

## PRAETER ORWELL: SUJETOS, ACCIÓN Y OPEN DATA CIUDADANA

### PRAETER ORWELL: SUBJECTS, ACTION AND CIVIC OPEN DATA

JAVIER DE LA CUEVA  
Escuela de Organización Industrial, Madrid  
jdelacueva@derecho-internet.org

RECIBIDO: 20/09/2011

ACEPTADO: 12/12/2011

**Resumen:** La aparición de las tecnologías de la información y comunicación ha supuesto la difusión de información con ámbito territorial universal por parte de los ciudadanos. En la actualidad, están surgiendo numerosas iniciativas de acción política en internet que abarcan desde la generación de una opinión pública al margen de los medios de comunicación tradicionales hasta proyectos de software cuyo objeto es el tratamiento de datos públicos promoviendo simultáneamente la apertura de los mismos. Los actores intervinientes en estas iniciativas adoptan organizaciones que no necesariamente corresponden con las formalizadas tradicionalmente (partidos políticos, sindicatos, asociaciones). En el presente trabajo se analizan los agentes intervinientes en este tipo de actividades y su acción, se describen los principios «open data» y se propone una estructura informacional básica que pueda ser tenida en cuenta por todas las iniciativas para facilitar en el futuro la coordinación de los intercambios de información entre ellas.

**Palabras clave:** Activismo, democracia, control del poder, internet, hacktivism, open data, scraping, web semántica.

**Abstract:** The advent of information and communications technology has led to global spread of information by citizens. Currently, many initiatives for political action are emerging on the internet, ranging from generation of public opinion outside the traditional media to software projects which purpose is to process public data proposing simultaneously its openness.

The actors involved in these initiatives use organizations that do not necessarily correspond to the ones traditionally formalized (political parties, unions, associations). This paper examines the agents involved in this type of activities and their actions, describes "open data" principles and proposes a basic information structure that can be taken into account by all initiatives to facilitate future exchanges of information and coordination between them.

**Key words:** Activism, democracy, power control, Internet, hacktivism, open data, scraping, semantic web.

## Introducción

Narro una situación real: El día 8 de septiembre de 2011, las webs del Congreso de los Diputados y del Senado de España publicaron<sup>1</sup> por primera vez en su

---

<sup>1</sup>La publicación trae causa de lo establecido en el artículo único de la Ley Orgánica 7/2011, de 15 de

historia las declaraciones patrimoniales de los parlamentarios. El formato que se utilizó para poner a disposición pública los archivos era portable document format (PDF) con imágenes incrustadas, igual formato en el que la Cámara de los Comunes británica publicó en junio de 2009 casi un millón de documentos que contenían las declaraciones de gastos efectuadas por los miembros del parlamento.<sup>2</sup> El problema de este tipo de formato con imágenes incrustadas es que los datos que contienen los documentos no pueden tratarse automatizadamente: no se puede realizar estadística con el contenido de unas fotografías ni seguir evoluciones históricas de los datos, por lo que hay que extraerlos y transcribirlos mediante mecanismos bien de reconocimiento óptico de caracteres (OCR), bien a mano.

En los casos español y británico hubo una diferencia: si bien en el Reino Unido el periódico *The Guardian* publicó un software<sup>3</sup> para que los lectores fueran consultando los documentos y los fuesen transcribiendo, en España esta labor se realizó por un grupo de voluntarios sin intervención de organización ninguna. Tras haber aplicado software OCR a dos archivos PDF publicados por el Senado y verificar la dificultad de la extracción de los datos por medios automatizados, el día 8 por la noche publique el tuit<sup>4</sup> «Algún periódico de este país se anima a copiar al *Guardian* con los pdfs del patrimonio de los parlamentarios? <http://mps-expenses.guardian.co.uk>»,<sup>5</sup> tuit al que inmediatamente me contestó David Cabo (miembro de la Asociación Pro Bono Público y conocido paladín del open data): «@jdelacueva si no tendremos que hacerlo nosotros». <sup>6</sup> Así que nos pusimos a ello.

Los servidores del Congreso y del Senado se hallaban saturados por el interés del público en las declaraciones así que, para no depender de conexiones a dichos servidores, David Cabo, que ya había descargado todos los archivos PDF, los comprimió, subió el archivo resultante al servidor rapidshare.com<sup>7</sup> y me

---

julio, de modificación del artículo 160 de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General.

<sup>2</sup><<http://mpsallowances.parliament.uk/mpslordsandoffices/hocallowances/allowances-by-mp/>> Fecha de última consulta: 15 de septiembre de 2011.

<sup>3</sup><<http://mps-expenses.guardian.co.uk/page/237840/>> Fecha de última consulta: 15 de septiembre de 2011.

<sup>4</sup>Un tuit es un mensaje enviado en la plataforma Twitter <<http://twitter.com>>. El término en inglés es Tweet, castellanizado como tuit.

<sup>5</sup><<http://twitter.com/#!/jdelacueva/status/111866169249632256>> Fecha de última consulta 12 de septiembre de 2011.

<sup>6</sup><<http://twitter.com/#!/dcabo/status/111916491867308032>> Fecha de última consulta 12 de septiembre de 2011.

<sup>7</sup><<http://rapidshare.com>> es un servidor muy utilizado para subir al mismo archivos de gran tamaño que luego pueden ser descargados por los usuarios que conozcan la URL (el enlace) a los mismos.

comunicó<sup>8</sup> la URL<sup>9</sup> desde donde todos podíamos descargarlos<sup>10</sup> en caso de seguir con los problemas de saturación. Además, creó una plantilla<sup>11</sup> de hoja de cálculo en *Google Docs* cuyos títulos de columnas correspondían con las divisiones del documento de declaración patrimonial de los parlamentarios y nos propuso a Pablo Soto, Alberto Abella y a mí iniciar la campaña «Adopta a un Senador».<sup>12</sup> Debe mencionarse que los mensajes que cruzamos David Cabo y yo son públicos por lo que en nuestra conversación entraron estas dos personas más, que son ampliamente conocidas por sus trabajos en el sector de los datos y de las tecnologías de la información.<sup>13</sup> Ya desde el inicio, en una de las celdas bien visibles de la hoja,<sup>14</sup> se anunció que la licencia de lo que se estaba construyendo era el dominio público.

En la madrugada del día 9 publiqué un pequeño artículo en mi blog<sup>15</sup> en el que explicaba el trabajo que debíamos realizar, «se trata de utilizar colectivamente una hoja de cálculo de *Google Docs* donde David ha diseñado una plantilla e ir transcribiendo los datos.», le dimos difusión en *Twitter* y a la mañana siguiente ocupó la portada en *Meneame.net*.<sup>16</sup> La tarde del día 9, el

---

<sup>8</sup> «@jdelacueva lo miraré. de momento me he bajado todos los pdfs de los senadores, por comodidad y por guardar registro [rapidshare.com/files/24839753...](http://rapidshare.com/files/24839753...)» En <http://twitter.com/#!/dcabo/status/111926994383540224> Fecha de última consulta 12 de septiembre de 2011.

<sup>9</sup> Acrónimo de Uniform Resource Locator, una cadena de caracteres única mediante la que se localiza un recurso en Internet. Es lo que se conoce coloquialmente como “enlace”.

<sup>10</sup> [https://rapidshare.com/#!/download|1cg|2483975347|Declaraciones\\_de\\_Bienes\\_Senadores\\_20110908.zip|184818|R~8B11F399A457CE7067502CA579325DCF0|0](https://rapidshare.com/#!/download|1cg|2483975347|Declaraciones_de_Bienes_Senadores_20110908.zip|184818|R~8B11F399A457CE7067502CA579325DCF0|0) Fecha de última consulta 12 de septiembre de 2011.

