

Los Metadatos como amenaza a la Privacidad

Por Juan F. Mancebo¹

jmancebo@infodmz.com

1. Introducción

Durante las Séptimas Jornadas Académicas del Instituto de Derecho Informático de la Universidad de la República del Uruguay realizadas en mayo de 2006, ha quedado de manifiesto la importancia que el derecho a la protección de los datos personales ha tomado en nuestro país. La justificada preocupación de los profesionales del Derecho y de la Informática por esta temática, antecede a la preocupación que a nivel del ciudadano comienza a percibirse por proteger su intimidad de los riesgos, que significan la recolección indiscriminada de datos personales y el cruzamiento de esta información en grandes bases de datos.

Una experta en el tema, la Dra. María José Viega en su ponencia “Protección de Datos y Delitos Informáticos” realizada en un congreso anterior, proporciona a este trabajo, un marco teórico adecuado, desde la perspectiva del Derecho cuando dice:

“En ambas circunstancias, surge la libertad informática como un nuevo derecho de autotutela de la propia identidad informática. Este derecho otorga a las personas la posibilidad de controlar sus propios datos. El derecho a la intimidad a evolucionado hacia un nuevo concepto: la privacidad, la cual se ve seriamente violentada por las nuevas tecnologías, al punto tal que se ha llegado a hablar de la existencia de un hombre de cristal.”²

Efectivamente el derecho a la privacidad se ve seriamente comprometido por el uso de las nuevas tecnologías de la información, pero quizás más que nada por el desconocimiento de todas las implicancias que su utilización tiene. En efecto, es la falta de aplicación de procedimientos de trabajo, -lo que en Seguridad de la Información se denominan Políticas de Seguridad- lo que ocasiona el problema y quizás no la tecnología en si misma. Una Computadora Personal es un conjunto de dispositivos de alta tecnología que no debería ser usado inexpertamente o en forma descuidada, a riesgo de comprometer inadvertidamente entre otras cosas, la seguridad de la información y la propia privacidad.

Con este trabajo pretendemos ejemplificar como el uso de un popular procesador de texto como lo es Microsoft Word, compromete seriamente la privacidad de nuestros datos y nos convierte por nuestra propia mano en verdaderos “hombres de cristal”. Imaginemos documentos confidenciales del gobierno, importantes contratos comerciales, fallos judiciales y documentos personales en los que quedan identificadas las personas y computadores utilizados en su redacción y revisión, donde quedan registrados y accesibles todos los cambios introducidos así como las palabras borradas intencionalmente del texto. Esto realmente ocurre y ha tenido importantes consecuencias a nivel mundial y quizás también pueda tenerlo en lo nacional. Para centrarnos en la raíz del problema debemos conocer que son los metadatos.

¹ El autor es Analista de Seguridad en Redes con certificación Enterprise Security Risk (ESR), Auditor UNIT en Sistemas de Gestión de la Seguridad de la Información (SGSI) y Consultor en Seguridad de la Información para infoDMZ, <http://www.infodmz.com>

² “Protección de Datos y Delitos informáticos”, Dra. Esc. María José Viega, 2003, <http://www.viegasociados.com/publicac/ProteccionDatosYDelitosInformaticos.pdf>

2. Los Metadatos.

Estadísticamente la aplicación ofimática más utilizada es el editor de textos y entre ellos seguramente lo es el producto Word de Microsoft que se integra en un conjunto de herramientas conocidas como Microsoft Office³. Este producto, que ha conocido distintas versiones, siendo la última la 2003 sin duda ha contribuido grandemente a popularizar el uso de los Computadores Personales en la empresa y en el hogar por su facilidad de uso y posiblemente también por el excelente marketing de la gigantesca corporación.

MS Word permiten crear y editar documentos que pueden ser grabados electrónicamente para su posterior impresión, almacenamiento y distribución en un formato propietario conocido como “.doc”. Si bien Word permite utilizar otros formatos de intercambio como el RTF⁴, no lo utiliza como opción preestablecida o por defecto, adoptando en cambio el formato DOC que posibilita la inclusión de los metadatos, dando con ello lugar a interesantes prestaciones del producto, pero también a un grave riesgo para la seguridad de la información que los contiene, así como para la privacidad del usuario, que sin saber que son, los utiliza inadvertidamente.

