JIM 5, 1 (2017) 4-13
HANDLE: http://hdl.handle.net/10216/103566

Cybersecurity - Personal Security Agents for People, Process, Atoms & Bits

Dr. Shoumen Palit Austin Datta

Senior Member, MIT Auto-ID Labs
Research Affiliate, Department of Mechanical Engineering
Massachusetts Institute of Technology, Room 35-203, Cambridge, Massachusetts 02139, USA
Senior Scientist, Medical Interoperability Program, MDPnP Labs
Massachusetts General Hospital, Harvard Medical School
Partners MGH Research Building, 65 Landsdowne Street, Cambridge, MA 02139, USA

shoumen@mit.edu, sdatta8@mgh.harvard.edu

Letter from Academia

It was predicted by experts. The DDoS attack using an internet device on October 21, 2016, generated a flurry of suggestions from vast number of pundits. What if the attacks were not limited to social media sites but instead targeted heart monitors to deliver shocks to patients with cardiac arrhythmia? In this article the potential of personal security agents (PSA) is suggested as a modular tool to model people, process, bits and atoms (objects) with layers to address trust, privacy and security. Can we explore the potential of creating wrappers within these layers to include cognitive firewalls?

Keywords. Cybersecurity, agents, cognitive firewall, IPv6, IoT, healthcare.

1 A simple path – distributed trust management in cybersecurity?

The recent distributed denial of service (DDoS) attack originating from IoT type devices infected by the Mirai botnet[1] was predicted[2] by experts. The recent news about tracking individuals by name and location[3] to profit from advertisements was also perpetrated by experts. Objects and people appear to be equally vulnerable to cyber-intrusion.

The concept[4] of personal security agent (PSA) may not be novel and the idea of PSA may not be restricted to individuals. Devices may have PSA linked to a registry for example, GHR[5] maintained by CNRI[6]. Other less robust examples include mobile device management tool and rosy predictions[7] about MDM which are attracting new[8] entrants.

In light of revelations[9] which continues to accumulate, it is perhaps worth considering the security function (security-as-a-service) in layers where manufacturers of devices provides a layer of security but other independent layer(s) may be procured, installed and managed by the user, based on source(s) trusted, by the user.

ISSN 2183-0606
http://www.open-jim.org

Perhaps, consider access to a secure box in a traditional bank vault which requires two keys to open. In its elemental form, this may be also viewed as a modular approach.

Let us extend this over-simplification to a traditional hotel room. The key card or mobile code to open the door may not allow the help to enter if the occupant dead-bolts or uses the security chain from inside. Imagine if neither the dead-bolt nor the security chain is made available to the hotel guest. The hotel advertises BYOD for security. On entering the hotel room or suite you find a bag of screws, a screw driver and four relevant holes on the door frame to insert and fix your own dead-bolt or chain. The digital equivalent (screws and holes) may be a *dynamic* API. The digital equivalent for dead-bolt (BYOD) is a digital certificate, with mobile duo authentication provided by a third party trusted vendor (of your choice).

2 A simple path – *Modus Operandi*

The suggestion involves the purchase of security-as-a-service (SECaaS) from a trusted vendor. This is, at this time, a hypothetical vendor. In the future it may manifest as a new line of business. SECaaS providers may include NGOs in the developing economies, government-academic-industry consortia, consumer watchdogs, global organizations or standards body (eg GS1, EAN, ITU[10], IEEE) which may not have a direct financial interest but offers an independent service for pay-per-use nano-fee-payment. The assumption is that an "external" agency removed from direct corporate influence is less likely to compromise security or squander privacy in line with their pedigree, credibility or brand recognition. Hence, we are more likely to trust their security-as-a-service (SECaaS) offer.

Vendors of devices (shop floor machines, healthcare robots, heart monitors[11], airplane turbines, valves for oil pipelines, refrigerators, automobile parts, prosthetics, phones) will offer APIs to digitally receive, install and activate the security service (SECaaS) layer. In future, devices which generate, transmit or acquire data may not be sold (FTC, FCC, FDA, UL type regulation) without APIs which may be deployed as an user-exit to install one or more layers (think containerization of the sand-box concept) of security protocols from one or more SECaaS providers, perhaps from different global regions.

Mobility makes it imperative that the security-as-a-service function is user controlled and calls for new software tools (CubeFog). The elements of this security may draw on a sub-tier of vendors specializing in offering a smorgasbord of dynamic security engines, for example, random number scheme, prime number cryptography, biometric coordinates.

