Preprints.org

Article

Not peer-reviewed version

`I Consent to These Terms': A Legal and Technical Approach for Obtaining Valid Consent in Solid

Marcu Florea and Beatriz Esteves *

Posted Date: 19 July 2023

doi: 10.20944/preprints202307.1344.v1

Keywords: personal information management systems; solid; semantic web; data protection; consent

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

'I Consent to These Terms': A Legal and Technical Approach for Obtaining Valid Consent in Solid

Marcu Florea [1,†], **Beatriz Esteves** [2,†,*]

[1] Security, Technology and ePrivacy Group (STeP), University of Groningen, Netherlands; m.o.florea@rug.nl
[2] Ontology Engineering Group (OEG), Universidad Politécnica de Madrid, Spain; beatriz.gesteves@upm.es
[*] Corresponding author. E-mail: beatriz.gesteves@upm.es. Address: Universidad Politécnica de Madrid, Ramiro de Maeztu 7, 28040 Madrid, Spain.
[†] These authors contributed equally to this work.

Abstract: Personal Information Management Systems (PIMS) are acquiring a prominent role in the data economy by promoting products and services that help individuals to manage and control their online identity and thus have more control over the processing of their personal data, in line with the European strategy for data. One of the highlighted solutions in this area is Solid, a new protocol which is decentralising the storage of data, through the usage of interoperable Web standards and semantic vocabularies, to empower its users to have more control over the agents and applications that can access their data. However, to fulfil this vision and gather widespread adoption, Solid needs to be aligned with the law governing the processing of personal data in Europe, the General Data Protection Regulation (GDPR). To assist with this process, we analyse the current efforts to introduce a policy layer in the Solid ecosystem, in particular, related to the challenge of obtaining consent focusing on the GDPR. Furthermore, we investigate if, in the context of using personal data for biomedical research, consent can be expressed in advance, discuss the conditions for valid consent and how it can be obtained in this decentralised setting, namely through the matching of privacy preferences, set by the user, with requests for data and whether this can signify informed consent. Finally, we discuss the technical challenges of an implementation that caters to the previously identified legal requirements.

Keywords: personal information management systems; solid; semantic web; data protection; consent

1. Introduction

The General Data Protection Regulation (GDPR) [1] has become the lighthouse to follow when it comes to the protection of personal data in the European Union (EU) and its effects are being globally felt, with Asia, Latin America and Africa taking similar stances to protect its citizens' private information [2]. However, this hasn't come without challenges in its interpretation and enforcement. The allocation of rights and obligations concerning the processing of personal data in GDPR is structured around different roles: data controllers, data processors and data subjects. Data controllers (the natural or legal person who determines the purposes and means of the processing of personal data) data must declare a lawful, fair and transparent *purpose* to justify the processing of personal data so that the data subject (the person whose data are processed) can make an informed decision when it comes to the usage of its personal data. The allocation of rights and responsibilities based on the concepts of controller and data subject is challenged by complex data flows where multiple parties govern the usage, storage and collection of personal data for both distinct or shared purposes.

In addition, there is a risk that information requirements described in GDPR's Articles 12 to 14 are being treated by personal data-processing companies as a 'tick-box' compliance exercise. Although Article 12 [1] provides that the information is presented in a concise, transparent, intelligible and easily accessible form, using clear and plain language, this is not easy to assess and implement in reality. Compliance with transparency obligations can be dealt with by providing lengthy, complex, ungraspable privacy notices, which place a significant burden on the data subjects [3,4].

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

Even if the requirements of clarity, and transparency are complied with, data subjects are offered the possibility to be informed, this involves a nearly impossible exercise: to understand the myriad of terms and conditions of all the services and applications that are used these days, from smartphone applications to social media websites and personalised streaming of digital content. Thus, the information provided to data subjects fails to be an efficient tool for data subjects to monitor how their data is used or how their rights can be exercised [5].

Moreover, as recognized by the European Commission in its 'strategy for data' communication, currently only a few *"Big Tech firms hold a large part of the world's data"*, a fact that is making smaller businesses struggle to grow and innovate in this digital era. To this end, the EU's vision encompasses the creation of a single European market for data, where access to personal and non-personal data from across the world is secure and can be used by an ecosystem of companies, governments and individuals to provide high-quality data-driven products and services for its citizens, while ensuring that *"EU law can be enforced effectively"* and data subjects are still in control of what happens to their personal data [6]. In this sense, novel data-related legislation with new data governance schemes, such as the Data Governance Act (DGA) [7], is being brought forward by the EU to improve the citizens' trust[1] in data-handling services and allow them to share their sensitive data for the 'public good'. One mode which is being largely discussed these days is the so-called *personal data sovereignty* governance scheme, which presents a radical change in the paradigm of data-sharing. In this new governance model, the access to data is decentralised — data subjects assume direct control over the usage and sharing of their data, a solution that promises to balance the power relationship between Web users and digital platforms by promoting the development of digital services focused on the users' needs [9–11].

In this context, the emergence of personal data spaces managed through Personal Information Management Systems (PIMS) is already being envisioned by the European Data Protection Supervisor (EDPS) as a mechanism to enable personal data sovereignty where *"Individuals, service providers and applications would need to authenticate to access a personal storage centre" and individuals can "customize what categories of data they want to share and with whom"* while keeping a track of *"who has had access to their digital behaviour"* and enabling data portability and interoperability [12]. Furthermore, these new user-managed systems represent the next step towards the matching of privacy terms between data subjects and data controllers and can actually play an important role in facilitating the exercise of data subjects' rights, including the rights of access, erasure, and data portability or the right to withdraw consent [13]. In this context, a set of different PIMS initiatives has been gaining prominence and adoption in the last few years, including the Solid project[2]. Solid is a free, open-source initiative that delivers on the promise of decentralising the storage of data by relying on Web standards and on Semantic Web vocabularies to promote data and services interoperability. To fulfil this vision, the Solid specification relies on authentication and authorization protocols to provide private, secure and granular access to data stored in Solid's personal online datastores, the so-called 'Pods'.

As such, there have been recent efforts to align the GDPR with personal datastores and in particular with Solid. One of the more discussed issues relies on the uncertainties generated by such decentralised systems in the definition of responsibilities under the GDPR [14,15] — while some defend that in such settings data subjects become data controllers of their own data [16], a view that clashes with the existing regulations [17], others maintain that the user remains the data subject and the providers and developers of such systems are data controllers.

It is, therefore, also important to make a distinction between what can be enforced technologically and what can only be legally enforced – while technically we can restrict the data that applications can have access to, and remove the access grant when we no longer want to use them, when an app can read data from a Pod, it can also copy it, even if with Solid they do not need to do it. At this point,

[1] Trust has been proven as an important factor that positively influences the perceived usefulness and ease of use of digital personal datastores [8].

[2] https://solidproject.org/

we enter the realm of the law — where processing must comply with several principles and rules. Although the data subject wishes, as declared by the policies that they have stored in the Pod, play an important role, their legal significance depends on how and when they are expressed [18].

In what concerns the requirement that processing is lawful (Article 6 [1]), the usage of other lawful grounds for processing beyond consent [15,19] or dealing with access to special categories of personal data [20] remain up for discussion.

In addition to the challenges around legal bases, when it comes to the alignment of Solid with data protection requirements, a number of relevant initiatives has been materialising in recent years, mainly through academic projects and publications. Pandit analysed this technology in terms of the involved actors, according to existing standards related to cloud technology, in order to identify GDPR issues that are still applicable in decentralised settings, such as the transparency of information, purpose limitation and exercising of data subject's rights [21]. Other researchers have been focused on adding a legally-compatible policy layer to Solid as a tool to express consent and determine access [22,23] and usage control [24] to data stored in Pods and on using the Verifiable Credential model to have an attribute-based access control mechanism [25].

Taking into consideration this 'law+tech' approach to the management of personal data in decentralised settings, in this work, we focus on the current efforts to introduce a policy layer to the Solid ecosystem, further developed in Section 2, as a tool to obtain informed and valid GDPR consent and, in particular, for the usage of GDPR's special categories of personal data for biomedical research. The following challenges were identified for the implementation of such a system:

Ch1. Users' policies as a precursor of consent -– Previous studies have shown that the current access control mechanisms supported by the Solid protocol are not enough to deal with GDPR requirements, however, there is work being developed to introduce a policy language – the Open Digital Rights Language (ODRL) – *"for Expressing Consent through Granular Access Control Policies"*[3] [22]. User policies can enable compliance with several requirements of the GDPR. Pursuant to Articles 13 and 14 [1], data controllers have the obligation to provide the data subject information about the processing of their personal data and users' policies can enable communication of this information. Furthermore, information about the processing of personal data is a prerequisite for obtaining valid consent pursuant to Articles 7 and 4 (11) [1].

Ch2. Automation of consent — Decentralised ecosystems, such as the one involving Solid Pods, rely on the existence of authorizations to provide access to (personal) data. Since its users are the ones specifying the access authorizations, said systems provide a fertile ground for research on the automation of access to resources – in this case, a data request might be automatically accepted, with no further action from the user, if the user had previously added a policy in its Pod stating that said access can be granted. Whether such automation can be considered consent under the GDPR is still up for debate. Even though there is no provision in GDPR prohibiting the expression of consent in advance, for it to be valid, the conditions set in Article 7 and Article 4 (11) [1] must also be met. In addition to the requirement of consent to be informed, the controller must be able to prove that consent was freely given, specific and explicit.

Ch3. Dealing with health data for biomedical research – The processing of GDPR's special categories of personal data, such as data concerning health, is prohibited by default and brings extra "burdens" to data controllers. In addition to identifying a legal basis under Article 6 [1], they must rely on an exception under Article 9 [1]. Also, at the national level, further limitations regarding the processing of health data can be introduced. There are however certain derogations when health data are processed for scientific research or for the management of public health (Recital 52 [1]).

[3] Beatriz Esteves is the lead author of this work.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

To tackle such challenges, we focus on addressing the following research question: ***Can the matching between user policies and data requests, in a decentralised setting such as the Solid project, signify consent?***

To address this question, as the main contributions of this paper, in Section 2 we provide an overview of Solid and relevant work in the area, in Section 3 we provide a legal overview of the distinction between providing consent and granting access to data, in Section 4 we discuss the automation of consent, in particular regarding the expression of consent in advance, the specificity of purposes, the disclosure of the identity of data controllers and the special requirements related with the usage of personal data for biomedical research, and in Section 5 we discuss future research directions and provide the concluding remarks of the work.

