

Self-sovereign identity for digital publication

by Juan Caballero

Self-sovereign identity¹ (SSI) could be a quantum leap for the development of digital publishing and journalism. SSI has the potential to greatly simplify many centuries-old identity problems in publishing, fortifying authorship and copyrights, expanding the toolkit of licensing in complex ways, and forging more sustainable funding models for publications.

Challenges for the digital publication industries

Compared with most other industries, the stakes are so much higher for the thorny issues of **information quality**, a defining factor that could drive adoption of SSI into publishing infrastructure. Journalism in particular, but also academic publishing and all other media businesses, are **information businesses**: the reputation of each media or information brand is staked on its demonstrated commitment to ethical standards and its track record of objectivity. Plagiarism, trolling, misattribution, “payola,” and other breaches of trust can be fatal to such businesses. If the success rate of bad actors spreading bad information, fraud, hate, or spam on an information platform gets too high, public trust in it can erode quickly. And the platforms themselves are not the only “brands” at risk, since these same platforms are crucial infrastructure for our democratic processes.

It is easy to forget how recently these problems grew so serious. **Trust-poor** information systems less than 30 years old were scaled up very rapidly, disrupting the reader-publisher relationship and displacing traditional gatekeepers to assume monopsonic² control of the industry’s advertising and discovery mechanisms. These systems’ resilience to disinformation and other forms of abuse has failed to keep pace with a user base that went global in just a few years’ time. They were even shielded (in the US at least) from the liabilities that had kept publishing more fragmented and local³. Revenue-starved journalism has been hit hard by all this: few industries have suffered worse losses of **public trust, social power, or accountability** during the Internet Era.

That said, few industries have more to gain from SSI’s toolkit for building complex **trust networks**, granular **transparency**, and scalable **accountability**. Anchored in a better

¹This work is conceptually indebted to the W3C verified credentials working group (and their use-case documentation format), the RWoT community & the German Blockchain Bundesverband.

²For a concise overview, see Antonio García Martínez, “Facebook Is not a monopoly, but it should be broken up” in Wired, March 2019, <https://www.wired.com/story/facebook-not-monopoly-but-should-be-broken-up/>.

³Many critics blame so-called “platform immunity” (and an anti-trust framework inappropriate to these new business models) for the major accountability deficits seen in the highly centralized information systems and online publishing models of recent decades. See Citron, Danielle Keats, “Section 230’s Challenge to Civil Rights and Civil Liberties” (Knight Institute Emerging Threats, Columbia University, 2018): <https://knightcolumbia.org/content/section-230s-challenge-civil-rights-and-civil-liberties>.

identity layer (see Chapter 1), far more rich and versatile “reputation systems” and information-grading systems⁴ can be built into the publishing ecosystem of the future web. Tracing all information to signed, original sources could greatly empower professional researchers, fact checkers, librarians, and many other kinds of information stewards to establish their trustworthiness. They could keep dubiously sourced or un-sourced content out of mainstream circulation. Naturally, some amount of abuse will evolve to exploit these new systems, but they will be slower, more robust, and more distributed than our current ones, making abuse or manipulation at scale inhibitably expensive. Scientific research, creativity, and independent thought could be shielded from the economic Goliaths of the publishing business once small outlets and distribution channels proliferate to create a more diverse, segmented, and democratic media landscape.

SSI-powered digital publishing

Journalism has many centuries-old identity and anonymity issues that could be transformed by integrating SSI. Other forms of publishing stand to be transformed just as drastically by the **new possibilities** and efficiencies in distribution, patronage, rights, and licensing opened up by SSI. These include copyright enforcement, direct micro-payments for long-term royalties, and more complex forms of media ownership. We have tried to illustrate some of these new social and legal forms with the following narrative vignettes imagining how different users might interact with a publishing ecosystem powered by SSI authorship.

*



Before going fulltime in journalism, Sahar worked many odd jobs and uncredited gigs in the publishing and movie industries: she was a **ghostwriter**, a **scriptfixer**, various kinds of **editor**, a **translator**, and more. Some of these gigs still pay her residuals years later in the form of tiny micropayments. When these now-classic works are reprinted or syndicated, the payments are automatically sent out, without anyone involved thinking much about who exactly the contract refers to as “anonymous ghostwriter X” or “translator of appendix C”.

Sahar, now a full time reporter, has had to negotiate increasingly delicate relationships with her sources since moving to the Politics section of her newspaper. Tobias, a **whistleblower** who is passing Sahar information for her story about unethical retaliations in the Kenyan military, needs to remain completely anonymous, so he provides her with a “**verifiable credential**” (an electronic affidavit from the personnel office) attesting his rank and security clearance but not his name. After confirming that nothing in the story can be traced back to him that might endanger his career in the military or his safety, he agrees to digitally sign the final story as an anonymous (yet credentialed) **contributor**.