<sup>11</sup> La plantilla son los títulos de las columnas de lo que finalmente es el archivo construido entre los voluntarios participantes en la actividad y se halla en la siguiente URL: <https://docs.google.com/spreadsheet/ccc?key=0AowzHU9kHzeudG9aSjVVOEQxVHpzR2E4ZDdhVXJLQIE&hl=es#gid=0>

<sup>12</sup> @pablomp2p @jdelacueva @abella chicos, qué os parece iniciar una campaña "adoptar un senador"? [docs.google.com/spreadsheet/ccc...](https://docs.google.com/spreadsheet/ccc...) (o diputado, vamos) En <http://twitter.com/#!/dcabo/status/111944699778043904> Fecha de última consulta 12 de septiembre de 2011.

<sup>13</sup> No es relevante si las personas que participan en este tipo de iniciativas se conocen personalmente, si bien existe un núcleo en el que existe un conocimiento personal y relaciones AFTK (away from the keyboard – fuera del teclado).

<sup>14</sup> En la celda E1, que se halla en la cabecera de la tabla, se destacó: «Licencia del fichero: dominio público»

<sup>15</sup> Por la transparencia: adoptar un senador <http://derecho-internet.org/node/569> Fecha de última consulta 12 de septiembre de 2011.

<sup>16</sup> *Meneame.net* es el agregador de noticias más visitado de la lengua española. Cualquier usuario envía una noticia con un enlace a una publicación electrónica. Esta noticia entra en la sección “menear pendientes” donde es votada por los lectores de *Meneame*, pasando a portada únicamente las más votadas. <http://www.meneame.net/story/transparencia-adoptar-senador> Fecha de última

diario *El País*<sup>17</sup> dedicó a esta iniciativa un artículo en su web y el día 10 un artículo en su edición de papel.

Muchas personas colaboraron<sup>18</sup> para transcribir los datos hasta tal punto que la plataforma utilizada, *Google Docs*, estuvo un tiempo saturada. Se produjo vandalismo y spam en los datos ya extraídos por lo que hubo que cambiar los permisos de la hoja de cálculo desde una configuración en la que todas las personas podían modificar la misma a otra en la que se debía solicitar permiso. Finalmente, el día 12 de septiembre, David Cabo verificó que los datos de los 261 senadores que habían presentado la declaración patrimonial se habían transcrito correctamente en la hoja de cálculo.<sup>19</sup> El archivo comprimido de los PDF de las declaraciones patrimoniales publicado por los órganos parlamentarios ocupa 184.818.156 bytes mientras el archivo de la hoja de cálculo<sup>20</sup> con todos los datos tratado por los ciudadanos ocupa 158.991 bytes, esto es, un 0,086% de aquél.<sup>21</sup> Ese mismo día 12, David Cabo envió un correo electrónico<sup>22</sup> al grupo Pro Bono Público en el que bajo el asunto «Adopta un senador: post-mortem» realizaba un resumen final de la experiencia compartiendo públicamente el aprendizaje obtenido.

Este tipo de acciones colaborativas no planificadas de antemano son frecuentes. En la narración anterior encontramos todos los elementos de un nuevo ejercicio de activismo político: (i) El sujeto es un grupo de personas no organizado en estructuras formales; (ii) La acción consiste en extraer el contenido de una documentación de los poderes públicos, transformarlo y publicarlo siguiendo los criterios de (iii) open data, cuidando que el formato y la licencia del producto final permita su uso, difusión y transformación libres. Partiendo de estas características, el presente artículo propone ir más allá

consulta: 12 de septiembre de 2011.

<sup>17</sup> «Transparencia política con carencias» <[http://politica.elpais.com/politica/2011/09/09/actualidad/1315584504\\_266528.html](http://politica.elpais.com/politica/2011/09/09/actualidad/1315584504_266528.html)> Fecha de última consulta: 12 de septiembre de 2011.

<sup>18</sup> Tuit de David Cabo: «Wow @aabella: Hay 50 personas trabajando simultáneamente en #adoptaunsenador docs.google.com/spreadsheets/cc... recuperando datos #transparencia #opendata» En <<http://twitter.com/#!/dcabo/status/112100394359996416>> Fecha de última consulta: 12 de septiembre de 2011.

<sup>19</sup> «Ya están los datos de todos los senadores revisados, gracias a todos bit.ly/qL9efc Ahora, #adaptaundiputado bit.ly/ruO6zn» <<http://twitter.com/#!/dcabo/status/113252014040948736>> Fecha de última consulta: 12 de septiembre de 2011.

<sup>20</sup> <<https://docs.google.com/spreadsheets/cc?key=0AowzHU9kHzeudG9aSjVVOEQxVHpzR2E4ZDd hVXJLQIE&hl=es#gid=0>> Fecha de última consulta: 15 de septiembre de 2011.

<sup>21</sup> A la hora de tratar la información, cuanto menos ocupe más fácil será de transmitir, trabajar en ella y almacenar.

<sup>22</sup> <<https://groups.google.com/d/msg/pro-bono-publico/euSNKX4C9QQ/0Sgi8rQdOIMJ>> Fecha de última consulta: 15 de septiembre de 2011.

mediante la descripción de un modelo que permita establecer campos comunes a todas las acciones de esta naturaleza que se realizan en la actualidad para así favorecer la interoperabilidad entre las iniciativas.

Por último, ha de señalarse que estos trabajos colaborativos buscan una mejor transparencia de los poderes públicos dada la existencia de déficits democráticos y el desplazamiento de la acción política que éstos producen (cfr. Marres 2005, 26-30). Sin embargo, no vamos a caer en la ingenuidad de pensar que mediante la puesta a disposición de los datos públicos más su posible análisis por los ciudadanos van a aflorar los verdaderos intereses y la realidad de las transacciones económicas subyacentes en el ejercicio del poder político. Si bien este paso por sí sólo no es suficiente, sin embargo, es totalmente necesario no sólo para el control político sino para perseguir la corrupción y proteger a quienes la denuncian.

## Sujetos

Las actividades descritas en la introducción fueron realizadas por un grupo de personas interconectadas entre sí. Se trata de un grupo heterogéneo de personas cuya afinidad es la de hallarse interesadas en el fenómeno de “open data” y dedicarse cada uno en la medida de su voluntad, a proponer iniciativas, ejecutar o difundir las mismas sin recibir retribución alguna a cambio. Los sistemas “tradicionales” basados en premisas tales como la unión<sup>23</sup> de sus integrantes y la existencia de un histórico en la adscripción de los miembros, mutan en redes:

Si la “fraternidad” implicaba una estructura preexistente que predeterminaba y predefinía las reglas vinculantes para la conducta, las actitudes y los principios de la interacción, las “redes” carecen de historia previa; nacen en el transcurso de la acción y se mantienen con vida (o, mejor dicho, son continua y repetitivamente recreadas/resucitadas) únicamente gracias a sucesivos actos comunicativos (Bauman, 2010, 173)

En el activismo online, se producen dinámicas de trabajo en red donde los sistemas de entrada y salida o adscripción a estos grupos informales son efímeras, aunque, añadimos, altamente eficaces siendo «el rasgo más trascendental de las redes [...] la inusitada flexibilidad con la que puede modificarse su composición: se suman o se restan elementos individuales sin

---

<sup>23</sup>La tradicional consigna que se corea en las manifestaciones de “El pueblo unido jamás será vencido” hace referencia a este concepto de unión como fuente de poder.

más esfuerzo que el preciso para introducir (o borrar) un número en la agenda de un teléfono móvil» (2010, 173). Se generan de esta manera relaciones con vínculos fácilmente disolubles, sin una duración determinada y sin los compromisos a largo plazo que el citado autor (2010, 29) caracteriza como enjambres: «los grupos (y sus líderes, jerarquías y escalafones característicos) tienden a ser reemplazados por los enjambres... se juntan, se dispersan y se vuelven a reunir en ocasiones sucesivas, guiados cada vez por temas relevantes diferentes y siempre cambiantes».