2. Background — Decentralising the Web with Solid

2.1. Solid overview

Solid presents a radical paradigm shift in relation to today's web – by detaching data from Web applications, users are given *control* over their data and *choice* over which apps they want to use with said data. This represents a major shift in power in relation to what users experience nowadays when they go online. By unlocking the storage of data from the hand of just a few storage providers, such as Google or Facebook, Solid gives its users the option of having a Pod — a *personal online datastore* — using their storage provider of choice or even hosting their own storage server [26]. While multiple users can use the same Solid server to host their data Pod, Solid's ultimate goal is to give its users the highest degree possible of decentralisation – one Pod per person, or even multiple Pods per person, with a granular access control mechanism where they can choose which people and apps have access to their Pod, to a particular container of resources stored in their Pod or even to an individual Pod resource. In this scenario, applications act as clients that can read and/or write data from/to different Pods, without storing it in their own servers. Therefore, beyond giving people control over their data, such an ecosystem *"fosters innovation and competition through separate markets for data and applications"* [18].

Solid's two main building blocks[4] are its authentication[5] and authorization protocols[6]. The authentication protocol is related to the identification of agents – the WebID specification[7] is used to identify agents through URLs, which when dereferenced, direct to a profile document that can contain information describing the agent it identifies. The authorization protocol deals with the server's responses to requests of particular agents, in other words, it is the access control mechanism of Solid. Furthermore, the current version of the Solid protocol[8] specification states that, for a Solid server to be compliant, it "MUST conform to either or both Web Access Control (WAC)[9] and Access Control Policy (ACP)[10] specifications". Further details on the authorization protocol will be given in Section 2.2. A third building block is now being developed – the Solid Application Interoperability specification[11]. Said specification details how agents and applications can interoperate and reuse data from different sources.

[4] https://solidproject.org/TR/
[5] https://solidproject.org/TR/oidc
[6] https://solid.github.io/authorization-panel/authorization-ucr/
[7] https://solid.github.io/webid-profile/
[8] https://solidproject.org/TR/protocol
[9] https://solidproject.org/TR/wac
[10] https://solidproject.org/TR/acp
[11] https://solid.github.io/data-interoperability-panel/specification/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

2.2. *Access control in Solid*

As pointed out in the previous section, access control in Solid can currently be determined with two different specifications, WAC and ACP. While the Solid protocol mandates that the servers where the Pods are hosted conform to only one of the WAC or ACP access authorizations, Solid applications must comply with both or else they take the risk of not being usable by half of the ecosystem. Both solutions rely on IRIs to identify resources and agents, while WAC uses Access Control Lists (ACLs) to store authorizations, defined per resource or inherited from the parent resources, and ACP uses Access Control Resources (ACRs) to describe who is allowed or denied access to resources and access grants to represent already authorised accesses.

As illustrated by Listings 1 and 2, neither WAC nor ACP have the coverage to model GDPR's information requirements (Articles 13 and 14 [1]) for the processing of personal data, in particular when it comes to the modelling of the purpose for processing, personal data categories, legal basis or even information on the identity of the data requester. To overcome this issue, research has been developed in the area of integrating the ODRL model into the Solid ecosystem.

As illustrated by Listings 1 and 2, neither WAC nor ACP have the coverage to model GDPR's information requirements (Articles 13 and 14 GDPR) for the processing of personal data, in particular when it comes to the modelling of the purpose for processing, personal data categories, legal basis or even information on the identity of the data requester. To overcome this issue, research has been developed in the area of integrating the ODRL model into the Solid ecosystem [22,23].

Listing 1. WAC authorization that makes a WebID profile, https://solidweb.me/besteves4/profile/card, readable by any agent.

```
PREFIX acl: <http://www.w3.org/ns/auth/acl#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

<#public>
    a acl:Authorization ;
        acl:agentClass foaf:Agent ;
        acl:accessTo <https://solidweb.me/besteves4/profile/card> ;
        acl:mode acl:Read .
```

Listing 2. ACP authorization that makes a WebID profile, https://solidweb.me/besteves4/profile/card, readable by any agent using any client application.

```
PREFIX acp: <http://www.w3.org/ns/solid/acp#>
PREFIX acl: <http://www.w3.org/ns/auth/acl#>

<#public>
    acp:grant acl:Read ;
    acp:context [
    acp:agent acp:PublicAgent ;
            acp:target <https://solidweb.me/besteves4/profile/card> ;
            acp:client acp:PublicClient ;
            acp:issuer <https://solidweb.me/>
    ] .
```

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

ODRL[12] [27] is a W3C standard for policy expression which includes an information model and a vocabulary of terms. It provides a convenient extension mechanism, through the definition of ODRL profiles[13], that can be used to create policies for different use cases, from software licences to access and usage control policies. Since ODRL is not domain specific, e.g., it can be extended to create policies for financial[14] or language[15] resources, it means that it is also not equipped to deal with legal requirements. To this end, the ODRL profile for Access Control (OAC)[16] makes use of ODRL's deontic representation capabilities and connects them with the Data Privacy Vocabulary (DPV)[17] [28] to invoke data protection-specific terms. DPV provides an ample set of taxonomies that can be used to specify entities, legal basis, personal data categories, processing activities, purposes, or technical and organisational measures. Therefore, by integrating the usage of ODRL and DPV, OAC allows Solid users to express their privacy preferences and requirements over particular types of data, purposes, recipients or processing operations at distinct levels of specificity – from broad, e.g., allow data use for scientific research, to narrow policies, e.g., prohibit sharing a particular resource with a particular application. Figure 1 presents a diagram with the main concepts defined in OAC to express such policies. Requests for access, either from other users or from applications or services, can be modelled in a similar manner and stored in the Pod to have a record of said requests. Listings 3 and 4 illustrate an example of a user policy as an `odrl:Offer` and an example of a data request as an `odrl:Request`, respectively.

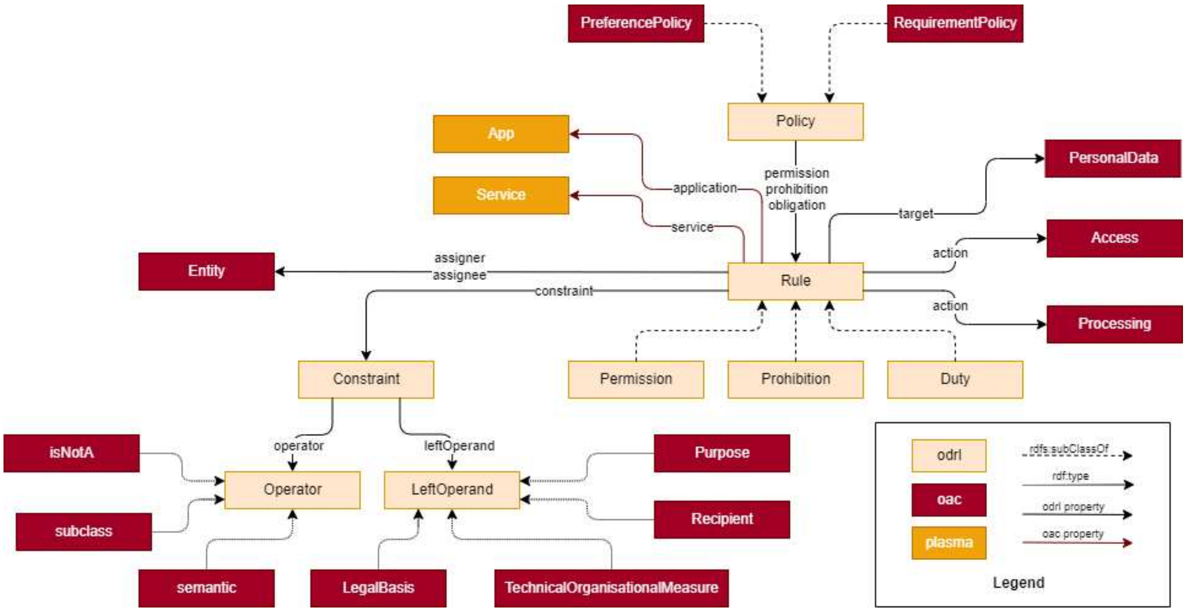


Figure 1. Core concepts of the ODRL profile for Access Control (OAC).

[12] http://www.w3.org/ns/odrl/2/
[13] https://www.w3.org/profile-bp/
[14] https://w3c.github.io/odrl/profile-temporal/
[15] https://rdflicense.linkeddata.es/profile.html
[16] https://w3id.org/oac
[17] https://w3id.org/dpv

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

Listing 3. An example ODRL offer policy generated by https://solidweb.me/besteves4/profile/card#me, stating that health records data can be accessed for the purpose of health, medical or biomedical research.

```
PREFIX odrl: <http://www.w3.org/ns/odrl/2/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX oac: <https://w3id.org/oac#>
PREFIX dpv: <https://w3id.org/dpv#>
PREFIX duodrl: <https://w3id.org/duodrl#>
PREFIX ex: <https://example.com>

<https://example.com/offer1> a odrl:Offer ;
  dct:description "Offer to read health records data for health, medical or biomedical
   ↪  research." ;
  dct:creator <https://solidweb.me/besteves4/profile/card#me> ;
  dct:issued "2023-05-30T17:26:35"^^xsd:dateTime ;
  odrl:uid ex:offer1 ;
  odrl:profile oac: ;
  odrl:permission [
        dpv:hasContext dpv:Optional ;
        odrl:assigner <https://solidweb.me/besteves4/profile/card#me> ;
        odrl:target oac:HealthRecord ;
        odrl:action oac:Read ;
        odrl:constraint [
          dct:title "Purpose for access is to conduct health, medical or biomedical (HMB)
           ↪  research." ;
          odrl:leftOperand oac:Purpose ;
          odrl:operator odrl:isA ;
          odrl:rightOperand duodrl:HMB ] ] .
```

Listing 4. An example ODRL Request policy made by https://solidweb.me/arya/profile/card#me, using the https://example.com/healthApp application, to use health records data from https://solidweb.me/besteves4/profile/card#me to conduct research on arterial hypertension disease.

```
PREFIX odrl: <http://www.w3.org/ns/odrl/2/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX oac: <https://w3id.org/oac#>
PREFIX duodrl: <https://w3id.org/duodrl#>
PREFIX dpv: <https://w3id.org/dpvl#>
PREFIX ex: <https://example.com>

<https://example.com/request1> a odrl:Request;
  dct:description "Request to use health records data for research on arterial
   ↪  hypertension." ;
  dct:creator <https://solidweb.me/arya/profile/card#me> ;
  dct:issued "2023-05-31T18:15:56"^^xsd:dateTime ;
  odrl:uid <https://example.com/request1> ;
  odrl:profile oac: ;
  odrl:permission [
        odrl:assignee <https://solidweb.me/besteves4/profile/card#me> ;
        odrl:assigner <https://solidweb.me/arya/profile/card#me> ;
        oac:application <https://example.com/healthApp> ;
        odrl:action oac:Use ;
        odrl:target oac:HealthRecord ;
        odrl:constraint [
          dct:title "Purpose for access is to conduct research on arterial hypertension." ;
          odrl:leftOperand oac:Purpose ;
          odrl:operator odrl:eq ;
          odrl:rightOperand ex:HypertensionResearch ] ] .

ex:HypertensionResearch a dpv:Purpose, rdfs:subclassOf duodrl:HMB ;
  rdfs:label "Conduct research on arterial hypertension disease." .
```

By integrating the usage of such a policy layer in the Solid ecosystem, the matching of users' preferences and requests for data is possible and can be automated. OAC's proposed matching algorithm consists of checking for subsumption between data requests and user policies — if the data request satisfies the users' policies, then access can be provided to the Pod. On the other hand, if any prohibitions are found in the users' policies that match the data request, access to the Pod is denied. The result of the matching is stored in the Pod for record keeping and future inspection. Thus, OAC will be used as our motivating scenario. While the decision to deny access based on user policies can be interpreted as the exercise of a data subject right, e.g., the right to object in Article 21 [1], this article focuses on whether the positive result of the matching can signify consent.