En Teoría de la Información se definen a los metadatos como datos que describen a otros datos, de algún modo son datos de datos. Estos metadatos están presentes en todos los productos Office de Microsoft al menos desde la versión ‘97 y continúan existiendo en la última versión. En un documento de Microsoft se dice acerca de los metadatos:

“Cuando crea, abre o guarda un documento en Microsoft Office Word 2003, el documento puede contener contenido que quizás no desea compartir con otros cuando distribuye el documento electrónicamente. Esta información se conoce como metadatos. Los metadatos se utilizan para diversas finalidades, como mejorar la modificación, presentación, archivo y recuperación de los documentos de Office. Algún metadatos es fácilmente accesible a través del interfaz de usuario de Word. Por ejemplo, otro metadatos sólo es accesible a través de medio extraordinario abriendo un documento en un editor de archivo binario de nivel bajo.”⁵

Es decir, **Microsoft define a los metadatos como información adicional contenida en un documento que quizás el usuario no desee compartir cuando lo distribuya electrónicamente.** Algunos de estos metadatos son el nombre de usuario, su compañía u organización, el de su equipo, el del disco duro donde se guardó el documento, las sucesivas revisiones, ...“texto oculto” (sic) y otros datos. Los metadatos se crean de varias formas en el documento de Word y “no hay ningún método único de quitar todos los dichos contenidos de sus documentos...”⁶. Debo señalar que ni en las especificaciones del producto ni en el procedimiento de instalación en el PC se advierte sobre estas características, ni tampoco se da la opción explícita de desactivarlas.

³ El autor se ha desempeñado como docente en cursos de Operador Windows y Office en el Instituto Superior de Informática (ISI) de la ciudad de Montevideo por varios años.

⁴ El formato RTF (Rich Text Format), propietario de Word, se ha convertido en un estándar para el intercambio de archivos de texto entre distintas plataformas, admitiendo texto enriquecido con distintos tipos de letra, colores, imágenes, etc. Lo que no admite son macros y otras prestaciones que pueden afectar la confidencialidad e integridad de la información.

⁵ “Cómo minimizar metadatos en Word 2003”, Microsoft, 2006, <http://www.support.microsoft.com>

⁶ idem.

3. El caso inglés.

Presentamos ahora un resonado caso en el que el uso descuidado o inexperto de Microsoft Word vulneró la confidencialidad de la información y la privacidad de sus usuarios con graves consecuencias de afectación del prestigio del gobierno británico.

En febrero de 2003 el gobierno inglés publicó el llamado “dodgy dossier” donde se exponía la evidencia de la existencia de armas de destrucción masiva en Irak, usado como argumentación ante las Naciones Unidas para apoyar la invasión a ese país. Posteriormente se demostró que este supuesto informe de inteligencia británico era realmente un plagio basado en una tesis que un estudiante californiano de tan solo 13 años de edad había hecho hacía algún tiempo.

Para peor se descubrió que el documento, redactado en Word, incluía oculto en su interior datos sobre las personas que lo habían editado vulnerando así la seguridad elemental que un documento de inteligencia debe tener. Entre esos datos se identificó al oficial Paul Hamill, a John Pratt: oficial de Downing Street, a Alison Blackshaw: asistente personal del secretario de prensa del Primer Ministro y a Murtaza Khan, oficial de prensa del Primer Ministro. Como resultado de esto Alastair Campbell, director de comunicaciones del gobierno británico y mano derecha de Blair tuvo que dar explicaciones ante el Parlamento. Entre los detalles que se vio obligado a revelar explico que John Pratt facilitó el informe en un disquete a Alison Blackshaw que a su vez lo entregó a Colin Powell para su presentación ante las Naciones Unidas. También quedó evidenciado como John Pratt pasó el documento de su disco duro a la disquetera. Como resultado el gobierno británico tomo la peor decisión, abandonar el uso de Microsoft Word en favor del formato PDF, en vez de elaborar y difundir Políticas de Seguridad de la información entre sus empleados.

Este caso ilustra claramente las consecuencias del uso del formato DOC, propietario de Microsoft Word que incluye en los archivos generados los metadatos de los que ya hemos hablado.

En ese mismo año, Simon Byers de la AT&T publicó una investigación que deja más aún en evidencia esta problemática, al recolectar de Internet más de 100.000 documentos Word y usando herramientas de análisis de texto de dominio público, encontró que más de la mitad de ellos contenían entre 10 y 50 palabras ocultas y el 10% más de 500.

