# Приложение к статье «Высокоуровневая семантическая интерпретация структуры статических моделей для русского языка»

Сериков Олег Алексеевич, Школа Лингвистики НИУ ВШЭ, МФТИ, Институт искусственного интеллекта AIRI, Лаборатория исследования и сохранения малых языков ИЯЗ РАН, srkvoa@gmail.com

Ганеева Вероника Александровна, НИУ ВШЭ, vaganeeva@edu.hse.ru

Аксенова Анна Александровна, ПАО "Сбербанк", aaaksenova2@gmail.com

Клышинский Эдуард Станиславович, НИУ ВШЭ, eklyshinsky@hse.ru

В таблицах ниже приведены списки слов, автоматически полученные методом, предложенным в статье. Слова представляют собой первые 30 слов, получивших минимальное или максимальное значение по полученной оси. Для соблюдения полноты картины, написание слов сохранено. Название разделам давалось исходя из анализа полного списка полученных слов.

## Модель ruscorpora\_upos\_cbow\_300\_20\_2019

Схема разделения модели на основе НКРЯ

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Абстрактное,  виртуальные референты | | | | Конкретное,  референты в физ. мире | | | |
| Духовное, внутренний мир человека, литература | | Индустрия, прикладные науки | | Архаичное | | Современное | |
| Духовное | Общественное | Развлечения | Наука | Книжное | Повседневное | Книжное | Повседневное |

Результаты разделения на этапе 1

|  |  |
| --- | --- |
| Максимальные значения по оси SVD | Минимальные значения по оси SVD |
| развитие, деятельность, отношение, условие, система, задача, организация, действие, процесс, правительство, значение, изменение, государство, исследование, вопрос, решение, влияние, период, возможность, закон | ладонь, стол, шея, изба, шапка, рубаха, спина, комната, палец, щека, колено, грудь, губа, голова, платок, глаз, рука, крыльцо, дверь, нога, окно |

Результаты разделения на этапе 2

|  |  |  |  |
| --- | --- | --- | --- |
| Ось 1 | | Ось 2 | |
| Максимальные значения по оси SVD | Минимальные значения по оси SVD | Максимальные значения по оси SVD | Минимальные значения по оси SVD |
| вера, истина, убеждение, чувство, любовь, добродетель, мысль, христианство, народ, жизнь, стремление, невежество, идеал, зло, поступок, страдание, воззрение, разум, религия, отрицание, страсть, презрение, отечество, заблуждение, эгоизм, порок, предрассудок, ненависть, величие, нравственность | топливо, транспорт, аэропорт, офис, скважина, стоимость, заявка, комбайн, аппаратура, температура, мощность, доставка, маршрут, тонна, доллар, зона, автомобиль, вертолет, компания, база, завод, самолет, участок, фирма, ремонт, станция, километр, район, метр, оборудование | козак, боярин, воевода, верста, стрелец, слобода, казак, купец, лях, государев, острог, барин, изба, поселянин, пан, деревня, разбойник, ямщик, лошадь, конь, холоп, ратник, село, воевод, кабак, слуга, управитель, староста, милость, божий | холл, блокнот, сантиметр, колготки, комбинезон, футболка, майка, сумочка, шприц, галстук, витрина, очки, флакон, тумбочка, лампочка, холодильник, свитер, плитка, блузка, пепельница, столик, сигарета, кнопка, папка, шкаф, резинка, джинсы, обложка, коробка, фотография |

