



OSSIR – 06/10/2003



La fuite d'informations dans les documents propriétaires



Patrick CHAMBET

Eric FILIOL

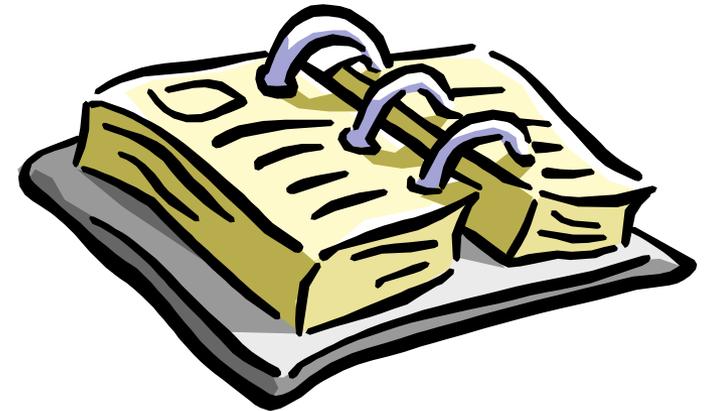
Eric DETOISIEN

patrick.chambet@edelweb.fr – <http://www.chambet.com>

efiliol@esat.terre.defense.gouv.fr

valgasu@rstack.org

- **Objectifs**
- **Généralités**
- **Quelques exemples**
 - Acrobat
 - Microsoft Word
 - Autres
- **Recommandations**
- **Conclusion**



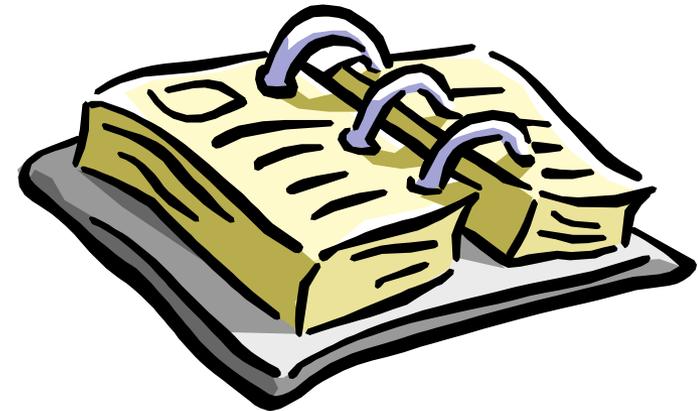
Objectifs



- **Présenter certaines caractéristiques importantes des documents propriétaires complexes**
- **Décrire les faiblesses de certains formats propriétaires**
- **Présenter quelques cas concrets de fuites d'informations**
- **Décrire les vulnérabilités liées à la signature numérique de documents complexes**
- **Présenter des recommandations permettant de contrer la fuite d'informations**
- **Conclure sur la sécurité liée aux documents propriétaires**



- Objectifs
- ✓ • Généralités
- Quelques exemples
 - Acrobat
 - Microsoft Word
 - Autres
- Recommandations
- Conclusion



Généralités (1/2)



- **Les formats propriétaires de documents sont de plus en plus complexes**
 - **Modèle objet élaboré**
 - **Non documenté**
 - **Reverse engineering partiel**

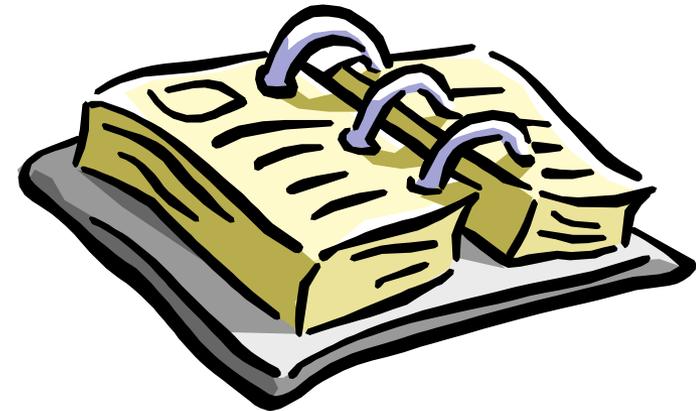
- **Ex: quelques headers Word**
 - **DC A5 65 00**
 - **DC A5 68 00**
 - **97 A6 68 00**
 - **EC A5 C1 00**

Généralités (2/2)



