



**МАЛЬЦЕВА
ДАРЬЯ ВАСИЛЬЕВНА**

к.с.н., заведующий лабораторией,
Международная лаборатория
прикладного сетевого анализа,
НИУ «Высшая школа экономики»



**ВАЩЕНКО
ВАСИЛИСА АНДРЕЕВНА**

стажер-исследователь,
Международная лаборатория
прикладного сетевого анализа,
НИУ «Высшая школа экономики»



**ЩЕГЛОВА
ТАМАРА ЕВГЕНЬЕВНА**

аспирант, младший научный
сотрудник, Международная
лаборатория прикладного сетевого
анализа, НИУ «Высшая школа
экономики»



**МОИСЕЕВ
СТАНИСЛАВ ПАВЛОВИЧ**

к.с.н., младший научный сотрудник,
Международная лаборатория
прикладного сетевого анализа,
НИУ «Высшая школа экономики»;
партнер, лидер практики
«Исследования и консалтинг»,
ООО Авентика

ЗДОРОВЫЕ ГОРОДА: ВЫДЕЛЕНИЕ АКТУАЛЬНЫХ ТРЕНДОВ ИССЛЕДОВАНИЙ В НАУЧНОЙ ЛИТЕРАТУРЕ И СОЦИАЛЬНЫХ МЕДИА

Авторский коллектив благодарит за помощь в подготовке материала коллег:



**ЧЕРНОГО
ВАСИЛИЯ АНАТОЛЬЕВИЧА**

директора по стратегическим
коммуникациям Brand Analytics



**ВЕБЕР
ВАЛЕРИЮ ВЯЧЕСЛАВОВНУ**

ведущего маркетолога
Brand Analytics

Резюме

Эта статья открывает рубрику аналитических обзоров предметных областей, становящихся темами для очередных выпусков журнала. Начать мы решили с темы городского здравоохранения.

Для обзора использовали методологию сетевого анализа библиометрических источников: перебрали 5597 статей из Web of Science, найденных по ключевым словам, и 2179 статей журнала Journal of Urban Health. Добавили анализ публичного дискурса по здоровым городам с помощью анализа публикаций в русскоязычных социальных медиа, выполненного с помощью системы Brand Analytics.

Рост интереса к исследованиям городского здоровья начался в 1970-х, тогда же вышли самые важные тексты по вопросу (см. списки в таблицах 1 и 2).

Самыми популярными из тем последнего года в русскоязычных соцмедиа стали мусорные полигоны (2,24 млн упоминаний с сентября 2021 года по август 2022 года) и семейное насилие (2,03 млн упоминаний). Следом идут спортивные площадки (1,42 млн упоминаний) и безбарьерная среда (1,41 млн упоминаний). Дальше идут экологические привычки горожан (909 тыс. упоминаний) и высотная застройка (772 тыс. упоминаний).

Тематика изучения городского здоровья подвижна и фокусирует внимание на актуальных проблемах: опасные заболевания, доступность медицины и маргиналы.

Введение

Поиск работ для аналитических обзоров — нетривиальная задача: велика вероятность сделать неверные выводы либо пропустить важный текст (Моисеев, Мальцева, 2018). Поэтому мы используем алгоритмический метод и библиометрический анализ: через них мы выделяем значимые работы и авторов, исходя из их цитирования, и делаем выводы на основе анализа библиографических сетей (цитирования, социтирования, библиографического сочленения, соприсутствия различных библиографических единиц.

Иной способ обзора трендов — анализ отражения тематик в публичном дискурсе через социальные медиа. Здесь работают алгоритмы количественного текстового и сетевого анализа. Тематики и тренды формируются общественностью: частые тематики и связи между ними показывают всплески внимания и позволяют понять причины этих всплесков.

Сегодня в нашем фокусе тематика здоровых городов и городского здравоохранения. Тренды и тематики мы выявляем через анализ данных научных публикаций и социальных медиа. База данных научных публикаций состоит из двух массивов, индексированных в WoS. Массив собран по характерным для тематики ключевым словам и массиву публикаций Journal of Urban Health (JoUH). Массив публикаций соцмедиа для анализа представляет собой посвященные городскому здравоохранению публичные сообщения русскоязычных пользователей социальных медиа, собранные и проанализированные с помощью системы Brand Analytics.

Методология

Сетевой анализ применяется для анализа публикаций из Web of Science и социальных медиа.

В рамках этого подхода сеть цитирования — это граф: его вершины есть публикации, а ребра — отношения цитирования между работами. Мы делаем подсчет базовых метрик и более продвинутые методы анализа: выделение основных путей (наиболее плотно связанных друг с другом цепочек публикаций).