*

Sophie, a colleague of Sahar’s from the business section of the same newspaper, writes a

⁴ The “reputation systems” of Web2.0 score individual authors or publishers as they do vendors in open marketplaces, but presumably in the future they could also score recommenders, aggregators, and social sharers as well, integrating social signals and market signals in a more nuanced (and anonymized) way. Similarly, information-grading systems of the pre-Web2.0 era scored academic and/or non-commercial publishers of information by their objectivity and their perceived adherence to disciplinary and ethical norms of production; these mechanisms, still not modernized much at time of press, could perhaps be updated at scale and incorporate younger publishing authorities in a Web3.0/SSI context, as hypothesized in the final section of this paper.

glowing review of a startup for which entrepreneur Charlie is currently raising funds. Her review is so glowing, in fact, that just to avoid suspicions of influence, Sophie's **editor**, Agnes, contacts Charlie for a public attestation that she has not influenced Sophie in any way, and that no one involved in her past or current fundraising has any substantial ownership interest in the newspaper. Charlie reviews the final version of the story and digitally signs it as a **contributor**, as is the norm in the business section, which runs all such disclosures and attestations in a footnote after each story.



*

In the middle of a heated political debate around an unpopular climate-change law coming to a vote soon in Congress, Valentina, previously an **expert witness** in court cases and a quoted source in newspaper articles, decides that too much politically motivated misinformation is circulating about the scientific basis for the law. She contacts Ona, an old friend from graduate school, **editor** of the opinion section of a widely-read newspaper, and who convinces her that an anonymous opinion piece alone will not do enough to dispel all the pseudo-science. Along with her anonymous essay, Ona includes a footnote that includes verified credentials about this illustrious-yet-anonymous author--the prestigious university and department that granted her doctorate (but not the year she received it), as well as the number of years she has been a member of the American Geophysical Union.



A few weeks later, an even more widely read paper runs a rebuttal in its opinion section, accusing Valentina of unscientific bias, and of being influenced by the cash-rich political party behind the unpopular law. Communicating anonymously, she is interviewed as "Climatologist X," proving by a digital signature that she is the anonymous author of the earlier piece, and rebutting the rebuttal, swearing not to have any ties to the unpopular law or to the party promoting it. "Climatologist X" reviews the final transcript of the interview, and signs it as a **contributor**, including an anonymous attestation from the electoral register that she was, in fact, registered to vote in the primaries for another party in the most recent election! (Luckily, in her state, primary affiliations are a matter of public record, so the electoral register provides such attestations online for a small fee.)

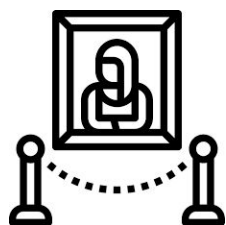


After all this controversy passes, Valentina can return to her academic publishing without her anonymous misadventures impacting the objectivity of her anonymously peer-reviewed research. Valentina uses her real name when she submits a new manuscript, a study on ice caps (for which she was head researcher), to a more technical journal, a non-profit, open-source review of academic research. Two blind peer **reviewers** of the initial draft, without knowing who wrote the paper, submit their feedback anonymously after signing it and attesting to their relevant degrees and current membership in the American Geophysical Union to prove they are peers. After Valentina addresses their critiques, they sign the updated version, as do Valentina's research colleagues (as **co-authors**), along with a **contributor** credit for the director of the university lab who oversaw some of the mathematical modeling. All of these contributions, signed to the authoritative version of the text, thus become verifiable records of contributions to pure research for the



respective résumés and tenure files of everyone involved⁵.

*



Over the course of his career working as a **rights librarian** for a museum, Kiowa has seen the prices for images of famous artworks and people vary wildly from decade to decade. The norms around licensing evolve rapidly, after all. When asked to explain his job at cocktail parties, he usually explains that some images, like his own face or a famous painting in the museum's collection, are "owned" and to reproduce them or use them commercially, fees must be paid or waived. Secure, decentralized identity makes it easier to track rights and fees, but also to trace unauthorized or unpaid reproductions, as well as to link to current ownership as artworks change hands. Before, Kiowa worked with similar systems handling **song royalties, stock photography, and television syndication**; his master's thesis in Library Science was actually about paying fair prices for **machine-learning datasets**.

*

Mat, a DJ and music professor in Berlin, founded a **dance club, record label, and "project space"** that operates as a **collective**.⁶ The actual building that houses the club and recording studio is itself owned collectively by Mat, the artists, and even some of the bartenders who bought a stake in the cooperative. All are treated equally and paid proportionally from the **pooled royalties** of ticket sales, live recordings, and original records produced there. No one gets rich, but some of the artists stayed in the collective after getting famous, so the collective proves itself fairly sustainable as its reputation deepens and grows.