Taken together, the management of trusted partners in the security-as-a-service ecosystem requires storage and connectivity with other domains (intruder detection, non-obvious relationship analysis, fraud monitoring). Data protection[12] rules and new policies[13] makes it essential for users to store and access their security-as-a-service data in their preferred nations and in clouds or fogs of their choice. Redundancies introduced by the user may make it difficult to penetrate all layers of security which did not originate from the device vendor or may not be hacked through the cloud storage.

3 A complex path – die is about IPSA

The device-linked PSA is one version of the personal security agent (PSA) for objects. But how an individual wishes to interact with the cyber-world and what data one wants to share or which information one chooses to keep private is inextricably linked to one's identity. Protecting this identity and keeping it secure is another function for the PSA. It helps to differentiate between object-PSA (OPSA) and individual-PSA (IPSA).

Starting with individual medical records[14] and continuing[15] to "DIE" promoted[16] by the World Bank and advocated[17] by GSMA, we must now deliver security and privacy at the level of individual citizens with a digital footprint based on their digital identity. The medium of delivery for IPSA is linked to OPSA because humans need a medium to interact with the cyber-world. *That* medium is provided by objects and may use various forms of IoT, by design, for industrial or consumer applications.

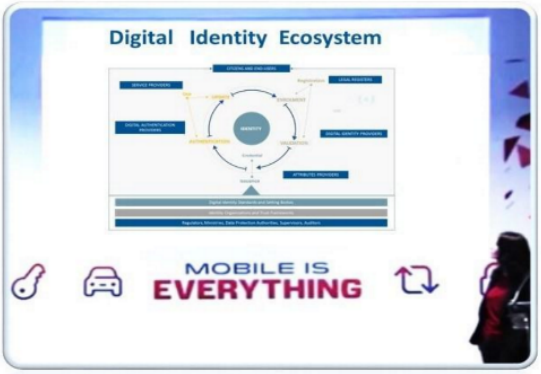

Fig. 1. The global momentum to endow each individual with a digital identity may be an appropriate vehicle to embed individual personal security agents (IPSA).

4 A complex path – *modus operandi*

There is very little debate to refute the need for national policy[18] and global tools. The latter must be redundant and distributed with very high fault tolerance. These tools must be capable of *ad hoc* dynamic composition when the *status quo* is challenged due to threats arising from breach of cybersecurity. Because *post hoc* security is useless in the IoT-by-design paradigm, the emphasis is on the dynamic composable architecture and its ability to "discover" atoms and bits in real-time related to the person or event.

IoT related cybersecurity may improve by implementing IPv6 which is the essence of connectivity. It is the fundamental routing layer in packet communications. The rapid diffusion of connectivity catalysed by IoT is increasingly inextricably linked with our lives and our digital twin[19] representations which includes information about information[20] may emerge as quintessential "avatars" to offer decision support.

Individual Personal Security Agents (IPSA) may protect privacy, regulate data sharing and communicate (including emergencies) between individuals and their digital world, which may include digital twins, in certain instances, especially for industrial systems.

IPSA must be globally unique and coupled with individual identification systems[21] eg social security number (USA) or Aadhaar (UIDAI, India). We propose a digital twin[22] approach handled by creating a mobile software intelligent agent for each citizen of the world (one may think of IPSA as an avatar popularized by "Second Life" games, 2003).

To tamper-proof the digital footprint and protect the digital records of this agent (IPSA) we may converge with tools from "blockchain" to document, authenticate and grant permissions to "handshake" for interoperability and multi-tasking across many diverse applications eg healthcare, banking, fintech, e-vote and remote 3D printing-on-demand. Using public key cryptography and personal agents to protect private keys, PSA may be equipped with tools to de-identify data in course of case-specific dynamic composition of responses. Hence, the potential to use IPSA in e-vote or sharing de-identified private data, for example, healthcare data collection, used for census or public policy surveys.

The agency of agents (APSA) acting on your behalf (IPSA) or on behalf of machine (OPSA) components or devices, associated with you, must be trained, updated and maintained to be in tune with your personal likes/dislikes/preferences for allowing or not allowing your data (location, medical data) to be shared (or not) when external agents or bots query your digital ecosystem. What if a mapping service queries your phone to share your location data? Who will protect you and offer privacy, if desired?

