>> Jorge Dávila

Verdades y postureos sobre la identidad digital soberana

Hay indicios claros de que cierto grado de postureo y mucha falsa apariencia ronda dentro de las lides de la Ciberseguridad (que es como se llama ahora). El abuso discursivo de vacuidades como "tecnologías disruptivas" está frenando que se resuelvan problemas muy serios y todavía pendientes, como es el caso de una Identidad Digital de verdad. Recientemente se ha sacado a la palestra el binomio Identidad y Blockchain, y va siendo hora de indagar qué es lo que se esconde en esa dualidad tan… "postural".

Sin saber claramente cómo, pero hemos llegado a un mundo, a una sociedad, en la que casi todo es postureo y mucho aparentar. Que eso ocurriese en ámbitos del ocio podría tener hasta su gracia. Que se infiltre y ocurra en entornos políticos es mucho más grave y probablemente haya sido algo inevitable en tanto y cuanto la política se degrada y se convierte en la mera recolección de votos para luego hacer cualquier otra cosa con el poder que otorgan. Que el postureo y el aparentar entren en los pagos de la tecnología es algo inaudito. ¿De qué sirve el aparentar frente a la necesidad real de resolver problemas?

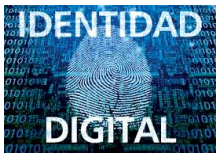
Hace poco y con gran sorpresa, en la pasada celebración del congreso Securmática pude oír como una prestigiosa institución española se lanzaba a casar la Identidad Digital, negocio en el que lleva muchas décadas, con la muy cacareada y algo cansina Blockchain. Sin profundizar en el esquivo concepto de la **Identidad**, sin concretar qué es realmente eso de **Blockchain** y a cuáles de sus múltiples y endebles sucedáneos se estaban refiriendo, los ponentes los juntaron y de ellos nació algo presentado como nuevo y rompedor, al que bautizan como "**Identidad Soberana**". ¡Qué desastre fue aquello! Y encima, hablar en nuestro país y estos días de "*Identidad*" y además calificarla de "*soberana*", ¡todo un atrevimiento!

Aquello que parecía un simple acto de publicidad corporativa para hacer creer al sector y a ellos mismos lo muy avanzada que está la I+D+i interna (aunque esté externalizada en una consultora al uso), se convirtió en un espectáculo dadaista[1] que sólo puso de manifiesto lo contaminado que está nuestro sector con el postureo y falsas apariencias.

¿Llegará algún día en que dejaremos de oír hablar de Blockchain? Yo espero que sí, ya que hay muchas cosas importantes todavía por hacer. El postureo pseudo-tecnológico sólo sirve para dificultar la resolución de problemas y para impedirnos avanzar. Llenarse la boca hablando del mucho dinero que algunos se han gastado y que otros más rezagados están gastándose ahora en "*tecnologías disruptivas*", lo único que consigue es que no haya dinero ni recursos para resolver problemas de verdad. Entre las muchas cosas que están todavía por resolver en Internet encontramos el magnífico problema de la **Identidad Digital**.

En 1997 la Administración española se lanzó a montar proyectos software de identidad digital, que consistieron esencialmente en comprar tecnología PKI americana[2] con el fin de montar la primera Autoridad de Certificación Española (ACE). Posteriormente, la Ley 59/2003 de Firma Electrónica asignó al Documento Nacional de Identidad, un objeto netamente físico, una dimensión digital que no tenía e intentó convertirse en el mecanismo de acreditación electrónica de la identidad de los ciudadanos españoles. Han pasado ya muchos años y todavía no lo han conseguido, a pesar de tener, a fecha de hoy, 69.511.297 DNIe expedidos.

Algún tiempo antes y por necesidades de la Administración de Hacienda en la recolección de las declaraciones del IRPF por Internet, la Fábrica Nacional de Moneda y Timbre y Real Casa de la Moneda (FNMT-RCM) consiguió poner en pie el denominado Proyecto Ceres para acuñar identidad digital mediante soluciones software que permiten autenticar personas (certificados x509v3) y, más adelante, garantizar la confidencialidad de las comunicaciones (certificados https) en Internet. Todavía hoy se utiliza profusamente esta solución que está construida sobre certificados digitales que fueron inventados en el verano del 1988[3].



