



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

Разбираемся вместе: рекомендации W3C по публикации данных в Сети

Вебинар | 15.02.2017



Алексей Ница

event-редактор
вебинары Теплицы

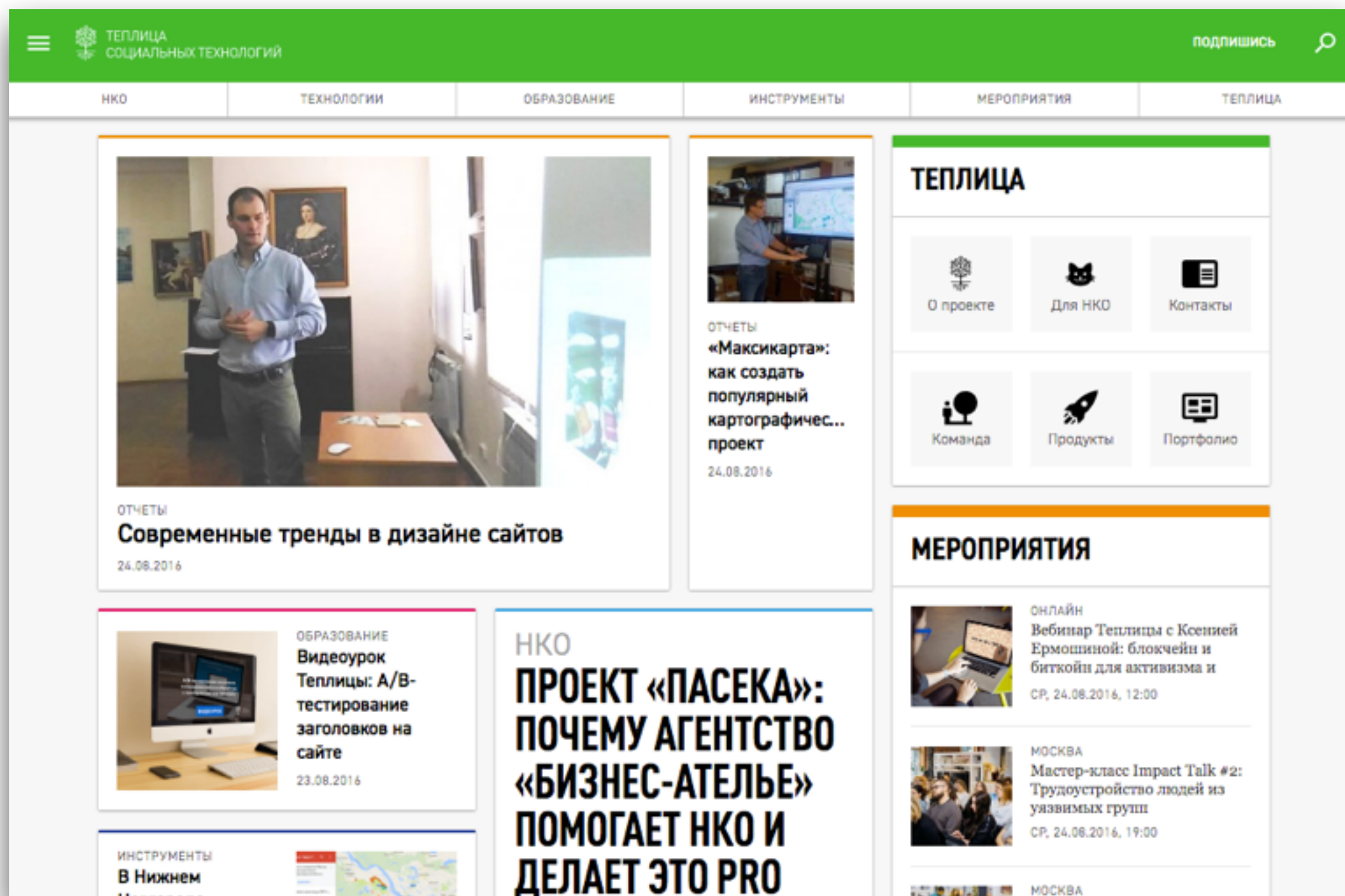
nitsa@te-st.ru

fb.com/mrnitsa



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

Сайт Теплицы



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

<https://te-st.ru>
социальные сети: /TeplitsaST

О чем поговорим сегодня?