Estos grupos son sujetos del ejercicio de acciones de reivindicación política a través de nuevos modelos de actividades. Ya no se trata de ejercer el derecho de manifestación, de asociación política o de generar una opinión pública a través de la repetición goebbeliana de la información, sino de generar datos que puedan ser utilizados por todos libremente y que tengan las menores limitaciones posibles (bien barreras por su formato, bien por su licencia legal) para así facilitar su distribución en soportes tangibles o su difusión mediante procedimientos intangibles.

La procedencia de los integrantes de estos grupos es variada, si bien se les puede caracterizar por su conocimiento y uso del Software Libre, el seguimiento del lema «information wants to be free»<sup>24</sup> y, como acertadamente señalan Carceler y López (2003, 256), la intersección de sus valores «con los que según Merton definen la actividad del científico –universalidad, comunidad de los conocimientos, escepticismo organizado y desinterés– no sería vacía en ningún caso», pudiendo incluso especularse que son continuación de la teoría política tecnocrática en la que «la política considerada como diálogo, conflicto o pluralidad, es sustituida por la certeza de la ciencia y de la técnica, convirtiéndose la eficacia en un criterio fundamental para la adopción de decisiones políticas» (Valencia, 1995, 418). Su método de trabajo consiste en la *peer production*, de la que Benkler (2006) se ha ocupado con profundidad y extensión.

Habremos de plantearnos si un sujeto informal como el anterior es susceptible de protección y amparo legal, lo que requerirá un estudio más extenso que no puede realizarse en este artículo. No obstante, se ha de señalar

---

<sup>24</sup>El origen del lema «la información quiere ser libre» proviene de una intervención de Stewart Brand en la primera conferencia hacker celebrada en otoño de 1984: «In fall 1984, at the first Hackers' Conference, I said in one discussion session: "On the one hand information wants to be expensive, because it's so valuable. The right information in the right place just changes your life. On the other hand, information wants to be free, because the cost of getting it out is getting lower and lower all the time. So you have these two fighting against each other." That was printed in a report/transcript from the conference in the May 1985 -Whole Earth Review, p. 49». <<http://www.rogerclarke.com/II/IWtbF.html>> Fecha de última consulta: 15 de septiembre de 2011.

que el Tribunal Constitucional español ha declarado reiteradamente que el ejercicio del derecho de información y de la libertad de expresión son la parte nuclear de la democracia pues permiten que los ciudadanos puedan hacerse una opinión fundada que utilicen luego a la hora de ejercitar el voto. Ahora bien, a nadie le es desconocida la quiebra de la confianza en los mecanismos tradicionales de generar opinión, que son los medios de comunicación social, y la aparición de nuevos canales de adquisición de conocimiento como lo son los de acudir a las fuentes primarias<sup>25</sup> ya que todos disponemos de la posibilidad de distribuir nuestra información a todos los confines del globo y por un precio despreciable cuando no nulo. Ya adelanto que la respuesta a la protección legal de estos grupos como contribuyentes de datos soporte de una fundada opinión pública ha de ser positiva por las mismas razones que la tienen los medios de comunicación. A mayor abundamiento, los trabajos de estos grupos son utilizados por los medios de comunicación tradicional en lo que se ha venido a denominar «Data-Driven Journalism».<sup>26</sup>

El Derecho no ha sido ajeno a contemplar la existencia de derechos y obligaciones dentro de grupos no formalizados legalmente. Sin pretender un carácter exhaustivo, podemos señalar los siguientes ejemplos en los que un colectivo indeterminado es susceptible de ejercitar derechos y responder a obligaciones:

1. En el ámbito del Derecho procesal civil, el titular legitimado para comparecer en el proceso ha sufrido una evolución histórica en la que, con Montero (1997, 51), «podemos decir que hemos pasado de la nada hasta un final que no se descubre todavía». En lo que nos interesa para este estudio, en los litigios en los que pueden concurrir pluralidad de partes, el autor citado diferencia entre los intereses colectivos y los intereses difusos:

Lo característico de los intereses colectivos es que los mismos corresponden a una serie de personas, más o menos numerosa, que están o pueden estar

---

<sup>25</sup>En el mundo jurídico, ya no es admisible la lectura online de una noticia sobre una resolución judicial si la noticia no va acompañada con un enlace a la resolución. Es más, en todos los procedimientos sobre propiedad intelectual que derivaron en la falta de infracción de la propiedad intelectual de las webs de enlaces y en los procedimientos contra el canon en los soportes digitales, las fuentes primarias han sido los blogs de los abogados intervinientes en los procedimientos, que siempre han enlazado a la resolución para su descarga por los lectores y posibilidad de comprobación de las afirmaciones del articulista. En otros sectores, basta el ejemplo del CERN suizo en cuya página web se va informando sobre los resultados del colisionador de partículas en su búsqueda del bosón de Higgs.

<sup>26</sup>Data-Driven Journalism se está traduciendo indiferentemente por «periodismo de datos» o por «periodismo basado en datos».

determinadas, o que por lo menos no son absolutamente indeterminables, entre las cuales existe un vínculo jurídico [...] existiendo una entidad que es persona jurídica a la cual se atribuye por ley la “representación institucional” para la defensa de ese interés (1997, 48),

mientras que los intereses difusos:

se caracterizan porque corresponden a una serie de personas que están absolutamente indeterminadas, no existiendo entre ellas vínculo jurídico alguno, de modo que la afectación a todas ellas deriva sólo de razones de hecho contingentes (como ser posibles consumidores de un mismo producto, vivir en el mismo lugar, ser destinatarios de una campaña de publicidad (1997, 49)

También, añadimos nosotros, la afectación de los intereses difusos deriva de ser activistas del open data y hallarse ejercitando acciones de significado político.

2. El Tribunal Constitucional inauguró una corriente jurisprudencial en el caso Violeta Friedman contra Leon Degrelle y el semanario Tiempo mediante la sentencia 214/1991 de fecha 11 de noviembre<sup>27</sup>. En dicha resolución se admitió la legitimación activa de una persona física en representación de colectivos sin personalidad jurídica, estableciéndose que también un colectivo es susceptible de tener honor. Señaló la sentencia en su Fundamento de Derecho 3 que:

Pero esta legitimación originaria no excluye, ni la existencia de otras legitimaciones [...] ni que haya de considerarse también como legitimación originaria la de un miembro de un grupo étnico o social determinado, cuando la ofensa se dirigiera contra todo ese colectivo [...]. En tal supuesto, y habida cuenta de que los tales grupos étnicos, sociales e incluso religiosos son, por lo general, entes sin personalidad jurídica y, en cuanto tales, carecen de órganos de representación a quienes el ordenamiento pudiera atribuirles el ejercicio de las acciones, civiles y penales<sup>28</sup>, en defensa de su honor colectivo, de no admitir el art. 162.1 b) C.E., la legitimación activa de todos y cada uno de los tales miembros, residentes en nuestro país, para poder reaccionar jurisdiccionalmente contra las intromisiones en el honor de dichos grupos, no sólo permanecerían indemnes las lesiones a este derecho fundamental que sufrirían por igual todos y cada uno de sus integrantes, sino que también el Estado español de Derecho

<sup>27</sup><[http://www.boe.es/boe/consultas/bases\\_datos/doc.php?coleccion=tc&id=SENTENCIA-1991-0214](http://www.boe.es/boe/consultas/bases_datos/doc.php?coleccion=tc&id=SENTENCIA-1991-0214)> Fecha de última consulta: 12-09-2011

<sup>28</sup>Sin cursivas en el original.

permitiría el surgimiento de campañas discriminatorias, racistas o de carácter xenófobo, contrarias a la igualdad [...]

