mauvaises photocopies (à l'époque la qualité des photocopieuses était déplorable). Il s'en suit que les archives correspondantes ont été numérisées au scanner et que le résultat n'est pas toujours probant, malgré le soin extrême avec lequel nous avons procédé. Autre très gros inconvénient, les pdf correspondants sont en mode "image" et non en mode "texte". Il est donc impossible d'y faire des recherches à l'aide de mots clefs.

Pour les mags du n°112 à aujourd'hui, nous possédons les originaux sous forme numérique. Les pdf correspondant sont au format "texte", d'emploi infiniment plus souple.

L'existence de ces différentes versions du Ceo-Mag ainsi que la nécessité de disposer d'outils d'exploration des archives ont entraîné la création de diverses méthodes complémentaires de recherche.

### 1) Le moteur de recherche de Simon

Pour le début de la version française (du n°1 de 1990 au n°77 de 1996), Simon a élaboré un moteur de recherche dans lequel il a indexé

manuellement les pdf "partiels" au format "image".

Cette version "Moteur de Simon" a été largement diffusée par le CEO. Les fichiers d'indexation et le moteur sont disponibles en téléchargement sur <http://andre.cheramy.net/ceomag/Index.htm>.

Pour la mettre en place sur votre disque dur (ou DVD ou clef USB), télécharger et décompresser le fichier "[MoteurSimon.zip](#)", puis placer dans les 7 répertoires "1990" à "1996" (qui sont livrés vides) les pdf "partiels", dont vous disposez déjà si vous êtes un lecteur de longue date du Ceo-Mag. Sinon, il vous faudra les télécharger sur le site [http://www.oric.org/ceomag\\_download/](http://www.oric.org/ceomag_download/) (munissez-vous de votre login et de votre mot de passe). Cette deuxième possibilité vous assure de disposer de la dernière version, corrigée d'éventuels défauts. Au final, vérifiez que vous avez bien 52 pdf pour le répertoire "1990", 99 pour "1991", 114 pour "1992", 113 pour "1993", 102 pour "1994", 105 pour "1995" et 85 pour "1996" (inutile de mettre en place les pdf correspondant aux numéros 78 à 82, qui ne sont pas, à ce jour, exploités par le moteur). Cela représente en tout 167 Mo et 670 fichiers.

Le point d'entrée se fait par le fichier test.htm.

Rendons hommage à Simon qui a relu, mot à mot, ces 77 Ceo-Mag et les a indexé dans sa base de données.

### 2) Moteur de recherche de Jérôme

Pour les mags présents sur [http://www.oric.org/ceomag\\_download/](http://www.oric.org/ceomag_download/) Jérôme a mis au point un moteur de recherche, capable d'explorer les pdf "partiels" (un article = un pdf). Ce système très performant connaît actuellement un problème technique, suite à un changement de serveur internet. Espérons que Jérôme pourra rétablir la situation. NDLR. Le site [http://www.oric.org/ceomag\\_download/](http://www.oric.org/ceomag_download/) est la seule référence en matière d'archives du Ceo-Mag. Ce site est constamment mis à jour afin de corriger les éventuels problèmes signalés par les adhérents du CEO.

**3) Utilisation de la fonction Recherche d'Acrobat Reader (ou équivalent)** sur tous les pdf d'un répertoire. Rappelons que, à part quelques exceptions (rééditions ponctuelles), tous les articles des mag n°1 à n°111 sont sous forme de PDF "image" et échappent donc à la recherche de mots. Par contre, c'est facile et agréable pour les mags postérieurs au n°111.

Pour pallier à cette limitation, nous avons tenté de passer à l'OCRisation des pdf en mode "image" pour les transformer en pdf en mode "texte". Ce procédé est impossible à finaliser proprement sans passer des années à corriger les inévitables erreurs d'OCRisation (aggravées par la mauvaise qualité des photocopies de départ). Mais une OCRisation "en gros" peut aider à sortir des infos avec Acrobat Reader (ou équivalent). Notons qu'un mag ou un article OCRisé pourra toujours être mis au propre en cas de besoin (récupération de listing etc.). Le chantier d'OCRisation est en cours. Patience.

