

Enrique Alonso

Resumen. La edad de oro de la lógica contemporánea parece asociada al estudio de la noción de *teoría formalizada* y a las consecuencias que se siguen de todo ello. En el momento actual esta noción no parece capaz de suscitar entre nosotros el interés y fascinación que en su día tuvo para los padres fundadores. Pese a ello aún seguimos haciendo girar nuestras investigaciones en torno herramientas directa o indirectamente dedicadas al estudio de esta noción. Quizá sea el momento de buscar otros conceptos necesitados de análisis formal y capaces de volver a conectarnos con las preocupaciones que la ciencia tiene en estos momentos. Pretendo defender que la noción general de *lenguaje etiquetado* puede ocupar ese lugar si sabemos entender adecuadamente su importancia en la construcción de la *Sociedad del Conocimiento*.

2.1. Algunas buenas razones

No es costumbre entre nosotros agradecer el trabajo de años haciendo aquello que mejor sabemos hacer, es decir, escribir acerca de las cosas que nos interesan. Esta es la primera vez que participo en una iniciativa como esta y me gustaría pensar que no va a ser la última. Cuando conocí a Ángel me temo que ninguno de los dos tenía claro su futuro. Yo aún tardé en resolver algo más el mio pero lo cierto es que sobrevivimos a momentos que, vistos en perspectiva, parecen aún más difíciles de lo que entonces nos permitimos considerar. Quizá por eso nos hemos sentido siempre bien cuando hemos coincidido en algún Congreso, por eso, y por una sana afición a relajarnos repasando los numerosos dichos y sucedidos cómicos que inevitablemente tienen lugar una comunidad como la nuestra¹.

Profesionalmente hemos mantenido siempre trayectorias adyacentes, no directamente relacionadas pero lo suficientemente afines al menos como para que siempre me haya interesado leer y escuchar su investigación. Quizá no es momento este para comentar sus aportaciones pero al menos sí me gustaría valorar algo que me parece importante. Vivimos una época dominada por modelos académicos altamente excluyentes, aquellos que no responden a la biografía tipo impuesta en lejanas instancias de nuestro sistema

¹Quizá no deberíamos dejar pasar mucho más tiempo antes de pasarlas al papel. Sería una auténtica lástima que algunas de ellas se perdieran o deformaran hasta lo irreconocible.

investigador suelen encontrar dificultades crecientes para acceder a los escasos recursos disponibles. Una de las cosas que suelen ser motivo de penalización es mostrar un currículum investigador dedicado a más de una materia. Hay otras cosas, por supuesto, pero no es el momento de hablar de ellas. Ángel ha tenido la valentía de tocar tantos palos como su curiosidad y decencia investigadora le ha sugerido y eso es algo bueno, muy bueno. Espero que siga así.

2.2. Un cambio de modelo

Hace tiempo que mis preocupaciones se reparten, casi a tiempos iguales, entre ciertas investigaciones que podrían caer perfectamente en el ámbito de la historia y la filosofía de la lógica contemporánea y otras, muy distintas, orientadas a la aplicación de la lógica en lo que se ha dado en llamar *Sociedad Digital*. Este giro hacia las aplicaciones tiene una justificación que creo que debo aportar. La era dorada de la lógica, aquella vivida desde finales del siglo xix hasta mediados del xx, ha dejado una cantidad ingente de trabajo de gran calidad y también la impresión de que aquello que la lógica aspiraba a resolver no tiene soluciones, o al menos no las esperadas. Es poco probable que en los próximos años asistamos a la presentación de resultados como los que en su día supusieron los teoremas de Löwenheim-Skolem, Tarski, Gödel, o Henkin, por poner solo algunos ejemplos. Tampoco parece posible ir mucho más allá en la definición de *nuevas lógicas*, carrera ésta, que además ha generado serias dificultades para saber en definitiva *qué es eso de la Lógica*². Agotados estos caladeros, ¿qué hacer?

La incorporación a Internet de las habilidades aportadas por el *HTML*³ contribuyó de forma decisiva a construir lo que ahora denominamos, no sin problemas, *Sociedad Digital o de la Información*. Pocos podrían haber supuesto entonces que la estructura de ese lenguaje fuera a contribuir de forma tan decisiva a una transformación tan radical de nuestros hábitos y costumbres. La lógica, esa es al menos mi opinión, tiene algo que decir en un ámbito hasta ahora dejado a manos poco expertas y desde luego nada sensibles a algunos de los problemas que nuestra disciplina tan bien conoce. La Red exige, cada vez más, formas de integrar la actividad que llevan a cabo los seres humanos con aquella que realizan ciertos sistemas de forma automática. Estas formas de integración están siendo gestionadas por lenguajes que de un modo u otro intentan añadir cierto contenido conceptual a los mensajes que los seres humanos producimos en la Red. Hubiera sido muy deseable que la IA clásica hubiera sido capaz de resolver el problema suministrando herramientas apropiadas para *entender* lo que los seres humanos pretendemos hacer con las palabras, pero es un hecho que esto no ha llegado aún y apenas sabemos si algún día será posible. Pero el futuro es muy capaz de encontrar por sí mismo vías para abrirse camino y en este sentido la integración del conocimiento en la Red es, a fecha de hoy, uno de los más prometedores.