- **La tendance est à l'inclusion d'informations de nature très diverse, à l'insu de l'utilisateur**
 - Informations personnelles
 - Informations "marketing"
 - Temps d'utilisation
 - Habitudes d'utilisation
 - Relations avec d'autres documents, applications, ressources réseau (y compris Internet)
- **Contenu actif, pouvant modifier l'apparence des documents en fonction de l'environnement dans lequel ils sont ouverts**
 - **Problème de la signature de documents actifs**

- Objectifs
- Généralités
- ✓ • Quelques exemples
 - Acrobat
 - Microsoft Word
 - Autres
- Recommandations
- Conclusion



- **Génération des documents PDF**
 - **Outils**
 - Acrobat Distiller
 - PDF Maker
 - Outils tiers (Fineprint PDF Factory, ...)
 - **Méthodes**
 - Conversion de documents en passant par PostScript
 - Scan direct de documents

- **Deux exemples de révélation d'informations**
 - **Recadrage d'un document scanné**
 - Permet de faire apparaître des parties supplémentaires du document qui auraient dû être éliminées
 - Dans *Document/Recadrer des pages*, cliquer sur *Remettre à zéro*
 - **Suppression de masques en surimpression**
 - Permet de révéler des informations masquées volontairement

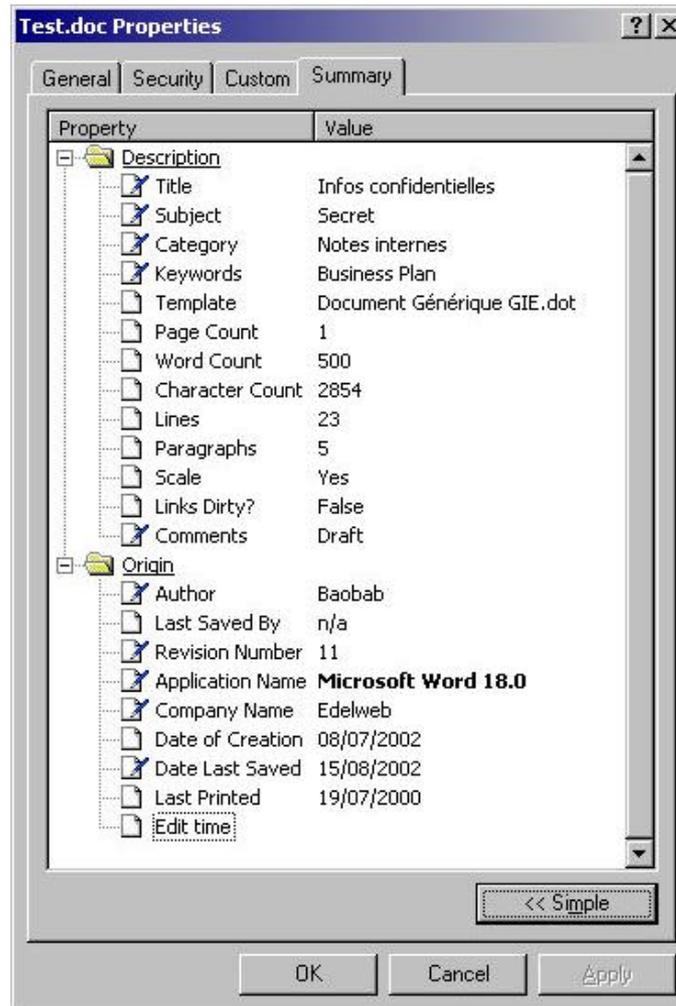
DEMONSTRATIONS



Microsoft Word (1/9)



- Propriétés des documents



- **Informations lisibles directement:**
 - Nom de l'auteur
 - Entreprise de l'auteur
 - Date et heure de création
 - Temps passé à l'édition
 - Heure d'impression
 - Etc...
- **Si un document de 100 pages a un temps d'édition de 5 minutes, c'est qu'il provient d'un copier-coller !**
- **Attention au mode de suivi des modifications: accès aux versions antérieures !**
 - Exemple: affaire Alcatel



DEMONSTRATION

Microsoft Word (3/9)



- **Ouverture du document avec un éditeur hexa**

- Nom des rédacteurs **successifs**

- Nom de la machine

- Chemin complet du document sur les disques des rédacteurs successifs

`C:\Documents and Settings\Stagiaire Dupont\
Confidentiel\Clients mauvais payeurs\Contrat.doc`

- Chemin complet du modèle de document

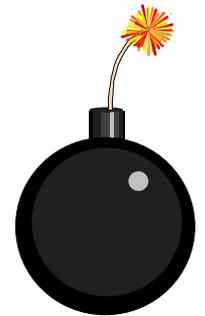
`\\SRV_FICH\PUBLIC\MODELES-WORD\Document
Générique GIE.dot`