Выделение путей показывает цепочки значимых работ в виде графа во времени. Их мы выделяем алгоритмом Search path count (SPC) (Batagelj et al., 2014)¹. Он вычисляет

¹ Для проведения анализа сеть должна быть ациклична (не должна содержать «циклы», или последовательно связанные ребра, где один из узлов является и началом, и концом цепочки) и не содержать «петли» (ссылки на саму себя), иметь только одну связь между парами узлов, вес которой равен единице. Сеть с циклами трансформируется посредством алгоритма Preprint Transformation.

индикатор веса проходов (traversal weights (Batagelj et al., 2014)) для каждой конкретной связи. К сети с рассчитанными весами применяются алгоритмы выделения основного или ключевых путей².

Если определить тему выделенных цитирований, мы увидим путь развития тематики либо предметной области в соответствующем контексте.

Также мы использовали двумодальную сеть работ и ключевых слов³: построили сеть совстречаемости ключевых слов в описании одной статьи через ее перемножение. В анализе сетей ключевых слов мы вели расчет степени входящей центральности и выделение подгрупп в сети с помощью подхода островов. Он помог выявить локально важные группы узлов.

Анализ данных библиометрического сетевого анализа мы делали в программе Pajek для анализа и визуализации больших сетей (De Nooy, Mrvar, Batagelj, 2018) и в программе R (R Core Team, 2020)⁴.

Выделяя тематики обсуждений в социальных медиа, мы применяли тематическое моделирование. Latent dirichlet allocation (LDA) — это подход, используемый в тематическом моделировании на основе вероятностных векторов слов, которые указывают на их релевантность текстовому корпусу (David et al., 2003).

Для поиска сетевой структуры обсуждений используется сетевой подход к анализу текстов с созданием «ментальной модели» — сети отношений между понятиями, встречающимися в текстах (Carley, 2020). Анализ данных социальных медиа выполнен с помощью языка программирования Python (пакеты nltk (NLTK Project, 2022) и networkx (NetworkX, 2022)), визуализация сети — в программе для визуализации сетевых данных Gephi (Bastian, Heymann et al., 2009).

База данных

Библиометрические данные

Мы искали публикации по ключевым словам в WoS, в отдельных журналах (например, Hummon, Carley, 1993) и по коллекциям публикаций конкретных тематик⁵. Использовано две стратегии сбора данных.

² При применении алгоритма основного пути (Main path) для каждой публикации «верхнего уровня» (имеющей ссылки на другие работы, но не цитируемой другими) на основе последовательного выбора вершин сети по определенному правилу конструируется цепочка, ведущая к публикациям «нижнего уровня» (цитируемым другими работами, но не имеющим цитирований). Цепочка с максимальным значением показателя отбирается в качестве основного пути. При применении алгоритма поиска ключевых путей (Key Routes) в сети выделяется не один, а несколько возможных путей — через увеличение количества включенных в основной путь связей.

³ В этой сети узлы одной части сети (работы) связаны с узлами другой части сети (ключевые слова), но не связаны друг с другом; вес каждой связи равен единице.

⁴ GitHub: <https://github.com/Daria-Maltseva/Sociodigger/wiki/UrbanHealth>.

⁵ Именно эта база использовалась для сбора в связи с тем, что входящие в нее работы имеют полные библиографические описания, включающие списки цитируемых источников (поле CR в описании), наличие которых является важным для построения сети цитирований. Входящая в WoS база отечественных публикаций, формируемая РИНЦ (RSCI, Russian Science Citation Index) такой информации не содержит, поэтому ее нельзя было использовать в качестве источника информации.

- ⇒ Первый массив получен с помощью поисковых запросов по ключевым словам на английском языке: «urban health» — 5508 статей, «health-saving environment*» — 4 статьи, «barrier-free environment*» — 85 статей; общее количество статей в массиве составило 5597 статей⁶.
- ⇒ Второй массив получен сбором всех статей флагманского для темы JoUH. Общее количество статей в массиве — 2178⁷.

Массивы переведены в сетевой формат с помощью программы WoS 2Pajek. Получая «на входе» файл формата .txt, программа выдает файлы с сетевыми данными о цитировании между работами и о связях между публикациями и авторами, журналами и ключевыми словами в формате .net. Сети цитирований по этим массивам обозначены как **CiteSQ** и **CiteJ**. Еще на выдаче формируются два файла формата .clu: один делит массив по годам выхода публикаций, другой делит на публикации с полным описанием (хиты) и цитируемые работы. Количество хитов равняется количеству работ, найденных при поиске в базе данных (за вычетом дублей) и составляет 5590 и 2177 публикаций в двух сетях соответственно (при этом второй массив полностью входит в состав первого), количество цитируемых ими работ составляет 144 152 и 54 126 соответственно (с учетом чистки данных от анонимных работ).