{Открытые} данные

- Есть организация, которая занимается маршрутами наземного транспорта
- Собирает данные о движении, остановках, потоке людей, пробках и др.
- Что делать организации с этими данными? Как эффективнее данные использовать другим?

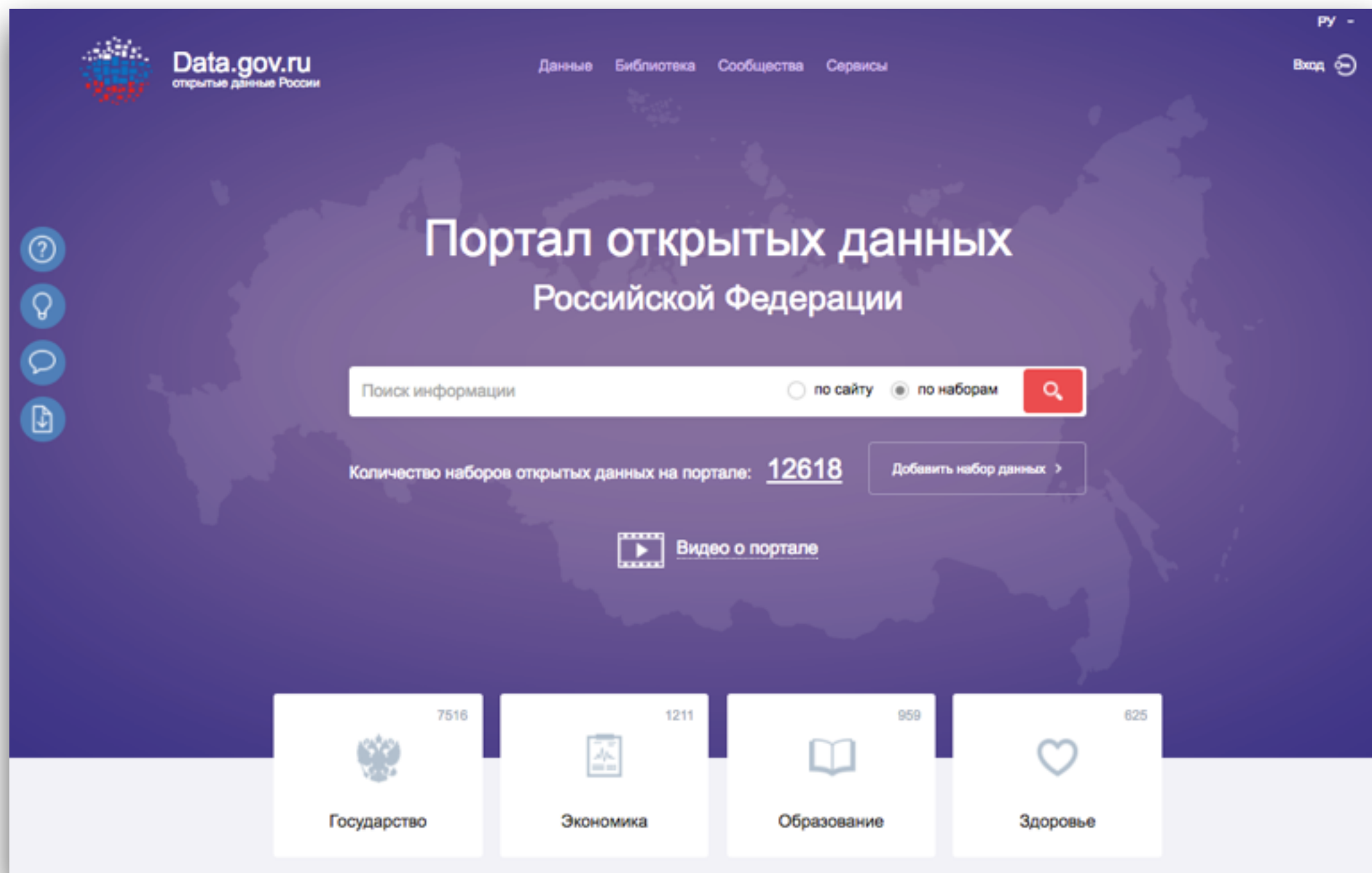
Открытые данные

англ. **open data** — концепция, отражающая идею о том, что определённые данные должны быть свободно доступны для машиночитаемого использования (например, в XLS-таблице или CSV-файле) и дальнейшей републикации без ограничений авторского права (например, с помощью свободной лицензии Creative Commons), патентов и других механизмов контроля.