=> On en déduit le nom du serveur de fichiers de l'entreprise

- Serveurs d'impression et imprimantes

`\\SRV_PDC\HPPCL5MS LaserJet 4 Plus`

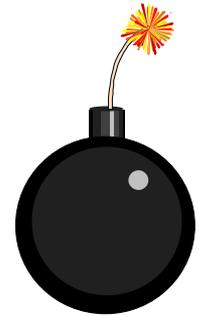
=> On en déduit le nom du PDC du domaine



Microsoft Word (4/9)



- **Noms des fichiers inclus dans le document**
 - **Ex: fichiers images**
- **GUID (Global Unique Identifier)**
 - **Chercher après « `_PID_GUID` » :**
`{F165CB92-D166-12D5-AB67-0010A41432AF}`
 - **Les 12 derniers chiffres sont l'adresse MAC de la carte réseau !**
 - **Présent dans les documents Office mais aussi Visual C++, les ActiveX, etc...**



L'outil « Analyse doc word »



```
C:\ Patrick's Shell

C:\>"Analyse doc word 2701.exe"

*****  ETAT DES SAUVEGARDES D'UN DOCUMENT WORD  *****
*****  pour les versions posterieures a Word 6  *****

Veuillez entrer le nom du fichier word a analyser (avec extension .doc)
Nota : il doit etre dans le meme repertoire et ne pas comporter d'espaces.
Nom du fichier : document1.doc

Les resultats de la recherche sont dans result.txt .

C:\>_
```

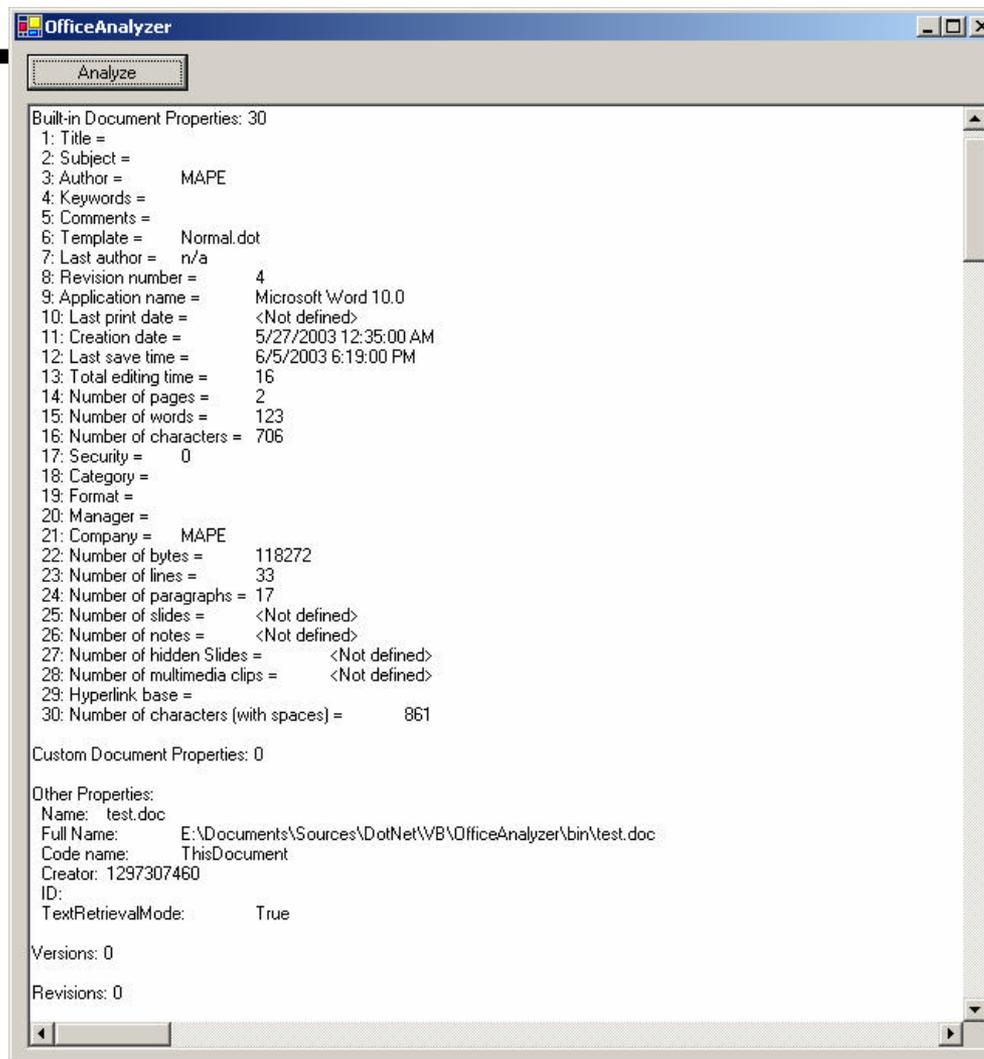


DEMONSTRATION

L'outil « OfficeAnalyzer »