Данные социальных медиа

Эти данные собраны и проанализированы с помощью системы Brand Analytics (<https://br-analytics.ru>). Экспертно отобраны шесть тем о здоровье города: безбарьерная среда, мусорные полигоны, семейное насилие, высотная застройка, спортивные площадки и эко-привычки горожан. Предметом более конкретного анализа стала безбарьерная городская среда.

0 Результаты

Результаты анализа библиометрических данных

Анализ цитирований

Ниже — распределение количества публикаций по годам выхода и двум анализируемым массивам для хитов (рис. 1). В первом случае распределение приведено с 1990 года, во втором — с 1950 года. Рисунок 1 показывает, что количество публикаций по нашим запросам возрастает с 2005 года — от 150 статей до 500 в год. Количество публикаций в JoUH держится на одном уровне. Рост числа релевантных публикаций наблюдается и по графику распределения всех публикаций (хитов и цитируемых ими работ, рис. 2). В массиве по журналу пик публикаций приходится на 2000-е годы, а в массиве по поисковому запросу видно плавное увеличение числа публикаций, достигающее максимума в 2013—2014 годах.

⁶ Для поиска рассматривались также такие ключевые слова, как «public health», «mental health» и «medical prevention*», однако в связи с их чрезмерно широким смыслом, выходящим за рамки городского здоровья, их использование без добавления уточняющих слов привело бы к слишком большому количеству статей в выдаче результатов и, что более важно, к смещению результатов анализа на более широкую тему. Использование этих ключевых слов вместе с указанными более узкими ключевыми словами привело к 1920 статьям в выдаче, все из которых входили в полученную выборку из 5597 статей.

⁷ Журнал выпускается Springer Science+Business Media и New York Academy of Medicine; его история ведется с 1851 года, а импакт-фактор в 2021 году составляет 5,801.

Рис. 1. Количество публикаций по годам в сети CiteSQ (Urban Health) и CiteJ (JoUH) – хиты

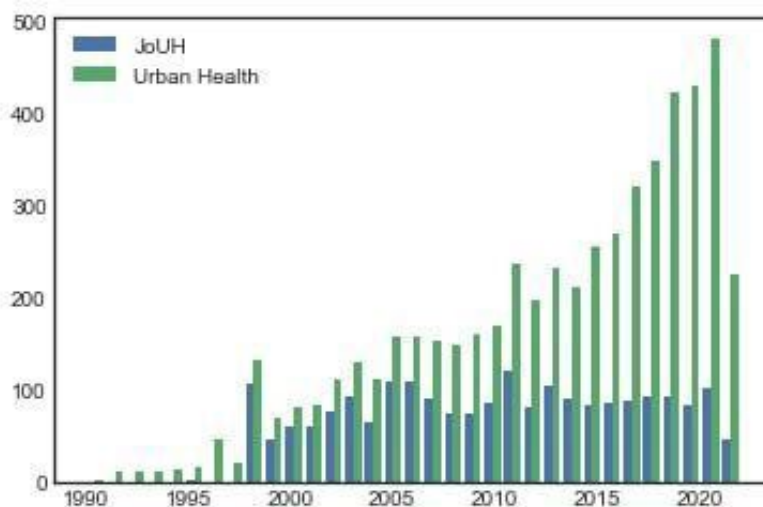
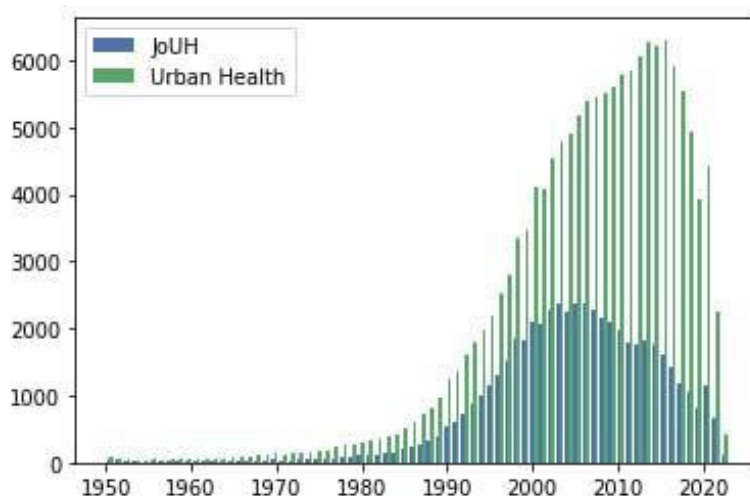