3. En Derecho fiscal, se estableció el reconocimiento de grupos informales como sujetos tributarios. Obviamente la finalidad era la de someter a los grupos informales a los supuestos de tributación, señalándose como ejemplo típico a los grupos de personas que organizaban una tómbola en la que parte de lo recaudado se destinaba a premios, y parte a beneficio sin tributar. En este sentido, la Ley 58/2003, de 17 de diciembre, General Tributaria, establece en su artículo 35 apartado 4 que «Tendrán la consideración de obligados tributarios, en las leyes en que así se establezca, las herencias yacentes, comunidades de bienes y demás entidades que, carentes de personalidad jurídica<sup>29</sup>, constituyan una unidad económica o un patrimonio separado susceptibles de imposición.»

En Derecho penal basta señalar el apartado 6 del artículo 177 del Código penal, que establece la imposición de una pena superior en grado cuando el culpable perteneciera a una organización o asociación de más de dos personas, «incluso de carácter transitorio», que se dedicase a la realización de las actividades de la trata de seres humanos. Otro ejemplo es el artículo 515, que contempla específicamente el delito de asociación ilícita.

En resumen, si bien el Derecho y la jurisprudencia no son ajenos a la regulación del ejercicio de derechos y obligaciones en ámbitos grupales no formalizados legalmente, las resoluciones judiciales todavía no han tenido la oportunidad de pronunciarse sobre las bondades que estos grupos efímeros, mediante sus trabajos espontáneos, aportan en favor de una opinión pública soporte de democracia. La jurisprudencia favorable a la función de los medios de comunicación social como constructores de prius de la democracia deberá actualizarse con el reconocimiento y defensa de este tipo de grupos, máxime dada la pérdida de confianza generalizada que están sufriendo los medios tradicionales por sus vinculaciones con los poderes económicos y políticos y que según Castells (2009, 262) «No son el Cuarto Poder. Son mucho más importantes: son el espacio donde se crea el poder [...] en conjunto no son neutrales»

## Acción

Las acciones llevadas a cabo por los grupos informales suponen una extracción de la información y una reestructuración de la misma bajo los principios, que

---

<sup>29</sup>Sin cursivas en el original.

posteriormente detallaremos, de los denominados «open data». La acción se halla finalísticamente determinada puesto que se conoce ab initio tanto el formato como la estructura en la que se archivará la información tratada. Las actividades que el grupo de voluntarios realiza consisten en la (i) extracción de la información, su (ii) transformación para adecuarla a una estructura y un formato determinados y, por último, su (iii) puesta a disposición pública (la publicación online).

En el caso concreto del ejemplo señalado en la introducción, las plataformas utilizadas para la realización de las diversas acciones fueron <http://twitter.com>, <http://derecho-internet.org>, <http://groups.google.com>, <http://meneame.net> y <http://docs.google.com>. Tal y como se ha relatado, en Twitter se intercambiaron los primeros mensajes y se obtuvo una primera difusión del proyecto. La web Derecho de Internet<sup>30</sup> se utilizó para una explicación más detallada del proyecto, lo que hubiera sido imposible en Twitter dada la limitación de sus mensajes a 140 caracteres. Sin concertarlo con nadie, el usuario GotMilk<sup>31</sup> publicó la noticia en Menéame, siendo votada por los usuarios pasando a portada, lo que extendió la difusión del proyecto. Por otra parte, Google Docs fue utilizada por los voluntarios para la transcripción simultánea de los archivos PDF a la hoja de cálculo y finalmente el análisis post mortem fue enviado por David Cabo mediante un correo electrónico al grupo Pro Bono Público.<sup>32</sup> Salvo la web Derecho de Internet, los participantes iniciales utilizaron recursos gratuitos accesibles por todos. Los gastos que ha costado la acción han sido cero euros.<sup>33</sup> En este tipo de acciones, la barrera de entrada no es la económica sino la del conocimiento.

## La extracción de la información

Las técnicas que se utilizan para la extracción de la información dependen de la capacidad de aplicar métodos automatizados al documento contenedor de la información inicial, siendo residual la técnica manual cuando no existe ninguna otra posibilidad más eficaz. La técnica manual consiste en la visualización del

---

<sup>30</sup><<http://derecho-internet.org>>

<sup>31</sup><<http://www.meneame.net/user/GotMilk/>>

<sup>32</sup>Un grupo de Google es una lista de correo de suscripción más o menos abierta. El grupo de Google de Pro Bono Público es una lista abierta a cualquier suscriptor.

<sup>33</sup>En pro del rigor ha de señalarse que el servidor que aloja Derecho de Internet tiene un coste de 50 USD mensuales, que han de repartirse entre diversas webs alojadas, siendo utilizado por el autor de este artículo para muchos más servicios que los de la mera publicación de un artículo.

documento y la transcripción en texto de los datos relevantes contenidos en el mismo.<sup>34</sup> Los formatos del documento inicial habitualmente son archivos PDF o páginas web, a los que habrá de aplicárseles técnicas diferentes:

**Documentos en formato PDF.** Se utiliza un software de reconocimiento óptico de caracteres, que en muchas ocasiones logra transcribir de una manera más o menos exacta la información, debiéndose siempre verificar y, en su caso, corregir los datos obtenidos (actividad que se conoce como «data massage», masaje de datos).

**Documentos en formato HTML.** Se aplican técnicas de scraping (raspado), consistentes en verificar el código fuente de una página web y escribir un programa de software que la descargue, la recorra y extraiga la información relevante. Para comprender cómo se utiliza esta técnica, pongamos un ejemplo partiendo de un documento HTML muy básico: una página web que contiene una lista con dos diputados. Tal y como puede observarse a continuación, el nombre y apellido de los diputados se halla entre las etiquetas <li> y </li>, que en lenguaje HTML se utilizan para señalar los elementos de una lista. La técnica de scraping consiste en construir un «parser» (analizador) que extraerá del archivo únicamente los datos «Abad Abad, José» y «Avila Avila, María» que son los que se hallan entre las etiquetas <li> y </li> indicadoras de los elementos de la lista.

```
<html>
  <header>
  ...
  <header>
  <body>
    <h1>Diputados</h1>
    <ul>
      <li>Abad Abad, José</li>
      <li>Avila Avila, María</li>
    </ul>
  </body>
</html>
```

Dado que cada página web tiene diferente contenido, para conocer qué ha de extraerse se verifica su código fuente y se escribe el programa de extracción. Para facilitar la verificación de los datos obtenidos mediante la extracción, es

---

<sup>34</sup>Coloquialmente: picar un documento.

aconsejable publicar no sólo los datos sino también el software creado para generarlos<sup>35</sup>

## Transformación

La transformación también se automatiza todo lo posible mediante la utilización de un software escrito ad hoc. Con la información extraída se construye un archivo en un formato que permita su reutilización y se verifica la corrección de los datos (ejemplos típicos a verificar son los diferentes formatos de fechas según los países «dd/mm/aa», «aa/mm/dd», y la utilización del punto o coma para señalar los decimales en los datos numéricos), si bien con elevada frecuencia nos encontramos con diferentes plantillas a través de las que se genera la información. Por ejemplo, el Boletín Oficial del Estado utiliza diferentes sistemas para formatear las fechas de una disposición según el período histórico de la norma. Así encontramos que en las órdenes ministeriales, hasta el 31 de diciembre de 2001 se utilizaba un formato, cambiando a partir del 1 de enero de 2002, lo que nos obliga a utilizar técnicas de scraping variables por fecha, formatos que vemos a continuación.

Antes del 31 de diciembre de 2001:

ORDEN de 28 de diciembre de 2001 sobre los Convenios de Promoción de Fondos de Titulización de Activos para favorecer la financiación empresarial.