4. Un caso uruguayo.

A los efectos de evaluar la magnitud que puede tomar esta situación en nuestro país hemos iniciado este trabajo a partir de una búsqueda en www.google.com.uy de documentos en formato DOC, habiendo encontrado 20.600 documentos publicados por organismos gubernamentales uruguayos en internet⁷. Debo hacer la salvedad que algunos de estos pueden ser copias unos de otros y por consiguiente un mismo archivo puede estar disponible simultáneamente en varios sitios. Asimismo es necesario considerar que en esta búsqueda no están incluidos los documentos “.doc” que puedan estar publicados por las Empresas Públicas en dominios “.com.uy” con lo que seguramente la cantidad de

⁷ Ver Anexo 1.

documentos del Estado uruguayo en este formato ha de superar largamente la cantidad de 25.000.

A efectos de proporcionar un interés especial a este trabajo y a título de ejemplo hemos seleccionado para estudiar el archivo “Informeplantasdecelulosa-uruguay.doc”⁸ publicado en el sitio web⁹ del Ministerio de Relaciones Exteriores (MRREE) bajo el hipervínculo: “Informe Uruguay sobre Plantas de Celulosa” y encontrado a partir de una búsqueda¹⁰ en el sitio del MRREE:

[doc| INFORME AMBIENTAL](#)

Formato de archivo: Microsoft Word - [Versión en HTML](#)

Plantas de Celulosa M'bopicuá y Botnia. Ubicación y características generales. Celulosas de M'bopicuá (grupo español ENCE) cuenta con Autorización Ambiental ... [www.mrree.gub.uy/mrree/Prensa/1erinforme_archivos/ Informeplantasdecelulosa-uruguay.doc](http://www.mrree.gub.uy/mrree/Prensa/1erinforme_archivos/Informeplantasdecelulosa-uruguay.doc)

Una vez abierto el archivo en Microsoft Word se aprecia que el título del documento es “Plantas de Celulosa M'bopicuá y Botnia” y en el que **no se identifica explícitamente ningún autor**, si embargo, consultando la ficha de propiedades del documento puede verse que el título es “Informe Ambiental” y el autor es “Diego Pastorin” del “MVOTMA” (Ministerio de Vivienda, Ordenamiento Territorial y Medio Ambiente). El mismo fue creado el jueves 2 de febrero de 2006 a las 9:46 am, consta de 2.232 palabras y fue guardado por Lucas Baresi. Esto no es poco para un documento en el que no se señala autoría de ninguna persona, pero además se aprecia como valor para el metadato “_AuthorEmailDisplayNam” a “Marcelo J. Cousillas”¹¹, proporcionando además su e-mail: mcabogado@adinet.com.uy.

También desde Word pero ahora observando los comentarios insertados podemos ver:

“[MC1] Me parece que M'Bopicuá también opera en ZF”

“[MC2] Esto está viejo”

“[MC3] Esto no es bueno. A lo sumo “apoyar”, pero DINMA tiene que tener el propio”

Entre paréntesis rectos aparece el autor de los mismos, que por sus iniciales podría ser Marcelo Cousillas.

⁸ Puede bajar el archivo de http://www.mrree.gub.uy/mrree/Prensa/1erinforme_archivos/Informeplantasdecelulosa-uruguay.doc o de <http://www.infodmz.com/documentos/Informeplantasdecelulosa-uruguay.doc>

⁹ Ver Anexo 2.

¹⁰ Ver Anexo 3.

¹¹ El Dr. Marcelo Cousillas de la DINAMA (Dirección Nacional de Medio Ambiente), se encuentra en este momento en La Haya como asesor de la delegación uruguaya ante ese Tribunal.

Hasta aquí la información que obtenida directamente de la observación del documento en Word y no contenida en el texto del mismo.

Para un segundo nivel de análisis se puede abrir el documento con cualquier editor ASCII como el propio Block de Notas de Windows. La primera cosa que se advierte es que hay mucho más información que el mero texto escrito en el documento y que esta aparece en forma de caracteres ininteligibles (código ASCII). En el final del archivo consta que el mismo fue creado con la versión 9.0 de Microsoft Word y un poco más arriba pueden verse los metadatos que ya conocemos, expuestos en las propiedades del documento.