Результаты разделения на этапе 3

|  |  |  |
| --- | --- | --- |
| Абстрактное | | |
| Духовное, внутренний мир человека | Духовное | сей, сие, добродетель, токмо, чувствование, наставление, прилежание, поелику, житие, кой, оный, слог, похвала, любовь, премудрость, благость, дар, изъяснение, совершенство, воображение, опытность, благодеяние, живость, естество, предмет, дарование, доверенность, несчастие, ученость, дружество |
| Общественное | лозунг, сторонник, конфликт, пропаганда, капитализм, фашизм, война, диктатура, кризис, социализм, элита, националист, империализм, коалиция, оппозиция, репрессия, реформа, страна, партия, восстание, контрреволюция, коммунист, политика, большевик, буржуазия, правительство, борьба, террор, демократия, революция |
| Прикладное | Развлечение | фестиваль, шоу, чемпионат, фильм, чемпион, компьютер, экран, дизайнер, турнир, гонщик, хит, сериал, рекорд, сайт, матч, тренер, телевидение, болельщик, пилот, кино, журналист, мюзикл, студия, продюсер, ролик, интервью, финиш, презентация, сборная, бренд |
| Производство | введение, хозяйство, доход, ведение, распоряжение, увеличение, принятие, предотвращение, перевозка, подвоз, усиление, платеж, ликвидация, постановление, насаждение, внесение, ссуда, скот, расход, уменьшение, пользование, продовольствие, улучшение, предоставление, вывоз, обеспечение, уплата, выдача, имущество, снабжение |
| Конкретное | | |
| Архаичное | Книжное | сердце, взор, чертог, душа, меч, темница, око, уста, богиня, мрак, небо, печаль, горесть, небеса, врата, пламень, веселие, бездна, лик, трепет, ангел, благоухание, пустыня, воин, трон, ужас, призрак, чудовище, гнев, воинство |
| Повседневное | починка, трактир, огород, картуз, девчата, писарь, лавка, сельсовет, фельдшер, артельщик, прииск, водка, овес, воз, лошадь, пуд, полушубок, мужик, староста, дворник, сено, извозчик, копейка, подвод, харчи, приказчик, контора, провизия, дрова, подвода |
| Современное | Книжное | волос, румянец, кудри, лицо, локон, платье, глаз, личико, бровь, бородка, чепец, ресница, улыбка, седина, морщина, губа, грация, бледность, фрак, складка, бакенбарда, усы, щека, кружево, цвет, талия, парик, пушок, платьице, кружева |
| Повседневное | фургон, электричка, метро, ангар, самосвал, камера, площадка, мусор, вездеход, кран, прицеп, вокзал, бункер, вагончик, ларек, лифт, поезд, багажник, рация, ящик, мотор, кабина, автомат, люк, будка, вагон, холодильник, автобус, гараж, грузовик |

## Модель araneum\_upos\_skipgram\_300\_2\_2018

Схема разделения модели на основе Araneum Russicum

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Люди, социальное регулирование | | | | Организация, технология,  правовые отношения | | | |
| Абстрактное | | Конкретное | | Технологии производства | | Управление и производство | |
| Оскорбления | Духовный мир | Диминутивы, имена | Природа | Химия и биология | Физика, автоматизация | Экономика | Политика и общество |

Результаты разделения на этапе 1

|  |  |
| --- | --- |
| Max SVD | Min SVD |
| пашка, малец, сашка, гришка, шурка, санька, женька, гриша, валерка, ванька, колька, старуха, эдик, борька, васька, прохвост, машка, митя, володька, поганец, братец, катька, ленька, митька, плут, генка, горбун, танька, мерзавец, дружок | котирование, поставка, пакет, планир, неиспользование, длительность, долгосрочность, авизование, обновление, минимизация, учет, периодичность, объем, формирование, видеотерминал, обеспечь, фотопродукция, закупка, телерадиопрограмма, бизнес-сектор, приоритизация, безопасность, видеорежим, доступ, стоимость, доукомплектация, поступление, тэгу, экспресс-груз, экспресс-обследование |

Результаты разделения на этапе 2

|  |  |  |  |
| --- | --- | --- | --- |
| иррационализм, догматизм, обскурантизм, отрицание, релятивизм, индивидуализм, рационализм, аморализм, мистицизм, интеллектуализм, безнравственность, агностицизм, спиритуализм, морализм, отступничество, пантеизм, односторонность, критицизм, этноцентризм, материализм, конформизм, идеализм, ортодоксия, изоляционизм, ригоризм, клерикализм, аскетизм, деизм, греховность, универсализм | аленушка, санька, дубок, анжела, андрюшка, девчата, рябинка, регин, данил, нюся, есения, мишина, олешек, сережа, фунтик, светланка, артем, синичка, машенька, тимоша, ира, машка, валек, пчелка, чижик, марийка, снежанна, маринка, маша, любава | смачивание, перемешивание, нагрев, вспенивание, натяжение, изгибание, прилипание, разбрызгивание, встряхивание, прижатие, смешивание, обжатие, склеивание, шлифование, термоэлемент, стекание, выглаживание, эмульгирование, засаливание, вдавливание, сминание, засасывание, всасывание, пенообразование, пропаривание, сваривание, слипание, вибрирование, отмачивание, скольжение | саратовэнерго, башавтотранс, бишкек, темиртау, ханты, караганда, октябрь, октябрьск, ставрополье, белогорск, сибакадемстрой, зеленодольск, краснодар, директор, армавир, жигулевск, пенза, гидрострой, братск, астана, худжанд, ржд, омскнефтехимпроект, тольятти, ноябрьск, череповец, югорск, волгодонск, филиал, кировский |