A partir del 1 de enero de 2002:

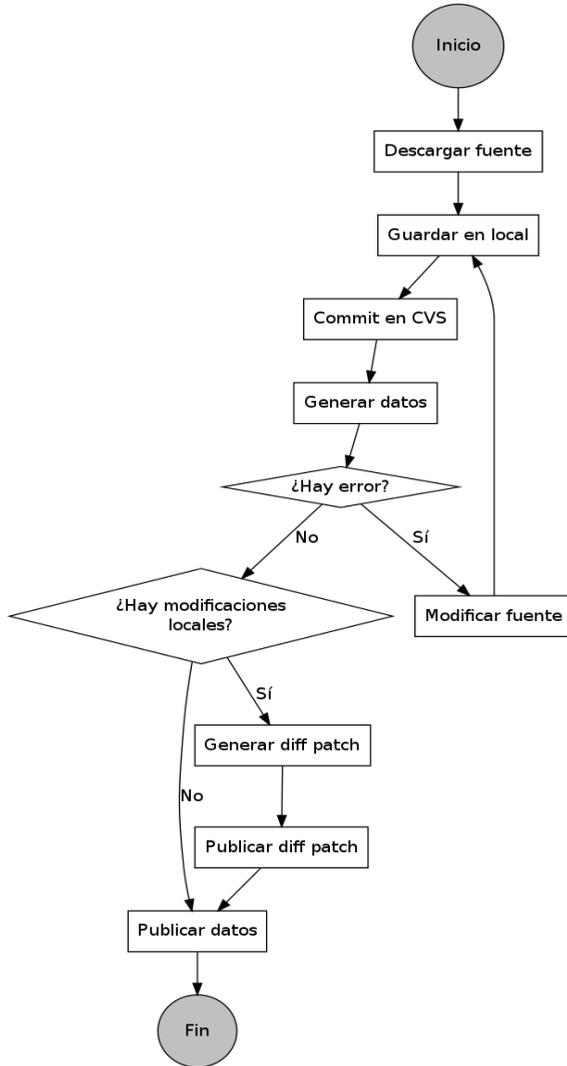
ORDEN APA/16/2002, de 2 de enero, por la que se establece un plan de pesca para la pesquería de arrastre de fondo en el Caladero Nacional del Cantábrico y Noroeste.

En cuanto a las omisiones de un dato o errores de tipografía (por ejemplo, diferencia tipográfica septiembre–setiembre), dos son los posibles sistemas de trabajo: refinar el analizador para resolver todas las diferencias (que admita ambos términos septiembre y setiembre) o corregir la alteración de las fuentes de los datos, documentar esta alteración y hacerla pública para que los organismos

---

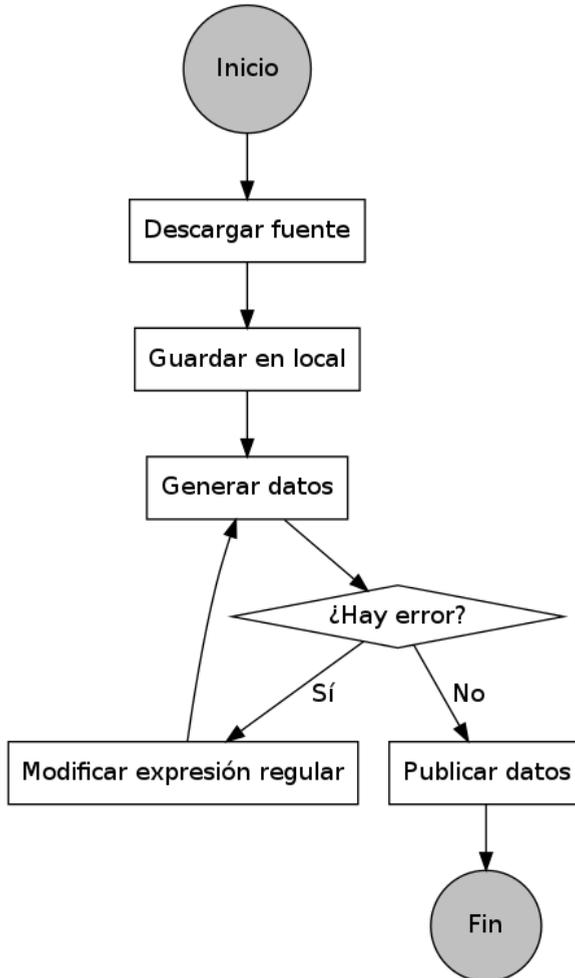
<sup>35</sup>Los scripts de scraping de los datos de los diputados del Congreso y de los senadores son públicos: «Script para extraer los datos de los senadores de la novena legislatura» <<http://derecho-internet.org/node/568>>, fecha de última consulta 13 de septiembre de 2011, y «Script para extraer los datos de correo electrónico, blog, Twitter, Facebook y Youtube de los diputados del Congreso» <<http://derecho-internet.org/node/567>>, fecha de última consulta 13 de septiembre de 2011.

oficiales, a su vez, puedan verificar estos extremos y corregir las fuentes en origen. Los dos sistemas de trabajo pueden observarse en las figuras 1 y 2.



**Figura 1**

En este primer sistema de trabajo, se modifican las fuentes, que hemos guardado en el ordenador, y se publican finalmente tanto los datos como las modificaciones en las fuentes en un formato estándar para diferencias –diff patch–.



**Figura 2**

En este segundo sistema de trabajo, se modifica el analizador, concretamente una parte del mismo denominada «expresión regular», que es el patrón para identificar los caracteres que se están buscando. Por ejemplo, para encontrar una fecha como «28 de diciembre de 2001», el patrón que se utiliza es: [uno o dos dígitos cualesquiera] [espacio] [preposición «de»] [el nombre de un mes] [espacio] [preposición «de»] [cuatro dígitos de los que los dos primeros han de ser «19» o «20»].

## La puesta a disposición de la información extraída: los principios de open data

Los días 7 y 8 de diciembre de 2007, treinta personas<sup>36</sup> se reunieron en Sebastopol, California, convocadas<sup>37</sup> por Carl Malamud y Tim O'Reilly y bajo el patrocinio de la Sunlight Foundation, Google y Yahoo. La cita tenía como objetivo discutir cómo deberían los órganos públicos almacenar la información electrónica para uso público. En esa reunión se establecieron los iniciales ocho principios<sup>38</sup> de open data,<sup>39</sup> luego más tarde ampliados a diez por la Sunlight Foundation, que señala que los principios son un continuo de apertura, teniendo

---

<sup>36</sup>Relación de asistentes en <[https://public.resource.org/open\\_government\\_meeting.html](https://public.resource.org/open_government_meeting.html)> Fecha de última consulta: 13 de septiembre de 2011.

<sup>37</sup>Texto de la convocatoria en <[https://public.resource.org/open\\_government\\_meeting.html](https://public.resource.org/open_government_meeting.html)> Fecha de última consulta: 13 de septiembre de 2011.

<sup>38</sup>Open Government Data Principles

Government data shall be considered open if it is made public in a way that complies with the principles below:

1. Complete *All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.*
2. Primary *Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.*
3. Timely *Data is made available as quickly as necessary to preserve the value of the data.*
4. Accessible *Data is available to the widest range of users for the widest range of purposes.*
5. Machine processable *Data is reasonably structured to allow automated processing.*
6. Non-discriminatory *Data is available to anyone, with no requirement of registration.*
7. Non-proprietary *Data is available in a format over which no entity has exclusive control.*
8. License-free *Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.*

Compliance must be reviewable.

<sup>39</sup><<http://www.opengovdata.org/home/8principles>> Fecha de última consulta: 13 de septiembre de 2011.

su listado carácter descriptivo y no normativo.<sup>40</sup> Los diez principios señalados son los que se resumen a continuación:<sup>41</sup>

*Datos completos:* Los datos puestos a disposición pública por el gobierno deben ser lo más completos posibles, debiendo incluir toda la información en bruto (obviamente respetando la privacidad de las personas), los metadatos y las fórmulas y explicaciones de cómo se han calculado los datos.