Your IPSA may be pre-programmed by you to respond according to your preferences which *you may change* using another remote device (OPSA) or modify associated dependencies using digital assistants (time of day, office or home, travelling for business or in clandestine meetings, medical status or trigger emergency *blue button*[23] over-ride). Perhaps similar approaches are necessary for Agents (OPSA) overseeing sensors, machine parts, medical devices, turbine blades, smart grids, automobile brake pads, water filters and trillions of non-human objects or things or processes using IoT as a digital by design metaphor.

APSA, IPSA, OPSA and other PSAs may be driven by standards. The road taken by trusted organizations may drive standards based operations for security-as-a-service to evolve.

However, standards or policies for every possible situation cannot be conceived *a priori*. Systems must be installed to trigger dynamic composition of *ad hoc* micro-directives[24]. Open data sharing may be as essential as selective de-identification schemes when anonymity and privacy are critical yet must be balanced in the best interest of the individual. For example, Sam collapses on the steps of Vittoriano[25] due to heat exhaustion and rushed to Ospedale Fatebenefratelli, nearby. But, they are unable to access her Epic-locked EHR from Massachusetts General Hospital. Sam is injected with a steroid and drifts to a comatose state. When staff in Rome speaks with Sam's physician in Cambridge, they learn that Sam is diabetic. Injecting steroid was a nail on her coffin.

In certain scenarios, IPSA and OPSA, if adequately tuned, may be a life saver. In other cases, the responses handled by the agents may be denuded of certain data or values to protect personal ID (convergence of public key encryption[26] with editable blockchain principles and IPv6 for mobile e-vote). Machines and devices (OPSA) may also want

data, information and metrics to remain cryptic to deter industrial espionage.

5 Why a Modular Approach may better optimize a dynamic connected path?

The ecosystem and/or community of PSAs must converge, connect and communicate to continuously monitor and curate diverse digital threads in order to synthesize the value and extract dividend from digital transformation.

The example of community as a *function* based on components as a *form* is a robust time-tested bio-inspired theme of modularity. One example of such a theme is anchored in the principles surrounding the evolution of nanobiomes with respect to life forms in the oceans[27]. Rather than combining all life functions into a single organism, the nano-biome works as a network of specialists, each with a special form (module), that can only exist as a community. The forms must converge in a spatio-temporal interplay to give rise to systemic interactions which, in turn, will manifest the desired function.

In the modular approach to cybersecurity, the form may be equivalent to agents, each specifically created to execute a particular task or role. When the agents aggregate to support a system, the overall outcome from convergence of these forms generates the function, in this case, the security of the system.

How do we determine that the individual agents and their related tasks are secure? This is where one begins to appreciate the value of modularity, convergence and the formation of agencies (groups of agents) which enhances the function (cybersecurity).

For example, you receive a message from your wife to warm up the lunch casserole for five minutes in the microwave. You proceed to perform this task and enjoy your lunch. What if you received a message from your wife to heat up the lunch for five hundred minutes in the microwave? You (the human) wouldn't comply with the command. Would you? Your sense of what is reasonable prevents the execution of the message and offers security. This action represents the concept[28] of a "cognitive" firewall which will raise an alarm based on what is reasonable (Joshua Eric Siegel, PhD thesis, MIT; Sanjay Sarma, personal communication).

Consider a simple command to a temperature sensor in a critical environment (cooling tower in a nuclear facility or combustion chamber in a jet engine or turbine). Usually, an external command may trigger the sensor to sense the temperature and report back to the data center every five minutes. An intruder-designed action or malware mimics the command but changes the time interval to five seconds. This action appears benign but the battery life of sensor may be depleted within a few hours. The sensor will cease to function.

If the temperature of the cooling tower or combustion chamber exceed the limits, it may result in a meltdown or some other form of catastrophe including loss of lives, injuries and contamination, due to failure of cybersecurity. The sensor and the micro-system functioned exactly in the manner it was designed - sense temperature on demand. But, the micro-system was not designed to reason that there may be a breach of security because the task demanded – sense and respond every five seconds – did not include

within its scope the validation process whether or not a particular command "makes sense" for the connected outcome and for the whole system.

A cognitive firewall and its "supervisor" function, if installed, as separate yet linked modules, could evaluate the system's process evolution and test incoming commands, to ensure that a particular command "makes sense" for the connected system. Supervisor may proactively monitor the system evolution and identify anomalous behavior due to system model breakdown, physical system failure, or other self-inflicted mechanisms.