Para lograr este fin hay muchos mecanismos, pero el que más me interesa es el uso de lenguajes de etiquetas -tag systems- como forma de añadir estructura conceptual a

²Expresión irónica donde las haya, ya que *Was ist das, die Philosophie?* es, uno de los títulos más celebrados de Martin Heidegger, personaje poco a nada afín a nuestros intereses teóricos.

³HTML significa, por si alguien aún lo dudada, "Hypertext Markup Language" y es el lenguaje que Tim Berners-Lee ideó en 1989-91 para ofrecer un acceso amigable a la información del CERN y en la actualidad constituye la infraestructura que aún se observa en las páginas web que habitualmente visitamos.

nuestras transacciones de información en la Red. Un lenguaje de etiquetas constituye una entidad formal a la que por el momento no hemos prestado la atención que merece. Es cierto que se ha hecho algún trabajo muy prometedor en el ámbito de las lógicas del conocimiento -independientes del origen de Internet- llevando a la definición de lenguajes como OWL, cuyo uso está bastante extendido hoy en día en la denominada *Web semántica*, pero creo que aún no tenemos clara su importancia en los nuevos modelos adoptados por la Red.

Los lenguajes de etiquetado pueden ser descritos como una colección de etiquetas -tags- capaces de acompañar a cualquier pieza de información de forma reconocible e independiente. Lo cierto es que no tenemos nada que sea mejor como definición, pero por el momento es suficiente. Las etiquetas suministradas por estos lenguajes son estructuras que toman como argumento una secuencia de datos -del tipo que sean- que pueden acotar de forma absolutamente precisa. Las etiquetas pueden admitir o no modificadores, los cuales tienen a su vez un rango de valores previamente definido extraídos de conjuntos de lo más variado. Quizá un ejemplo aclare el punto. Supongamos que la estructura de datos la ofrece un texto escrito en lenguaje ordinario al que deseamos añadir cierto formato tipográfico. “Hoy hace un día precioso” es una oración a la que queremos dotar de un aspecto algo más cuidado, no importan los matices, ni las intenciones, sólo el formato. Para que la palabra “precioso” aparezca en cursiva dando lugar a “Hoy hace un día *precioso*” es preciso indicarlo mediante alguna etiqueta que exprese esa acción y que sea capaz de indicar de forma exacta el fragmento de texto al que se aplica. La estructura real de esta nueva pieza de información podría ser, dependiendo del sistema de etiquetado elegido, algo como esto: “Hoy hace un día `<itálicas>precioso</itálicas>`”. En aquellos casos en los que la etiqueta admita modificadores, como por ejemplo, en el caso en que el texto aumente de tamaño punto a punto, la pieza de información podría adquirir un aspecto similar a lo siguiente: “Hoy hace un día `<talla pt=18>precioso</talla>`” dando como resultado “Hoy hace un día PRECIOSO”.

Aquellos que estén familiarizados con la estructura de los documentos HTML, es decir, con las páginas web que pueblan la Red, reconocerán perfectamente la sintaxis característica de este lenguaje, pero existen otras opciones, todas ellas válidas. El lenguaje de etiquetado en el que he compuesto este texto, el \LaTeX , emplea una sintaxis completamente diferente pero igualmente eficaz. Por otra parte, no es cierto que sea HTML el primer lenguaje en hacer uso de la sintaxis basada en las etiquetas de apertura y cierre, es decir, en servirse del formato básico dado por:

```
<etq atrib1 = .....atribn = > </etq>
```

El origen que parece más cierto para esta peculiar sintaxis se remonta al lenguaje SGML estandarizado por una norma ISO de 1986 pero procedente en realidad de un desarrollo de los ingenieros de IBM que dataría de los años 60 y que fue empleado para producir informes técnicos legibles desde cualquier plataforma.

Los lenguajes de etiquetado son sistemas como este, pero, ¿qué más sabemos de ellos?

