

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 Великородная, 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...

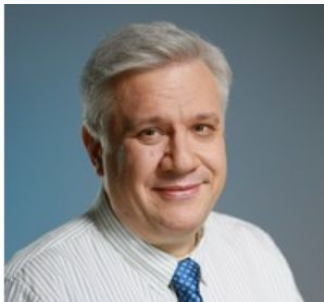
Структура Датасета 2

- 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



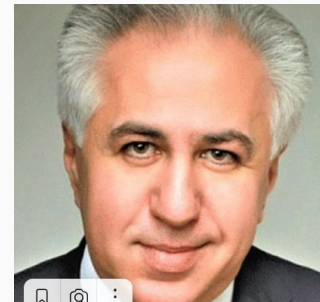
Руслан Гатаров

Единая Россия



Рашид Исмаилов

Вымпелком



О Чем говорили в 2021

вопрос интернет мочь наш весь говорить свой большой какой-то человек цифровой государство год спасибо платформа должный ещё хотеть компания данные страна право сегодня поэтому сказать управление дело работа первый форум российский собственно сторона дать решение принцип международный просто знать мир правило слово нужно действительно проблема работать наверное россия общий также регулирование интеллект пользователь бизнес развитие возможность уровень система например разный являться технология думать самый именно время понимать участник новый иметь доверие организация точка важно информация сделать день глобальный число достаточно видеть искусственный сеть делать коллега важный тема больший безопасность заниматься зрение процесс соответственно общество вообще хороший

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

это интернет который очень вопрос всё человек наш данные мочь год сегодня весь должный говорить дать свой какой-то большой страна знать российский форум компания ещё спасибо управление хотеть государство право сказать пользователь первый регулирование россия международный развитие являться время глобальный поэтому дело самый разный сторона нужно сеть ваш например правило думать просто мир жизнь новый технология решение закон понимать работа организация национальный действительно защита цифровой уровень давать коллега важно будущее казаться общество больший любой процесс домен принцип дискуссия точка персональный день друг информация существовать безопасность возможность общий хороший проблема часть важный использовать наверное европа вместе именно вещь участник рынок зрение

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

О Чем говорили в 2017

это который интернет всё очень вопрос мочь год говорить музыка какой-то человек наш весь большой
должный правительство сказать ещё государство управление новый страна свой сегодня хотеть сторона
работать развитие электронный значит просто поэтому знать самый первый спасибо проблема далее дело
являться сделать стандарт россия российский вещь работа форум время нужно сеть общий организация
технология делать например возможность видеть уровень думать понимать точка разный данные право
безопасность собственно действительно наверное решение компания тема пожалуйста глобальный стать
некий достаточно именно слово дать система каждый принцип больший идти сообщество случай мир
обсуждать образ друг казаться технический вообще кстати иметь информация экономика

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 2016

это который интернет всё [] музыка вопрос мочь весь говорить очень наш свой сказать год ещё какой-то человек управление собственно принцип хотеть должный инфраструктура работа самый большой спасибо институт сеть например новый страна развитие сторона просто уровень россия точка российский работать значит организация правительство государство разный сегодня форум европа проблема связь далее соответственно домен знать время каждый система дело сообщество являться хороший оператор поэтому думать национальный действительно процесс иметь важный общий первый зрение именно закон также достаточно технический общество возможность момент документ понимать безопасность мир право угроза стать участник европейский структура пожалуйста рамка нужно происходить подход ваш группа критический

В сегодняшнем обзоре новых и самых интересных новостей, которые пройдут в России в ближайшее время, мы попытаемся выяснить, что происходит в стране в настоящее время и в будущем.

О Чем говорили в 2014

это который интернет всё [] музыка вопрос мочь весь говорить очень наш свой сказать год ещё какой-то человек управление собственно принцип хотеть должный инфраструктура работа самый большой спасибо институт сеть например новый страна развитие сторона просто уровень россия точка российский работать значит организация правительство государство разный сегодня форум европа проблема связь далее соответственно домен знать время каждый система дело сообщество являться хороший оператор поэтому думать национальный действительно процесс иметь важный общий первый зрение именно закон также достаточно технический общество возможность момент документ понимать безопасность мир право угроза стать участник европейский структура пожалуйста рамка нужно происходить подход ваш группа критический

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