La identidad es increíblemente compleja, multifacética y cambia con el tiempo, y dado que las personas no respondemos a ningún modelo suficientemente preciso y universal, es muy difícil establecer un nivel de confianza suficiente entre muchas partes esencialmente desconfiadas.

La identidad es la "*circunstancia de ser una persona o cosa en concreto y no otra, determinada por un conjunto de rasgos o características que la diferencian de otras*" o, dicho de otro modo, identidad es el "*conjunto de rasgos o características de una persona o cosa que permiten distinguirla de otras en un conjunto*"[4]. Desde un punto de vista más formal, la identidad *es la relación que toda entidad mantiene sólo consigo misma*, pero no es este el aspecto que preocupa en Internet, sino más bien el de la **identidad legal**. A fin de cuentas, lo que mueve la Internet que conocemos es hacer dinero y conseguir poder, y para ello hay que comerciar. Comprar y vender no es posible sin seguridad jurídica y esa sólo la dan las Leyes y los jueces de los Estados que están detrás de ellas y ellos.

La identidad realmente es increíblemente compleja, multifacética y cambia con el tiempo, y dado que las personas no respondemos a ningún modelo suficientemente preciso y universal, es muy difícil establecer un nivel de confianza suficiente entre muchas partes esencialmente desconfiadas. De lo que estamos realmente hablando en Internet es de la identidad que reconocen los jueces a la hora de establecer propietarios y resolver conflictos.

El otro componente del problema son los efectos de más de una década de la Web 2.0 o colaborativa, más conocida como Redes Sociales. Con el paso de los años ya son pocos los que no están convencidos de que el negocio de las RRSS no es otro que vender nuestra vida privada particular en crudo o destilada a través de algoritmos de AI y de Big Data. Toda esta monitorización sólo es posible porque existen otras identidades (no oficiales) que ya nos delatan (direcciones IP y MAC, perfilado de los agentes y navegadores de acceso, *smartphones*, *smartTVs*, coches conectados, *wearables*, GPS, WiFi, IoT, etc.). Por todo ello, las necesidades de anonimato o de una gestión legal de las identidades de todos y cada uno de nosotros apremian para descubrir una nueva forma de identidad[5]

[1] Ver https://es.wikipedia.org/wiki/Dadaísmo
[2] Ver https://en.wikipedia.org/wiki/Entrust
[3] Ver https://en.wikipedia.org/wiki/X.509#History and usage
[4] Ver https://www.google.com/search?q=que+es+la+identidad
[5] Ver https://youtu.be/3CWj9TqMzaU

que permita el negocio económico en Internet, pero que no lleve asociado la firma, como la del Doctor Fausto[6], de un cheque en blanco a los muchos demonios que operan y sostienen Internet.

Identidad Auto-Soberana

En este desolador escenario surge la idea, quizás utópica, de que pueda existir una identidad en la que el titular de los datos, y sólo el titular de los datos, tenga el control de los mismos. En ese ideal, se imagina a las personas (o incluso a los dispositivos) entregando sólo aquellos datos que son realmente pertinentes para la transacción u operación que desean realizar. A ese ideal, algunos lo llaman **Identidad Auto-Soberana** o **SSI**[7], que son sus siglas en inglés.

Es curioso ver lo rebuscadas o excesivamente simplistas que son las informaciones que uno puede encontrar en Internet cuando se interesa en esto de la SSI. Lo más sorprendente es que muchas de ellas lo relacionan con Blockchain y en realidad no tiene nada que ver.