Если набор данных не связан открытой лицензией, дающей права на свободное повторное использование, то такой набор данных не считается открытым, даже если он выложен в машиночитаемом виде.

Материалы сайта Теплицы по открытым данным.

Открытые данные России



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

data.gov.ru

В широком смысле – издатели данных стремятся обмениваться данными либо открыто, либо с контролируемым доступом.

Потребители данных хотят легко находить и использовать данные, особенно если они точны, регулярно обновляется и гарантированно доступны в любое время.

Это создает фундаментальную необходимость общего понимания между издателями данных и потребителями данных.

**Не всю информацию безопасно
публиковать как открытые данные:
персональные данные, IP-адреса и
др.**

Лучшая практика публикации данных в Интернете – использовать единые стандарты

Data on the Web Best Practices

W3C Recommendation

Data on the Web Best Practices

W3C Recommendation 31 January 2017



This version:
<https://www.w3.org/TR/2017/REC-dwbp-20170131/>

Latest published version:
<https://www.w3.org/TR/dwbp/>

Latest editor's draft:
<http://w3c.github.io/dwbp/bp.html>

Implementation report:
<http://w3c.github.io/dwbp/dwbp-implementation-report.html>

Previous version:
<https://www.w3.org/TR/2016/PR-dwbp-20161215/>

Editors:
Bernadette Farias Lóscio, [CIn - UFPE, Brazil](#)
Caroline Burle, [NIC.br, Brazil](#)
Newton Calegari, [NIC.br, Brazil](#)

Contributors:
Annette Greiner
Antoine Isaac
Carlos Iglesias
Carlos Laufer
Christophe Guéret
Deirdre Lee
Doug Schepers
Eric G. Stephan



Немного о W3C



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

Data on the Web Best Practices

The screenshot shows the W3C website with a blue header containing the W3C logo. Navigation links include 'STANDARDS', 'PARTICIPATE', 'MEMBERSHIP', and 'ABOUT W3C'. A search bar is located on the right. The main content area features a news article titled 'W3C Welcomes IDPF as Organizations Officially Combine to Develop Roadmap for Future of Publishing' dated 1 February 2017. The article includes a 'PUBLISHING@W3C' logo and text about the collaboration between W3C and IDPF. A sidebar on the left lists categories like 'TECHNICAL REPORTS', 'WEB AND INDUSTRY', and 'WEB FOR ALL'. The right sidebar contains a 'W3C BLOG' section with recent posts.

Views: desktop mobile print

W3C By Region Go

STANDARDS PARTICIPATE MEMBERSHIP ABOUT W3C

TECHNICAL REPORTS

By date

By group

WEB AND INDUSTRY

Automotive

Publishing

Entertainment (TV and Broadcasting)

Web and Telecommunications

Web of Data

Web of Things

Web Payments

Web Security

WEB FOR ALL

W3C Welcomes IDPF as Organizations Officially Combine to Develop Roadmap for Future of Publishing

1 February 2017 | [Archive](#)

PUBLISHING@W3C W3C announced on 1 February 2017 that the World Wide Web Consortium (W3C) and the International Digital Publishing Forum (IDPF) have combined organizations to better align Publishing and Web technologies and to create a new roadmap for the future of publishing.

"W3C is thrilled to gain the expertise of the publishing industry with its rich tradition of excellence in developing many forms of content for books, magazines, journals, educational materials and scholarly publications," said Dr. Jeff Jaffe, W3C CEO. "Working together, Publishing@W3C will bring exciting new capabilities and features to the future of publishing, authoring and reading using Web technologies."

More than 75 organizations provided [EPUB 3.1 Member Submission](#) as a Member Submission to ensure that the EPUB 3.1 standard, developed by IDPF, remains royalty-free. To maintain EPUB and advance its further adoption, W3C is setting up a separate EPUB (3.1) Community Group that is free and open to anyone to participate.