2.3. Other related works

The issue of control and privacy in Solid has been further explored by academia and industry. Beyond access, research on usage control has also been developed [24,29], with the main goal of creating tools to enforce policies and ensure that data is being used according to the users' preferences after the access has been provided. In addition, the exercising of GDPR's data subject rights, in particular of the Right to Data Portability [30] and the Right of Access [31], has been proven to be facilitated through the usage of Solid. Digita, a Belgium-based startup commercialising Solid solutions[18], also published a research report reflecting on the applicability of GDPR's requirements to Solid implementations, in particular, regarding data exchange with consent [19]. Recent efforts also promoted a tool to generate and store OAC policies in Solid Pods [32] and evaluated the usage of the Solid Application Interoperability specification to create a User Interface for users to evaluate data requests [33]. Hochstenbach et al. are developing RDF Surfaces[19], a Notation3 language which intends to bring first-order logic to the Semantic Web and therefore can be used to *"provide enforcement of data policies using logic-based rules"* [34].

In the particular field of health research, a Solid-powered platform has been developed to manage data requests and provide consent for health-related research using DPV [35]. Solid is also being tested by the United Kingdom's (UK's) National Health Service (NHS) to collect and process patient data from several systems, which is then hosted in individual patient Pods owned by the patients, who can authorise their healthcare professionals to have access to the data [36].

3. Describing the distinction between consent and granting access to a resource

From a legal perspective, the governance rules of Solid do not have to always rely on individual choice to access resources in a Pod. While control is an important principle of the GDPR [37,38], it is achieved by both empowerment (consent, the right to object, the right to be forgotten) and protective

[18] https://www.digita.ai/
[19] https://w3c-cg.github.io/rdfsurfaces/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

measures (such as conducting a balancing exercise, data protection impact assessments (DPIAs), or by implementing the requirements on Article 25 related to data protection by design and by default (DPbDD)). Empowerment measures can be expressed *ex-ante*, e.g., consent, but also *ex-post*, e.g., the right to object or the right to be forgotten.

The processing of personal data must respect several principles mentioned in Article 5 [1]. Article 5 (1) a) concerns the principle of lawfulness, fairness and transparency. In order for the processing to be lawful, it must rely on one of six legal grounds: (a) consent (b) processing is necessary for the performance of a contract (c) processing is necessary for compliance with a legal obligation (d) processing is necessary in order to protect the vital interests (e) processing is necessary for the performance of a task carried out in the public interest (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party.

Apart from consent, there are five grounds that do not involve the active choice of the data subject [39]. For example, if the data subject entered into a contract with the entity that requests access (the data controller) and the processing of personal data is necessary for the performance of that contract, there is no need for the consent of the data subject [40].

Apart from the requirements for lawfulness, GDPR provides information obligations that aim to respect the principle of transparency. Articles 12 and 14 [1] require data controllers to take appropriate measures to provide the data subject with information regarding the processing of their personal data and users' policies can function as an information mechanism. The result of the matching exercise is stored in a known location in the Pod, where the data subjects can easily access and check the result, enabling them to easily comprehend whether the agreed specific conditions for processing personal data differ from the pre-set of preferences stored in the Pod. Listing 5 presents the ODRL agreement generated as the result of the matching exercise between the data request in Listing 4 and the user policy in Listing 3. Since the personal data type in the user policy, i.e., health records, matches the requested personal data type, and the requested purpose for processing is "to conduct research on arterial hypertension", which is a subclass of health, medical or biomedical research — the purpose allowed by the user policy —, an agreement between the data subject and controller is generated and stored in the Pod, allowing the controller to access said data. Such a record also allows the user to check which entities requested access to the data.

Listing 5. An example ODRL Agreement policy that grants https://solidweb.me/arya/profile/card#me read access to health records data from https://solidweb.me/besteves4/profile/card#me to conduct research on arterial hypertension disease.

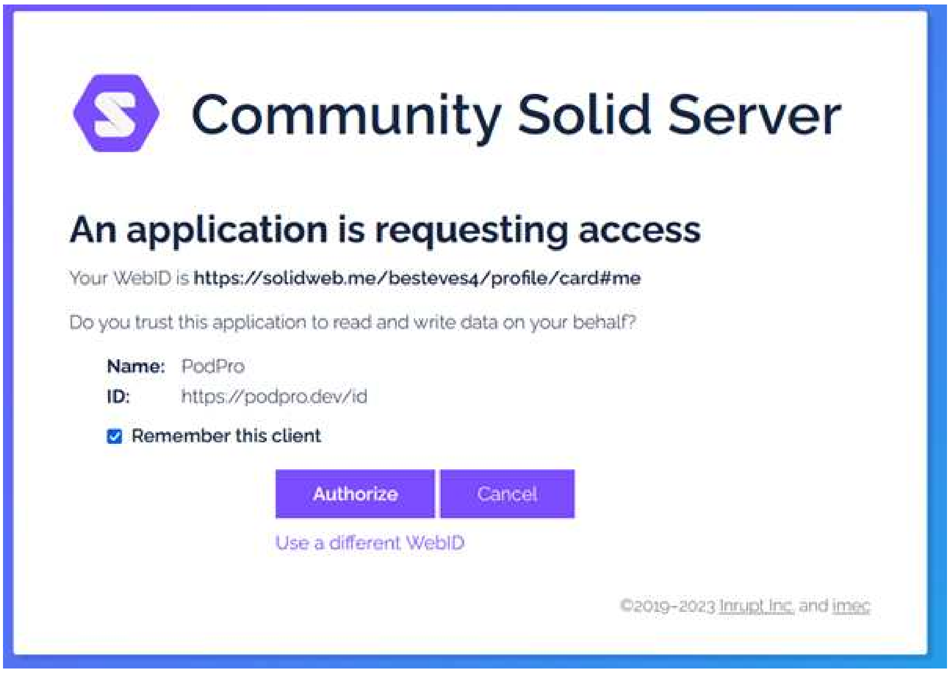
```
PREFIX odrl: <http://www.w3.org/ns/odrl/2/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX oac: <https://w3id.org/oac#>
PREFIX dpv: <https://w3id.org/dpv#>
PREFIX duodrl: <https://w3id.org/duodrl#>
PREFIX ex: <https://example.com>

<https://example.com/agreement1> a odrl:Agreement;
  dct:description "Agreement to read health records data for research on arterial
   ↪  hypertension." ;
  dct:issued "2023-05-31T18:20:06"^^xsd:dateTime ;
  odrl:uid <https://example.com/agreement1> ;
  odrl:profile oac: ;
  dpv:hasDataSubject <https://solidweb.me/besteves4/profile/card#me> ;
  dpv:hasDataController <https://solidweb.me/arya/profile/card#me> ;
  dpv:hasLegalBasis dpv:Consent ;
  dct:references ex:offer1, ex:request1 ;
```

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

```
odrl:permission [
      odrl:assignee <https://solidweb.me/arya/profile/card#me> ;
      odrl:assigner <https://solidweb.me/besteves4/profile/card#me> ;
      oac:application <https://example.com/healthApp> ;
      odrl:action oac:Read ;
      odrl:target oac:HealthRecord ;
      odrl:constraint [
        dct:title "Purpose for access is to conduct research on arterial hypertension."
          ↪  ;
        odrl:leftOperand oac:Purpose ;
        odrl:operator odrl:eq ;
        odrl:rightOperand ex:HypertensionResearch ] ] .
```

Data controllers have an obligation to provide the information in a concise, transparent, intelligible and easily accessible form (Article 12 [1]), using clear and plain language and according to Recital 59 [1], modalities should be provided for facilitating the exercise of the data subject's rights. The matching between preferences and requests can enable compliance with these requirements.

As previously explained in Section 2.2, access to resources in Solid depends on authentication and authorization protocols. Access can be granted by data subjects when they start using a new application, through an authorization dialogue, such as the example provided in Figure 2[20], or can be pre-set in the Pod in advance, such as the example provided in Figure 3[21]. Both options can take various forms, depending on the technical access control mechanism that is implemented in the server where the Pod is hosted, as the servers are not obliged to implement both authorization protocols (WAC and ACP) promoted by Solid, as was previously discussed in Section 2.2, and on the interfaces used to interact and manage the data and the access grants on the Pod.



Community Solid Server

An application is requesting access

Your WebID is https://solidweb.me/besteves4/profile/card#me

Do you trust this application to read and write data on your behalf?

Name: PodPro
ID: https://podpro.dev/id

☑ Remember this client

Figure 2. Screenshot of the authorization dialogue of the Community Solid Server (CSS) Pod provider.