#### **4) Indexation au format html du Ceo-Mag version anglaise**

C'est l'ensemble de tous les mags traduits par Laurent. Il s'agit de pdf "full" (un n° du mag = un pdf) en mode "image". Outre la lecture page par page, il n'existe qu'un seul mode d'exploration (très primitif) basé sur quelques index au format html.

Cette version anglaise a été largement diffusée par le CEO. Les fichiers d'indexation sont disponibles en téléchargement sur <http://andre.cheramy.net/ceomag/Index.htm>.

Pour mettre le tout en place sur votre disque dur (ou DVD ou clef USB), télécharger et décompresser le fichier "[CeoMagUK.zip](#)", puis placer dans les 7 répertoires "1990" à "1996" (qui sont livrés vides) les pdf "full", dont vous disposez déjà si vous êtes un lecteur de longue date du Ceo-Mag. Sinon, il vous faudra les télécharger sur le site [http://www.oric.org/ceomag\\_download/](http://www.oric.org/ceomag_download/) (munissez-vous de votre login et de votre mot de passe). Je rappelle que cette deuxième possibilité vous assure de disposer de la dernière version, corrigée d'éventuels défauts. Au final, vérifiez que vous avez bien 6 pdf pour le répertoire "1990", 10 pour "1991", 11 pour "1992", 11 pour "1993", 11 pour "1994", 11 pour "1995" et 11 pour "1996". Cela représente en tout 50 Mo et 71 fichiers.

Le point d'entrée se fait par le fichier

"English\_Index.htm".

Rendons hommage à Laurent qui a traduit ces 71 Ceo-Mag du français en anglais et a dressé un inventaire des principaux articles publiés.

#### **5) Indexation au format html du Ceo-Mag version française**

Cette indexation porte sur l'ensemble hétérogène de tous les mags français du n°1 à aujourd'hui, dans leur meilleure version. Tous les fichiers pdf sont sous forme "full" (un mag = un pdf). Les noms de fichier sont de la forme mmm.pdf où mmm représente les numéros de mags (3 digits). Par exemple 123.pdf.

Cette indexation a été mise à jour du n°1 (1990) au n°272 inclus (fin 2012).

MAIS, il faut savoir que cette indexation est très artisanale, basée sur des versions antérieures, péniblement mises à jour sans remise en cause de sa structure et de sa logique. Le résultat est une usine à gaz qui n'est pas facile à utiliser. Un conseil, si vous ne trouvez pas ce que vous cherchez, effectuez une recherche du mot-clef (du texte) désiré dans le contenu des fichiers \*.htm. La fonction "Rechercher" de Windows, jusqu'à XP inclus, était particulièrement facile à utiliser. Pour les versions ultérieures de Windows, il faudra ruser : au lieu de taper votre mot-clef dans un champ spécial, comme avant, il faut taper System.Keywords : "mot-clef" dans la petite fenêtre recherche de l'explorateur de fichiers).

Vous trouverez ce système d'indexation html à l'url <http://andre.cheramy.net/ceomag/Index.htm>. Pour le mettre en place sur votre disque dur (ou DVD ou clef USB), télécharger et décompresser le fichier "[CeoMag.zip](#)", puis placer dans les répertoires "1990" à "2012" (qui sont livrés vides) les pdf "full" dont vous disposez déjà si vous êtes un lecteur de longue date du Ceo-Mag. Sinon, il vous faudra les télécharger sur le site [http://www.oric.org/ceomag\\_download/](http://www.oric.org/ceomag_download/) (munissez-vous de votre login et de votre mot de passe). Ici encore cette deuxième possibilité vous assure de disposer de la dernière version, corrigée d'éventuels défauts. Au final, vérifiez que vous avez bien 6 pdf pour le répertoire "1990", 10 pour "1991", 11 pour "1992" à "2006", "2009" et "2011", 12 pour "2007", "2008", "2010" et "2012". Cela représente en tout 773 Mo et 251 fichiers.

L'entrée dans le système d'indexation se fait par [GOCD.htm](#). Bonne recherche...