Desbrozar Blockchain

Empezando a desbrozar por el lado del Blockchain, es necesario recordar que lo que hoy llamamos así, nació con otros nombres a finales de la década de los 80s y principios de los 90. En 1989, Leslie Lamport desarrolló el Protocolo Paxos[8], y en 1990 mandó el artículo titulado "*The Part-Time Parliament*" a la revista ACM Transactions on Computer Systems y tuvo que esperar hasta 1998 para que apareciese publicado. Ese artículo describe un modelo de consenso para llegar a acuerdos entre ordenadores que no necesariamente tienen por qué confiar entre sí. Por otro lado, en 1991, se propuso crear una cadena de informaciones digitalmente firmadas, organizadas a modo de registro digital, de modo que fuese muy fácil demostrar que ninguno de ellos había cambiado[9]. Esa misma irreversibilidad se propuso como elemento imprescindible de los servicios de Sellado de Tiempos (*Timestamp services*) aunque no llegó a ser incluido en el estándar[10].

Las cadenas de bloques son registros secuenciales e irreversibles compuestos por bloques. Cada bloque tiene una cabecera con metadatos acerca de lo que hay en él (transacciones), y cada cabecera incluye un enlace criptográfico irreversible (valor *hash*) relacionado con el bloque anterior. Cada una de las transacciones del bloque involucra a uno o varios usuarios de la red, y un registro de lo que declaran que ha ocurrido o han hecho. Cada transacción está digitalmente firmada por el usuario o usuarios que envían la transacción para que sea incluida dentro de la secuencia irreversible que representa esta cadena de bloques. En resumen, **las Blockchain son secuencias irreversibles que no generan identidad alguna sino que las utilizan en tanto en cuanto sus transacciones vayan digitalmente firmadas**, y eso es opcional.

Identidad y anonimato

Una vez olvidado este incomodo pasajero que son las Blockchain en las discusiones actuales, lo que sí es más importante es el **antagonismo esencial que hay entre Identidad y Anonimato**; la presencia de uno excluye al otro. **Todo sistema que permita relacionar unos hechos con otros, unas características con otras, constituye un sistema de Identidad Digital**. En muchos escenarios una identidad digital es la IP desde la que te conectas, o el teléfono que llevas encima y que casi siempre respondes tú. Una identidad puede ser una dirección postal física, una dirección de correo electrónico, una cuenta en alguna red social o en cualquier otro servidor. Una identidad digital puede ser también una pauta dinámica; alguien que va a una determinada mezquita o iglesia los viernes o los domingos a la misma hora, alguien que desayuna los mismos días, a las mismas horas y en el mismo bar, etc. El anonimato sólo se puede conseguir si desaparecen todas esas "constantes" o "hábitos" que se convierten en identidad por el mero hecho de practicarlas.

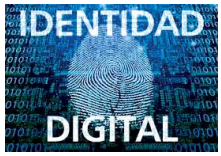
Tipos de identidades

Las identidades digitales difieren entre sí según del **tipo** que son y por el nivel de **detalle** que permiten. Las hay que se distinguen por el **tipo** (nombre, edad, género, fecha de nacimiento, peso, color del iris, etc.), por su **precisión** (1,87 m de altura, 95 kg y 32 años) y por el grado de **abstracción** que implican (nombre propio frente a nombre completo en el DNI + domicilio postal + la fecha de nacimiento, o bien la edad exacta frente a si es o no mayor de edad).

Casi todas las identidades cambian con el tiempo[11] (edad, peso, dirección postal, cuentas corrientes, modelo de coche, pareja, etc.), pero algunas lo hacen muy frecuentemente (dirección IP usada, teléfono móvil, dirección de correo electrónico, cuentas en servicios y RRSS, usuarios y contraseñas, posición geográfica, etc.). El número de identidades que definen a una persona pueden contarse en centenares, y se hacen millares si se incluyen sus relaciones sociales (hermano/a de, padre/madre de, hijo/a de, amigo/a de, votante de, etc.).

Cualquier sistema de Identidad Auto-Soberana que se precie, debe conseguir: **1)** eliminar factores que se repitan y **2)** aumentar hasta el infinito, y más allá, el número de posibles descriptores efímeros de una persona. Esa es la esencia de los métodos que se han propuesto pero que todavía están en desarrollo.

Hoy ya son pocos los que no están convencidos de que el negocio de las RRSS no es otro que vender nuestra vida privada particular en crudo o destilada a través de algoritmos de AI y de Big Data. Toda esta monitorización sólo es posible porque existen otras identidades (no oficiales) que ya nos delatan (direcciones IP y MAC, perfilado de los agentes y navegadores, smartphones, smartTVs, coches conectados, llevables, GPS, WiFi, IoT, etc.)



