БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи Код \_\_\_\_

УДК

Коберник-Березовский Юрий Алексеевич

Изучение спецификации Solid и разработка недостающих библиотек

Реферат по  
«Основам информационных технологий»

Магистранта кафедры системного анализа и компьютерного моделирования факультет радиофизики

Специальность: 1-13.80.08 – физическая электроника

Рецензент:   
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Минск, 2018

# Оглавление

[Оглавление 2](#_Toc530320224)

[Перечень условных обозначений 3](#_Toc530320225)

[Введение 4](#_Toc530320226)

[Общая характеристика работы 5](#_Toc530320227)

[Глава 1 Аналитический обзор 6](#_Toc530320228)

[Глава 2 Методика исследования 7](#_Toc530320229)

[2.1 Представление концепции Solid 8](#_Toc530320230)

[2.2 Основные технологии и стандарты Solid 8](#_Toc530320231)

[2.3 Хранение личных данных и выбор системы управления 10](#_Toc530320232)

[2.4 Solid POD как личное хранилище 11](#_Toc530320233)

[Глава 3 Результаты 13](#_Toc530320234)

[Заключение 14](#_Toc530320235)

[Библиографический список 15](#_Toc530320236)

[Приложение А 16](#_Toc530320237)

[Приложение В 17](#_Toc530320238)

# Перечень условных обозначений

Solid – Social Linked Data

MIT – Massachusetts Institute of Technology

W3C – World Wide Web Consortium

API – Application programming interface

REST – Representational state transfer

JSON – JavaScript Object Notation

HTML – Hyper-Text Markup Language

URI – Uniform Resource Identifier

HTTP – Hyper-Text Transfer Protocol

FOAF – Friend of a Friend

POD – Personal Online Data storage

# Введение

Информационная безопасность в настоящее время играет важную роль в жизни каждого человека, имеющего доступ к сети интернет и регулярно предоставляющего свои данные на всеобщее обозрение.

В последнее время участились случаи кражи личных данных, нецелевое использование данных своих клиентов компаниями, на серверах которых хранится личная информация о миллионах пользователях. Возникают случаи, когда компании продают информацию о своих клиентах третьим без прямого согласия на это самих клиентов. Также в современном мире существует множество информационных порталов и ресурсов, для входа на которые необходимо каждый раз вводить и передавать один и тот же набор данных.

Таким образом, у человека возникает множество копий идентичной информации о себе на различных ресурсах, что создаёт проблему контроля этих данных для пользователя. Социальные сети и их разнообразие - наиболее яркий пример описанных проблем. Одним из возможных решений можно представить программное обеспечение на основе спецификации Solid, по которой доступ к данным контролируется непосредственно её владельцем.

В рамках этой спецификации имеется ограниченный набор реализаций не универсального назначения. Соответственно требуется разработка и структуры, и реализации инструментов создания программного обеспечения на основе данной спецификации.

# Общая характеристика работы

Концепция Solid была представлена миру сравнительно недавно, в августе 2018 года. Её основоположником являются британец сэр Тим Бернерс-Ли, при поддержке американского университета MIT и международной организации W3C. На сегодняшний момент существует ограниченный набор реализаций, предоставленный авторами Solid, подобной системы только на основе технологии NodeJS, позволяющей наглядно изучить принцип работы и создания системы на основе Solid спецификации.

Цель данной исследовательской работы заключается в рассмотрении концепции распределённого хранения данных в сети интернет между пользователями. В данной работе описывается принцип распределённого хранения и передачи личных данных человека в сети интернет через специальные системы, основанные на правилах и технологиях, описанных в спецификации Solid.

Мы рассмотрим основные концепции, технологии и способы разработки Solid систем, ознакомимся с авторами данной идеи, их готовой реализацией и технологиями, которые были использованы в процессе разработки.

Данная работа носит ознакомительных характер и не содержит личной реализации подобной системы, а лишь описывает стандарты и технологии, используемые в реализации Solid систем.

# Глава 1 Аналитический обзор

На сегодняшний момент существует проблема того, как данные хранятся в сети интернет и как они используются теми, на кого возложена задача по хранение этих данных.

