

Context

In this chapter, we will describe the principal elements needed to understand the context and motivation of this thesis. First, we will go through a quick recall of the history of the web and its current challenges. Then we will talk about the motivation behind Solid and describe its main components.

History and current state of the World Wide Web

In the 1980s, the European Organization for Nuclear Research, also known as CERN, was looking for a way to make physicists easily share their work and data around the globe. Internet was at its beginning, the TCP/IP protocol was freshly invented, but at this time, sharing data across multiple computers was laborious and inefficient¹. A layer was lacking to write, transmit and store information easily through the network. In 1989, Sir Tim Berners Lee submitted a proposal for an information management system to tackle the former issue with new concepts such as hypertext links, web browsers, the HTML language, and the HTTP protocol^{2 3}.

On 1991 August 6, the first website available on the world wide web, “<http://info.cern.ch>” was born.

Today’s web might be far from the decentralized digital utopia imagined by Sir Tim Berners Lee in the early 90s. Currently, most services, data, and patents related to web technologies belong to only a few companies, commonly referred to as GAFAM⁴. This oligopoly is in part due to a phenomenon called the “vendor lock-in”: users tend to stay with the same platform since it is too hard to switch to another one. Furthermore, The centralization of web services and data has been at the heart of the latest internet scandal such as Cambridge Analytica or the worldwide NSA information gathering revealed by Edward Snowden⁵.

Solid: Tim Berners-Lee proposal to Re-Decentralise the web

To tackle those problems, Tim Berners-Lee has started the Solid project in 2016 as a mean to re-decentralize the web⁶. Solid take advantage of the Semantic Web (also known as Web 3.0 but not to be confused with web3), a new set of standards extending the current World Wide Web to make the internet’s data machine-readable. It is made machine-readable through Linked Data, a new way of formatting structured data to ensure its integration, interpolation, and interpretation by machines in the Semantic Web. Solid - sometimes stylized

¹<https://eu.usatoday.com/story/tech/news/2019/03/12/world-wide-web-turns-30-berners-lee-contract-thoughts-internet/3137726002/>

²<https://www.vox.com/2019/3/12/18260709/30th-anniversary-world-wide-web-google-doodle-history>

³<https://www.w3.org/History/1989/proposal.html>

⁴https://www.cairn.info/article.php?ID_ARTICLE=COMLA_188_0061#

⁵<https://www.theguardian.com/technology/2018/sep/08/decentralisation-next-big-step-for-the-world-wide-web-dweb-data-internet-censorship-brewster-kahle>

⁶[https://en.wikipedia.org/wiki/Solid_\(web_decentralization_project\)](https://en.wikipedia.org/wiki/Solid_(web_decentralization_project))

SoLiD - stands for **S**ocial **L**inked **D**ata. It aims to give back data control to internet users and improve their privacy online. Solid is not a technology in itself, but rather a set of specifications that allows decoupling of web applications' authentication, data, and app logic⁷. The end goal is to give internet users complete control over their data⁸. However, the web needs a couple of new artifacts and vocabulary to allow this decoupling; the two most important are the Pod (decoupling data storage) and the Identity Provider (decoupling authentication).

A Pod (Personal Online Datastore) is where Solid users store their data online. Like a Unix filesystem, the WAC (Web Access Control) controls who can read, write or delete files stored in a Pod. Each of these permissions is defined in `.acl` "Access Control List" files[^wiki solid].

WebIDs are the standardization of a universal identifier used for authentication. More than replacing the traditional username, it is a full URI that, once dereferenced, can give more information on the end user. It usually has the form of: `https://my-webid-host.net/my-username/profile/card#me`. The Identity Provider (IDP) permits the user to login into various applications with their WebID in the fashion of today's "Sign in with Google/Facebook". It uses The Solid-OIDC protocol - built on top of the Open ID Connect - that manages authentication in a Solid environment.

Identity Provider and Pod Provider are two decoupled services, but a Solid Server usually serves both.

Let us illustrate all those new terms with an example: Alice has a pod hosted on `http://alice.pod.org` where she stores her data, including a holiday picture at `http://alice.pod.org/picture/holiday.png`. By associating the `READ` permission to Bob's WebID `https://bob.anotherpod.org/profile/card#me` in the ACL file `http://alice.pod.org/picture/holiday.acl`, now Bob, if authenticated with his WebID will be authorized to access the picture at `http://alice.pod.org/picture/holiday.png`.

Those specifications are developed by the World Wide Web Consortium (W3C), an organization of web standard founded by Tim Berners-Lee, which promote technologies compatibility around the web. Tim Berners-Lee has also created a startup called Inrupt that aims to build a Solid commercial solution. Inrupt is financing a team of researchers from Ghent University to develop the Community Solid Server.

⁷<https://solid.github.io/specification/protocol>

⁸[https://en.wikipedia.org/wiki/Solid_\(web_decentralization_project\)](https://en.wikipedia.org/wiki/Solid_(web_decentralization_project))