A new Publishing Business Group will be the focal point for the community to address new needs and requirements and to serve as a forum for industry discussions. The [Publishing@W3C](#) roadmap includes plans to charter new

The World Wide Web Consortium (W3C) is an international community that develops open [standards](#) to ensure the long-term growth of the Web. W3C operates under our [Code of Ethics and Professional Conduct](#). [Become a Friend of W3C](#) to support the [W3C mission](#) and free developer tools.

W3C BLOG

W3C expresses concerns on visa suspension that may hurt our worldwide collaboration
7 February 2017 by [Coralie Mercier](#)

Data on the Web? Here's How
31 January 2017 by [Phil Archer](#)

DPUB-ARIA 1.0 is Released as a Candidate Recommendation
15 December 2016 by [Tzviya Siegman](#)

JOBS



W3C

Консорциум Всемирной паутины (от англ. World Wide Web Consortium) – организация, разрабатывающая единые технологические стандарты для Интернета, которые затем внедряются производителями программ и оборудования.

Так достигается совместимость между программными продуктами и аппаратурой различных компаний, что делает Интернет более совершенным.

Рекомендации W3C

- рекомендации открыты – не защищены патентами и могут внедряться любым человеком без всяких финансовых отчислений консорциуму
- W3C не имеет программ сертификации и не планирует их вводить, поэтому рекомендации W3C получили большее распространение
- из-за отсутствия сертификации многие производители следуют рекомендациям лишь частично
- некоторые популярные рекомендации имеют несколько степеней внедрения

Процесс стандартизации W3C

Любой стандарт W3C проходит 5 стадий согласования:

1. **Черновик спецификации** (англ. Draft)
2. **Рабочий проект** (англ. Working Draft)
3. **Последний созыв** (англ. Last Call)
4. **Возможная рекомендация** (англ. Candidate Recommendation);
5. **Предлагаемая рекомендация** (англ. Proposed Recommendation)

Только после этого официально становится рекомендацией W3C

Процесс стандартизации W3C

Любой стандарт W3C проходит 5 стадий согласования:

1. **Черновик спецификации** (англ. Draft)
2. **Рабочий проект** (англ. Working Draft)
3. **Последний созыв** (англ. Last Call)
4. **Возможная рекомендация** (англ. Candidate Recommendation);
5. **Предлагаемая рекомендация** (англ. Proposed Recommendation)

Только после этого официально становится рекомендацией W3C

Структура W3C

Общую администрацию консорциума осуществляют 3 организации:

- Массачусетский технологический институт в США;
- Европейский консорциум по исследованиям в области информатики и математики во Франции;
- Университет Кейо в Японии.

Членом консорциума может стать юридическое или частное лицо, занимающееся веб-технологиями и заинтересованное в развитии Интернета. Основным источником финансирования консорциума – членские взносы. Членами консорциума уже являются более 350 организаций.

Успехи W3C

Консорциум был создан в 1994 году. Первым крупным успехом консорциума стала стандартизация языка гипертекстовой разметки **HTML** (англ. HyperText Markup Language) в 1996 году.