*Primariedad:* Debe tratarse de fuentes primarias, esto es, las fuentes e información original recolectadas por el gobierno, así como los detalles de cómo se recolectaron.

*Oportunidad:* Siempre que sea posible, la información deberá ser puesta a disposición tan rápidamente como es recogida y reunida. Deberá darse prioridad a los datos cuya utilidad sea sensible al tiempo.

*Facilidad de acceso físico y electrónico:* Los conjuntos de datos publicados por el gobierno deberán ser lo más accesibles posible. La accesibilidad es la facilidad con la que se puede obtener la información, ya sea a través de medios físicos o electrónicos. Se evitarán las barreras para el acceso físico (por ejemplo, visitar personalmente una oficina) y para el acceso electrónico (por ejemplo, la imposición de una tecnología o la previa presentación de formularios).

*Lectura por máquinas:* La información debe ser almacenada en formatos que, siendo ampliamente utilizados, permitan el procesamiento por máquinas.<sup>42</sup>

*No discriminación:* Cualquier persona debe poder acceder a los datos en cualquier momento sin necesidad de identificarse o dar ninguna justificación por realizar dicho acceso.

*Utilización de estándares abiertos:* Los datos deben estar almacenados en un formato que corresponda a un estándar no privativo, esto es, que no sea necesaria una licencia de software para acceder a los mismos.

*Licencia:* Los datos no deben tener condiciones legales, requisitos de atribución de autoría, restricciones de difusión que actúen como barreras para su uso público.

---

<sup>40</sup>«Each one exists along a continuum of openness, and the list writ large is intended as a guidebook, not a rulebook.» en <<http://sunlightfoundation.com/blog/2010/08/11/a-principled-look-at-open-data/>> Fecha de última consulta: 13 de septiembre de 2011.

<sup>41</sup><<http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>> Fecha de última consulta: 13 de septiembre de 2011.

<sup>42</sup>El caso descrito en la introducción vulnera este principio, puesto que los archivos PDF con imágenes incrustadas no pueden tratarse por las máquinas. De ahí el trabajo colaborativo para extraer la información y escribirla en una hoja de cálculo.

*Permanencia:* La permanencia es la capacidad de encontrar información a través del tiempo. La información debiera estar accesible a perpetuidad mediante archivos en línea en los que, además, se indiquen las fechas de actualización, cambio o borrado, cuestiones también aplicables a cualquier flujo de información.

*Costes de uso:* El coste de uso debe ser el mínimo posible, sin que deban trasladarse al usuario de los datos los costes que hubiera tenido que pagar el gobierno en cualquier caso.

Si bien los principios señalados por la *Sunlight Foundation* señalan cómo el gobierno ha de poner a disposición la información pública, estos principios deben entenderse aplicables a todo ejercicio de transparencia no sólo de cualquier organismo público (incluyendo los de cualquier poder constitucionalmente diseñado) sino también a la información que deben publicar por obligación legal las entidades privadas (por ejemplo declaraciones ante la Comisión Nacional de Valores o cualquier declaración sobre responsabilidad social corporativa).

Además de los requisitos que acabamos de enumerar, los activistas se preocupan de las condiciones legales aplicables en la jurisdicción desde la que operan. Particularmente en Europa existe sobre las bases de datos el denominado derecho «sui generis» que consiste en un derecho del fabricante de la base de datos por el tiempo y los recursos efectuados en la construcción de la misma. De esta manera, en la obra producida por estos grupos existe una doble regulación: la de la propiedad intelectual por la originalidad de la obra y la del derecho «sui generis» en favor del fabricante de la base de datos. Para garantizar la libertad del producto resultante, se especifican tres cuestiones legales: (i) se adhiere una licencia libre sobre la obra, (ii) se establece una cláusula liberando el derecho «sui generis» y (iii) se especifica si los datos contenidos en la obra son meras cuestiones fácticas, en cuyo caso no son susceptibles de regulación por propiedad intelectual.

Es un hecho que los principios de open data se incumplen sistemáticamente<sup>43</sup> e incluso en España se produce una situación peor, ya que es uno de los cinco países europeos que no dispone de una ley de transparencia de los datos

---

<sup>43</sup>Son excepciones a fecha de escritura de este artículo las iniciativas de las administraciones del Principado de Asturias, Islas Baleares, Ayuntamiento de Barcelona, Cataluña, Ayuntamiento de Córdoba, Ayuntamiento de Gijón, Navarra y País Vasco. El Ayuntamiento de Zaragoza, si bien publicita su portal como de open data, sin embargo requiere que el usuario rellene un formulario electrónico donde ha de ingresar sus datos personales y manifestar el propósito de su tratamiento de la información, por lo que no puede considerarse open data.

públicos<sup>44</sup> por lo que se dificulta sobremanera el acceso ciudadano a los datos hasta en formato papel. Sin embargo, los principios de open data suponen una guía para los activistas que extraen y liberan el material documental. Se produce de esta manera la paradoja de que funciones y competencias que debieran ser obligación de los organismos públicos se están realizando por parte de ciudadanos que emplean su tiempo y sus recursos personales en ello, sin coste alguno para el Estado. La motivación de la ciudadanía ha sido estudiada<sup>45</sup> por los economistas ARIELY et al. (2005, 21) quienes hallaron que los incentivos no tienen efectos a partir de un determinado nivel de ingresos e incluso, dado que los incentivos suponen costes para quienes los entregan, la subida de los incentivos a partir de un determinado punto puede ser una propuesta perdedora. Se trata de la misma motivación que la que encontramos entre los voluntarios que escriben en la Wikipedia o producen Software Libre y que nunca será evaluable económicamente para quienes contribuyen. Tal y como se señaló, la actividad de los grupos de voluntarios se enmarca en los procesos de «peer production» estudiados por Benkler (2006) y si bien no existe ninguna obligación legal para los participantes en estos procesos de realizar sus acciones conforme unos determinados principios, sin embargo su cumplimiento se realiza debido a que suponen la mejor *lex artis* y tienen en cuenta dinámicas de reutilización de la información en las que se intenta evitar cualquier obstáculo para los siguientes reutilizadores. En definitiva, se procura enriquecer ese «common pool of information» estudiado por Hess y Ostrom (2001) desde una perspectiva no economicista sino de necesidad de transparencia democrática, conscientes de que para insistir en dicha transparencia los datos han de ser de acceso libre y deben poderse tratar de manera automatizada.

## **Praeter Orwell**

Es bien conocida la premonición orwelliana de la utilización de la tecnología por parte del poder político para controlar a los ciudadanos.<sup>46</sup> Sin embargo, hasta fechas recientes no se ha tomado conciencia de que la tecnología puede cumplir

---

<sup>44</sup> Los otros cuatro países que, junto con España, no disponen de ley de transparencia son: Chipre, Grecia, Luxemburgo y Malta.

<sup>45</sup>Federal Reserve. Bank of Boston: Large Stakes and Big Mistakes <<http://www.bos.frb.org/economic/wp/wp2005/wp0511.htm>> Fecha de última consulta: 14 de septiembre de 2011.

<sup>46</sup>La Constitución española es legataria de este sistema de pensamiento orwelliano al establecer en su artículo 18.4 que «La ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos».

la función inversa, praeterorwell, y que seamos los ciudadanos los que la utilicemos para ejercer un control sobre las personas y organismos que gestionan la res publica. Hay que mencionar como punteras las iniciativas de la *Sunlight Foundation*<sup>47</sup> y de la *Open Knowledge Foundation*<sup>48</sup> en los Estados Unidos y Gran Bretaña, respectivamente, quienes se hallan desarrollando software para el control de las actividades públicas. Los proyectos que realizan abarcan los ámbitos ejecutivo, legislativo y judicial pero no existe una vinculación de los datos entre las diferentes iniciativas que realizan. Por ejemplo, la iniciativa sobre el Congreso Abierto<sup>49</sup> no contempla cruzar los datos de un congresista con los de la Administración pública, cruce de datos que podría permitir conocer si ha existido un sistema de «puertas giratorias» entre los poderes ejecutivo y legislativo.