Результаты разделения на этапе 3

|  |  |  |
| --- | --- | --- |
| Человек, социальное | | |
| Абстрактное | Сниженная оценочная лексика | подонок, жулик, отморозок, недоумок, прихлебатель, уголовник, подлец, лизоблюд, прощелыга, сброд, выродок, голодранец, дегенерат, кровопийца, бездарь, шантажист, кретин, садист, извращенец, идиот, политикан, майданщик, буржуй, пустобрех, тупица, саботажник, наемник, шпион, нытик, насильник |
| Духовный мир | эмпиризм, книжность, позитивизм, христианство, теодицея, рационализм, деизм, брахманизм, буддизм, индуизм, вероучение, гегельянство, бытие, антропософия, теизм, иконография, структурализм, нераздельность, атомизм, умозрение, историзм, богочеловечество, изобразительность, монизм, антропологизм, предметность, трансцендентализм, исихазм, герменевтика, платонизм |
| Конкретное | Диминутивы, антропонимы | люся, близняшка, милочка, мамуля, девчушка, подруга, рэйчел, анджела, алисий, катя, девченка, доченька, леночка, милая, лиля, мадам, сестричка, сашенька, девчонка, эшли, мама, кэт, братик, дашка, докторша, памела, мамзель, катенька, стажерка, сабрина |
| Описание ландшафта  географические термины и топонимы | сельцо, анюй, полуверста, кубенское, ореховка, городец, кологрив, новоселка, камыш, устье, заимка, сулак, мелколесье, дубовка, матвеевка, слобода, воробьевка, вознесенское, копань, увал, верховье, березовка, вязовка, плавни, мезеконь, никольское, низовье, пойма, суходол, переволока |
| Организация, технология, правовые отношения | | |
| Технологии производства | Химия и биология | метаболит, инозин, цитохром, аутоантитело, соматостатин, резистентность, лиганд, этанол, кумуляция, секретин, изофермент, протромбин, флюоресценция, акриламид, биорегулятор, ингибирование, тромбопластин, макромолекула, орнитин, тромбин, радиоизотоп, фагоцитоз, гадолиний, канцерогенность, аденозинтрифосфат, архебактерия, автолиз, радионуклид, монослой, технеций |
| Физика, автоматизация | прочистка, пневмоподушка, стартер, выпиливание, лоток, редуктор, талреп, сервопривод, глушитель, пантограф, чехол, щиток, торцовка, поручень, склиз, гидроподъемник, электрошнур, унитаз, люлька, завертка, привод, люверс, ремень, коробка, перекидка, суппорт, доводчик, прикуриватель, подставка, крепеж |
| Управление и индустрии | Экономика | облкоммунэнерго, башкирэнерго, ростовэнерго, горэлектросеть, татэнерго, пермэнерго, салаватнефтеоргсинтез, самаранефтегаз, транснефть, желдорреммаш, химпром, мосводоканал, дальполиметалл, волгограднефтемаш, энергосбыт, юганскнефтегаз, кировэнерго, тюменьэнерго, тепловодоснабжение, удмуртнефть, кубаньэнерго, камгэсэнергострой, строммашина, смоленсктеплосеть, техперевооружение, промгаз, тепловик, комиэнерго, волгоградэнерго, славнефть |
| Политическая география и общество | конвент, софтбол, дананг, яхтинг, фиджи, поклонник, футбол, мастер-класс, кинопрограмма, приглашение, цюрих, дискотека, сиэтл, викторин, лусака, кампус, регби, монако, коллоквиум, сан-пауло, инсбрук, тайланд, марракеш, джулианс, монте-карло, кинопросмотр, ток-шоу, турне, фестиваль, мальта |

## Модель, обученная на научных статьях

Схема разделения модели на основе научных статей

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Более повседневный дискурс | | | | Более научный дискурс | | | |
| абстрактное | | предметное | | Научные термины | | места и люди | |
| Общество и политика | Литература и духовный мир | Объекты- специальные термины | Бытовые предметы | Внешнее научное | Внутреннее научное | Фамилии и имена учёных | Места и организации |

Результаты разделения на этапе 1

|  |  |
| --- | --- |
| Max SVD | Min SVD |
| мочь, образ, место, вода, значение, лицо, день, сцена, ребёнок, действие, женщина, количество, комната, девушка, желание, линия, площадь, смерть, группа, эпизод, сын, ночь, гора, лес, рис, возможность, жена, двор, слой, храм | музыкознание, прагматик, дейк, рецепция, диахронии, реферирование, социолингвистика, кемерово, симпозиум, новосибирск, востоковедение, информ, лексикология, доцент, семиотика, методы, религиоведение, мгу, типология, семинара, формирования, болотнов, тюркология, рггу, инион, вопр, дис, языкознание, вып, спб |