Рис. 2. Количество публикаций по годам в сети CiteSQ (Urban Health) и CiteJ (JoUH) – все работы



В таблице 1 (приложение 3) приведены топ-50 наиболее цитируемых работ в сети цитирований **CiteSQ**⁸. Максимально цитируемая статья — Sampson, Raudenbush, & Earls в журнале Science 1997 года о сплоченности в соседских сообществах и умень-

⁸ В список топ-50 наиболее цитируемых публикаций попали работы формата «WORLD_H(2022):», которые сложно идентифицировать в связи с тем, что используемое название может обозначать разные работы, например: World Bank, 2022, LIFE EXPECTANCY BIRT, World Bank, 2022, MORTALITY RATE UNDER, World Bank, 2022, HOSP BEDS PER 1000 P, World Bank, 2022, MATERNAL MORTALITY R. В связи с таким форматом программа WoS 2Pajek считывает первое слово в названии организации как фамилию автора, выбирает из второго слова первый знак как инициал его имени, указывает в скобках год публикации и не указывает после никаких данных, так как информации о выпуске журнала и странице публикации в этих описаниях нет. В связи с тем, что под короткое описание может подходить несколько публикаций, показатель его цитирования равен сумме цитирований разных работ. Более корректно исключить такие работы из списков (что было сделано в обоих массивах).

шении насилия. Три следующие работы имеют почти равное количество цитирований (72, 71 и 70).

Большинство работ с высокими показателями цитирований имеют программный характер. Это статьи о связи здоровья с урбанизацией и жизнью в городах в целом (18, 33), соседскими общинами и сообществами (10, 13, 21, 40) и социальными условиями как причинами заболеваемости (23, 41), публичными пространствами (50), зелеными насаждениями (34), объектами физической активности (46), распространением ВИЧ (45), жизнью в трущобах (44), об улучшении условий жизни в городах для снижения заболеваемости и различий в состоянии здоровья (8, 17, 30), о планировании здоровой городской среды (9, 38), в том числе в развивающихся странах (11, 39). Кое-где фокус ставится на американских проблемах: политика апартеида (36), городское планирование (ведущее к упадку многих кварталов) (35).

Итоги анализа по сети **CiteJ** приведены в таблице 2 (приложение 3). Больше всего цитируется работа Neckathorn 1997 года с описанием выборки, движимой респондентами, для изучения скрытых популяций. Следующие публикации топа цитирований — о сплоченности в городских сообществах и насилии, шкале депрессивности и расовой сегрегации по месту жительства. 21 из 47 представленных в таблице 2 работ присутствует также в таблице 1 приложения 3 (пересекающиеся в двух таблицах статьи выделены серым цветом).