Labelling this function as "cognitive" may invite justifiable criticism from scientists because the science of cognition is far more complex. Use of "cognition" is at best like using a new language without a dictionary or knowledge of grammar. It may be analogous to learning a language like a child - through imitation and trend identification in complex examples of linguistic usage. Similarly, claims of "intelligence" by corporate marketing departments are vapid, trite and shoddy if one considers the rigor of neural underpinnings of intelligence and compares them with the pedestrian apocalyptic momentum attributed to artificial intelligence[29] by those who may be biologically uninformed or inspired only by profit.

Connectivity of these forms and agents via IPv6 is one mechanism by which each entity level unit/model may communicate when the systems are distributed or where long range interactions are essential. One idea is rooted in the concept of *cube-on-cube* proposed by Marvin Minsky, MIT. See page 315 (Appendix: Brain Connections) from *Society of Mind* by Minsky (page 311 in this[30] PDF). The convergence of Marvin Minsky's cube-on-cube with connectivity between the cubes using IPv6 is a concept promoted by companies [31] professing that "containerization" as a software tool is an innovative new dimension.

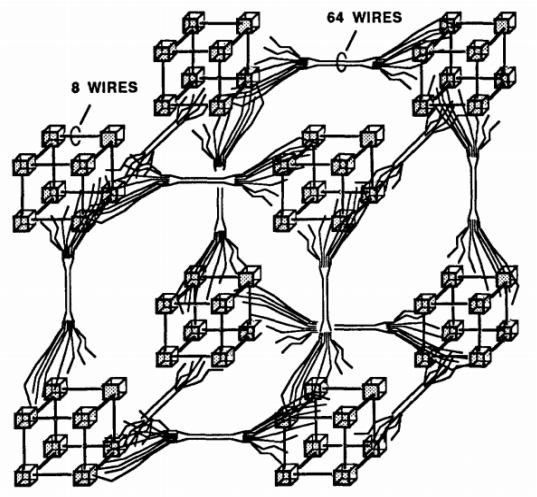

Fig. 2. Marvin Minsky's cube-on-cube concept [30] is extrapolated from and may be a representation of the topological connectivity between neurons or neuronal circuits.

Connectivity using IPv6 draws on an earlier (2006) proposal[32] advocating use of the internet protocol version six (IPv6) as a format to uniquely identify not only things (objects in the IoT) but also processes, relationships (syntax, semantics) and interfaces (sensors). The design key in the earlier paper (2006) relied on using the 128-bit IPv6 global standard which offers $3.4x10^{38}$ unique "names or addresses" to uniquely identify every instance of any transaction and follow their trail even when distributed (routed). By extrapolation, the 2006 paper appear to contain a few elements related to the concept[33] of blockchain[34] triggered[35] by the bitcoin principle[36] which highlights the principle of digital ledger. Implementing digital ledgers may monetize PSA but there aren't any low hanging fruits.

6 A path to monetization?

New business growth from security-as-a-service (cybersecurity software) may serve about 10 billion consumers by 2050 and trillions of B2B operations, much sooner. One must create and sell/lease/rent/train the Agents (IPSA, OPSA) for security-as-a-service.

Monetization of software for security-as-a-service (SECaaS) calls for innovation using the principles of the "digital ledger" to innovate a creative digital ledger practices.

The management of micro-payments from a pay-per-use model needs service request and service delivery documentation as well as QoS compliance. The trusted vendor must be dissociated from the service delivery once the service is activated, in order to reduce threat point of entry. The sales of the software license by the trusted vendor may not be the only transaction (in the old model a fixed payment was offered for a fixed term). If the trusted vendor wishes to charge by usage, then the ability of the trusted vendor to *monitor* the use of the service is pivotal. Dissociating the trusted vendor from keeping tabs on your security service is essential to improve security for the user. The latter introduces loss of opportunity for the trusted vendor because the vendor is in the dark about how many times the user is accessing the security-as-a-service (SECaaS) application.

Consider the camera on the front door of your house. You are in Princeton, NJ and your smartphone screen lights up. It shows FedEx Fiona walking up your driveway to deliver a package. Fiona rings the bell. You open the door to your home in Cambridge, MA while visiting Tom (you are in 203 Lewis Thomas Lab) using remote key pad on your Iris app connected to the Schlage digital lock on your door. Fiona goes inside the house. She exits from the house. The door locks behind her. You see Fiona walk down the drive.