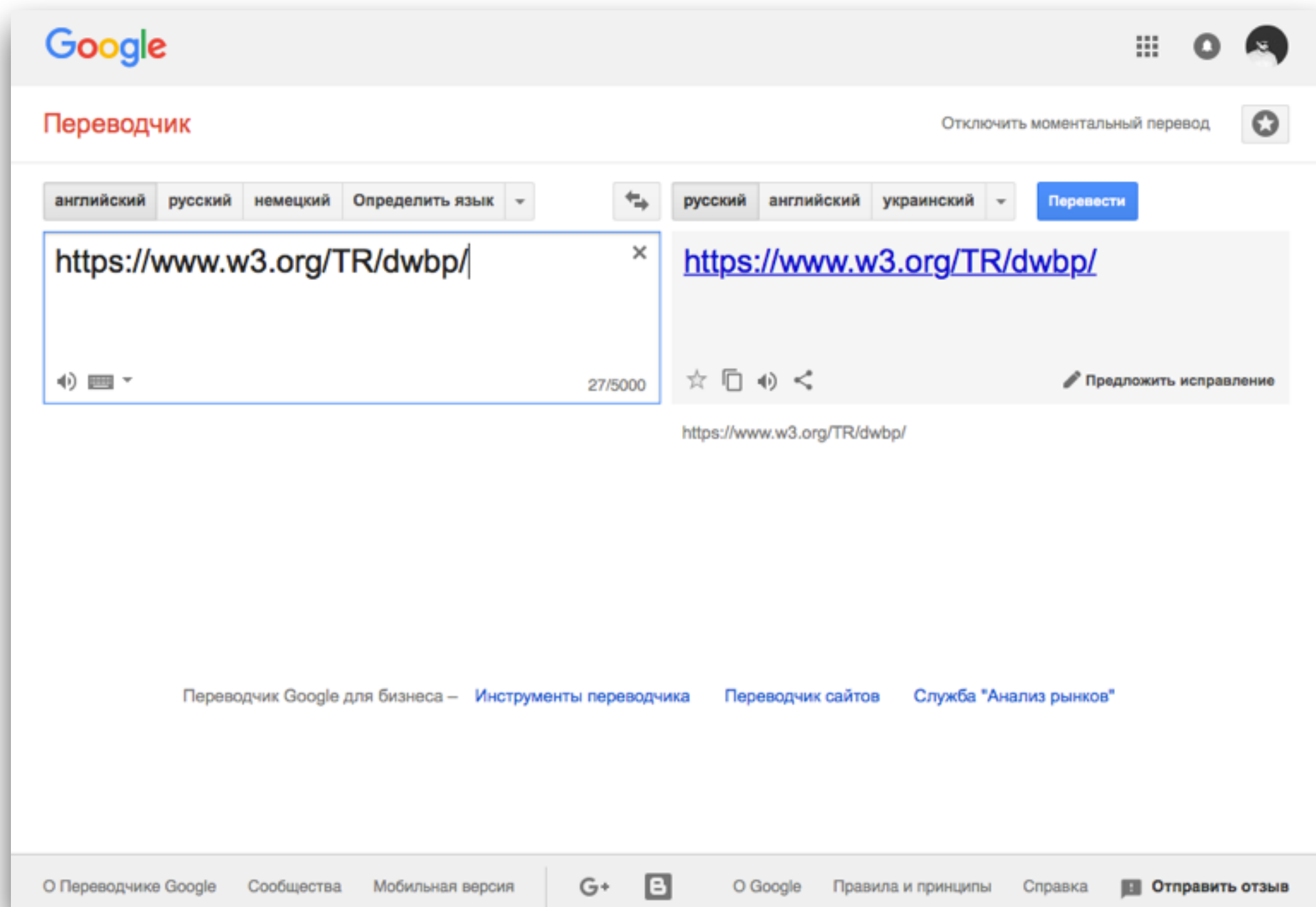
В середине 1990-х годов ряд крупнейших производителей программного обеспечения планировал выпустить каждый свою версию языка HTML со своими названиями тегов. В результате веб-страница одной компании была бы размечена совершенно не так, как страница другой компании. Из-за этого веб-браузер одной компании не мог бы отображать страницы, созданные по правилам другой компании.

Полного совпадения тегов и атрибутов HTML достичь не удалось до сих пор.

Переходим к рекомендациям W3C по публикации данных в Сети



Google Переводчик в помощь



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

translate.google.ru

Google Переводчик в помощь

The screenshot shows a Google Translate interface. The address bar displays the URL <https://www.w3.org/TR/dwbp/>. The language settings are set to "С языка: английский" (From language: English) and "На: русский" (To: Russian). The "Показать:" (Show:) dropdown is set to "Перевод" (Translation). The translated text is displayed in Russian, with the title "Данные о Web Best Practices" and the subtitle "Рекомендация W3C 31 января 2017". The W3C logo is visible in the top right corner. A vertical blue bar on the left side of the page contains the text "W3C Recommendation". The translated text includes links to the latest version, previous versions, and authors of the document.