Ahora bien, las preguntas que debemos hacernos son: ¿Es posible sistematizar la información de tal manera que permita la conexión entre iniciativas de control político? Y, en el supuesto de que la sistematización fuera posible, ¿cómo podrían implantarse unos estándares que fueran seguidos por las diversas iniciativas?

A continuación, realizo un esbozo de propuesta como respuesta a la primera cuestión. Con respecto a la segunda de las preguntas, será necesario desarrollar el código para la interoperabilidad y darle la suficiente difusión para intentar su implantación por parte de los activistas políticos.

Para sistematizar el control del poder que pudieran intentar los ciudadanos utilizando la tecnología a su alcance, los datos necesarios inicialmente se indican en la siguiente tabla:

<b>Poder</b>	<b>Personas</b>	<b>Órganos</b>	<b>Información producida</b>
Constituyente	Pueblo soberano	Cortes constituyentes	Constitución
Ejecutivo	Políticos y funcionarios	Gobierno y órganos administrativos	Actos administrativos y reglamentos
Legislativo	Parlamentarios	Parlamentos	Leyes
Judicial	Jueces y Magistrados	Juzgados y Tribunales	Resoluciones

<sup>47</sup><<http://sunlightfoundation.com/>> Fecha de última consulta: 17 de septiembre, 2011.

<sup>48</sup><<http://okfn.org/>> Fecha de última consulta: 17 de septiembre, 2011.

<sup>49</sup>Open Congress <<http://opencongress.org/>> es una iniciativa conjunta de la Sunlight Foundation y la ONG <<http://www.participatorypolitics.org/>> Fecha de última consulta: 17 de septiembre, 2011.

Poder	Personas	Órganos	Información producida
			judiciales
Otros órganos constitucionales	Miembros (Magistrados, Consejeros ...)	Órganos (Tribunal Constitucional, Tribunal de Cuentas...)	Nombre de la información (sentencias, informes...)
Legislativos autonómicos	Parlamentarios	Parlamentos	Normas autonómicas
Ejecutivos autonómicos	Políticos y funcionarios	Gobierno y órganos administrativos	Actos administrativos y reglamentos
Administración local	Diputados provinciales, alcaldes, concejales	Diputaciones provinciales y municipios	Normativa local
<b>Organización territorial</b>			

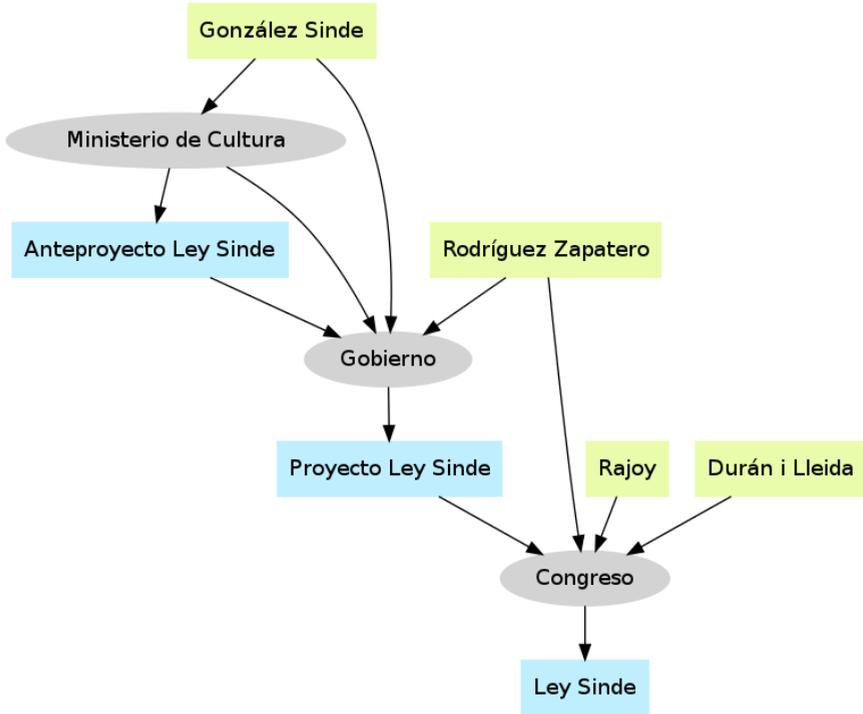
Los datos anteriores no son suficientes para ejercer un control político, pero sí son necesarios para cualquier desarrollo informático que se quiera programar. Los datos se han clasificado en función de una tríada compuesta por dos elementos subjetivos (personas y órganos formados por personas) que generan un elemento objetivo, que es la información. La relación entre los componentes se realiza mediante grafos, que se muestran en la Figura 3:



**Figura 3**

Toda información producida por el Estado puede ser clasificada utilizando estos tres parámetros: la (i) persona que compone el (ii) órgano para generar la (iii) información. En la Figura 4, se representa una hipotética denominada «Ley Sinde» para señalar un ejemplo de relaciones entre personas que usan órganos para generar una información. Los grafos corresponden a una hipotética norma ya que la denominada «Ley Sinde» se tramitó en comisión, no en el pleno del

Congreso. Lo relevante de la representación es ver las relaciones entre personas, órganos e información generada por las personas utilizando los órganos.



**Figura 4**

Toda la información open data que se extraiga por los ciudadanos puede incluirse en alguna de las 24 categorías propuestas: una relación de diputados son las personas pertenecientes al poder legislativo y que son parlamentarios, su relación de bienes se vincula a los mismos y expresándose tal vinculación, pertenecerá al agregado informacional predicable de un miembro del parlamento que integre el sumatorio de toda la información que pueda difundirse en línea sobre él. Imaginemos que el parlamentario es propietario de, por ejemplo, un establecimiento farmacéutico; en ese caso puede verificarse qué votaciones ha realizado en normas jurídicas de ese sector. La estructura propuesta, al contemplar a la persona como primer elemento de la tríada, permite vincular a la misma todos y cada uno de los órganos con los que ha estado relacionado y, a su

vez, la información que dichos órganos han generado hallándose la persona vinculada al órgano.

Las categorías propuestas sirven de poco sin las herramientas informáticas que las utilicen. El modelo ha de completarse con los siguientes elementos:

Una múltiple estructuración de los datos a fin de facilitar cualquier utilización de los mismos. Inicialmente, los datos han de archivar en formato de valores separados por comas (CSV)<sup>50</sup> que los estructura por filas y columnas, pero ha de hacerse una propuesta para bases de datos relacionales y otra para tripletes de la web semántica.<sup>51</sup> Se busca de esta manera evitar los problemas de integración de las fuentes en las bases de datos, liberando al usuario de tareas tales como encontrar las fuentes relevantes (Borgida, Lenzerini y Rosati, 2003, 517).

Un directorio de enlaces a los archivos de datos extraídos, ya operativo en la URL <http://opengov.es>.<sup>52</sup>

Una ontología jurídica libre para modelar las 24 categorías y la estructura de la información (por ejemplo, las partes de una norma jurídica) conforme los requisitos de la web semántica. Esta ontología se halla actualmente en desarrollo en la URL <http://purl.org/derecho>.

Por último, aplicaciones de gestión y visualización de los datos extraídos.

## Límites de la propuesta

Ahora bien, pensar que mediante la estructura propuesta se puede lograr un control del poder significa que se desconocen las relaciones económicas u otras relevantes e intervinientes en la toma de decisiones. La experiencia nos demuestra que el poder no se manifiesta de una manera transparente, sino ocultando las relaciones verdaderamente importantes: por ejemplo las relaciones de corrupción, de intereses o incluso las sentimentales cuando pudieran recibir un reproche público. Tampoco se incluye en las categorías propuestas actores fundamentales en el ejercicio del poder como lo pudieran ser los partidos políticos o los sindicatos. Como ya se ha mencionado anteriormente, los datos que se proponen son necesarios, pero no son suficientes y es de sentido común

---

<sup>50</sup> Acrónimo de «Comma Separated Values».