What did FedEx Fiona do inside your home? You assume that she left the package on the table. What else did she do? Did she re-apply her lipstick? Did she use the toilet?

You wouldn't know what happened inside your home unless you have a camera inside. Trusted vendor for security-as-a-service needs the equivalent of a "camera outside" to know when you ping SECaaS. It chooses to remain oblivious of your use (what you accomplish). The latter would be the data from "camera inside" the house but trusted vendor is dissociated from that function. Trusted vendor does not need "inside" data because the trusted vendor charges micro-payments based on pay-per-use each time you ping the SECaaS app. It does not matter what you do but what matters is the

duration of the use (unit rate or cost) and the time of the day (traffic volume, bandwidth, latency may be factors in the quality of service or priority queue determination). Thus, the time-stamp is important for monetization and it is equivalent to house entry/exit data captured by the external camera. Privacy inside the house (data) remains unshared.

Distributed digital ledger, in practice, if combined with IPv6 transaction identity, can guarantee authenticity and auto-process time-based micro-payment for service delivery. Each unit of this distributed digital ledger is in the form of an agent module. The nature of the service can "drag and drop" the selected service units (modules) necessary to complete the function. The same modular principle which applies to each distributed task sub-unit is applicable in the distributed digital ledger paradigm. Concurrent execution, co-location and semantic interoperability between standards/platforms are key elements of this vision if we wish to transform the suggestions to implementations.

7 A path less travelled – the road not taken?

Returning to the discussion of cybersecurity, one wonders if these concepts threaded with IPv6 may be extended to propose a potential mechanism to improve cybersecurity by engineering design. Is there room for convergence between IPv6 and blockchain with selective use of public key encryption for digital object architecture[37] and IPSA?

How can we (can we?) use the 40 bits or an extension of the security domain in the current IPv6 design to serve as a cybersecurity base in the engineering design? Digital crypto-tokens concealed in the alphanumeric stretch may be connected with software security agents to authenticate (handshake) transactions, data transmissions or user activated action (the nature of which may be immensely diverse and vast in number).

The hypothetical concept of a set of *cascading locks*, is suggested. Only the header of the lock may be part of the 40-bit design of IPv6. Data related to the lock and its functional activation (I/O, open/close) may reside in a separate agent. It may mimic how RFID[38] EPC[39] contains a reference[40] to the location where the actual data[41] (or modular agent) is stored.

The locks may only open (allow, activate) with a digital key or digital token which must be generated in real-time (if triggered) using reference data (authentication?) secured by an agent in another location (potential for network verification at the edge).

The "open lock" status in tier-x could trigger the process to open the lock in tier-y (next lock in the cascade) using information (dependencies) from tier-x. This hypothetical cascade of locks and the sequential effect (outcome) may offer the ability to trap an intruder, in time. The system may sacrifice a few locks but eventually the aberration due to the intrusion or anomaly (if detected) can turn off the cascade (remaining locks, sequences in queue) to prevent the remaining steps (remaining locks are still locked). This is the type of function one may also expect from the supervisory layer of the cognitive firewall.

The hypothetical idea of cascading functions (with lagged dependency) is an attempt to theoretically propose protocols to prevent or contain an attack (intruder detection, repulsion, protection) if cascading functions were implemented (not known if it is

possible). If feasible and deployed, this system of security may be useful for autonomous transportation (prevent vehicle from being hacked) or machines (could prevent turning off turbine when a plane is in flight) or healthcare (prevent over-dose of morphine in post-operative surgical care) or energy grid brown-outs (time spoofing synchrophasor by creating anomalies in time-sensitive networking or causing protocols to malfunction).

8 Temporary conclusions

Being Digital[42] may not benefit from any hasty conclusions about people, bits and atoms surrounding cybersecurity. Questions will continue to accumulate and good answers may be few and far between. The patch work of solutions are not optimum but dreams of a final solution does not call for a dystopian or utopian classification. Those who may need to assign cybersecurity a category may wish to consider – what is being protopian?

Being protopian is balanced view and implementation of security and cybersecurity. Talking about security by design may be modified to security by engineering design to reflect that "baking" in security at the foundation is more robust than the after-thought of adding it to application layers, in many instances. The inclusion of PSA in the form of security as a service is not a common trend and may not gain momentum for some time.