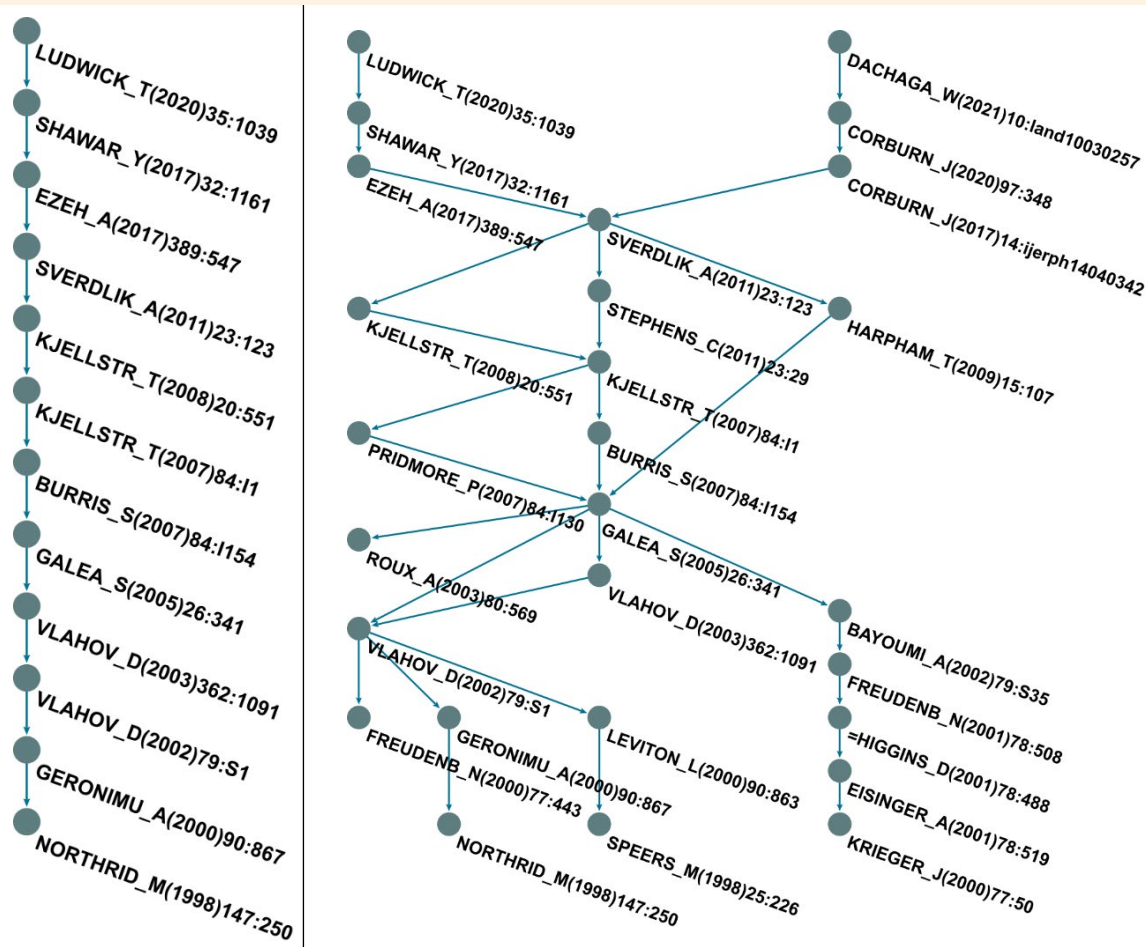
На рисунке 3 мы видим путь в сети **CiteSQ**, состоящий из 12 работ. По 3 работы опубликованы в *JoUN* и *Health Policy and Planning*, по 2 работы — в *Environment and Urbanization* и *Lancet*. Первая по времени работа (Ludwick) 1998 года посвящена влиянию курения на смертность жителей Гарлема. Новая статья того же автора (2000) — программная. Опираясь на эмпирические исследования, автор говорит о важности изучения структурных детерминант состояния здоровья и смертности в гетто. Три следующие работы авторов Vlahov и Galea (2002, 2003, 2005) описывают городское здравоохранение как самостоятельную дисциплину, опирающуюся на экологические, эпидемиологические и социологические наработки. Детерминанты здоровья описываются в контексте социальной и физической сред, доступа к медицинским и социальным услугам. Следующие статьи посвящены стратегиям управления здоровым городом (2007) и достижению справедливости в отношении здоровья в городах (2007, 2008). Работы основного пути формируют городское здравоохранение как самостоятельную область исследований.

Тематику основного пути можно дополнить через анализ работ, входящих в ключевые пути (таблица 1 в приложении 1, где работы, входящие в основной путь, отмечены кодом 1, а публикации из ключевых путей — кодом 2). Это 27 работ, из которых 12 включены в основной путь. Ранние работы (1998—2000) — о предотвращении болезней и пропаганде ЗОЖ. Публикации 2001—2002 годов — об исследовательских центрах на базе городских сообществ. Работы средней части рисунка — о связке городского здоровья с социальным неравенством, проживании в трущобах и влиянии медицинского обслуживания для решения этих проблем.

Основной путь сети **CiteJ** из 8 работ в журнале *JoUN* показан на рисунке 4 (слева). Он начинается с работ 2001 года о влиянии пенитенциарных учреждений на здоровье

населения. Три следующие работы (2001, 2005 и 2006) — о вопросах реинтеграции выходцев из ИУ в городские сообщества и их влиянии на общественное здоровье.