EdelWeb



DEMONSTRATION

- **Inclusion d'informations**

- La directive INCLUDETEXT peut être utilisée pour inclure automatiquement le contenu d'un document du disque dur dans le document courant

```
{ IF { INCLUDETEXT { IF { DATE } = { DATE }  
"C:\\confidentiel.txt" "C:\\confidentiel.txt"  
} \* MERGEFORMAT } = "" "" \* MERGEFORMAT }
```



DEMONSTRATION

Microsoft Word (6/9): Contenu actif



- Exemple de contenu actif: modification furtive de l'apparence d'un document
- Test:

```
{ IF { FILENAME \* MERGEFORMAT { DATE } }  
= "contrat.doc" "blanc" "noir" \*  
MERGEFORMAT }
```

Microsoft Word (7/9): Signature numérique



- **Problématique en cas de signature numérique du document**
 - Alice et Bob établissent un contrat qu'ils signent tous deux
 - Plus tard, Bob veut se référer au contrat, qui paraît alors ne plus contenir du tout les même clauses
 - Pourtant, la signature du document est toujours valide
- **La signature d'un tel document équivaut à signer un chèque en blanc !**

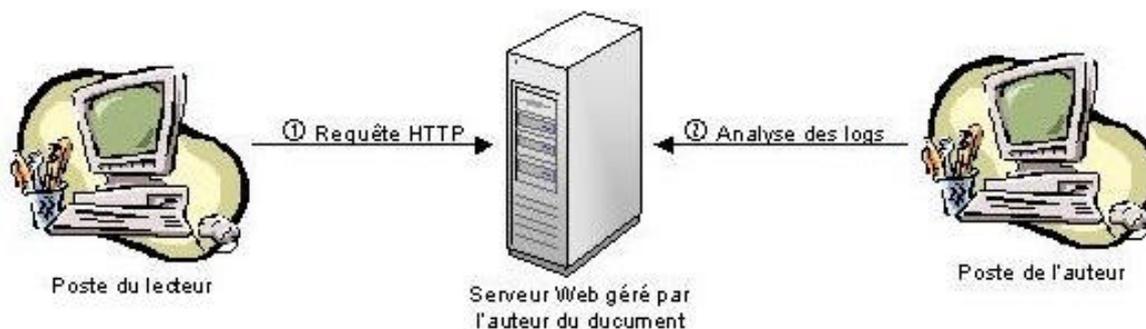


Microsoft Word (8/9): Signature numérique



- **Le contenu brut en hexa du document propriétaire est inchangé**
 - **La signature du document est donc toujours valide**
- **L'apparence du contenu du document propriétaire dans son éditeur a changé**
- **L'apparence à l'écran et l'apparence à l'impression peuvent être volontairement différentes**

- Les Word bugs



- **Permettent à l'auteur de recueillir de l'information sur les *lecteurs* des documents (processus inverse)**
 - Moment de la lecture
 - Lieu de la lecture (adresse IP)
 - Informations diverses sur l'identité et sur l'environnement du lecteur (logiciel utilisé, langue, etc...) et sur son type de connexion Internet

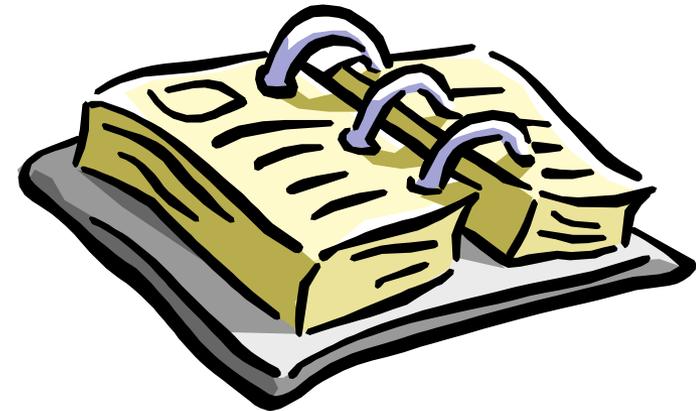
Autres exemples



- **WordPerfect enregistre toutes les étapes de création des documents**
 - On peut revenir aux états antérieurs d'un document en annulant les dernières modifications !
- **MS Outlook et MS Exchange**
 - Sous certaines conditions, envoi d'un fichier winmail.dat révélant le chemin complet de la boîte aux lettres de l'expéditeur (fichier .PST)
- **Les spywares**
 - Ex: MediaPlayer et RealPlayer, Windows XP, ...
 - Cf TCPA/Palladium et la FAQ de Ross Anderson