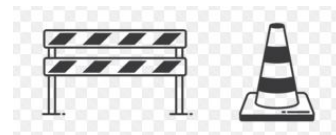


Research in a decentralized, trust-ranked archive

The preceding vignettes were straightforward extensions of existing services and structures within the publishing sector. Going one step further, we will end with a more detailed use-case of a more **speculative** corner of SSI publishing. What new ranking **systems** and trust-sensitive **search** capabilities might be available to future researchers?

A few decades into the SSI era, we can imagine that our journalist **Sahar** will start her research much as journalists do today, casually and organically, on platforms that are some combination of social-recommendation engine, search engine, and commercial space: "social media". Trends, weak signals, and clues abound there, and many story-ideas and rumors will always circulate wherever opinion and social information are aggregated with commercial and non-commercial publications.

But that "commingling" will look quite different after a few decades of the internet organizing itself according to a worldwide trust network. Being a professional, Sahar will probably use a **highly-customized interface** for interacting with her social networks, much as online journalists today might juggle many different user accounts and third-party hacks on each platform, with customized display



⁵Actually, many pre-existing organizations working on these topics are embracing SSI as its standards get more legally sophisticated and codified. For rigorous theories of academic self-publishing workflows, see the Elephant in the Lab blog (<https://elephantinthelab.org/distributed-organisations-for-collaborative-research/>) and for making publishing and library science more democratically available for pure research, see ScholarlyHub (<https://www.scholarlyhub.org/>).

⁶This hypothetical use-case is not exactly original: see the real-life **Mat** describing his vision for sustainable business models for contemporary music publishing here: <https://medium.com/@matdryhurst/protocols-duty-despair-and-decentralisation-transcript-69acac62c8ea>

preferences and organized subsets of contacts. She might click between multiple views, **filtering out** or **focusing on** mainstream content, independent news, or fringe self-publications, depending on the context. She might disable or override the sophisticated **guardrails** that media sites use to filter out hate-speech, suspicious authors, spam, paid content, or misinformation. When studying communities with unorthodox views or “filter bubbles” of their own making, she might sift carefully through little-seen content signed by less-reputable journalists or news-outlets, made invisible for most users by **default trust settings and source-filters**.

In one such fishing expedition in deep, unfiltered waters, Sahar finds an intriguing article about a conspiracy that puts forth some controversial ideas about the motivations of some elected officials. Realizing that few professionals or casual readers will ever have seen this obscure piece, she feeds it into a complex **source-analysis system** that she can access through her newsroom’s subscription. This directs her to a version of the article with the most controversial assertions made in the piece highlighted by a collaborative **fact-checking service**; she doesn’t know who paid for it to be vetted, or what the results were, but they probably were not very favorable because the piece has very low circulation statistics. Hovering over the highlighted claims, a panel on her interface pulls up relevant **citation-trails**: like audit trails, these show who cited whom, all the way back to a few original articles in newspapers and publications. She disagrees with one of the citation trails, and flags it for review within the system; because she has been using it for years and her stories are rarely flagged for factual or citation errors, it is an easy way to keep her **contributor score** high.



Clicking around through these citation trails, Sahar uses her interface to compile a list of **root sources** at the end of the trails. Then, she pulls up the list of sources directly cited or linked in the article, comparing the two lists to get a sense of where the author’s information came from in writing this piece, including which newspapers, universities, schools of thought, and countries. In the direct list, a few unfamiliar publications catch her eye-- she looks up which are still publishing today to get a sense of where they fall in the scheme of things and orders a few pieces from an immutable archive for “**digital microfiches**” of the sources no longer online. A few have gained in reputation since, but mostly these are hobbyist publications, self-funded blogs, and shoestring historian journals on the fringes of the ecosystem.



Some authors from the root-source list are unfamiliar to her. Intrigued, she looks them up by their unique-author numbers (what in today’s terms we call *digital object identifiers*, or DOIs). Three, it seems, have never published anything else, ever: maybe this is why the article has such a low **aggregate trust score**! One of these intrigues her, so she searches more and turns up some contact info signed by the same credentials on a contact list for a community research organization. She contacts the author, Simon, and miraculously, he responds: it turns out his piece was pseudonymous because he was an elected official at the time! He had published the piece at a friend’s request and forgotten all about it.

Sahar interviews this obscure pseudonymous historian and writes her own controversial and provocative piece in a small journal of independent political analysis. She registers it under a pseudonym of her own to keep it separate from her professional publication record, since these “**censorship-resistant**” non-commercial channels could bring down the trust scores of her professional work.

Inspired by their conversation, Simon digs up an old laptop and proves his authorship of the original piece and re-registers it under his real name with some publication indexes and fact-checking systems. Enough time has passed that it his old pseudonym seems

unnecessary now, and Sahar's piece (and the scrappy politics journal that publishes it to a small network of devoted subscribers) both benefit from the **trust-rating boost** of having one less shadowy source in its citation trail.