Переводчик С языка: английский На: русский Показать: Перевод Оригинал

Данные о Web Best Practices

Рекомендация W3C 31 января 2017

Эта версия:
<https://www.w3.org/TR/2017/REC-dwbp-20170131/>

Последняя опубликованная версия:
<https://www.w3.org/TR/dwbp/>

Проект последнюю версию редактора:
<http://w3c.github.io/dwbp/bp.html>

Доклад об осуществлении мер:
<http://w3c.github.io/dwbp/dwbp-implementation-report.html>

Предыдущая версия:
<https://www.w3.org/TR/2016/PR-dwbp-20161215/>

редакторы:
Бернадетт Фариас Lôscio, [CIn - UFPE, Бразилия](#)
Кэролайн Бурлев, [NIC.br, Бразилия](#)
[Ньютон Calegari](#), [NIC.br, Бразилия](#)

Авторы:
Аннет Greiner
Антуан Исаак
Карлос Ислас



**Этот документ содержит Рекомендации,
связанные с публикацией и использованием
данных в Интернете.**

**Данные должны быть обнаруживаемы и
понятны как для людей, так и машин.**

**Эти Рекомендации помогут способствовать
взаимодействию между издателями и
потребителями данных.**

**Рекомендации содержат
12 основных разделов
и 3 дополнительных блока**



Всего предлагается

35

**рекомендаций
(лучших практик)**



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

35 лучших практик

- [Best Practice 1](#): Provide metadata
- [Best Practice 2](#): Provide descriptive metadata
- [Best Practice 3](#): Provide structural metadata
- [Best Practice 4](#): Provide data license information
- [Best Practice 5](#): Provide data provenance information
- [Best Practice 6](#): Provide data quality information
- [Best Practice 7](#): Provide a version indicator
- [Best Practice 8](#): Provide version history
- [Best Practice 9](#): Use persistent URIs as identifiers of datasets
- [Best Practice 10](#): Use persistent URIs as identifiers within datasets
- [Best Practice 11](#): Assign URIs to dataset versions and series
- [Best Practice 12](#): Use machine-readable standardized data formats
- [Best Practice 13](#): Use locale-neutral data representations
- [Best Practice 14](#): Provide data in multiple formats
- [Best Practice 15](#): Reuse vocabularies, preferably standardized ones
- [Best Practice 16](#): Choose the right formalization level
- [Best Practice 17](#): Provide bulk download
- [Best Practice 18](#): Provide Subsets for Large Datasets
- [Best Practice 19](#): Use content negotiation for serving data available in multiple formats
- [Best Practice 20](#): Provide real-time access
- [Best Practice 21](#): Provide data up to date
- [Best Practice 22](#): Provide an explanation for data that is not available
- [Best Practice 23](#): Make data available through an API
- [Best Practice 24](#): Use Web Standards as the foundation of APIs
- [Best Practice 25](#): Provide complete documentation for your API
- [Best Practice 26](#): Avoid Breaking Changes to Your API
- [Best Practice 27](#): Preserve identifiers
- [Best Practice 28](#): Assess dataset coverage
- [Best Practice 29](#): Gather feedback from data consumers
- [Best Practice 30](#): Make feedback available
- [Best Practice 31](#): Enrich data by generating new data
- [Best Practice 32](#): Provide Complementary Presentations
- [Best Practice 33](#): Provide Feedback to the Original Publisher
- [Best Practice 34](#): Follow Licensing Terms
- [Best Practice 35](#): Cite the Original Publication



Каждая из рекомендаций имеет преимущества (часть 1)



- **Reuse** – данные будут повторно использоваться потребителями
- **Comprehension** – люди смогут лучше понимать данные и их значение
- **Linkability** – можно будет создавать связи между ресурсами данных
- **Discoverability** – возможность проще найти сами наборы данных или информацию в наборе данных

Каждая из рекомендаций имеет преимущества (часть 2)



- **Trust** – уверенность в том, что доверие потребителей к данным усилится
- **Access** – люди и машины будут иметь доступ к современным данным в различных форматах
- **Interoperability** – станет легче достичь договоренности между издателями и потребителями данных
- **Processability** – машины будут иметь возможность автоматически обрабатывать данные

35 рекомендаций

Лучшая практика 1 : Предоставление метаданных

Лучшая практика 2 : Предоставление описательных метаданных

Лучшая практика 3 : Обеспечить структурные метаданные

Лучшая практика 4 : Предоставление сведений о лицензии данных

Лучшая практика 5 : Предоставление информации о происхождении данных

Лучшая практика 6 : Предоставление информации о качестве данных

Лучшая практика 7 : Обеспечить индикатор версии

Лучшая практика 8 : Обеспечить историю версий

Лучшая практика 9 : Используйте постоянные идентификаторы URI в качестве идентификаторов наборов данных