[20] https://communitysolidserver.github.io/CommunitySolidServer/6.x/
[21] https://docs.inrupt.com/user-interface/podbrowser/

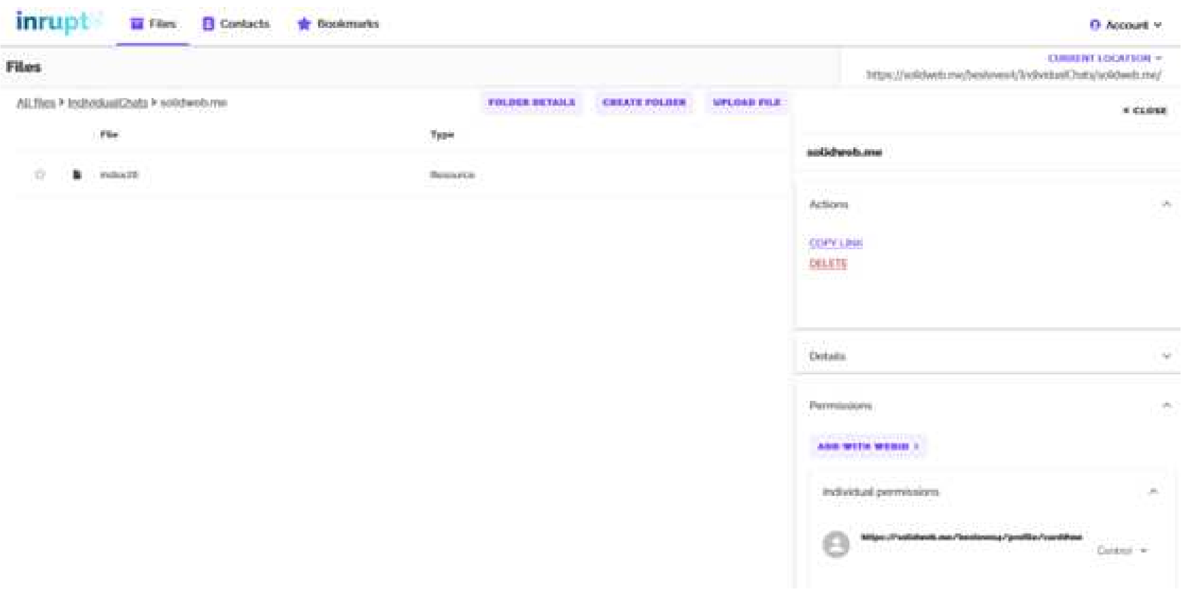


Figure 3. Screenshot of Inrupt's PodBrowser app to manage data and access grants.

We can conclude at this point that setting preferences and matching them with access requests can lower the burden of data subjects in reading and comprehending the information related to the processing of their personal data.

But can the legal effects of this matching go beyond information and signify consent? Expressing consent pursuant to GDPR requires compliance with rather strict conditions. Consent is defined in Article 4 (11) [1] as the '[...] indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'. Consent must be freely given, specific, informed and unambiguous. These requirements are further developed in the European Data Protection Board (EDPB) and WP29 Guidelines [41–43].

Failing to comply with one or more of the requirements for obtaining valid consent will render the authorization effective from a technical perspective, but this will not signify consent from a legal perspective. Especially problematic are the requirements of consent to be specific, informed and unambiguous. To comply with these conditions, consent must be granular, the purposes of processing cannot be bundled, and the data subject must be informed of the consequences of refusal. A mere acknowledgement or an agreement for sharing a resource does not always signify consent. In the next section, we will develop whether and how these requirements can be complied with in the context of Solid. Furthermore, in the context of biomedical research, when data concerning health are processed, consent must comply with an additional requirement: it must be specific.

4. Can consent be automated?

Because of cognitive limitations (skewed decision-making) and structural limitations (scale, aggregation, assessing harm), it is difficult to obtain meaningful consent for the processing of personal data [44].

To avoid a dilution of consent, GDPR requires strict conditions for expressing valid consent. This results in an abundance of choices, overburdening individual data subjects. Can technology solve some of the problems by helping the individual in decision-making and in expressing consent? Sharing the task of decision-making between a human and a machine may be the way forward for consent.

Consent is generally viewed as binary: an individual expresses her/his wishes representing agreement or disagreement with the processing of her personal data. The standard model of consenting involves the following steps: the data controller sends a request for consent and the data subject accepts or rejects it.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

Pre-setting access permissions in the Pod in advance switches this order. The data subjects make the first move and set their preferences concerning the processing of personal data in advance. Even though the GDPR does not provide a framework for the interaction between the data subject and a technology-based system in expressing consent, the possibility of expressing consent by choosing technical settings is suggested in Recital 32 [1]: 'Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement [...]. This could include [...] choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. [...]'

As explained in [22], there are two possible levels of automation and both start with a request for accessing the personal data. However, what follows after this is different. In the first case, the result of the matching is presented to the data subject and the data subject is requested to provide her/his consent. In this case, the matching process enables information and helps the data subject to make an informed choice, but consent is only expressed after the moment when the app provider/developer requests permission to process the personal data, i.e., when the user wishes to use the app. In the second case, access to personal data is granted automatically, based on the preferences of the data subject, which are expressed in advance. The first type of automation affirms the principle of transparency and helps the data subject to comprehend the information in the privacy policy and to apply it to the choice that she/he is requested to make. It partially solves the cognitive problems that data subjects are facing, but it does not provide a solution for the problem of scale as it requires new consent for each instance of access and might overwhelm the data subject with too many requests. The second one grants automated access to entities that request permission to access resources in a Pod, if the preferences of the individual user match the request for consent. The problem that we will analyse is whether, from a legal perspective, this second instance of matching can signify consent.

4.1. Expressing consent in advance

There is no provision in the text of the GDPR that forbids the expression of consent in advance. However, in order to be valid, consent must refer to specific circumstances that may arise in the future [45].

The idea of automated implementation of user preferences was heavily discussed in the context of cookies and behavioural-based advertising. The Do Not Track Initiative[22] and the Platform for Privacy Preferences[23] are two examples in this respect. Neither succeeded in being uptaken on a large scale, but both can constitute helpful examples in developing a system for consent based on Solid. WP 29 commented on this initiative [46]. Furthermore, in a rather recent development, the Commissioner for Justice and Consumers, Didier Reynders, launched a reflection on how to better empower consumers to make effective choices regarding tracking-based advertising models[24]. The problem that it tries to solve is similar to the difficulties around accessing content in Solid Pods.

The possibility of pre-configured choices was also mentioned in the context of the implementation of Article 5 (3) of the ePrivacy Directive [47], which refers to the storage of information or gaining access to information stored in the terminal equipment of a subscriber or users of publicly available electronic communications services. Some national laws implemented the text by allowing the expression of consent via technical means. For example, the implementing law in Romania [48] refers to "using the settings of the Internet browsing application or other similar technologies" for expressing consent,

[22] https://www.eff.org/issues/do-not-track

[23] https://www.w3.org/P3P/

[24] https://commission.europa.eu/live-work-travel-eu/consumer-rights-and-complaints/enforcement-consumer-protection/cookie-pledge_en

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

while in Finland, the ombudsman and the Transport and Communications agency expressed different views on the validity of consent to cookies given through browser settings[25].

The next sections will discuss what are the requirements for expressing consent in advance and whether Solid can be adapted to meet them.

4.2. Expressing specific consent

To answer the question of whether consent can be automated, it is relevant to first discuss the level of granularity required by the law and whether and how this can be achieved in a technical implementation. Pursuant to Article 4 (11) [1], consent must be a specific indication of the data subject's wishes that signifies agreement to the processing of personal data. The expression 'indication of wishes' is rather indeterminate. One might question if the wishes of the data subject refer to the categories of data, the purpose of processing, the processing operations, the identity of the data controller (s) or a correlation between them. Dammann and Simitis (apud Eleni Costa) interpret the requirement for specific consent as one that refers to (a) a specific processing of personal data, (b) by a specific data controller (c) for a specific purpose. According to Eleni Kosta, specificity is respected when the relation between the personal data and the processing of which the data subject wishes to consent, are clearly defined and the conditions surrounding the processing are explained. The consent statement should be so specific that it safeguards the right to informational self-determination [45].

In the following subsections, we will focus on two elements that are relevant for expressing valid consent: the purpose of processing and the identity of the data controller and we will discuss how they relate to the specific character of consent.

4.2.1. Specificity of purpose and processing operations

Article 6 (1) a) [1] pinpoints the specificity requirement in relation to the *purpose* of processing. 'Processing shall be lawful only if and to the extent that at least one of the following applies: (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes.' However, the preamble of the GDPR suggests that consent is expressed not only for separate purposes but also for separate data processing operations. Recital 42 [1] reads as follows: 'Where processing is based on the data subject's consent, the controller should be able to demonstrate that the data subject has given consent to the processing operation [. . .]' and Recital 43 [1] provides that 'Consent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case'. There is a categorical difference between purposes and processing operations. While purposes describe the reason or the objective for processing personal data, processing operations refer to the action performed in relation to the personal data.

Article 4 (2) [1] provides several examples of operations: "collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction". Several operations might be necessary to reach a purpose. Several purposes can be reached by the same processing operation, e.g., use, which has a very broad scope).

The WP 29 provides in its opinion on consent that *specific* consent is intrinsically linked to the fact that consent must be informed [42]. There is a requirement of granularity of consent with regard to the different elements that constitute the data processing: it can *not* be held to cover "all the legitimate purposes" followed by the data controller [42]. Consent should refer to the processing that is reasonable and necessary in relation to the purpose. However, this does not address the question of how broadly

[25] https://cookieinformation.com/resources/blog/finland-changes-cookie-rules/, https://tietosuoja.fi/-/apulaistietosuojavaltuutettu-maarasi-yrityksen-muuttamaan-tapaa-jolla-se-pyytaa-suostumusta-evasteiden-kayttoon

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

the purposes can be defined. While the WP29 gives a negative example of how consent could fail to be specific, it does not provide a comprehensive set of tools for assessing specificity [43].

In its opinion on electronic health records (EHR) [49], WP 29 provides that 'specific' consent must relate to a well-defined, concrete situation in which the processing of medical data is envisaged. Again, the meaning of 'concrete situation' is rather vague. The specificity of consent, according to the same Opinion, also means that if the reasons for which data is processed by the controller change at some point in time, the user must be notified and put in a position to consent to the new processing of personal data. The information supplied must discuss the repercussions of rejecting the proposed changes in particular.

One question that can be relevant at this point is whether any change in the purpose (no matter how insignificant) requires re-consent. Minor changes will probably not require new consent, while important ones will. But how can one decide between the two categories?

The problem of matching preferences with requests poses similar challenges. If the preferences do not perfectly match the requests, it might be necessary to ask the data subject for consent. Unless the preferences are specified with clear use cases in mind, e.g., with clear relations between categories of data – (specific) purpose – (specific) data controller, the matching operation will always involve assessing the compatibility between expressed preferences and the specifics of the processing of personal data.

Thus, the reasoner that implements the matching transforms a preference into a choice between the available options as it materialises the preferences and produces legal effects towards third parties (the data controller).

WP 29 touches upon the relation between purposes and processing operations: 'it should be sufficient in principle for data controllers to obtain consent only *once* for *different operations* if they fall within the *reasonable expectations* of the data subject.' It is unclear though how these expectations are determined. One proposal for further research is to apply a framework such as the contextual integrity theory of privacy developed by Helen Nissenbaum [50].

4.2.2. Assessing compatibility between preferences and requests

How can one assess compatibility between two purposes of processing? For example, if a participant in research expresses her will to have her data processed for Alzheimer's disease, which is a type or subcategory of 'degenerative disease', will this consent be valid for a request for using the personal data for a study on 'dementia', also a category of degenerative diseases?