2.3. Lecciones de la Historia

La edad de oro de la lógica está íntimamente asociada al interés por la noción de *teoría* o más precisamente, de *teoría formalizada*. Quien conozca un poco la historia de aquellos años maravillosos -el último tercio del siglo XIX hasta la primera mitad del XX- sabe de sobra lo difícil y tentativo que pudo resultar en un principio delimitar esa noción básica y

sobre todo establecer sus propiedades fundamentales. No me refiero siquiera al hecho de si ésta o aquella teoría satisfacía o no alguna de las propiedades características previamente identificadas, sino al hecho mismo de entender con precisión cuáles eran las propiedades que correspondían al concepto de teoría formalizada. Pensemos, por ejemplo, en la *completitud* predicada de una teoría: dependiendo de si se trata de una teoría totalmente interpretada o un sistema formal, significará una cosa u otra. Este y otros matices fueron durante mucho tiempo objeto de controversia y no pocos malentendidos, algunos de los cuales se prolongan hasta nuestros días. ¿Qué hace, por ejemplo, que un procedimiento determinado de evaluación de la consecuencia lógica sea considerado como perteneciente al dominio de los *cálculos* y otro, totalmente próximo, se ubique en el dominio de la *semántica* o *teoría de modelos*?

Estas y otras cuestiones muestran la dificultad inicial para delimitar con precisión qué es una teoría formal, qué nociones son importantes a la hora de caracterizarlas adecuadamente y qué relaciones se dan entre estas propiedades. Costó mucho resolver estos problemas y creo que se puede decir que fue, precisamente, al abordarlos cuando la lógica se cobró alguna de sus mejores presas.

Líneas más arriba he dado una somera descripción de qué es un *lenguaje de etiquetas* indicando que ésta podía tomarse a lo sumo, como una aproximación al tipo de problemas que estas entidades plantean. Es obvio que nos falta mucho aún para llegar a una definición, aunque no me atrevo a decir que tal cosa no exista ya en alguna esquina de la academia o la industria de la cual no nos ha llegado noticia. Lo que sí puedo decir es que en este punto aprecio un cierto desinterés, perplejidad tal vez, por parte de la comunidad que tradicionalmente se ha dedicado al fomento de la lógica. Los lenguajes de etiquetado no tienen el aspecto de los sistemas que hemos aprendido a estudiar y caracterizar con precisión de cirujano, no son el tipo de cosas de las que uno entiende, y sin embargo qué cerca están, en realidad, de las que quizá nos deberían interesar. Tal vez sea excesivo pretender comparar la importancia que para la ciencia de la lógica tuvo en su día el análisis del concepto de teoría formalizada con lo que ahora podría suponer el estudio de la noción de *lenguaje de etiquetas* o *marcas*, pero no me resisto a pensar en ello como una posibilidad. Si la lógica contemporánea, la que indirectamente he venido a considerar como la versión 1.0, pudo y supo prosperar mediante el estudio de la noción de teoría formalizada, ¿por qué no imaginar que la lógica 2.0 tenga alguna oportunidad con la de lenguaje de etiquetas? No pretendo, ni mucho menos, hacer de esta temeraria afirmación una tesis que defender, sino tan solo una oportunidad de trabajo, sólo eso.

Los lenguajes de etiquetas, o marcas, son sistemas formados por etiquetas que toman como argumentos datos extraídos de un dominio previamente fijado. Cabría preguntarse, por ejemplo, cuál es el estatuto de la entidad resultante, es decir, de la pieza de información obtenida al etiquetar una cadena de datos. Cabría preguntarse, igualmente, cuántos lenguajes de etiquetado pueden concurrir simultáneamente sobre un mismo fragmento de información, o qué relaciones cabe esperar entre etiquetados de diverso tipo. Pero quizá la más importante no sea sino aquella que se preguntaría por la relación entre la entidad etiquetada y el resultado de ejecutar las acciones que tales etiquetas aportan. En el ejemplo anterior parece claro que el objeto que resulta de aplicar “< *italica* >< /*italica* >” a la palabra -el dato en este caso- “precioso” es “*precioso*”. Ahora bien, no siempre es tan fácil. En muchas ocasiones el etiquetado no tiene como efecto una expresión tan fácil de reconocer. Hay ocasiones en las que se trata de un contenido alojado en un espacio que previamente se ha definido y sobre el que se han establecido relaciones y propiedades

específicas. Estos son los problemas a los que me refiero.

2.4. ¿De qué está hecho el futuro?