Más arriba se encuentra la siguiente información¹²:

```
[1] "diegopastorin"  
[2] "diegopastorin_Marcelo Cousillas_Gringo"  
[3] "diegopastorin%C:\Gringo\Resumen_Papeleras_FINAL.doc"  
[4] "diegopastorin+C:\Diego\Dinama\Resumen_Papeleras_FINAL.doc"  
[5] "atorres TC:\Documents and Settings\atorres\Escritorio\Informe plantas de celulosa-uruguay.doc"  
[6] "Ariel.Bidegaray\C:\Documents and Settings\ariel.bidegaray\Escritorio\Informeplantasdecelulosa-uruguay[2].doc"  
[7] "dipr14kC:\Documents and Settings\dipr14\Mis documentos\PLANTAS DE CELULOSA\Informeplantasdecelulosa-uruguay[2].doc"  
[8] "Lucas BaresiN\Srviis\MRREE\Prensa\1erInforme_archivos\Informeplantasdecelulosa-uruguay.doc"
```

Estas ocho líneas revelan los usuarios que han abierto el documento así como la ubicación lógica del archivo en el disco duro de cada uno de ellos. Recordemos que esto fue lo que mostró el origen del documento en el caso del informe del gobierno inglés que ya comentamos y que justificadamente causó tanto escándalo. Word conserva la información de los diez últimos usuarios. Podemos ver aquí a los usuarios que ya conocíamos: Diego Pastorin y Marcelo Cousillas de la DINAMA así como Lucas Baresi del MRREE y a tres nuevos usuarios: "atorres", "Ariel Bidegaray"¹³ y un misterioso "dipr14" que no es otro que la Secretaría de Prensa del MRREE¹⁴.

Finalmente en un tercer nivel de análisis que preferimos no ilustrar en este momento es posible, con el auxilio de software especializado pero de dominio público como el utilizado por Byers, conocer todos los cambios y correcciones hechas al documento, incluyendo todas las palabras que han sido suprimidas o cambiadas intencionalmente.

5. Conclusiones.

No es la finalidad de este trabajo evaluar la magnitud del riesgo que tiene la exposición de información no contenida explícitamente en el texto del documento que hemos examinado, sino meramente presentar el caso como ejemplo, a efectos de llamar la atención sobre el problema. Lo que sin duda

¹² Los paréntesis rectos han sido agregados para facilitar el análisis.

¹³ Ariel Bidegaray es Secretario Personal del Ministro de RREE.

¹⁴ Ver Anexo 4.

podemos concluir es que ese riesgo existe y que puede tener consecuencias no deseadas para el usuario de Word. Creemos también que el caso presentado no es la excepción, sino más bien, a juzgar por otros casos similares analizados, un caso común. Tampoco pretendemos descalificar al producto ya que se trata de un editor de texto de excelentes prestaciones, quizás mucho más que lo que necesita el usuario medio. Pero no podemos dejar de advertir, más aún, siendo el editor más popular en todo el mundo, que su empleo puede comprometer la privacidad del usuario. No se comprende como las reiteradas advertencias de la comunidad de seguridad de la información ha hecho sobre esta problemática no han causado resultados apreciables en los productos Office que Microsoft produce. El gobierno uruguayo por su parte, debería adoptar lo antes posible Políticas de Seguridad de la información que incluyan formatos estándar de intercambio de información, aplicables a todas las oficinas públicas a riesgo de encontrarse en cualquier momento en una situación indeseable. En tanto para el usuario individual, la solución es muy simple, use Word o cualquier otro editor, pero guarde sus archivos en formato RTF que no incluye metadatos que comprometan su seguridad. Puede que con ello pierda alguna funcionalidad pero seguramente estará resguardando mejor su privacidad a menos que naturalmente, Ud. prefiera ser un “hombre de cristal”.

Anexo 1.

Google [La Web](#) [Imágenes](#) [Grupos](#) [Directorio](#) [Noticias](#) [más »](#)

[Búsqueda](#) [Búsqueda Avanzada](#)
[Preferencias](#)

Búsqueda: [↶](#) la Web [↷](#) páginas en español [↶](#) páginas de Uruguay

La Web Resultados **1 - 10** de un total de **20.600** páginas en **español** del dominio **gub.uy** que contienen **filetype:doc**.