As detailed in Section 2.2, OAC's proposed matching algorithm is currently based on subsumption, i.e., if a user policy allows access for purpose A and a data request for purpose B, which is a subcategory of A, comes in, then the access should be permitted. The same is valid for the other matching operations that can be made, e.g., on processing operations or personal data categories. Continuing with the example presented above, this means that if the participant has a user policy that states that her data can be used for research on Alzheimer's disease and the request is for research on degenerative diseases, then access is not allowed as the purpose of the user policy is more specific than the purpose of the request. On the other hand, if the participant has a user policy that states that her data can be used for research on degenerative diseases and a request for research on Alzheimer's comes in, then access is permitted since the purpose of the request is more specific, i.e., is a subcategory, than the purpose of the user policy. However, OAC does not consider yet the matching of "compatible purposes", e.g., if the participant has a user policy that states that her data can be used for research on Alzheimer's, and a request comes in to use the data for research on dementia, does it mean that the access should be allowed since dementia is a subcategory of degenerative disease such as Alzheimer's? Therefore, the introduction of a "compatibility matching" would improve the model, if permitted by law. The question is how to know that a purpose is compatible with another and if the user wishes to use such a compatibility model to deal with data requests. Moreover, in order to enable this matching for the particular use case of biomedical research, the work of Pandit and Esteves [51] of having ODRL

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

and DPV policies for health data sharing can be re-used, as it provides a taxonomy of health-related research purposes, connects to other ontologies with taxonomies of diseases and reuses the matching algorithm of OAC. However, for the reasoner to be able to decide on compatibility, this information should be embedded in the used taxonomies of purposes, for instance, by adding a triple statement expressing that `:purposeX :isCompatible :purposeY`.

In the GDPR, compatibility is discussed in relation to the purpose limitation principle in Article 5 (1) b) and the criteria for assessing it are mentioned in Article 6 (4) [1], as follows:

(a) any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
(b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
(c) the nature of the personal data, in particular, whether special categories of personal data are processed [...];
(d) the possible consequences of the intended further processing for data subjects;
(e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.

There are certain practical and legal challenges in including these criteria in the matching exercise. From a technical perspective, the first two criteria (a) the link between purposes and (b) the context of processing, might be determined by automated, as described above. However, the third and fourth criteria involve an evaluation that connects the nature of data and the possible consequences of use to the data subjects. The last criterion, the existence of technical safeguards can be partially verified automatically on the basis of certifications. However, the appropriateness of the safeguards is an assessment that is difficult to automate, as it involves an assessment of the risks to the rights and interests of the data subjects.

In addition to this, there are two legal challenges for using these criteria to the matching between user preferences and requests from third parties (app providers and developers) for using the personal data stored in a Pod.

Firstly, the purpose of these criteria is not to assess the compatibility of a request for access with privacy preferences. The criteria are developed for compliance with the purpose limitation principle, more specifically for assessing the compatibility between (a) collection and (b) further use of personal data. This is appropriate in a setting where the data is collected by centralised datastores that collect the data for a purpose and then process it further for different purposes, Solid does not collect the data for a specific purpose, it merely stores it and enables the use of the data for certain purposes by third parties. However, these criteria might be applied by analogy to assess the compatibility between privacy preferences and requests for the use of data. The suitability of applying these criteria depends on the connection between the specific character of consent and the purpose limitation principle, which requires further legal research.

Secondly, at first read, consent is excluded from the compatibility assessment. According to Article 6 (4) [1]: 'Where the processing for a purpose other than that for which the personal data have been collected *is not based on the data subject's consent* or on a Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard the objectives referred to in Article 23(1), the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia [...]' followed by the criteria mentioned above. In one reading, the compatibility criteria cannot be applied when personal data was collected on the basis of consent, as a legal basis. In such cases, consent must be requested for each new purpose, even if the purpose for usage is compatible with the purpose of collection. This reading is also supported by Recital 32 [**?**]: '[...] Consent *should cover all processing activities* carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them. [...]' However, in another reading, the compatibility assessment can be undertaken even if the legal basis for the collection was consent. The apparent

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

exclusion of consent actually affirms the role of consent as a compensation for the incompatibility between purposes.

This interpretation would follow the approach in the OECD Guidelines from 1980 that read as follows:

Purpose Specification Principle

Article 9. The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

Use Limitation Principle

Article 10. Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except

(a) *with the consent of the data subject*; or
(b) by the authority of law.

Therefore, a compatibility assessment can be performed irrespective of the legal ground of collection. If the result of the assessment is that the purposes are not compatible, consent can be asked again and expressed to mitigate the lack of compatibility. This subsection presented general directions for assessing the compatibility between purposes, but there is a need for further research on the connection between the principle of purpose limitation, the principle of lawfulness and consent in order to develop a legal framework that can serve as a basis for a Solid *policy layer*.

4.2.3. The fine line dividing expressing and delegating consent

Is there a difference between privacy or data protection preferences and decisions? In their paper "An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs" Jessica Colnago and others [52] make a difference between privacy preferences (outcome for a specific privacy-related situation) and decisions (what an individual chooses to do in a specific privacy related situation *among available options*). The focus of the paper is on using these concepts in empirical studies about attitudes towards privacy, but the distinction can serve in the discussion about the matching process proposed in the context of Solid. One can ask whether setting privacy preferences also means making privacy decisions. If this is not the case, does the reasoner take a decision on behalf of the data subject? These questions shall be discussed in connection with the roles of the Pod provider and the data subject, their involvement in decisions regarding the processing of personal data and the effects on the delineation of the roles of the data controller and data processor.

Sheehan [53] discusses ethical consent and draws a difference between first-order and second-order decisions. Second-order decisions are fundamentally different from first-order decisions in that the details relevant to the decision-maker concern the decision-making process and not the subject matter of the choice. Sheehan provides the example of placing an order in a restaurant. He is imagining a scenario in which a few friends go for dinner and, before the waiter can get their order, one of them has to leave for a few moments to pick up a call. He asks one of the persons at the table, the one who is aware of his general preferences, to order for him. When taking decisions to delegate decision-making, the individual makes choices and gives preference to something that he/she values: the trust in his companions, their knowledge about his taste in food, the information that they hold about the approximate value that he wants to spend, etc. The arrangement can contain the option to withdraw, in case the food does not match the preferences of the person who was away, the friends can swap the dishes.

How to separate between matching and interpretation? Research for the public interest is a concept that is interpretable: what does public interest mean? Can the reasoner make such assessments? The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

reasoner algorithm cannot tell if accessing a certain type of data for a certain purpose is of public interest. It cannot make such interpretations as of now, because currently, it does not have access to this information. However, if we had this type of information in the ontology, then the reasoner would be able to do it, as discussed before with the compatibility of purposes. Also, we can discuss that instead of adding more information to the ontology, the reasoner algorithm can "learn" to make such inferences, however, in this case, Article 22 [1] (Automated decision-making) might have to be considered. If consent is not expressed by the user, but rather delegated to an agent (defined in the Solid protocol) that acts on behalf of the user, it is necessary to inquire into the relationship between the data subject and the agent (from a contract law perspective) and the validity of consent expressed through a proxy (from the perspective of data protection law).

4.3. Specific data controllers? Or categories of recipients?

With OAC [22], data subjects can express explicit authorization for specific data controllers that are known at the moment when preferences are set. However, in order to grant such an explicit authorization in advance, information needs to be conveyed to the users when first wanting to use a Solid application or should be available somewhere for the Solid user to check, e.g., through metadata present on an app store. This is currently missing from Solid implementations – as it is visible in Figure 2 which illustrates the current consent dialogue shown to Solid users when they want to use an app – the name and the ID of the app are shown, but nothing else, no contact details, no policies or links to policies defined somewhere else. Also, there is no established marketplace for Solid apps with information about the provider and/or developers of said apps.

In this context, an authorization can take different forms and degrees of specificity. In the first example, the data subject identifies the data controller by name and contact details and grants them access to the personal data in the Pod. This case does not pose problems in terms of specificity. However, the downside of this option is that the data subject would have to approve each new data controller.

A second option consists of authorising controllers based on relevant criteria, such as industry, country of incorporation or sector. The data subject would set a list of preferences, but would not identify the entities by name and contact details. While the second option has the advantage of flexibility, there are questions as to whether it complies with the legal requirements for expressing valid consent.

4.3.1. The moment when the identity of the controller is disclosed

The first challenge is represented by the moment when the data subject is informed about the identity of the data controllers. In the first example, the identity of the data subject is revealed in advance.

In the second example, the user sets the criteria that a requester should comply with and the identity of the data controllers becomes available at the moment when the reasoner decides that the requester complies with the criteria set by the data subject (industry, country of incorporation or sector). The identity and contact details are not explicitly acknowledged or approved by the data subject before access is granted, but they are available for consultation in the Pod.

As far as the requirements concerning transparency are concerned, Article 13 [1] (which regulates information to be provided where personal data are collected *from the data subject*) requires that the information, including the identity and the contact details of the controller, are disclosed *at the time when personal data are obtained*. Article 14 [1] (which refers to information to be provided where personal data have *not been obtained from the data subject*), provides that, if a disclosure to another recipient is envisaged, the information has to be provided at the latest *when the personal data are first disclosed*. Therefore, in both cases, the information must be provided *at the moment when the requester starts* to process the personal data. In what concerns the validity of consent, the EDPB Guidelines 05/2020 on consent under Regulation 2016/679 [41] note: 'in a case where the consent is to be relied upon

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

by multiple (joint) controllers or if the data is to be transferred to or processed by other controllers who wish to rely on the original consent, *these organisations should all be named*.' The guidelines on transparency under Regulation 2016/679 [54] emphasise the importance of the identity of the controller by stating that changes to a privacy statement/notice that *should always be communicated* to data subjects include inter alia: a change in processing purpose; a change to the *identity* of the controller; or a change as to how data subjects can exercise their rights in relation to the processing.

Although connected, the specific character of consent and the informed character of consent are mentioned separately in Article 4 (11) [1]: 'consent' of the data subject means any freely given, *specific*, *informed* and unambiguous indication of the data subject's wishes. According to Recital 42 [1], for consent to be informed, the data subject 'should be aware at least of the identity of the controller and the purposes of the processing for which the personal data are intended.' If we consider that consent is expressed when the preferences are set, then this requirement is not respected. The data subject does not know, at this point, the identity and contact details of the data subject. If, on the contrary, consent is formed when the preferences match the characteristics of the data controller, the identity of the controller can be disclosed at the moment when personal data are collected.

