

Датамайнинг RU-IGF

Получение датасетов и структурированных данных по материалам русскоязычного форума IGF

Российский форум по управлению интернетом

- Ежегодное событие, на котором обсуждаются вопросы governance русскоязычного Интернета
- Форум проходит с 2010 года
- Хороший источник для датамайнинга
- О чем говорили и кто влияет на повестку

Структура Датасета 1

- Выступления
- Спикеры
- Транскрипции выступлений
- Summaries выступлений
- Sentiment Analysis выступлений

	youtube_id	persons	title	summary	year	video_transcript
0	248b1oXAvB0	[Максут Шадаев, Александр Хинштейн, Татьяна Ма...	Церемония открытия	6 Всероссийский молодежный конкурс работ по п...	2021	друзья доброе утро здравствуйте дорогие участн...
1	aJfXUKudzwU	[Рашид Исмаилов, Мэнди Карвер, Вольфганг Клейн...	Секция 1. Управление интернетом	В июле 2018 года Генеральный секретарь ООН объ...	2021	российский форум по управлению интернетом и се...
2	pdTQahZV5TQ	[Эльза Ганеева, Андрей Игнатьев, Презентация, ...	Секция 2. Новые технологии. Искусственный интеле...	В соответствии с «Концепцией развития регулиро...	2021	секция называется новые технологии искусственн...
3	q4R98A4Eh84	[Милош Вагнер, Наталья Великородняя, Николай Д...	Секция 3. Суверенитет данных	Суверенитет данных – это идея о том, что данны...	2021	доброе утро уважаемые участники 11 российского...
4	XQZTCxwWlVE	[]	Вручение награды Virtuti Interneti и традицион...	В 2010 г. Координационный центр доменов .RU/.P...	2021	дорогие друзья мы переходим следующему этапу 1...

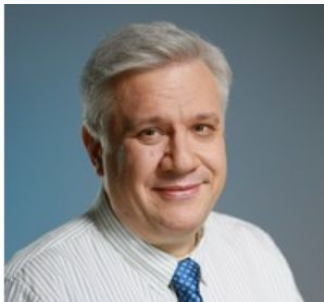
Структура Датасета 2

- 111 выступлений
- 171 докладчик
- 29 докладов с транскрипциями
- [DRAFT] Для некоторых выступлений отсутствуют транскрипции

Персоны оказывающие влияние

Андрей Ярных

Лаборатория Касперского



Георгий Грицай

Минсвязь / Минпромторг



Олег Демидов

МГИМО / РАЗК



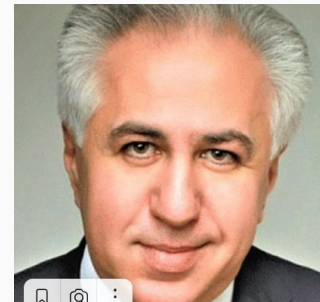
Руслан Гатаров

Единая Россия



Рашид Исмаилов

Вымпелком



О Чем говорили в 2021

управление интернет искусственный интеллект точка зрения цифровой платформа очень важно персональный данные форум управление первый очередь пользовательский соглашение российский федерация собственно говорить очень важный социальный сеть хотеть сказать интеллектуальный собственность цифровой экономика друг друг право человек год назад сеть интернет очень интересный информационный безопасность искусство интеллект весь участник добрый день прошлый год совет европа весь мир молодёжный форум сегодняшний день персональный дать очень часто правило игра заинтересованный лицо мочь сказать огромный количество суверенитет данные передать слово большой большой координационный центр предоставить слово 20 год добрый утро летний школа гражданский общество иметь право речь идти цепочка поставка международный право национальный домен

В сегодняшнем обзоре самых интересных и обсуждаемых технологий в мире мы поговорим о том, что нужно сделать для того, чтобы общество стало более прозрачным и успешным.

О Чем говорили в 2019

управление интернет точка зрения право человек искусственный интеллект форум управление персональный данные 25 год развитие интернет год назад социальный сеть российский федерация регулирование интернет прошлый год очень важно домен ru первый очередь хотеть сказать уважаемый коллега совет европа использование данные защита данные национальный домен друг друг гражданский общество ваш вопрос очень важный очень интересно высокий уровень речь идти 10 год российский форум очень интересный координационный центр наш страна организация объединить объединить нация весь мир глобальный сеть наш жизнь jedi pr защита персональный персональный дать большой данные наш уважаемый сеть интернет федеральный агентство международный сообщество наш форум российский интернет крупный компания

В современном мире все больше и больше людей хотят знать о том, что происходит в их жизни.

О Чем говорили в 2017

музыка управление интернет точка зрения интернет вещь электронный правительство совет европа развитие интернет друг друг заинтересованный сторона цифровой экономика год назад форум управление сеть интернет право человек хотеть сказать российский федерация доменный имя музыка аплодисменты информационный технология очень важный очень важно данный случай соединить штат прошлый год сей пора весь мир вопрос управление глобальный форум речь идти рабочий группа первый очередь собственно говорить наш страна российский форум аплодисменты музыка очень интересный республика татарстан большой большой сегодняшний день очень очень гражданский общество аплодисменты аплодисменты доступ интернет большой количество новый технология круглый стол очень часто любой случай государство должный евразийский повестка

В современном мире интернет-технологий очень много и очень много, и очень много интересного, и очень мало того, что можно сказать о нём.

О Чем говорили в 2014

совет европа управление интернет российский федерация парламентский ассамблея год назад персональный данные 20 год добрый утро наш форум первый очередь совет федерация ассамблея совет форум управление координационный центр дама господин очень большой домен ru наш открытие хотеть предоставить предоставить слово андрей колесников хотеть сказать домен . очень важно офис совет сеть интернет заниматься интернет соединить штат российский форум 25 год точка ру наш ребёнок утро уважаемый действительно очень большой праздник очень полезный интернет самый самый главное 20-летие точка сегодня хотеть . ребёнок детский интернет наш сообщество хотеть представить продолжать работать информационный офис конец прошлое прошлое год оказаться сегодня весь поздравить

В Москве отмечают 20-летие международного сотрудничества между Россией и Евросоюзом.

Заключение

- Получен датасет выступлений Russian IGF
- Испробованы доступные методы обработки текста
- Выявлены ключевые decision makers в русскоязычном IGF