Лучшая практика 10 : Используйте постоянные идентификаторы URI в качестве идентификаторов в пределах наборов данных



35 рекомендаций

Лучшая практика 11 : Назначение идентификаторы URI для набора данных версий и серии

Лучшая практика 12 : Использование машиночитаемых стандартизованных форматов данных

Лучшая практика 13 : Использование локально-нейтральных представлений данных

Лучшая практика 14 : Представление данных в различных форматах

Лучшая практика 15 : Повторное использование словарей, предпочтительно стандартизированных

Лучшая практика 16 : Выбор правильного уровня формализации

Лучшая практика 17 : Обеспечение объемной загрузки

Лучшая практика 18 : Обеспечение подмножеств для больших объемов данных

Лучшая практика 19 : Использование согласования содержания для обслуживания данных , имеющих в различных форматах

Лучшая практика 20 : Обеспечение доступа в режиме реального времени



35 рекомендаций

Лучшая практика 21 : Представление данных в актуальном состоянии

Лучшая практика 22 : Обеспечение объяснения данных, которые не доступно

Лучшая практика 23 : Открытие доступа к данным через API

Лучшая практика 24 : Использование веб-стандартов как основы API

Лучшая практика 25 : Обеспечение полной документации для API

Лучшая практика 26 : Не создавать критические изменения в своем API

Лучшая практика 27 : Сохранение идентификаторов данных

Лучшая практика 28 : Оценка охвата набора данных

Лучшая практика 29 : Сбор обратной связи от потребителей данных

Лучшая практика 30 : Создание канала получения обратной связи

35 рекомендаций

Лучшая практика 31 : Пополнение данных путем генерирования новых данных

Лучшая практика 32 : Обеспечение дополнительных презентаций

Лучшая практика 33 : Обеспечение обратной связи с первоисточником данных

Лучшая практика 34 : Следование условиям лицензирования

Лучшая практика 35 : Цитирование оригинальной публикации

Перейдем к рекомендациям

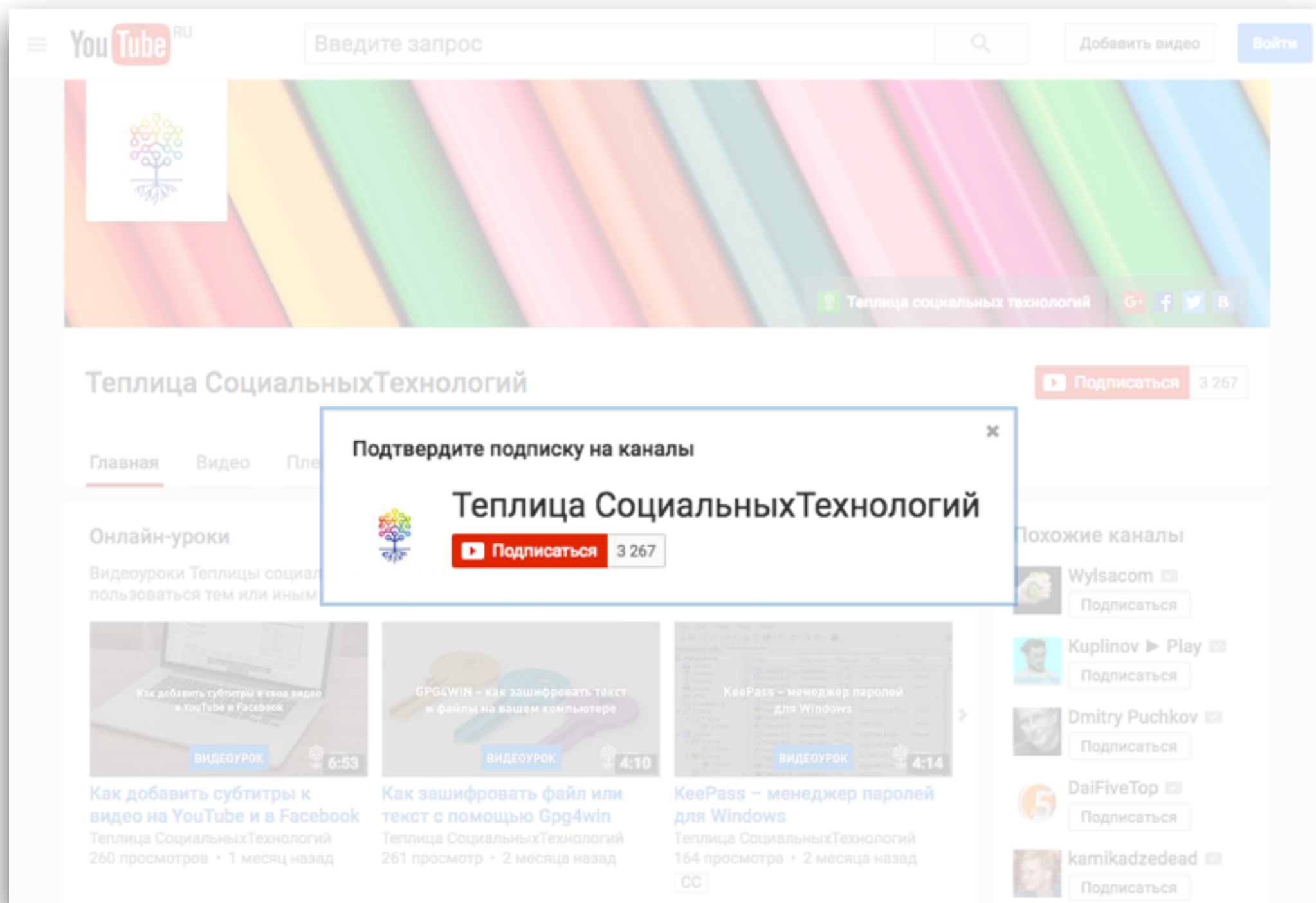


Вопросы и ответы



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