Recital 39 [1] refers to the principle of lawfulness, fairness and transparency and provides that transparency 'concerns, in particular, information to the data subjects on the *identity of the controller* and the purposes of the processing and further information to ensure fair and transparent processing [...]'. Transparency is also referred to in Recital 58 [1], providing that transparency is of particular relevance in situations where the proliferation of actors and the technological complexity of practice make it difficult for the data subject to know and understand whether, *by whom* and for what purpose, personal data relating to him or her are being collected. These texts are suggesting that the identity of the data controllers is an important element for compliance with information and transparency obligations.

Recital 39 [1] refers to the rationale of transparent communications, stating that 'Natural persons should be made aware of risks, rules, safeguards and rights in relation to the processing of personal data and *how to exercise their rights* in relation to such processing'. The matching between criteria and actual requests enables individuals to exercise their rights. Even though the data subjects do not know the identity and contact details of the requesters, this information is available in the Pod, and they can use it in order to exercise their rights, such as the right of access (Article 15 [1]), the right to rectification (Article 16 [1]), the right to erasure (Article 17 [1]) or the right to withdraw consent (Article 7 (3) [1]).

4.3.2. The difference between controllers and recipients of personal data

Articles 13 and 14 [1] provide a list of elements that must be communicated to the data subject. Both articles separate the information on(i) 'the identity and the contact details of the controller and, where applicable, of the controller's representative' (Article 13 (1) a) and 14 (1) a) [1]) from the information on (ii) 'the recipients or categories of recipients of the personal data, if any' (Article 13 (1) e) and 14 (1) e) [1]). What is then the difference between data controllers and recipients, in particular in a decentralised setting? Are the requesters *controllers* or *recipients*? And what is the relevance of this distinction?

Pursuant to Article 4 (9) [1], 'recipient' means a natural or legal person [...]to which the personal data are disclosed, whether a third party or not'. A third party is, according to Article 4 (10) [1], 'a natural or legal person, [...] other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data'.

EDPB's guidelines on controller and processor [55] provide that, as opposed to the concepts of controller and processor, the Regulation does not lay down specific obligations or responsibilities for recipients and third parties. These can be said to be *relative concepts* in the sense that they describe a relation to a controller or processor from a specific perspective, e.g., a controller or processor discloses data to a recipient. The EDPB gives the example of a controller who sends data to another entity. Irrespective of its role as controller or processor, this party is considered a recipient.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

The definition seems to be adapted to the current web, where entities in charge of centralised datastores collect the data and share it afterwards with other entities (as recipients). However, in the structure intermediated by Solid, the sharing takes place directly between the data subject (enabled by the Solid Pod) and app providers, developers or other users. This raises the question of whether, at the moment when the data is collected by the Pod provider, these entities can be considered recipients and, as a consequence, mentioned as categories and not individually identified by name and contact details. This identification would take place at a later stage when the user preferences match the request for access.

The WP 29 Guidelines on transparency under Regulation 2016/679 [54] state that the actual (named) recipients of the personal data, or the *categories of recipients*, must be provided. To comply with the principle of fairness, controllers must provide information on the recipients that is *most meaningful for data subjects*. This Guidance further clarifies this requirement and states that 'in practice, this will generally be the *named recipients*, so that data subjects know exactly who has their personal data'. However, WP 29 mentions the possibility to also inform the data subjects on the categories of recipients, but requires to identify them as specifically as possible by indicating the type of recipient, i.e., by reference to the activities it carries out, the industry, sector and sub-sector and the location of the recipients.

To conclude this subsection, it is unlikely that agreeing to the processing of personal data without identifying the providers of Solid services, specifically in terms of their identity and contact details, will be regarded as valid consent under the GDPR. However, it might serve as an information mechanism that can enable compliance with Articles 13 and 14 GDPR.

4.4. The special case of biomedical research

As explained in the previous sections, the strict requirements for obtaining consent under the GDPR impose burdensome obligations on the data subjects. A requirement for separate agreements for each app provider and for each specific purpose results in repeated requests for consent. In the biomedical context, the likelihood of individuals involved in decision-making regarding their data might be lower compared to other sectors because the incentives for individuals to participate in biomedical research are not connected to immediate benefits. Furthermore, these benefits usually do not reflect on the personal situation of the individual but have broad effects on society. There are several provisions that suggest a more flexible approach to the requirements related to consent or even to move away completely from consent.

Recital 33 [1] suggests that broad consent is acceptable for research. Under certain conditions, data subjects can express consent to 'certain areas of scientific research', if 'recognised ethical standards for scientific research' are observed. However, this possibility is limited, as individuals shall have the opportunity to 'give their consent only to certain areas of research or parts of research projects'. The concepts of 'areas of research' or 'part of research projects' are domain-specific notions that are not defined in the GDPR, which is an omnibus omnibus regulation. The work of Pandit and Esteves on DUODRL[26] [51], inspired by the work of the Global Alliance for Genomics and Health[27] on the Data Use Ontology [56], can be reused by data subjects and data controllers to create policies for health data sharing, as it provides a taxonomy of health-related research purposes, connects to other ontologies with taxonomies of diseases, and includes the concepts to model projects and duties to use data, e.g., requirement to have ethical approval, requirement of collaboration with the study's investigator(s), or requirement to return the generated results of the study. Listing 6 provides an example of a user policy that states that the dataset https://example.com/Dataset can be used for the purpose of health, medical or biomedical research, identified with the `duodrl:HMB` concept, in the context of Project X,

[26] https://github.com/besteves4/duo-odrl-dpv/

[27] https://www.ga4gh.org/

`ex:ProjectX`, provided that the user or app requesting access can provide proof of having ethical approval, identified with the `duodrl:ProvideEthicalApproval` concept.

Listing 6. An example ODRL offer policy generated by https://solidweb.me/arya/profile/card#me, stating that a dataset can be accessed for the purpose of health, medical or biomedical research in the context of Project X, provided that the entity requesting data provides documentation of ethical approval.

```
PREFIX odrl: <http://www.w3.org/ns/odrl/2/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX duodrl: <https://w3id.org/duodrl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX oac: <https://w3id.org/oac#>
PREFIX gist: <https://ontologies.semanticarts.com/gist/>
PREFIX ex: <https://example.com>

<https://example.com/offerForBiomedicalResearch> a odrl:Offer ;
    dct:description "Request to use data for biomedical research on Project X." ;
    dct:creator <https://solidweb.me/arya/profile/card#me> ;
    dct:issued "2023-06-01T18:15:56"^^xsd:dateTime ;
    odrl:uid <https://example.com/offerForBiomedicalResearch> ;
    odrl:profile oac: ;
    dct:source duodrl:DUO_0000006, duodrl:DUO_0000021, duodrl:DUO_0000027 ;
    odrl:permission [
        odrl:assigner <https://solidweb.me/arya/profile/card#me> ;
        odrl:action oac:Use ;
        odrl:target <https://example.com/Dataset> ;
        odrl:duty [
            odrl:action duodrl:ProvideEthicalApproval
        ] ;
        odrl:constraint [
            odrl:and ex:purposeConstraint, ex:projectConstraint
        ]
    ] .

ex:purposeConstraint a odrl:Constraint ;
    dct:title "Purpose for access is to conduct health, medical or biomedical (HMB)
     ↪ research." ;
    odrl:leftOperand oac:Purpose ;
    odrl:operator odrl:isA ;
    odrl:rightOperand duodrl:HMB .

ex:projectConstraint a odrl:Constraint ;
    dct:title "Data can be used in the context of Project X." ;
    odrl:leftOperand duodrl:Project ;
    odrl:operator odrl:eq ;
    odrl:rightOperand ex:ProjectX .

ex:ProjectX a gist:Project .
```

This provision of Recital 33 [1] is only present in the preamble of the Regulation, which does not have binding force and is not mirrored in the text of the GDPR. Furthermore, it was interpreted narrowly by the EDPS in its preliminary opinion on Scientific Research (2020) [57].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

Looking beyond the European Union, the UK government proposed a prominent role for broad consent in medical research in its proposal for a reform of the Data Protection Act in the UK [58] and the proposal was well received, although some concerns were voiced regarding its lack of certainty and potential for abuse.

Moreover, the European Commission proposed a Regulation instrument governing the processing of health data, the European Health Data Space [59], which proposes to completely move away from consent for secondary use of personal data for biomedical research. Data holders have an obligation to disclose personal and non-personal data, under certain conditions and for a restricted range of purposes, including scientific research (Article 34 (1) (e) [59]) without the consent of the data subject. Article 33(5) of the EHDS Proposal seems to also overrule national laws that require consent 'where the consent of the natural person is required by national law, health data access bodies shall rely on the obligations laid down in this Chapter to provide access to electronic health data.' It remains to be seen what the final versions of this proposal would be, whether consent will play a role and whether it will be broad or specific. The proposal was criticised by the EDPB and EDPS in a joint opinion [60] which requires further clarity on the interplay between national laws requiring consent and the proposed European instrument.

In the GDPR, biomedical research poses challenges because it combines a stricter regime (because research involves processing health data, which are part of GDPR's special categories of data) with a series of derogations (aiming at supporting research because of its importance for society).

4.4.1. A stricter regime

The processing of special categories of data (including data concerning health) is forbidden pursuant to Article 9 (1) [1]. There are ten exceptions from this rule, one of which is the explicit consent of the data subject. However, as mentioned in the EDPB Opinion on Consent [41], it is unclear what the explicit character refers to, since expressing a "statement or clear affirmative action" is a prerequisite for 'regular' consent. As the requirements for expressing consent in the GDPR, it needs to be clarified what extra efforts a controller should undertake in order to obtain the explicit consent of a data subject in line with the GDPR [41]. EDPB provides several examples of expressing explicit consent. The Guidance [41] mentions a written statement or in the digital or online context: filling in an electronic form, by sending an email, by uploading a scanned document carrying the signature of the data subject, or by using an electronic signature. Two-stage verification of consent is another option for expressing explicit consent. For example, a data subject may get an email informing them of the controller's intention to process a medical data record. In the email, the controller states that he is requesting permission to use a specific collection of information for a specified reason. If the data subject agrees to the use of this information, the controller requests an email response with the words 'I agree'. Following the receipt of the response, the data subject receives a verification link that must be clicked or an SMS message.