Las necesidades de anonimato o de una gestión legal de las identidades de todos y cada uno de nosotros apremian para descubrir una nueva forma de identidad que permita el negocio económico en Internet, pero que no lleve asociada la firma, como la del Doctor Fausto, de un cheque en blanco a los muchos demonios que operan y sostienen Internet.

[6] Ver https://en.wikipedia.org/wiki/Johann_Georg_Faust
[7] Ver https://scholar.google.es/scholar?q=self+sovereign+identity
[8] Ver Lamport, Leslie. "The Part-Time Parliament". ACM Transactions on Computer Systems, vol. 16, no. 2, Jan. 1998, pp. 133–169., en https://dl.acm.org/citation.cfm?doid=279227.279229
[9] Ver Narayanan, A.; Bonneau, J.; Felten, E.; Miller, A.; Goldfede, S.: "Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction", Princeton University Press, 2016
[10] Ver https://tools.ietf.org/html/rfc3339
[11] En un ser humano sólo no cambia significativamente con el tiempo la secuencia de su ADN, sus huellas dactilares y plantares, y la estructura de su iris ocular. Todo lo demás cambia, y bastante deprisa.

Los identificadores descentralizados (DIDs)

La esencia de estos nuevos sistemas de identidad son los **identificadores descentralizados** (**Decentralized Identifiers** o **DIDs**). Actualmente su definición se encuentra en la versión 0.12 de su modelo de datos y sintaxis [12]. Un DID es dos cosas: un identificador único y un documento asociado a él (DID Document). El documento asociado al DID es un objeto JSON que está almacenado en un servidor conocido y accesible del cual es fácil descargarlo.

Los documentos asociados con un DID incluyen **1)** su fecha de creación, **2)** una prueba criptográfica de que su contenido es válido (firma digital), **3)** una lista de claves criptográficas, **4)** una lista de nodos en lo que ese DID puede utilizarse para autenticar, **5)** una lista de servicios donde ese DID puede utilizarse, y **6)** cualquier número de extensiones definidas externamente.

En los esquemas propuestos se espera que el DID sea "persistente e inmutable", lo cual creo que es un error imperdonable, y que esté fuera de la influencia de cualquiera excepto su propietario por lo que, es de esperar, que lo que cambie sean los documentos asociados a ese DID.

Si quisiéramos también demostrar que somos propietarios de una casa de un tipo concreto, contactaríamos con el catastro, nos identificaríamos completamente ante él, y le pediríamos que generase un documento asociado al mismo DID en el que firmase digitalmente "*que ese DID es propietario de una casa de un tipo concreto*". Con esos dos documentos, **1)** sólo se informa de que "*es mayor de edad y es propietario de una determinada casa*", y **2)** no desvela absolutamente nada sobre quién está detrás de ese DID.

En principio, los proveedores de confianza no saben a qué otros proveedores hemos ido, por lo que no conocen nuestra "identidad compuesta" pero si conspirasen contra nosotros sí podrían descubrir nuestra identidad compuesta ya que hemos utilizado el mismo identificador aleatorio DID en ambos casos. Este proceder sería anónimo y confiable frente a cualquiera que verificase públicamente la validez de los dos documentos asociados, pero no sería anónimo frente a una confabulación de proveedores que saben para cada DID utilizado, quién realmente (Identidad legal) solicitó que se expidiera el correspondiente documento asociado al DID.

Simplificando un poco, estas aproximaciones hacia sistemas de Identidades Auto-Soberanas se basan en la expedición de certificados conteniendo aquellos datos (*Verifiable claims*) que le solicita su titular y de los que el servidor públicamente aceptado sabe fehacientemente. En esos certificados sólo aparece un identificador elegido al azar por el titular del dato que se certifica y de ahí emana su control absoluto.