Канал Теплицы на YouTube



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru

[Подписаться на канал Теплицы](#)

Записи вебинаров Теплицы



Программа ИнфоДонор и Google Grants для продвижения сайта НКО

Вебинар 1 27.12.2016

Вебинары Теплицы

Теплица Социальных Технологий • 55 видео • 1 830 просмотров • Обновлено 28 дек. 2016 г.

Записи вебинаров Теплицы социальных технологий <https://te-st.ru>

▶ Воспроизвести все < Отправить + Сохранить

1	 Программа ИнфоДонор и Google Grants для продвижения сайта НКО (вебинар #55) @TeplitsaST Теплица Социальных Технологий	1:04:01
2	 Quora и TheQuestion: как и зачем использовать и как создать свою платформу (вебинар #54) @TeplitsaST Теплица Социальных Технологий	57:40
3	 VR – виртуальная реальность для социальных проектов (вебинар #53) @TeplitsaST Теплица Социальных Технологий	1:30:55
4	 GitHub для новичков и непрограммистов (вебинар #52) @teplitsast Теплица Социальных Технологий	59:14
5	 Таргетинг в Facebook для НКО (вебинар #51) @teplitsast Теплица Социальных Технологий	1:16:18



Видеоуроки Теплицы



Онлайн-уроки

Теплица Социальных Технологий • 121 видео • 33 642 просмотра • Обновлено 2 дня назад

Видеоуроки Теплицы социальных технологий - это небольшие видео, цель которых - показать, как пользоваться тем или иным онлайн-сервисом или программой для повышения эффективности вашей работы ещё

▶ Воспроизвести все ➦ Отправить ✓ Сохранено

1	 Как отслеживать ссылки с помощью UTM меток Теплица Социальных Технологий4:09
2	 Видео на YouTube: как добавить ссылку на ваш сайт в видео и создавать иерархию роликов Теплица Социальных Технологий4:49
3	 Как обезопасить личную переписку с помощью виджета Mailvelope Теплица Социальных Технологий3:03
4	 Как с помощью Readymag сделать красивый сайт Теплица Социальных Технологий4:49
5	 Как установить и настроить WordPress на стандартном хостинге Теплица Социальных Технологий4:50



Спасибо!



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru



ТЕПЛИЦА
СОЦИАЛЬНЫХ ТЕХНОЛОГИЙ
te-st.ru