In Solid, this can be implemented in different ways. Depending on the Solid server where the users choose to host their Pod, an *inbox* container, similar to the email inboxes available in other ecosystems, might be present by default when the user creates the Pod and can be used to receive these special requests. This *inbox* container has a special access control authorization — other users, beyond the data subject of the Pod, can only write to the container, in order to ensure that only the data subject can read the resources in said container. However, since the presence of this container is not standardised across the Solid ecosystem, it cannot always be found or might be called something else, causing an interoperability problem and hence applications cannot rely on its existence. A more refined solution relies on a graph-centric interpretation of a Pod, where *'each Solid pod is a hybrid, contextualized knowledge graph, wherein "hybrid" indicates first-class support for both documents and RDF statements, and "contextualized" the ability to associate each of its individual documents and statements with metadata such as policies, provenance, and trust'* [61]. With the proper recording of metadata, including context and provenance metadata, multiple views of the Pod can be generated as required by the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

different applications that the data subject wishes to use. In this case, the requests can be simply added to the graph, with no need to *hardcode* in the app where the requests should be written, and such requests can be visualised by the data subject using a Solid app or service compatible with this graphic-centric approach. Moreover, the work of Braun and Käfer [62] can be leveraged to sign and validate resources that carry the "I agree" statement of the data subject.

We can conclude that it is difficult to express explicit consent by pre-set preferences. Matching user preferences (expressed in advance) with requests for processing personal data will, most likely, fail to comply with the explicit character of consent. While the matching can increase transparency and help the individual make a decision, a second action of approving the use of personal data is necessary in order to comply with the explicit character of consent.

4.4.2. A series of derogations

A. The purpose limitation principle

The principle of purpose limitation and the compatibility assessment were discussed in Section 4.2.2 on "Assessing compatibility between preferences and requests". To remind the reader, pursuant to Article 5 (1) b) [1], 'data shall be collected for specific, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes'. When data is processed for scientific research purposes, there is a presumption of compatibility between the purpose of collection and further use, if the processing is conducted in accordance with appropriate safeguards for the rights and freedoms of the data subject (as provided under Article 89 (1) [1]). However, the effects of this presumption are not straightforward and depend on the relation between the purpose limitation principle and the principle of lawfulness. The WP29 Opinion on Purpose Limitation [63] provides that purpose limitation and the requirement to have a legal ground are two separate and cumulative requirements. Article 8 of the Charter of Fundamental Rights of the European Union (CFREU) [64] mentions specificity of purpose separate from the requirement for consent or another legitimate basis. Therefore, irrespective of compatibility, the data controller would have to rely on consent or another legitimate basis laid down by law. However, one provision in the preamble of the GDPR questions the separation between the two requirements. Recital 50 [1] reads as follows: 'The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected. In such a case, no legal basis separate from that which allows the collection of the personal data is required.' Thus, the intersection between the two principles, described above, and the effect on Solid requires further legal research.

B. The obligation to provide information

We discussed the information obligations in Articles 13 and 14 [1] in Section 4.3, 'Specific data controllers? Or categories of recipients?' focusing on the moment when the information needs to be provided to the data subject. If personal data is processed for (biomedical) research purposes, Article 14 [1] provides an exception for cases when personal data has not been obtained from the data subject. This may apply to Solid if we consider that not all personal data stored in Solid Pods comes directly from the data subject, e.g., it can be generated by app providers, Pod providers or other users or agents. Pursuant to Article 14 (5) [1] if (i) the provision of information proves impossible or would involve a disproportionate effort or it is likely to render impossible or seriously impair the achievement of the objectives of the processing and (ii) the conditions and safeguards in Article 89 [1] are respected, the general information obligations in Article 14 do not apply. The data controller "shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available". Further research is necessary to discuss whether the notification system of Solid could serve as an appropriate measure to protect the data subject's rights.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

C. Alternative legal bases beyond consent

Besides explicit consent, Article 9 (2) [1] provides other exceptions from the prohibition to process special categories of data. Article 9 (2) j) is especially relevant for this section because it refers to processing personal data for health research. This provision allows the processing of data concerning health when it is necessary for scientific research in accordance with Article 89(1) [1], based on Union or Member State law, which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. Therefore, the application of this exception depends on the identification of a Union or Member State law that can serve as a basis for the processing of personal data. If the processing falls under the scope of such a law, the explicit consent of the data subject is not necessary.

It results from this section that the derogations for processing personal data for scientific research depend on the implementation and application of appropriate safeguards. Pursuant to Article 89 (1) [1], these safeguards refer to respect for the principle of data minimization and consist of, for example, pseudonymisation and techniques that do not permit the identification of data subjects. Future research can determine whether PIMS such as the matching system presented in this paper can play a role as a safeguard.

5. Future Research & Concluding Remarks

There is still much to be done when it comes to the alignment of Solid with legal requirements. In this article, we focused on the current efforts to have a policy layer in Solid for the expression of consent and started by identifying key challenges that need to be overcome for this alignment to occur. First off, there is the debate around the usage of user policies as a means for users to express their consent in advance, and how specific these need to be in order to provide the information required by the law. Further, we explored whether matching user policies with requests for accessing personal data can signify consent as a ground for lawfulness. Automated consent implemented in decentralised environments, such as the one presented by Solid, can help with the users' "burden" of reading the terms and conditions and consenting to dozens of requests. However, there are several legal challenges connected to the automation of consent. There is a need for further legal research to clarify to what extent it can be expressed by individuals with the help of technologies, such as the Solid protocol.

To bring light to the identified challenges, we started by providing an overview of Solid, and in particular of its access control mechanisms, and the legal and technical explorations that have been considered so far. We then developed on the important distinction between providing consent and granting access to a resource and how the integration of a policy layer, such as the one provided by OAC, can help data controllers to actually get explicit consent from the data subjects. The main body of this work was then devoted to answering our research question of how *Can the matching between user policies and data requests, in a decentralised setting such as the Solid project, signify consent?* To this end, we explored a set of criteria to express specific consent, in particular, related to the specificity and compatibility of purposes, to the disclosure of the identity of the data controller and third-party recipients, and to the special requirements of biomedical research, and which technical solutions can help to deal with such requirements within the Solid ecosystem.

As future work, we highlight the need to (i) study the specificity of purposes and processing operations provided in taxonomies, such as the ones available in DPV, to check whether their labelling is enough for both data controllers to declare their activities and for data subjects to understand what is happening to their data, (ii) have tools to assess the compatibility of purposes to put less burden in the users to access similar data requests, (iii) develop a taxonomy of recipients, e.g., by industry, sector, etc., to express which recipient categories can, cannot or are receiving a copy of the personal data of the users, (iv) research on the additional legal requirements of using other legal basis beyond consent and the use of PIMS as safeguards for the data subject's rights and freedoms and legitimate interests, (v) implement a stricter access control mechanism for special categories of data, for instance

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

using VCs, and (vi) look at the requirements of new data-related laws being discussed and approved in the EU, such as the Data Governance Act, Data Act or the European Health Data Space proposal.

Funding: This article is partially funded by the COST Action on Distributed Knowledge Graphs (CA19134), supported by COST (European Cooperation in Science and Technology). Beatriz Esteves was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813497 (PROTECT). Marcu Florea is an Early Stage Researcher within the KnowGraphs Project, the work of which is supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Innovative Training Network, grant agreement No. 860801.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2018.
2. Bradford, A. *The Brussels Effect: How the European Union Rules the World*; Oxford University Press, 2019. https://doi.org/10.1093/oso/9780190088583.003.0002.
3. Terpstra, A.; Schouten, A.P.; Rooij, A.d.; Leenes, R.E. Improving privacy choice through design: How designing for reflection could support privacy self-management. *First Monday* **2019**, *24*. https://doi.org/10.5210/fm.v24i7.9358.
4. Linden, T.; Khandelwal, R.; Harkous, H.; Fawaz, K. The Privacy Policy Landscape After the GDPR. In Proceedings of the Proceedings on Privacy Enhancing Technologies, 2020, Vol. 1, pp. 47–64. https://doi.org/https://doi.org/10.2478/popets-2020-0004.
5. Mohan, J.; Wasserman, M.; Chidambaram, V. Analyzing GDPR Compliance Through the Lens of Privacy Policy. In Proceedings of the Heterogeneous Data Management, Polystores, and Analytics for Healthcare; Gadepally, V.; Mattson, T.; Stonebraker, M.; Wang, F.; Luo, G.; Laing, Y.; Dubovitskaya, A., Eds. Springer International Publishing, 2019, Lecture Notes in Computer Science, pp. 82–95. https://doi.org/10.1007/978-3-030-33752-0_6.
6. European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A European strategy for data*. Brussels, COM(2020) 66 final ed., 2020.
7. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance), 2022. Legislative Body: CONSIL, EP.
8. Mariani, M.M.; Ek Styven, M.; Teulon, F. Explaining the intention to use digital personal data stores: An empirical study. *Technological Forecasting and Social Change* **2021**, *166*. https://doi.org/10.1016/j.techfore.2021.120657.
9. Craglia, M.; Scholten, H.; Micheli, M.; Hradec, J.; Calzada, I.; Luitjens, S.; Ponti, M.; Boter, J. *Digitranscope: The governance of digitally transformed society*; Publications Office of the European Union, 2021.
10. Ilves, L.K.; Osimo, D. A roadmap for a fair data economy. Policy Brief, Sitra and the Lisbon Council, 2019.
11. Verbrugge, S.; Vannieuwenborg, F.; Van der Wee, M.; Colle, D.; Taelman, R.; Verborgh, R. Towards a personal data vault society: an interplay between technological and business perspectives. In Proceedings of the 2021 60th FITCE Communication Days Congress for ICT Professionals: Industrial Data – Cloud, Low Latency and Privacy (FITCE), 2021, pp. 1–6. https://doi.org/10.1109/FITCE53297.2021.9588540.
12. Supervisor, E.D.P. TechDispatch #3/2020 - Personal Information Management Systems. Technical report, 2021.
13. Janssen, H.; Cobbe, J.; Singh, J. Personal information management systems: a user-centric privacy utopia? *Internet Policy Review* **2020**, *9*.
14. Janssen, H.; Cobbe, J.; Norval, C.; Singh, J. Decentralized data processing: personal data stores and the GDPR. *International Data Privacy Law* **2020**, *10*, 356–384. https://doi.org/10.1093/idpl/ipaa016.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