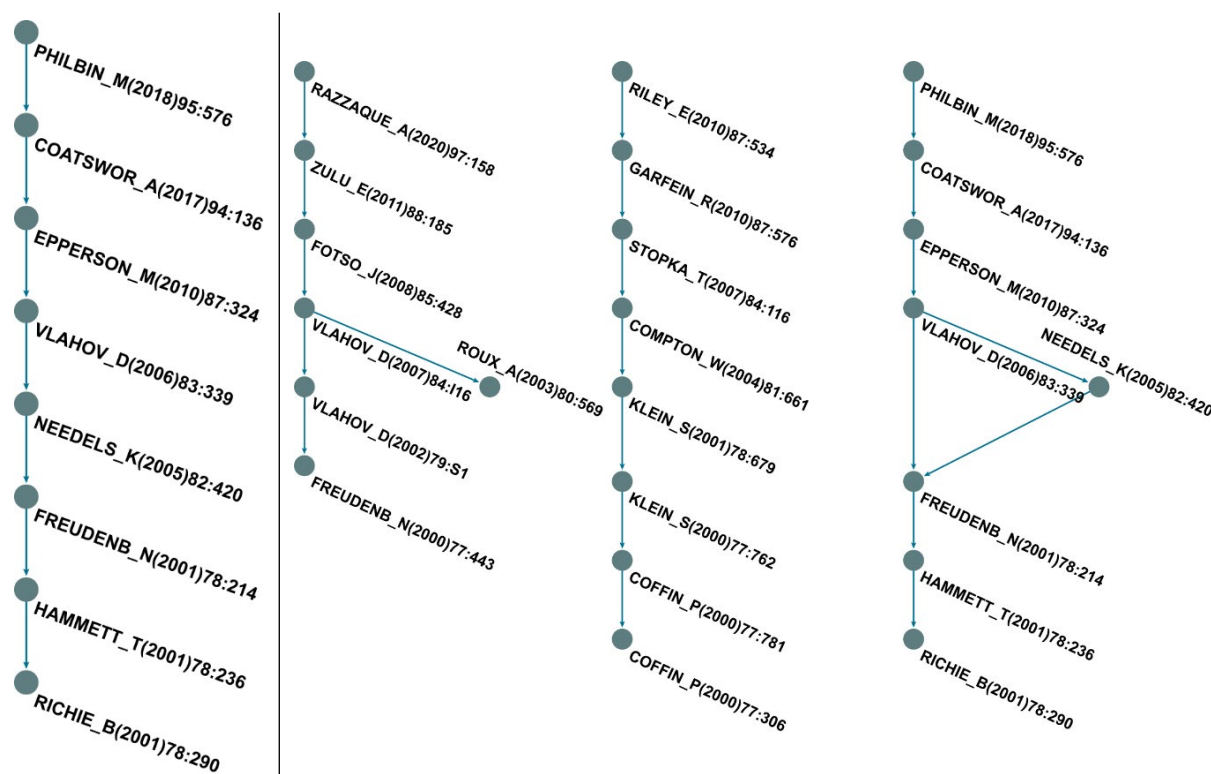
Рис. 3. Основной путь (слева) и ключевые пути (справа) в сети CiteSQ



В работы второго основного пути (таблица 2 в приложении 1) вошли 23 статьи, из которых 8 также входят в первый основной путь. Как видно на рисунке 4, ключевые пути разбиваются на три части. Структура справа повторяет основной путь, центр рисунка — о безрецептурной продаже шприцев и прочих вопросах жизни маргиналов (2000—2013). Структура слева — работы о связи урбанизации, здоровья и становления городского здравоохранения как самостоятельной дисциплины, также присутствуют в списке тем материнское и детское здоровье.

В каждой сети цитирований мы выделили «острова» — плотно связанные друг с другом узлы, отражающие локально важные участки сети. В сети **CiteSQ** основной остров состоит из 98 узлов и включает 26 из 27 работ основного и ключевых путей (код 3 в таблице 1 в приложении 1). Первые публикации острова посвящены общественному здравоохранению в развивающихся странах, социальному доверию и социальным связям в контексте здоровья, употребления наркотиков. С середины 2000-х все больше публикаций — о связи здоровья с жизнью в городе, соседскими общинами, жизнью в трущобах, развитии городской политики и планирования в области здоровья; работы по построению индекса здоровья населения и COVID-19.

Рис. 4. Основной путь (слева) и ключевые пути (справа) в сети CiteJoUH



В сети **CiteJ** основной остров состоит из 92 узлов и включает 14 из 23 работ основного и ключевых путей (код 3 в таблице 2 в приложении 1). Дополнительные темы острова — об исследовательских центрах на базе сообществ (встреченных в ключевых путях в сети **CiteSQ**), употреблении наркотиков, репродуктивном здоровье женщин и здоровье детей.

Интересны работы островов с меньшим количеством узлов.

В сети **CiteSQ** проявлены контроль и определены заболеваний жителей городских агломераций: контроль астмы в городских пространствах (кластер 3)⁹, профилактика рака (10), гепатит В и ВИЧ (11), выявление миомы матки (16). Тут же — тема здорового старения (12). Другая группа тем — о здоровье бездомных (18) и заключенных (19), последний кластер совпадает с основным путем сети **CiteJ**; в контексте бедности и бездомности обсуждаются и азартные игры. Представлены в числе тем наркотики в мегаполисах (13), программы продажи стерильных шприцев (14), связи потребления наркотиков с распространением ВИЧ (6) и выселением из жилых помещений (15).