Результаты разделения на этапе 2

|  |  |  |  |
| --- | --- | --- | --- |
| отношение, власть, интерес, идея, деятельность, позиция, поведение, убеждение, жизнь, признание, мысль, закон, характер, идеал, государство, смысл, норма, отказ, доверие, мочь, требование, осуждение, принцип, влияние, угроза, действие, политик, степень, противостояние, вера | бляшка, улитка, галька, лунка, рулон, коврик, палочка, кусочек, солонка, поднос, серьга, ящик, подол, войлок, доска, палатка, веточка, кувшин, трубка, яма, бугор, прялка, ручка, сумка, ямка, сноп, ножка, дощечка, фигурка, жердь | подход, формирование, дискурс, функционирование, мышление, коммуникация, синтез, взаимосвязь, интеграция, актуальность, лингвистика, адаптация, становление, синкретизм, идентификация, усложнение, адекватность, идентичность, разработка, развёртывание, постижение, раскрытие, рациональность, размывание, рецепция, детерминация, кодирование, предпосылка, дискретность, неоднозначность | вилайет, петропавловск, глазго, маадыр, осло, георги, танзания, летие, онон, огайо, абакан, анадырь, калевала, зао, фио, сургут, тоо, йов, тегеран, элиста, сидней, аран, проректор, берген, гтрк, краевед, анкара, сост, владикавказ, дата |

Результаты разделения на этапе 3

|  |  |  |
| --- | --- | --- |
| Более повседневный дискурс | | |
| Абстрактное | Общество и политика | предприятие, правительство, документ, комитет, управление, должность, работник, постановление, льгота, руководство, требование, расход, заявка, работа, санкция, выполнение, кредит, доход, регион, заведение, заём, бизнес, издание, инициатива, депутат, количество, милиция, министр, получение, заседание |
| Литература и духовный мир | творец, добродетель, пророк, хаос, восхищение, романтик, гений, фантазия, прелесть, доброта, благородство, меланхолия, жизнь, идеал, творчество, счастие, героиня, кротость, символ, творение, истина, переживание, смирение, смерть, плоть, мир, христос, природа, мудрость, любовь |
| Конкретное | Объекты-  специальные термины | поверхность, ширина, венчик, сечение, рельефе, скол, ареал, купол, складками, высота, конус, стенка, пояс, возвышенность, рис, туловище, карниз, долина, плита, полость, спираль, терраса, бугорок, вырез, ямка, перемычка, холм, полоска, многоугольник, медальон |
| Бытовые предметы | чемодан, бутылка, лавка, тракторист, трактор, офис, контора, рыбалка, мальчишка, табак, дом, дрова, мыло, завтрак, машина, вагон, холодильник, колбаса, самогон, картошка, автомобиль, девчонка, конфета, поезд, мороженое, фабрика, водка, вечер, шофер, билет |
| Более научный дискурс | | |
| Научные термины | Внешнее научное | изучение, наука, изыскание, совершенствование, специальность, освоение, ознакомление, воспитание, языкознание, информатика, апробация, диссертация, пересмотр, социология, менеджмент, литературоведение, культурология, психолингвистика, методология, источниковедение, журналистика, наследие, география, философия, автоматизация, методист, историография, модернизация, становление, познание |
| Внутреннее научное | разнородность, аналитичность, стереотипность, субъективность, предшествование, обязательность, разорванность, асимметрия, локальность, монолитность, официальность, однотипность, стандартность, дробность, предельность, многомерность, прерывность, вариативность, мозаичность, соотнесённость, непрерывность, предсказуемость, аномальность, связность, фрагментарность, гомогенность, иерархичность, спонтанность, связанность, линейность |
| Места и люди | Фамилии и имена учёных | дэвид, эдвард, стивен, уолтер, морган, стив, оуэн, адамс, джонатан, бернард, говард, грей, рид, герберт, эрнест, дональд, джей, бертран, уоллес, гордон, винсент, брук, дороти, перри, гарольд, ганди, камерон, киплинг, бестселлер, цру |
| Места и организации | сср, очерки, череповец, аспирант, башкортостан, сургут, новокузнецк, мчс, удмуртия, бийск, ооо, бюллетень, брянск, филиал, оренбург, чувашия, летию, вологда, сост, симферополь, тюмень, сыктывкар, белгород, тамбов, екатеринбург, омск, красноярск, чебоксары, махачкала, краснодар |