Миллионы пользователей предоставляют информацию о себе, своих близкий, предпочтения и взглядах, в открытый доступ не задумываясь о том, что всё что они рассказывают о себе, попадая в сеть, становится все доступным и хранится у третьих лиц, которые могут заниматься накоплением и продажей этих данных.

Интернет позволил нам создать более лучший и связанный мир, однако сеть превратилась в двигатель несправедливости и деления, раскачиваемый мощными силами, которые используют его в своих собственных интересах.

Сегодня мы достигли критического переломного момента, когда контроль над данными практически не осуществим со стороны пользователя, а контролируется сторонними компаниями и лишь изретка предоставляя видимость контроля за их распространением истинному владельце.

В качестве примера, можно упомянуть недавний скандал о краже данных пользователей социальной сети Facebook британской аналитической компанией Cambridge Analytica. Лондонская компании, предоставляющая политические консалтинговые услуги, воспользовалась невнимательностью и доверчивостью пользователей при помощи викторины “Это ваша цифровая жизнь”. Люди, не чуя подвоха, соглашались на любые условия работы приложения и отдавали множество личной информации о себе и своих друзьях. Таким образом, Cambridge Analytica при помощи Facebook смогла собрать достаточно информации для предоставления консалдинговых услуг политикам во время предвыборной компании.

Разработка ПО на основе Solid спецификаций позволит создавать системы, в которых для общения, хранения персональных данных, документов и фотографий не надо будет предоставлять свои данные в открытый доступ, а конкретно указывать кому предоставлять доступ и на какой период времени.

Solid изменяет текущую модель Интернета, где пользователи должны передавать персональные данные цифровым гигантам в обмен на доступ к интернет ресурсам. Как мы понимаем, подобное положение не в наших интересах. Solid – это концепция того, как мы развиваем сеть, чтобы восстановить равновесие – путем предоставления каждому из нас полного контроля над данными.

# Глава 2 Методика исследования

Solid – это технология, которая лежит в основе движения, которое возглавляет сэр Тим Бернерс-Ли, чтобы переориентировать сеть на свое первоначальное видение. Чтобы понять «почему» Solid, важно понять эволюцию Интернета и то, как она расходится с оригинальным и вдохновляющим видением сэра Тима.

Первый веб-браузером был также редактором веб контента. Идея заключалась в том, что пользователи могли не только читать контент в Интернете, но и помогать создавать его. Это должно было стать совместным пространством для всего человечества. И когда появился первый браузер, который популяризировал Интернет, он назывался Mosaic, он включал в себя мультимедиа и редактирование. Однако редактирование оказалось избыточным для рядового пользователя и было убрано в последующих версиях. Это изменение стало первым сокращением общения в сети и породило усилия Тима и его единомышленников вернуть функциональность. Концепция была названа «сетью для чтения и записи» и приводит к оригинальной статье Ричарда Макмануса, опубликованной в 2003 году.

Проблема с редактирование и наполнением данных, таких платформ как Википедия, заключается в том, что вам нужна определенная степень контроля над тем, кто может что-либо написать и насколько достоверна информация редактора контента. Это означает, что вам нужно создать набор прав и правила – что могут делать люди с данными. Чтобы предоставлять разрешения на модификацию данных, вам нужно иметь систему для идентификации редакторов – способ однозначно подтвердить, что человек – тот, кем себя представляет.

В то же время сеть всегда была предназначена для хранения данных, а также документов. Группа Linked Data Platform формализовала это, создав стандарт для работы с файловыми системами, который в сочетании с философией Unix является мощным способом достижения эффективного хранения. Эти два принципа, возможности чтения и записи и доступность управляемых данных являются ключевыми для оригинальной и мощной концепции Интернета. Они, к сожалению, потерялись на этом пути по мере развития сети.

Следовательно, Solid, основанный на существующих веб-стандартах и ​​результатах десятилетия работы, имеет функцию чтения и записи, включающую разрешения и идентификацию, а также возможность управления данными и обновления в реальном времени через websocket. Он реализует сеть, как первоначально предполагалось, и обеспечивает платформу для следующего поколения действительно расширяющих возможностей и инновационных приложений.

## 2.1 Представление концепции Solid

Solid (образованный из “social linked data”) представляет собой предлагаемый набор конвенций и инструментов для создания децентрализованных веб-приложений на основе принципов связанных данных. Solid является модульной и расширяемой системой. Он опирается на существующие стандарты и протоколы W3C настолько, на сколько возможно это реализовать.