В сети **CiteJ** работы посвящены влиянию городской среды на физическое состояние горожан. К профилактике же отнесем модель укрепления здоровья стареющего населения (14), здоровое старение в городах, вовлечение пожилых людей в волонтерскую деятельность для укрепления здоровья (20).

⁹ Кластеры 2 и 5 были исключены из описания ввиду их небольшого размера.

Видны темы расовых различий в состоянии здоровья в городских сообществах (10), самооценки состояния здоровья (13). Связаны с ними темы здоровья бездомных людей в контексте изменения климата (17) и жилищной нестабильности (18).

Отличительной особенностью анализа данной сети является получение кластера, в котором обсуждаются методологические вопросы исследований в области городского здравоохранения — оценка выборки, используемой для изучения труднодоступных групп (целевая выборка, выборка, управляемая респондентом) (7).

Анализ ключевых слов

Для анализа ключевых слов используется двумодальная сеть работ и ключевых слов **WK**, полученная в результате трансформации исходных данных массива, сформированного по поисковому запросу, в программе WoS 2Pajek, которая состоит из 149 740 публикаций и 8254 ключевых слов.

Ввиду малого количества работ с выделенными ключевыми словами до 1990 года внимание к различиям в частоте появления ключевых слов в работах было направлено к публикациям, вышедшим не ранее 1990 года. Публикации были разделены на 4 периода по десятилетиям: 1990—2000, 2001—2010, 2011—2020 и 2021 и далее. Для каждого периода была построена редуцированная версия сети работ и ключевых слов и подсчитана метрика входящей центральности *indegree*, позволяющая оценить, насколько часто в указанном периоде конкретное слово появлялось в списке ключевых слов.

Таблица 3. Топ-20 наиболее популярных ключевых слов с 1990 по 2022 год

	1990—2000		2001—2010		2011—2020		2021—2022	
	Value	Слово	Value	Слово	Value	Слово	Value	Слово
1	245	health	795	health	2133	health	528	health
2	162	urban	538	urban	1558	urban	407	urban
3	106	care	416	risk	658	risk	159	city
4	69	new	301	HIV	600	city	134	care
5	65	city	282	drug	593	care	112	study
6	57	York	266	care	434	community	108	risk
7	49	child	262	use	432	use	107	covid
8	48	risk	259	city	430	social	107	19
9	47	state	215	behavior	410	environment	106	environment
10	46	HIV	202	user	391	study	99	social
11	44	disease	198	prevalence	355	neighborhood	98	impact
12	37	service	183	infection	346	population	88	disease
13	35	mortality	182	new	337	public	88	disparity
14	33	woman	176	woman	333	disease	80	community
15	33	infection	176	prevention	327	disparity	79	use
16	31	aid	175	state	318	woman	78	public
17	31	policy	170	community	318	adult	76	mortality
18	31	public	164	factor	312	state	76	neighborhood
19	30	drug	157	population	312	physical	72	analysis
20	30	use	157	York	308	mortality	72	association

Топ-20 ключевых слов по 4 периодам приведен в таблице 3. Наглядно изменения в темах публикаций для обозначенных периодов показывают «облака слов» (рис. 5–6). Они позволяют получить «срез» инфоповодов за десятилетие.

Рис. 5. Облака слов для периода 1990–2000 гг. (слева) и 2001–2010 гг. (справа)



Рис. 6. Облака слов для периода 2011–2020 гг. (слева) и 2021– гг. (справа)



В 1990–2000-е годы наблюдается рост обеспокоенности распространением ВИЧ. Нередко можно увидеть разговор о рисках и организованных формах здравоохранения, причем авторы часто фокусируются на наркопотреблении как детерминанте распространения ВИЧ. В 2010-е годы виден фокус на социальных аспектах здравоохранения и концептах коммунальности. Чаше встречаются понятия сообществ, публичности и неравенства. Это долгий тренд, однако описание эпидемии COVID-19 его перебивает — как и популярную ранее гендерную тематику (рис. 7).

Распределение частот популярных ключевых слов (рис. 8) показывает, что многие из них отражают инфоповод либо академическую моду.

Сеть соприсутствия ключевых слов по всем годам состоит из 8254 ключевых слов. Построенная сеть нормализована с использованием фракционного подхода (Batagelj, 2020). С помощью подхода островов был выделен основной кластер связанных друг с другом ключевых слов, представленный на рисунке 9.