<sup>51</sup> La web semántica utiliza un sistema de tripletes para modelar la información. Cada triplete consta de un sujeto (S), un predicado (P) y un objeto (O). Por ejemplo: (S) Cervantes (P) escribió (O) El Quijote.

<sup>52</sup> Actualmente, el servidor sito en <<http://opengov.es>> donde se indican los enlaces a los datos que se van extrayendo.

que si se desea programar aplicaciones para conocer el voto de los parlamentarios, lo primero que se necesita es la relación nominal de los mismos y, en estructuras políticas como las de un estado central y varios estados federados (o un sistema con Comunidades Autónomas o Länder), lo más básico es una relación de parlamentos más la relación de parlamentarios por cada una de las cámaras.

Además de lo que nos muestra el sentido común, el actual sistema político se halla lleno de lo que podríamos denominar lagunas informacionales no documentadas, de las que simplemente señalamos dos: En primer lugar, los textos de las enmiendas transaccionales en el procedimiento de aprobación de una norma jurídica no aparecen más que en su redactado final, no existiendo la documentación de la discusión en comisión, lo que es fundamental para la explicación de la ratio legis y privándonos de ella cuando queremos investigar los antecedentes de una norma para encontrar su hermenéusis. En segundo lugar, existe información relevante relativa al diseño constitucional que no es competencia de ninguna de las administraciones públicas y sobre la que no existe documentación online, por lo que escapa a cualquier técnica manual o automatizada de obtención de información y ha de ser construida ex novo, siendo el mejor ejemplo el de la relación de domicilios y teléfonos de los parlamentos autonómicos, relación que ha de construirse a mano.

## **Conclusión**

En la actualidad, tenemos unos actores, unas herramientas, una forma de trabajar en red y unos contenidos que puestos a disposición universal pueden lograr que cualquiera pueda acceder al conocimiento con independencia de su lugar de nacimiento y origen social. Sin embargo, se siguen estableciendo barreras físicas y legales al acceso a la información y al conocimiento, perpetuándose el statu quo de un modelo de sociedad que cada día más nos demuestra que es esencialmente injusto. Hemos de recurrir a la tecnología, pero WINNER (2008, 134) nos advirtió contra los entusiasmos de la «tecnología adecuada»: «Son evidentes las insuficiencias de estas ideas. Los tecnólogos adecuados no estaban dispuestos a enfrentar los hechos del poder social y político organizado. Fascinados por el sueño de una revolución espontánea y de base, evitaron todo análisis profundo de las instituciones que controlan la dirección del desarrollo tecnológico y económico. En medio de esta feliz confianza en sí mismos, no se preocuparon por elaborar las estrategias que podrían haberlos ayudado a superar los evidentes puntos de resistencia». Coincidiendo con él, no debemos pensar

que basta la mera existencia de una tecnología para que surja la Justicia o se evite la desigualdad.

Por ello, como ciudadanos debemos ir más allá. No es aceptable éticamente un sistema en el que la ciudadanía financia la creación de datos a los que luego se le impide el acceso o se le entregan por los poderes públicos de una manera desestructurada y en formatos inverosímiles. Tampoco puede aceptarse gerencialmente tamaña incompetencia: si nuestros poderes públicos no saben tratar los datos correctamente, ¿cómo pueden pretender el ejercicio de una correcta gestión? Es un sinsentido del mundo al revés que los poderes públicos hagan mal las cosas, teniendo luego que venir los ciudadanos a reparar en cuatro días lo que los organismos públicos no han sabido hacer. Además, se permiten legislar sobre temas que desconocen, por lo que generan normas injustas e ineficaces: «si se quiere legislar sobre tecnología, habrá que estudiar tecnología».<sup>53</sup> Y, por último, políticamente se le está hurtando a la ciudadanía un capital civil que le pertenece (Tauberer, 2009).

Y como no puede aceptarse ni ética ni gerencial ni políticamente la solución actual, la toma de la bastilla digital consistirá en que si el poder no nos da los datos, los ciudadanos los retomaremos en el legítimo ejercicio de la vieja desobediencia civil y cuya responsabilidad tan adecuadamente nos justifica Rawls (1995, 354):

Empero, si la desobediencia civil justificada parece amenazar la concordia cívica, la responsabilidad no recae en los que protestan, sino en aquellos cuyo abuso de poder y de autoridad justifica tal oposición, porque emplear el aparato coercitivo del Estado para mantener instituciones manifiestamente injustas es una forma de fuerza ilegítima a la que los hombres tienen derecho a resistir

## Referencias

ARIELY, Dan; GNEEZY, Uri; LOEWENSTEIN, George y MAZAR, Nina. (2005) “Large Stakes and Big Mistakes. Working Paper 05-11”. *Working Papers of the Federal Reserve. Bank of Boston*. Accesible en línea: <<http://www.bos.frb.org/economic/wp/wp2005/wp0511.htm>> Fecha de última consulta: 14 de septiembre de 2011.

BAUMAN, Zygmunt (2010). *Mundo consumo. Ética del individuo en la aldea global*. Paidós contextos. Barcelona.

<sup>53</sup><<http://twitter.com/#!/jdelacueva/status/114370344537702401>> Fecha de última consulta: 18 de septiembre, 2011.

BENKLER, Yochai (2006). *The Wealth of Networks. How Social Production Transforms Markets and Freedom*. Yale University Press. New Haven y Londres. Accesible en línea: <[http://www.benkler.org/Benkler\\_Wealth\\_Of\\_Networks.pdf](http://www.benkler.org/Benkler_Wealth_Of_Networks.pdf)> Fecha de última consulta: 15 de septiembre de 2011.

BORGIDA, Alex; LENZERINI, Maurizio y ROSATI, Riccardo (2003). *Description Logics for Databases, en The Description Logic Handbook*, Cambridge University Press. Cambridge.

CARCELER HONTORIA, Víctor y LOPEZ ARNAL, Salvador (2003). *Ética hacker-científica contra ética mercantil: del software propietario al software libre*, en Popper/Kuhn. Ecos de un debate. Editorial Montesinos. Barcelona.

CASTELLS, Manuel (2009). *Comunicación y poder*. Alianza Editorial. Madrid.

HESS, Charlotte y OSTROM Elinor (2001). "Artifacts, Facilities, And Content: Information as a Common-pool Resource". Accesible en línea: <<http://www.law.duke.edu/pd/papers/ostromhes.pdf>> Fecha de última consulta: 14 de septiembre de 2011.

MARRES, Noortje S. (2005). "No issue, no public: democratic deficits after the displacement of politics". Amsterdam, 2005. Accesible en línea: <<http://dare.uva.nl/document/17061>> Fecha de última consulta: 14 de septiembre de 2011.

MONTERO AROCA, Juan (1997). *La legitimación colectiva de las entidades de gestión de la propiedad intelectual*. Editorial Comares. Granada.

RAWLS, John (1995). *Teoría de la Justicia*. Fondo de Cultura Económica. México D.F.

TAUBERER, Joshua (2009). "Open Data is Civic Capital: Best Practices for Open Government Data". Accesible en línea: <<http://razor.occams.info/pubdocs/opendataciviccapital.html>> Fecha de última consulta: 13 de septiembre 2011.

VALENCIA SAIZ, Angel. (1995). "La teoría política en la era de la tecnocracia", en VALLESPIN OÑA, Fernando (ed.), *Historia de la Teoría Política*, 6. Alianza Editorial. Madrid.

WINNER, Langdon (2008). *La ballena y el reactor. Una búsqueda de los límites en la era de la alta tecnología*. Editorial Gedisa. Barcelona.