Supongo que aún debo una explicación acerca de por qué creo que los lenguajes de etiquetas son un objeto digno de nuestra atención. Como ya mencioné de pasada líneas atrás, alguno de estos lenguajes forman la infraestructura en que está escrita la Red. HTML, XML, RDF, OWL, etc son todos ellos lenguajes de etiquetas o marcas. Creo que esto bastaría por sí solo para explicar parte de mi interés en este asunto, pero soy de la opinión de que aún puede decirse algo más.

La gran revolución habida en la Red en los últimos años, la asociada a lo que solemos denominar Web 2.0, no se produce tanto por un cambio radical en las infraestructuras que la gobiernan como en la forma de usarlas. La Web 2.0 es interactiva, dinámica, capaz de entrar en diálogo con el usuario, esto es al menos lo que se dice, pero ¿qué significa que la Red sea capaz de interactuar con el usuario? La Web 1.0, en su diseño original, es una colección de documentos escritos en HTML conectados entre sí mediante un recurso característico, los enlaces -links-. Los enlaces entre páginas se introducen mediante una etiqueta que toma como argumento un fragmento de texto, u otra entidad perteneciente al rango de los datos del documento, y le asocia la acción de conducir a otro objeto, generalmente otra página. Lo único que se puede hacer en la Web 1.0 es ir de documento en documento siguiendo la red de enlaces que estos contienen. La idea de “navegación” procede de esta experiencia, aunque no me resulta del todo claro por qué se habla de “navegar” y no más bien de otra cosa. En la Red tradicional no existe la posibilidad de que el usuario entre en contacto con el sistema que suministra los datos por la sencilla razón de que no hay ningún canal de comunicación definido. La Web 2.0 nace cuando un recurso, disponible de antiguo, se empieza a usar con ese fin. Me refiero a los formularios y botones que con mayor frecuencia se han ido incorporando en las páginas que ahora son de uso común. Un formulario o un botón producen, aunque no sea evidente en un principio, el etiquetado de un dato. Quizá sea más claro en el primer caso. Cuando relleno un formulario, en realidad, cuando relleno cualquier cuadro de texto o cuando elijo una opción entre aquellas que se me ofrecen, o cuando marco una casilla, estoy produciendo un contenido etiquetado que el sistema -servidor- que me ha suministrado la página original está en condiciones de interpretar. Los formularios toman ese nombre de su uso original como hojas de pedido en las transacciones comerciales o solicitudes que empezaron a tramitarse en la Red hace ya algún tiempo y que proceden a su vez de los modelos vigentes en la sociedad predigital. En la actualidad los formularios han perdido su aspecto inicial para convertirse en estructuras típicas, a menudo irreconocibles, en nuestras páginas. Las redes sociales, o los primeros modelos de *realidad aumentada* son el resultado del uso de toda la potencia que el uso de lenguajes etiquetados aporta a la Red como medio de comunicación.

Cuando relleno un cuadro de texto en el que figura una etiqueta de “Apellido” no sólo estoy suministrando la información solicitada, sino que lo hago además de forma que el sistema que la recibe pueda gestionarla *sabiendo* que se trata de un apellido y no de otra cosa. Cuando escribo mi nombre y apellidos en una hoja en blanco he aportado la misma información si el destinatario es otro ser humano, es decir, alguien capaz de aportar todo el aparato cognitivo que nos permite entender una redacción en el lenguaje ordinario. Los sistemas artificiales no han sido capaces de avanzar demasiado en esa dirección pese

a todos los intentos hechos durante décadas por diversos programas en IA. La Red, sin embargo, ha encontrado una forma sutil, ingeniosa y rápida de suplir esa deficiencia. Ha encajado la interacción entre usuarios y sistemas artificiales en canales predefinidos que aportan, mediante el uso de sistemas de etiquetado, el valor cognitivo que el sistema no es capaz por sí mismo de obtener y que solo otro ser humano puede gestionar de forma adecuada.

Los lenguajes de etiquetado son, si mi evaluación es correcta, el medio del que se sirve la Red para establecer una comunicación *cognitivamente eficiente* entre seres humanos y sistemas automáticos de forma indistinta. Y esta razón sí justifica por sí sola el interés que este tipo de entidades formales tiene para alguien que como yo, se ha formado en las técnicas y conceptos de la lógica contemporánea.

Me gustaría terminar confesando algo que posiblemente Ángel sabe, pero que no está de más volver a recordar ahora. Es muy difícil que me hubiera decidido a exponer estas tesis de forma tan escueta y tentativa si el destinatario de esta iniciativa hubiera sido otra persona. Hay ideas que uno solo acepta compartir con aquellos interlocutores con los que se siente agusto, a los que cuenta entre sus amigos y con los el diálogo sincero surge sin dificultades. Algo tan raro merece ser cuidado.