В частности, Solid это:

* Технический стек - набор дополнительных стандартов и форматов данных/словарей, которые вместе предоставляют возможности, доступные в настоящее время только через централизованные службы социальных сетей (таких как Facebook/Twitter/LinkedIn и многие другие), такие как идентификация, аутентификация, авторизация и списки уровней доступа к системе, управление контактами, обмен сообщениями и уведомлениями, объединение каналов и подписок на информационные ресурсы, комментарии, форумы и многое другое.
* Документ спецификации, в котором описывается REST API, который расширяет существующие стандарты, содержит примечания по дизайну для отдельных используемых компонентов и предназначен для разработчиков, планирующих создавать серверы или приложения.
* Набор серверов, которые реализуют эту спецификацию и имеют в своей реализации все необходимое для корректной работы данной концепции.
* Экосистема социальных приложений, поставщиков удостоверений и вспомогательных библиотек (таких как solid-auth-client), которые работают на платформе Solid. Поддержка систем как со стороны разработчиков Solid, так и различных объединений сторонников данной концепции.
* Сообщество, предоставляющее документацию, форумы, учебники и официальные презентации основоположников концепции, готовые реализации на технологии NodeJS c использованием языка JavaScript.

## 2.2 Основные технологии и стандарты Solid

Выбор используемых стандартов в первую очередь основан на возможности распределённого хранения данных между пользователями без централизованного хранения. Спецификация Solid обязывает использовать следующие стандарты:

* RDF 1.1 (Resource Description Framework) – широко используется в моделях Solid Data. По умолчанию предпочтительным форматом сериализации RDF является Turtle. Документ Turtle представляет собой текстовое представление RDF графа. Turtle позволяет записывать RDF в компактной текстовой форме. RDF граф состоит из списка инструкций, состоящих из субъекта, предиката и объекта. Также, могут использоваться альтернативные форматы сериализации, такие как JSON-LD и RDFa. JSON-LD – специальный формат JSON для сериализации связанных данных. Синтаксис разработан, чтобы легко интегрироваться в существующте системы, которые уже используют JSON, и обеспечивает удобный переход от JSON до JSON-LD формата передаваемых данных..
* Стандарт WebID 1.0 (Web Identity and Discovery) используется для предоставления универсальных имен пользователей/идентификаторов для Solid приложений и для создания ссылки на уникальные ресурсы: людей, организации, устройства. Глобальная распределенная социальная сеть требует, чтобы каждый человек мог контролировать свою личность, чтобы эта личность могла быть привязана к сайтам – помещая каждого человека в специальную сеть взаимоотношений с возможностью аутентифицировать глобально с помощью идентификаторов. WebID описывает простой, универсальный механизм идентификации, который свободно распространяем, открыт к расширениям, улучшает конфиденциальность, безопасность и контроль над тем, как каждый человек может идентифицировать себя, чтобы обеспечить более точный контроль доступа к их информации в Интернете. Он делает это, применяя лучшие практики веб-архитектуры, основываясь на хорошо зарекомендовавших себя широко распространенных протоколах и стандартах, включая HTML, URI, HTTP и RDF Semantics.
* По стандарту WebID, при доступе получают документы профиля WebID Profile (в Turtle и других RDF форматах). WebID Profile – основа работы и использования WebID: его RDF график содержит краткое описание WebID пользователя, таким образом, что этот подграф формирует определенное описание WebID референта(тот, на кого идёт отссылка), то есть описание, которое отличает референт этого WebID от всех другие ресурсов в сети. Документ может, например, содержать отношения к другому документу, имеющему вид WebID референта, или он может связывать WebID с принципами, используемыми различными протоколами аутентификации.
* FOAF – это проект, нацеленный на связывание людей и информации в интернет сети. Словарь FOAF используется как в профилях WebID, так и при создании списков по контролю доступа. Независимо от того, находится ли информация в головах людей, в физических или цифровых документах, или в виде фактических данных, ее можно связать. FOAF объединяет три вида информационный сетей: социальные сети сотрудничества, дружбы и ассоциации; репрезентативные сети, которые описывают упрощенное представление о фактических терминах и информационные сети, которые используют привязку к интернету, для совместного использования независимо опубликованных описаний этого взаимосвязанного мира. FOAF не конкурирует с социально ориентированными веб-сайтами, он обеспечивает подход, при котором разные сайты могут разделять общие данные, при помощью которых пользователи могут сохранять контроль над своей информацией в непатентованном формате.
* Аутентификация (для входа в систему, персонализация страницы) выполняется по протоколу WebID-TLS. WebID-TLS расширяет профили WebID Profile, чтобы подставлять ссылки на открытые ключи субъекта в виде сертификатов X.509, используя словарь Cert Ontology 1.0. Последовательность аутентификации выполняется с использованием протокола HTTP по протоколу TLS. В отличие от обычных случаев использования HTTPS, WebID-TLS выполняется без ссылки на иерархии сертификатов и вместо этого поддерживает сертификаты, подписанные сервером (или самозаверяющие).
* В Solid создание сертификата обычно выполняется в браузере с помощью элемента HTML5 keygen, чтобы обеспечить одноэтапное создание и публикацию публикации сертификатов.
* Списки авторизации и доступа реализуются с помощью базовой онтологии контроля доступа. WebAccessControl – это децентрализованная система для предоставления различным пользователям и группам различных форм доступа к ресурсам, где пользователи и группы идентифицируются с помощью URI HTTP. Таким образом вы можете предоставлять доступ к документу на одном сайте пользователям и группам, размещенным на других сайтах. Пользователям не нужно иметь профиль на сайте, чтобы иметь доступ к документам на нем.
* Solid использует стандарт Linked Data Platform (LDP) в качестве стандартного способа чтения и записи общих ресурсов связанных данных. Платформа связанных данных (LDP) определяет набор правил для HTTP-операций на веб-ресурсах, некоторые из которых основаны на RDF, чтобы обеспечить архитектуру для чтения-записи связанных данных в Интернете.