- Objectifs
- Généralités
- Quelques exemples
 - Acrobat
 - Microsoft Word
 - Autres
- ✓ • Recommandations
- Conclusion



Recommandations (1/3)



- **Ne pas diffuser un document ayant été retouché (très contraignant)**
- **Recréer les documents ex nihilo avant diffusion publique**
- **Opter pour un traitement de texte libre et compatible avec les leaders du marché**
 - **StarOffice**
 - **OpenOffice**
- **Utiliser un firewall personnel pour interdire certaines applications d'accéder à Internet**

Recommandations (2/3)



- **Dans MS Word**
 - **Désactiver l'enregistrement rapide**
 - **Désactiver le suivi des modifications**
 - **Désactiver toutes les macros (y compris signées)**
 - **Configurer les fichiers modèles (.dot) en read-only**
- **Depuis Office XP/2002**
 - **Cocher « Supprimer les informations personnelles de ce document lors de la sauvegarde »**
 - **Cocher « Avertir avant d'imprimer, de sauvegarder ou d'envoyer un fichier qui contient du suivi de modifications ou des commentaires »**

Recommandations (3/3)



- **Signature numérique**
 - Ne pas signer un document potentiellement dynamique: DOC, XLS, ...
 - Préférer pour la signature des documents à la structure moins riche: RTF, ...
 - Etudier avec prudence les formats que vous ne connaissez pas avant d'accepter de les signer ou d'en accepter la signature
- **Sensibiliser les utilisateurs**
- **Il est recommandé d'inclure la problématique des documents complexes dans les politiques de sécurité**

Conclusion

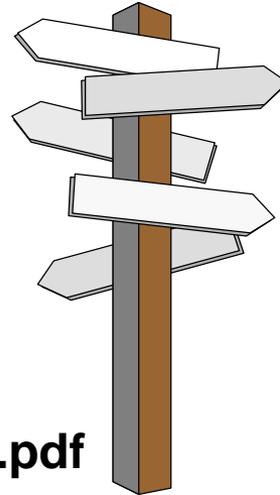


- **La tendance n'est pas à la simplification des formats propriétaires (malgré le succès mitigé de XML)**
- **Les spywares et le recueil de données marketing ont de beaux jours devant eux**
- **Les cas évoqués sont graves car ils se produisent aussi dans des environnements sensibles (entreprises, administrations)**
- **Toute organisation doit donc considérer la gestion des documents propriétaires en fonction du degré de confidentialité de ses informations**
- **Il est recommandé d'inclure cette problématique dans les politiques de sécurité**

Pour aller plus loin... (1/2)



- **Affaire Alcatel**
 - <http://www.landfield.com/isn/mail-archive/2001/Apr/0096.html>
 - Document Alcatel:
http://web.morons.org/external/CPE_statement.doc
- **Fichiers d'exemples**
 - <http://www-rocq.inria.fr/codes/Eric.Filiol/SSI/AdobeTestFile.pdf>
 - http://www-rocq.inria.fr/codes/Eric.Filiol/SSI/misc7_word.zip
- **Exchange**
 - <http://support.microsoft.com/default.aspx?scid=kb;en-us;298917>
 - <http://support.microsoft.com/default.aspx?scid=kb;en-us;259037>
 - <http://support.microsoft.com/default.aspx?scid=kb;en-us;138053>



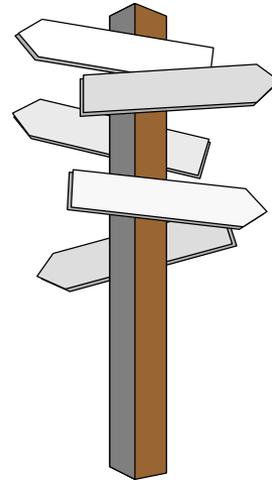
Pour aller plus loin... (2/2)



- **TCPA**
 - <http://www.cl.cam.ac.uk/~rja14/tcpa-faq.html>
 - En Français: <http://www.lebars.org/sec/tcpa-faq.fr.html>
- **MISC** (premier journal technique français sur la sécurité des SI)



- <http://www.miscmag.com>



Questions