The concept of agents may not have a mainstream following, yet, even though it is about 50 years old. One reason may be due to the obnoxious phrase "low hanging fruit" often used by the corporate world. Harvesting low hanging fruits require only low level skills.

---

[1] http://hub.dyn.com/static/hub.dyn.com/dyn-blog/dyn-statement-on-10-21-2016-ddos-attack.html
[2] http://www.politico.com/agenda/story/2015/06/internet-of-things-privacy-risks-security-000096
[3] www.propublica.org/article/google-has-quietly-dropped-ban-on-personally-identifiable-web-tracking
[4] http://bit.ly/IoT-Policy-Socio-Economic
[5] https://itu4u.wordpress.com/2014/01/06/lost-something-on-the-internet-never-again-with-new-digital-object-do-architecture/
[6] http://www.cnri.reston.va.us/papers/Architectural_Evolution_Internet_17Nov10.pdf
[7] http://www.researchandmarkets.com/research/lght2x/global_mobile
[8] http://cio.economictimes.indiatimes.com/news/mobility/tata-teleservices-enters-mobile-device-management-space-differentiates-on-unique-features/54709229
[9] http://www.theverge.com/2013/6/6/4403868/nsa-fbi-mine-data-apple-google-facebook-microsoft-others-prism
[10] https://itu4u.wordpress.com/2014/01/06/lost-something-on-the-internet-never-again-with-new-digital-object-do-architecture/
[11]http://www.qmed.com/mpmn/article/teardown-look-inside-st-jude-medical-merlinhome-transmitter
[12] http://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf
[13] http://ec.europa.eu/justice/data-protection/
[14] https://www.ncbi.nlm.nih.gov/pubmed/11734380
[15]https://openknowledge.worldbank.org/bitstream/handle/10986/20752/912490WP0Digit00Box385330B00PUBLIC0.pdf?sequence=1
[16] http://pubdocs.worldbank.org/en/959381434483205387/WDR16-Spotlight-on-Digital-ID-May-2015-Mariana-Dahan.pdf

[17] http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/07/Towards-Shared-Principles-for-Public-and-Private-Sector-Cooperation.pdf
[18] http://bit.ly/NOT-NIST
[19] https://arxiv.org/ftp/arxiv/papers/1610/1610.06467.pdf
[20] https://itu4u.wordpress.com/2014/01/06/lost-something-on-the-internet-never-again-with-new-digital-object-do-architecture/
[21] http://bit.ly/DID-PLATFORM-WB
[22] https://arxiv.org/ftp/arxiv/papers/1610/1610.06467.pdf
[23] http://www.va.gov/bluebutton/
[24]http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1067&context=law_econ
[25] http://www.ilvittoriano.com/
[26]https://dash.harvard.edu/bitstream/handle/1/12362600/Determenistic%20Pubic-Key.pdf?sequence=1
[27] http://pubs.acs.org/doi/pdf/10.1021/acsnano.5b07826
[28] https://dspace.mit.edu/handle/1721.1/104456
[29] https://arxiv.org/ftp/arxiv/papers/1610/1610.07862.pdf
[30] http://www.acad.bg/ebook/ml/Society%20of%20Mind.pdf
[31] https://docs.docker.com/engine/userguide/networking/default_network/ipv6/
[32] http://esd.mit.edu/WPS/2007/esd-wp-2007-17.pdf
[33] https://bitsonblocks.net/2015/09/09/a-gentle-introduction-to-blockchain-technology/
[34] https://www.weforum.org/agenda/2016/06/blockchain-explained-simply/
[35] http://www.mail-archive.com/cryptography@metzdowd.com/msg09959.html
[36] https://bitcoin.org/bitcoin.pdf
[37] https://itu4u.wordpress.com/2014/01/06/lost-something-on-the-internet-never-again-with-new-digital-object-do-architecture/
[38] https://ocw.mit.edu/courses/engineering-systems-division/esd-290-special-topics-in-supply-chain-management-spring-2005/lecture-notes/lect11.pdf
[39] http://cocoa.ethz.ch/downloads/2014/06/None_MIT-AUTOID-WH-002.pdf
[40] http://cocoa.ethz.ch/downloads/2014/06/None_MIT-AUTOID-WH-001.pdf
[41] http://cocoa.ethz.ch/downloads/2014/06/None_MIT-AUTOID-WH-003.pdf
[42]http://web.stanford.edu/class/sts175/NewFiles/Negroponte.%20Being%20Digital.pdf