## 2.3 Хранение личных данных и выбор системы управления

В Solid вы самостоятельно решаете, где хранить свои данные. Фотографии, которые вы берете, комментарии, которые вы пишете, контакты в своей адресной книге, события календаря, сколько километров вы пробегаете каждый день, от вашего фитнес-трекера ... все они хранятся в вашем Solid POD. Solid POD может находиться в вашем доме или на рабочем компьютере или онлайн-провайдере, имеющим реализацию Solid POD согласно вашему выбору. Поскольку вы владеете своими данными, вы можете перемещать, видоизменять и удвлять их в любое время без перерыва в обслуживании.

Вы даете людям и вашим приложениям право на чтение или запись части вашего Solid POD, поэтому всякий раз, когда вы открываете новое приложение, вам больше не нужно заполнять свои данные: они читаются с вашего POD предварительно попросив вашего разрешения. Вещи, сохраненные в одном приложении, доступны в другом: вам никогда не придется синхронизировать, потому что ваши данные остаются с вами.

Такой подход защищает вашу конфиденциальность, а также отлично подходит для разработчиков: они могут создавать классные приложения, не собирая вначале огромные объемы данных. Любой может создать приложение, которое использует то, что уже есть.

## 2.4 Solid POD как личное хранилище

POD – это как безопасные информационные накопители для Интернета, к которым вы можете получить доступ из любого места. Когда вы даете другим доступ к частям вашего POD, они могут реагировать на ваши фотографии и делиться своими воспоминаниями с вами. Вы решаете, какие приложения могут видеть приложения и люди.

Можно представить свой личный Solid POD как о частном веб-сайте пользователя сети, за исключением того, что ваши данные взаимодействуют со всеми вашими приложениями, а это значит, что у вас есть собственный персональный API, который согласуется с ним, умея предоставлять необходимые данные по строго заданным операциям. При публикации пользователем комментариев или видеороликов онлайн, друзья могут просматривать их с помощью любого приложения, которое предпочитает каждый из них, например, для просмотра альбома фотографий или для общения. Это ваши данные, которые могут быть сформированы любым способом или формой.