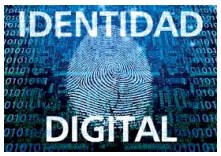
Si se usa el mismo DID en varias ocasiones, lo que se consigue es crear un "perfil" en el cual se han desvelado públicamente (nunca se sabe dónde pueden terminar los documentos asociados a ese DID una vez los hemos presentado en alguna ventanilla digital), una serie de hechos ciertos que, de ser demasiados o demasiado precisos, pueden terminar delatándonos. La solución más sencilla sería **1)** cambiar muy frecuentemente (o siempre) ese "alias" que representa el DID, **2)** no volver a utilizarlo nunca, y **3)** minimizar el número de descriptores (documentos asociados) que solicitamos/creamos con un mismo DID.

Sería como realizar cualquier proceso administrativo utilizando un pseudónimo diferente en el que, el verificador de las firmas contenidas en los documentos asociados a DID, sólo sabría **1)** que todos los datos son ciertos, y **2)** que describen parcialmente a una misma persona.

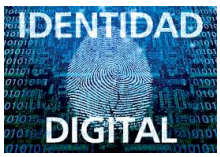
Con todo, el sistema no es perfecto porque **1)** esas asociaciones de datos ciertos deben considerarse eternos (Internet no olvida), **2)** los datos son ciertos, por lo que se ha creado una foto "muy parcial" pero cierta, de nuestra realidad, y **3)** el solapamiento parcial de diferentes conjuntos de documentos asociados (aunque tengan diferentes DIDs) podrían acabar desvelando nuestra identidad como si de la resolución de un puzle se tratara. Además de eso, los proveedores y firmantes de los documentos asociados con cada DID, conocen nuestra identidad legal y pueden guardar un historial de valores DID.

Otros problemas por resolver

Independientemente de que sea el titular el que "controla y dosifica" su identidad a través de la creación de DIDs, todavía quedan problemas serios que resolver como, por ejemplo, qué legislación se aplica en cada caso cuando los procesos son internacionales ya que no está claro "dónde" se ha hecho realmente la transacción [13].



Hay un antagonismo esencial entre Identidad y Anonimato; la presencia de uno excluye al otro. Todo sistema que permita relacionar unos hechos con otros, unas características con otras, constituye un sistema de Identidad Digital.



Las identidades digitales difieren entre sí según del tipo que son y por el nivel de detalle que permiten. Las hay que se distinguen por el tipo, por su precisión y por el grado de abstracción que implican.

Los DID vienen a ser pseudónimos o alias generados al azar por el titular del dato, y que va utilizando con los distintos proveedores de confianza para conseguir documentos asociados que están firmados por ellos y que establecen algún descriptor concreto (*Verifiable claims*). Por ejemplo, si quisiéramos probar que somos mayores de edad, contactaríamos con un proveedor confiable que realmente sepa cuál es nuestra edad (nos tendríamos que identificar completamente ante él), y pedirle que firme un documento en el que se indique "*que ese DID es mayor de edad*".

Está claro que el problema de la identidad, el anonimato, la autenticidad, la trazabilidad y las medidas anti-monitorización (**P**rivacy **E**nhanced **T**echnologies) son temas todavía no resueltos y que cada día son más necesarios si queremos utilizar Internet con un mínimo de seguridad. Sin embargo, si seguimos relacionándolo todo a ciegas con tecnologías que nada tienen que ver, como es el caso de Blockchain, y no sabemos distinguir lo esencial, que es utilizar siempre un pseudónimo distinto e irrepetible a la vez, y eliminar cualquier aspecto "repetitivo" en nuestros modos, frecuencias y hábitos en el proceder. Si no conseguimos esto tan simple, nada habremos avanzado.

De no hacerlo así, el posturreo acabará con nosotros y los gigantes tecnológicos (Google, Facebook, Youtube, WhatsApp, Wechat, QQ, Instagram, Tumblr, Skype y LinkedIn) habrán ganado y serán dueñas de nuestras almas. ■

Jorge Dávila
Consultor independiente
Director
Laboratorio de Criptografía
LSIIS – Facultad de Informática – UPM
jdavila@fi.upm.es

[12] Ver https://w3c-ccg.github.io/did-spec/

[13] En este punto el GDPR utiliza el principio de Lexloci, que fija la nacionalidad de la transacción en las nacionalidades de los participantes en ella.