

# Conservando el e-mail

## Más allá de la preservación digital

ANTONIO PALMA GÓMEZ

El correo electrónico ha pasado a ser uno de los medios de comunicación más usados por las empresas, las administraciones y los ciudadanos, y a menudo son el reflejo de sus actividades e interacciones. Cuando este es el caso, aparte del valor primario, a medio plazo permitirá dar fe del cumplimiento con las obligaciones legales, apoyar la toma de decisiones de las organizaciones, su gestión del conocimiento o, algunos de ellos conservados de manera permanente, preservar su memoria histórica y ciertos valores culturales.

Ya de por sí, conservar documentos electrónicos plantea ciertos retos por la obsolescencia de los formatos, la aparición, desaparición y evolución de programas informáticos y la corrupción de los datos. Frente a ello, algunas medidas integradas en las estrategias de preservación digital recomiendan adoptar formatos estándares o mantener la documentación del software para facilitar futuros procesos de migración o emulación.

Sin embargo, la naturaleza semiestructurada del correo electrónico puede dificultar aún más estos procedimientos de conservación. Su cabecera ('header') contiene los principales metadatos, su

cuerpo ('body') puede adoptar un formato texto, HTML u otra forma gracias a las especificaciones MIME. Además, cualquier correo electrónico puede tener ficheros adjuntos de cualquier tipo, cuya preservación supone un desafío en sí mismo así como el mantenimiento del vínculo con el resto del mensaje.

fichero propietario diferente. Por ejemplo, Microsoft Office Outlook almacena en formato msg mientras que Mozilla Thunderbird usa eml, aunque otras muchas aplicaciones pueden utilizar extensiones menos conocidas y difíciles de exportar.

El objetivo de la migración es transformar los formatos cerrados y propietarios en formatos abiertos o estándares que garanticen su acceso a medio o largo plazo. El proceso ha de evitar la pérdida significativa de información de su contenido o metadatos. Por este motivo, el PDF o PDF/A no parece ser una opción viable, menos aún ante la dificultad de codificar algunos tipos de ficheros adjuntos.

Aplicaciones como *Aid4Mail* o *Emailchemy*, que exportan e importan mensajes para migrarlos a eml u otros formatos más extendidos, ofrecen posibilidades para eliminar correo duplicado y basura y acceder a archivos adjuntos. También la aplicación *Xena*, desarrollada por los Archivos Nacionales de Australia, puede transformar tres tipos de formatos de e-mail a un esquema XML para su conservación a largo plazo. De modo parecido, en Estados Unidos el proyecto CERP ha permitido desarrollar un modelo que transforma en ficheros XML grupos de mensajes.

Por otro lado, junto con el formato, otros problemas surgen con respecto a la accesibilidad al e-mail a medio o largo plazo. La existencia

Así, los escollos con los que los archiveros topan al intentar recuperar cuentas de e-mail comienzan por el acceso a su contenido en correo web o en programas clientes ya que, aparte de requerir los permisos legales y técnicos, su lectura puede verse dificultada por la evolución de dichas aplicaciones. En este sentido, aunque el formato de transferencia de mensajes de e-mail está normalizado a través del RFC 5322, no pasa lo mismo con su formato de almacenamiento, que cambia de unos programas clientes a otros adoptando a menudo un tipo de

de datos personales, sensibles o con copyright, dificulta determinar los permisos y restricciones de acceso, cosa que a menudo sucede porque una cuenta de correo se usa tanto para fines profesionales como personales.

A nivel personal, una adecuada gestión del correo de manera proactiva y regular puede facilitar su preservación. La Biblioteca del Congreso de Estados Unidos recomienda a los ciudadanos seleccionar los mensajes a conservar, hacer copias en un formato abierto junto con sus metadatos y ficheros adjuntos, dar nombres descriptivos a los ficheros y carpetas, y organizarlos estructuralmente según categorías.

Estas prácticas pueden parecer tediosas, pero administraciones y empresas han considerado necesario desarrollar políticas de e-mail para que sus empleados actúen en pro de la reducción de costes de mantenimiento y del espacio de almacenamiento a la vez que la disminución de los riesgos por mal uso o eliminación. Junto con las políticas, es necesario un apoyo logístico y, dado que los mensajes de correo electrónico pueden formar parte del archivo de la organización, los sistemas de gestión de documentos electrónicos (SGDE) han sido habilitados para integrar los mensajes de correo con otros documentos para su clasificación, valoración, eliminación o, si cabe, su conservación en el repositorio digital de manera permanente. Los administradores pueden determinar y establecer límites automáticos (o no) de preservación o eliminación, y de acceso, según qué actividades o procesos o mediante la identificación de determinadas cuentas de correo de la organización.

Pese a ello, es el usuario final quien debe realizar la inserción de mensajes enviados y recibidos en el sistema de acuerdo con las actividades o funciones correspondientes. Así, ciertos estudios han revelado que no siempre los empleados entienden y aplican estas políticas sobre su correo, por lo que su integración en el SGDE no siempre está garantizada (Howard, Stephen *Why preserving email is harder than it sounds...* [en línea] [http://www.dpconline.org/component/docman/doc\\_download/641-emailhowardjul2011](http://www.dpconline.org/component/docman/doc_download/641-emailhowardjul2011)).

Como alternativa, existen otras aplicaciones que copian los mensajes de correo fuera de los sistemas de gestión de documentos electrónicos. En estos sistemas, las copias se hacen de todos los mensajes en el momento de su transacción, y no una vez que ya se han almacenado, evitando así los formatos propietarios de almacenamiento. Las copias se conservan en dispositivos ex-

evitando así la sobrecarga de los servidores locales a la vez que garantizando su acceso. Soluciones como *MailArchiva* aplican este método y ofrecen otras posibilidades como el establecimiento de restricciones de acceso o facilidades de búsqueda y recuperación.

Pese a que este método permite conservar los mensajes al margen de otros documentos con los que mantienen relaciones archivísticas, algunas veces indican que es la solución más viable hasta que aparezcan sistemas que disminuyan o faciliten la intervención del usuario y así garantizar su gestión dentro de un SGDE. Algunos esperan que la publicación en el futuro de un módulo para e-mail dentro de *MoReq2010* pueda facilitar esta tarea. ■