Существует возможность иметь столько версий POD, сколько необходимо для использования, и они размещаются на Web-серверах, поддерживающих Solid спецификацию. Устанавливая систему Solid Server на свой собственный сервер в вашем доме или на рабочем месте, пользователь получите Solid POD с самоличным контролем собственной размещённой на ней информации. Чтобы подтвердить право собственности на данные, пользователю нужен способ идентифицировать себя. Вместо того, чтобы полагаться на третью сторону, можно использовать Solid POD, чтобы предоставлять описание самого себя. Поэтому, вместо необходимости множество раз проходить операцию авторизации на различных ресурсах, достаточно однажды зайти в свой собственный Solid POD и в дальнейшем получать доступ к сторонним сайтам и сервисам без повторных операций регистраций.

# Глава 3 Результаты

Подводя итоги, можно сказать, что Solid – это новый этап развития межсетевых коммуникаций, развитие которого позволит создать более защищенные системы с распределённой системой хранения информации на просторах сети Интернет. Для её реализации необходимо владеть навыками и знаниями работы современной мировой сети и веб-приложений. Также необходимо изучить предложенные в спецификации типы данных и технологии, без которых соблюсти правила спецификации невозможно.

В качестве первоначального этапа разработки, мне будет необходимо ознакомиться с представленными в спецификации технологиями и подходами программирования. Наиболее важными технологиями являются Linked Data, RDF, FOAF, WebID, при помощи которых разрабатывает ядро Solid-ориентированных информационных систем.

Вторым этапом будет выбор технологии, на основе которой актуально разработать инструменты реализации спецификации. Последнее время набирает рост и становится более популярным технология .NET Core. Каждый год число программистов, предпочитающих данную технологию для реализации свои продуктов, растёт, сообщество разработчиков хорошо отзывается о скорости работы и производительности систем, написанных на её основе.

Третьим этапом является размещение моего решения в качестве свободно распространяемой библиотеки(модуля) на веб-ресурсах, специализированных на их распространении среди разработчиков. Наиболее популярным для размещения является www.nuget.org. Данный ресурс на сегодняшний момент хранит более 133 000 уникальных пакетов решений, более 1 400 000 версий размещённых на ней библиотек и имеет боле 13 млн. скачиваний.

Вклад в развитие Solid поможет сделать разработку приложений на основе спецификации проще и разнообразнее, а мировую сеть интернет более безопасной и свободной для её пользователей. От использования подобного подхода разработки станут безопаснее корпоративные и государственные сети, где необходимо высокий уровень безопасности.

# Заключение

Подводя итоги, можно сказать, что Solid – это новый этап развития межсетевых коммуникаций, развитие которого позволит создать более защищенные системы с распределённой системой хранения информации на просторах сети Интернет, корпоративных и государственных сетях.

На сегодняшний день существует единственный пример реализации приложения по спецификации Solid и передо мной стоит задача разобраться в спецификации и предложить собственные решения для других платформ и языков программирования.

Собственная реализация, отличных от имеющихся на сегодняшний день решений, позволит популяризировать изучаемую мной концепцию и упростить всем начинающим знакомство с ней на языковой платформе программирования, для которой будет решено создавать реализацию инструментов разработки систем с Solid ориентированным подходом.

При разработке, важно соблюдать принципы реализуемых технологий и изучать патентную базу во избежание патентных споров с правообладателями. Также я не имею права выдавать саму технологию за свою собственную, поскольку я лишь создаю своё решение на основе чужой концепции.

# Библиографический список

1. Berners-Lee, T. One Small Step for the Web... / T. Berners-Lee // Inrupt [Электронный ресурс] – 2018 – Режим доступа – https://solid.inrupt.com/how-it-works – Дата доступа : 11.11.2018.
2. Solid - Re-decentralizing the web (project directory) / T. Berners-Lee [и др] ; под общ. ред. T. Berners-Lee // Github [Электронный ресурс] – 2018 – Режим доступа – https://github.com/solid/solid – Дата доступа : 10.11.2018
3. The Solid spec and architecture / T. Berners-Lee [и др] ; под общ. ред. T. Berners-Lee // Github [Электронный ресурс] – 2018 – Режим доступа – https://github.com/solid/solid – Дата доступа : 10.11.2018
4. Berners-Lee, T. Linked Data / T. Berners-Lee // международная организация W3С [Электронный ресурс] – 2009 Режим доступа : https://www.w3.org/DesignIssues/LinkedData.html - Дата Доступа : 10.11.2018.

# Приложение А

# Приложение В