[doc] [Síntesis de las propuestas elaboradas en cada uno de los talleres ...](#)

Formato de archivo: Microsoft Word - [Versión en HTML](#)

Síntesis de las propuestas elaboradas en cada uno de los talleres. de la Asamblea

Permanente la Cultura. Salto, 1 y 2 de abril de 2006. ...

www.mec.gub.uy/Sintesis.doc - [\[filetype:doc\]](#)

[doc] [INTENDENCIA MUNICIPAL DE RIVERA](#)

Formato de archivo: Microsoft Word - [Versión en HTML](#)

INTENDENCIA MUNICIPAL DE RIVERA. LICITACIÓN ABREVIADA Nº 17 /2006. PLIEGO DE
ESPECIFICACIONES PARTICULARES. PARA LA CONTRATACIÓN DE: MAQUINARIA ...

www.rivera.gub.uy/licitacionretroexcavadorabuldoxer.doc - [\[filetype:doc\]](#)

Anexo 2.



Escala de Visitas

Contáctenos @

Viernes, 9 de junio de 106

- ◆ Notas enviadas por el gobierno uruguayo ⁺
- ◆ Actas de las reuniones del Grupo Técnico de Alto Nivel binacional ⁺
- ◆ Cortes de ruta y daños ⁺
- ◆ Informe sobre la instalacion de dos plantas de celulosa en el río uruguay ⁺
- ◆ Informe Uruguayo sobre Plantas de Celulosa ⁺

Anexo 3.

Google [La Web](#) [Imágenes](#) [Grupos](#) [Directorio](#) [Noticias](#) [más »](#)

[Búsqueda Avanzada](#)
[Preferencias](#)

Búsqueda: ☐ la Web ☒ páginas en español ☐ páginas de Uruguay

La Web Resultados 1 - 6 de un total de 6 páginas en **español** del dominio **mrree.gub.uy** que contienen **filetype:doc**.

[doc] [EMBAJADA DE URUGUAY EN VENEZUELA](#)

Formato de archivo: Microsoft Word 2000 - [Versión en HTML](#)

3.Ha trabajado anteriormente con el mercado venezolano? Si es así, agradécese informar el año y el producto comercializado. ...

www.mrree.gub.uy/mrree/dpc/formulario.doc -

[doc] [DU/1/03-08-05](#)

Formato de archivo: Microsoft Word - [Versión en HTML](#)

Anexo B. Listado de documentos entregados por la delegación uruguaya a la delegación argentina en el marco de las reuniones del Grupo Técnico Binacional: ...

www.mrree.gub.uy/mrree/Prensa/1erinforme_archivos/anexob.doc -

[doc] [Borrador de Informe de la Delegación Uruguaya al GTAN](#)

Formato de archivo: Microsoft Word - [Versión en HTML](#)

Anexo C. BORRADOR CON LO ACORDADO ENTRE AMBAS DELEGACIONES. Grupo Técnico Alto Nivel – Subgrupo Aguas (Emisiones líquidas). EMISIONES DE EFLUENTES LÍQUIDOS ...

www.mrree.gub.uy/mrree/Prensa/1erinforme_archivos/anexoc.doc -

Anexo 4.

Google

[La Web](#) [Imágenes](#) [Grupos](#) [Directorio](#) [Noticias](#) [más »](#)

dipr14

Búsqueda

[Búsqueda Avanzada](#)
[Preferencias](#)

Búsqueda: [la Web](#) [páginas en español](#) [páginas de Uruguay](#) [Personaliza](#)

La Web

Resultados **1 - 10** de aproximadamente **482** páginas en **español y inglés** de **dipr14** .

[DBGET Result: C.diphtheriae **DIPr14**](#) - []

Corynebacterium diphtheriae: **DIPr14**, Help. Entry. **DIPr14** rRNA C.diphtheriae ... Other DBs,
GeneDB: **DIPr14** NCBI-GeneID: 2650075. LinkDB, PDB · All DBs ...

[www.genome.ad.jp/dbget-bin/www_bget?cdi+DIPr14](#) - 12k -

Untitled Document

Dirección de Prensa, **dipr14**, 9022312, 9080620, **dipr14**@mrree.gub.uy · Instituto "Artigas"
del Servicio Exterior, iase15, 9022216, iase15@mrree.gub.uy ...

[www.mrree.gub.uy/mrree/informacion_general/organigrama/organigrama.htm](#) - 35k -