15. Van Damme, S.; Mechant, P.; Vlassenroot, E.; Van Compernolle, M.; Buyle, R.; Bauwens, D. Towards a Research Agenda for Personal Data Spaces: Synthesis of a Community Driven Process. In Proceedings of the Electronic Government; Janssen, M.; Csáki, C.; Lindgren, I.; Loukis, E.; Melin, U.; Viale Pereira, G.; Rodríguez Bolívar, M.P.; Tambouris, E., Eds. Springer International Publishing, 2022, Lecture Notes in Computer Science, pp. 563–577. https://doi.org/10.1007/978-3-031-15086-9_36.
16. Edwards, L.; Finck, M.; Veale, M.; Zingales, N. Data subjects as data controllers: a Fashion(able) concept? *Internet Policy Review* **2019**.
17. Chomczyk Penedo, A. Self-sovereign identity systems and European data protection regulations: an analysis of roles and responsibilities. In Proceedings of the Open Identity Summit 2021. Gesellschaft für Informatik e.V., 2021, pp. 95–106. Accepted: 2021-05-20T13:12:14Z ISSN: 1617-5468.
18. Verborgh, R. Paradigm shifts for the decentralized Web, 2017-12-20.
19. De Bot, D.; Haegemans, T. Data Sharing Patterns as a Tool to Tackle Legal Considerations about Data Reuse with Solid: Theory and Applications in Europe. Digita research reports, 2021.
20. Lodge, T.; Crabtree, A.; Brown, A. Developing GDPR Compliant Apps for the Edge. In Proceedings of the Data Privacy Management, Cryptocurrencies and Blockchain Technology; Garcia-Alfaro, J.; Herrera-Joancomartí, J.; Livraga, G.; Rios, R., Eds. Springer International Publishing, 2018, Lecture Notes in Computer Science, pp. 313–328. https://doi.org/10.1007/978-3-030-00305-0_22.
21. Pandit, H.J. Making Sense of Solid for Data Governance and GDPR. *Information* **2023**, *14*. https://doi.org/10.3390/info14020114.
22. Esteves, B.; Pandit, H.J.; Rodríguez-Doncel, V. ODRL Profile for Expressing Consent through Granular Access Control Policies in Solid. In Proceedings of the 2021 IEEE European Symposium on Security and Privacy Workshops (EuroS PW), 2021, pp. 298–306. ISSN: 2768-0657, https://doi.org/10.1109/EuroSPW54576.2021.00038.
23. Debackere, L.; Colpaert, P.; Taelman, R.; Verborgh, R. A Policy-Oriented Architecture for Enforcing Consent in Solid. In Proceedings of the Companion Proceedings of the Web Conference 2022. Association for Computing Machinery, 2022, WWW ’22, pp. 516–524. https://doi.org/10.1145/3487553.3524630.
24. Akaichi, I. Semantic Technology based Usage Control for Decentralized Systems, [2206.04947 [cs, eess]]. https://doi.org/10.48550/arXiv.2206.04947.
25. Braun, C.H.J.; Käfer, T. Attribute-based Access Control on Solid Pods using Privacy-friendly Credentials. In Proceedings of the Proceedings of the Poster and Demo Track and Workshop Track of the 18th International Conference on Semantic Systems Co-Located with 18th International Conference on Semantic Systems (SEMANTiCS 2022), 2022.
26. Sambra, A.V.; Mansour, E.; Hawke, S.; Zereba, M.; Greco, N.; Ghanem, A.; Zagidulin, D.; Aboulnaga, A.; Berners-Lee, T. Solid: A Platform for Decentralized Social Applications Based on Linked Data. Technical report, 2016.
27. Iannella, R.; Villata, S. ODRL Information Model 2.2, URL: https://www.w3.org/TR/odrl-model/, 2018.
28. Pandit, H.J.; Polleres, A.; Bos, B.; Brennan, R.; Bruegger, B.; Ekaputra, F.J.; Fernández, J.D.; Hamed, R.G.; Kiesling, E.; Lizar, M.; et al. Creating a Vocabulary for Data Privacy: The First-Year Report of Data Privacy Vocabularies and Controls Community Group (DPVCG). In Proceedings of the On the Move to Meaningful Internet Systems: OTM 2019 Conferences; Panetto, H.; Debruyne, C.; Hepp, M.; Lewis, D.; Ardagna, C.A.; Meersman, R., Eds. Springer International Publishing, 2019, Vol. 11877, pp. 714–730. https://doi.org/10.1007/978-3-030-33246-4_44.
29. Havur, G.; Sande, M.; Kirrane, S. Greater Control and Transparency in Personal Data Processing:. In Proceedings of the Proceedings of the 6th International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and Technology Publications, 2020, pp. 655–662. https://doi.org/10.5220/0009143206550662.
30. De Mulder, G.; De Meester, B.; Heyvaert, P.; Taelman, R.; Dimou, A.; Verborgh, R. PROV4ITDaTa: Transparent and direct transferof personal data to personal stores. In Proceedings of the Companion Proceedings of the Web Conference 2021. Association for Computing Machinery, 2021, WWW ’21, pp. 695–697. https://doi.org/10.1145/3442442.3458608.
31. Esteves, B.; Rodríguez-Doncel, V.; Longares, R. Automating the Response to GDPR’s Right of Access. In *Legal Knowledge and Information Systems*; IOS Press, 2022; pp. 170–175. https://doi.org/10.3233/FAIA220462.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

32. Esteves, B.; Rodríguez-Doncel, V.; Pandit, H.J.; Mondada, N.; McBennett, P. Using the ODRL Profile for Access Control for Solid Pod Resource Governance. In Proceedings of the The Semantic Web: ESWC 2022 Satellite Events; Groth, P.; Rula, A.; Schneider, J.; Tiddi, I.; Simperl, E.; Alexopoulos, P.; Hoekstra, R.; Alam, M.; Dimou, A.; Tamper, M., Eds. Springer International Publishing, 2022, Lecture Notes in Computer Science, pp. 16–20. https://doi.org/10.1007/978-3-031-11609-4_3.
33. Bailly, H.; Papanna, A.; Brennan, R. Prototyping an End-User User Interface for the Solid Application Interoperability Specification Under GDPR. In Proceedings of the The Semantic Web; Pesquita, C.; Jimenez-Ruiz, E.; McCusker, J.; Faria, D.; Dragoni, M.; Dimou, A.; Troncy, R.; Hertling, S., Eds. Springer Nature Switzerland, 2023, Lecture Notes in Computer Science, pp. 557–573. https://doi.org/10.1007/978-3-031-33455-9_33.
34. Hochstenbach, P.; De Roo, J.; Verborgh, R. RDF Surfaces: Computer Says No. In Proceedings of the 1st Workshop on Trusting Decentralised Knowledge Graphs and Web Data, 2023.
35. Sun, C.; Gallofré Ocaña, M.; van Soest, J.; Dumontier, M. ciTIzen-centric DAta pLatform (TIDAL): Sharing distributed personal data in a privacy-preserving manner for health research. *Semantic Web* **2023-01-01**, *14*, 977–996. Publisher: IOS Press, https://doi.org/10.3233/SW-223220.
36. Janeiro Digital at Solid World: NHS Personal Health Stores with XFORM Health and Solid, 2021.
37. Ausloos, J.; Ausloos, J. *The Right to Erasure in EU Data Protection Law*; Oxford Data Protection & Privacy Law, Oxford University Press, 2020.
38. Lynskey, O. *The Foundations of EU Data Protection Law*; Oxford Studies in European Law, Oxford University Press, 2015.
39. Kranenborg, H.R. Article 8 – Protection of Personal Data. In *The EU Charter of Fundamental Rights*; Hart Publishing, 2014.
40. European Data Protection Board. Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects, 2019.
41. European Data Protection Board. Guidelines 05/2020 on Consent under Regulation 2016/679 Version 1.1, 2020.
42. Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.
43. Article 29 Data Protection Working Party. Guidelines on consent under Regulation 2016/679.
44. Solove, D.J. Privacy Self-Management and the Consent Dilemma. *Harvard Law Review* **2012**, *126*.
45. Kosta, E. *Consent in European Data Protection Law*; Martinus Nijhoff Publishers, 2013.
46. Article 29 Data Protection Working Party. Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 24 April 2014, Tracking Preference Expression (DNT).
47. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), 2002.
48. Legea 506/2004 Privind Prelucrarea Datelor cu Caracter Personal si Protectia Vietii Private in Sectorul Comunicatiilor Electronice.
49. Article 29 Data Protection Working Party. WP 29 Working Document on the processing of personal data relating to health in electronic health records (EHR).
50. Nissenbaum, H. Privacy as Contextual Integrity. *Washington Law Review* **2004**, *79*, 119.
51. Pandit, H.J.; Esteves, B. Enhancing Data Use Ontology (DUO) for Health-Data Sharing by Extending it with ODRL and DPV. *Under Revision in the Semantic Web Journal* **2023**.
52. Colnago, J.; Cranor, L.F.; Acquisti, A.; Stanton, K.H. Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs. In Proceedings of the 18th Symposium on Usable Privacy and Security (SOUPS 2022), 2022, pp. 331–346.
53. Sheehan, M. Can Broad Consent be Informed Consent? *Public Health Ethics* **2011**, *4*, 226–235. https://doi.org/10.1093/phe/phr020.
54. Article 29 Data Protection Working Party. Guidelines on Transparency under Regulation 2016/679, 2018.
55. European Data Protection Board. Guidelines 07/2020 on the concepts of controller and processor in the GDPR, 2020.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2023 doi:10.20944/preprints202307.1344.v1

56. Woolley, J.P.; Kirby, E.; Leslie, J.; Jeanson, F.; Cabili, M.N.; Rushton, G.; Hazard, J.G.; Ladas, V.; Veal, C.D.; Gibson, S.J.; et al. Responsible sharing of biomedical data and biospecimens via the "Automatable Discovery and Access Matrix" (ADA-M). *npj Genomic Medicine* **2018**, *3*, 1–6. Number: 1 Publisher: Nature Publishing Group, https://doi.org/10.1038/s41525-018-0057-4.
57. European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research, 2020.
58. UK Government. Consultation outcome - Data: a new direction - government response to consultation, 2022.
59. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Health Data Space, 2022.
60. EDPB-EDPS Joint Opinion 03/2022 on the Proposal for a Regulation on the European Health Data Space, 2022.
61. Dedecker, R.; Slabbinck, W.; Wright, J.; Hochstenbach, P.; Colpaert, P.; Verborgh, R. What's in a Pod? – A knowledge graph interpretation for the Solid ecosystem. In Proceedings of the Proceedings of the 6th Workshop on Storing, Querying and Benchmarking Knowledge Graphs; Saleem, M.; Ngonga Ngomo, A.C., Eds., 2022, Vol. 3279, *CEUR Workshop Proceedings*, pp. 81–96.
62. Braun, C.H.J.; Käfer, T. Self-verifying Web Resource Representations Using Solid, RDF-Star and Signed URIs. In Proceedings of the The Semantic Web: ESWC 2022 Satellite Events; Groth, P.; Rula, A.; Schneider, J.; Tiddi, I.; Simperl, E.; Alexopoulos, P.; Hoekstra, R.; Alam, M.; Dimou, A.; Tamper, M., Eds. Springer International Publishing, 2022, Lecture Notes in Computer Science, pp. 138–142. https://doi.org/10.1007/978-3-031-11609-4_26.
63. Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013.
64. Charter of Fundamental Rights of the European Union, 2000.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.