Рис. 7. Распределение частоты употребления ключевого слова «woman»

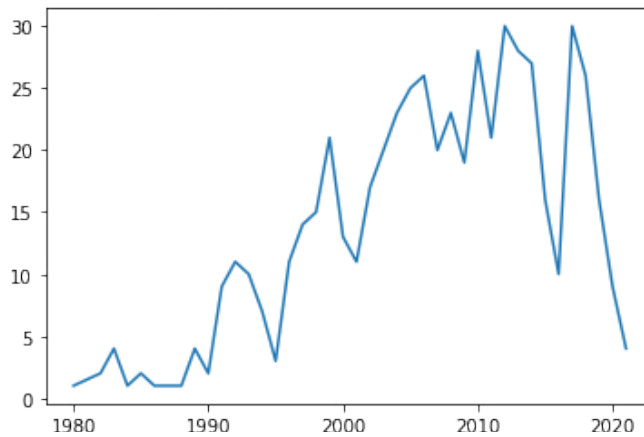


Рис. 8. Распределение частот упоминания наиболее популярных за весь исследуемый диапазон ключевых слов

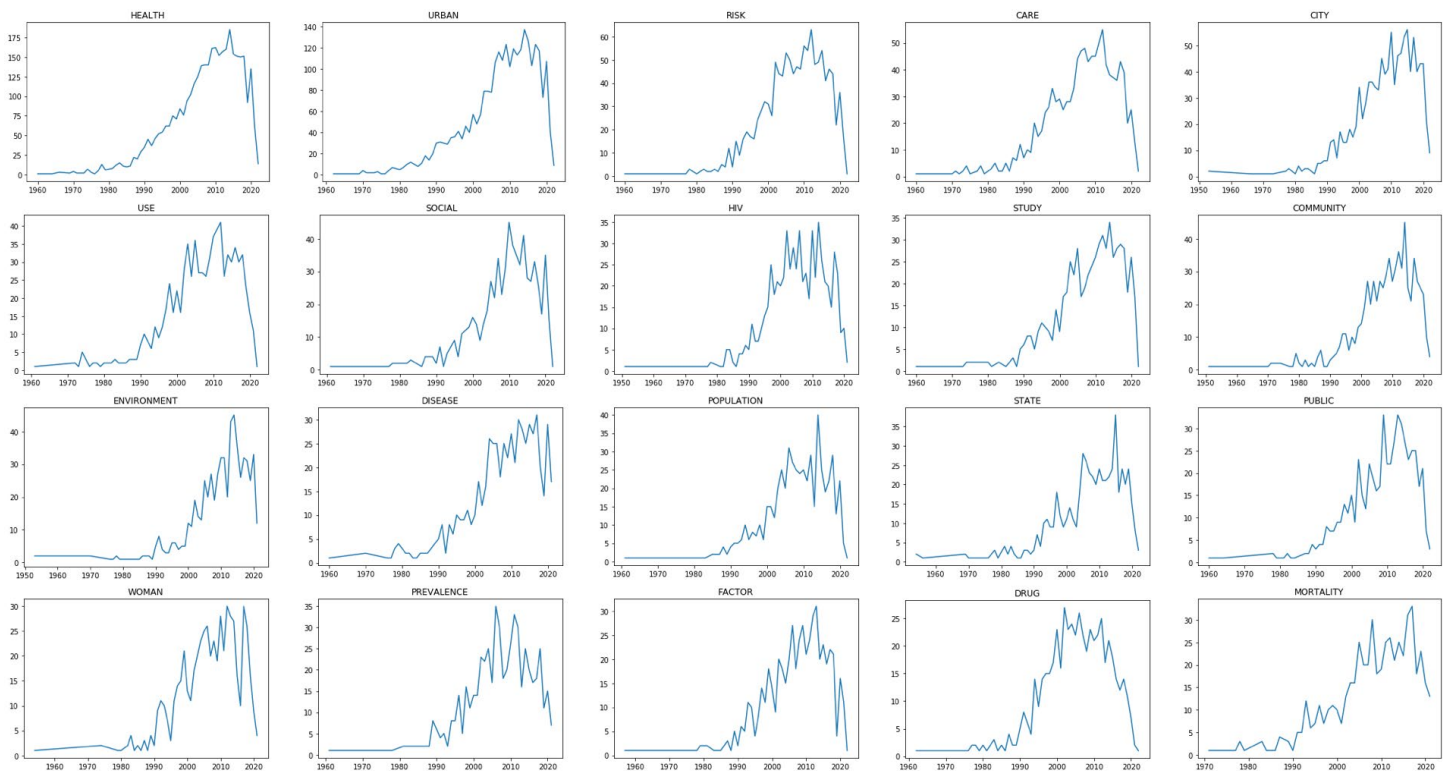
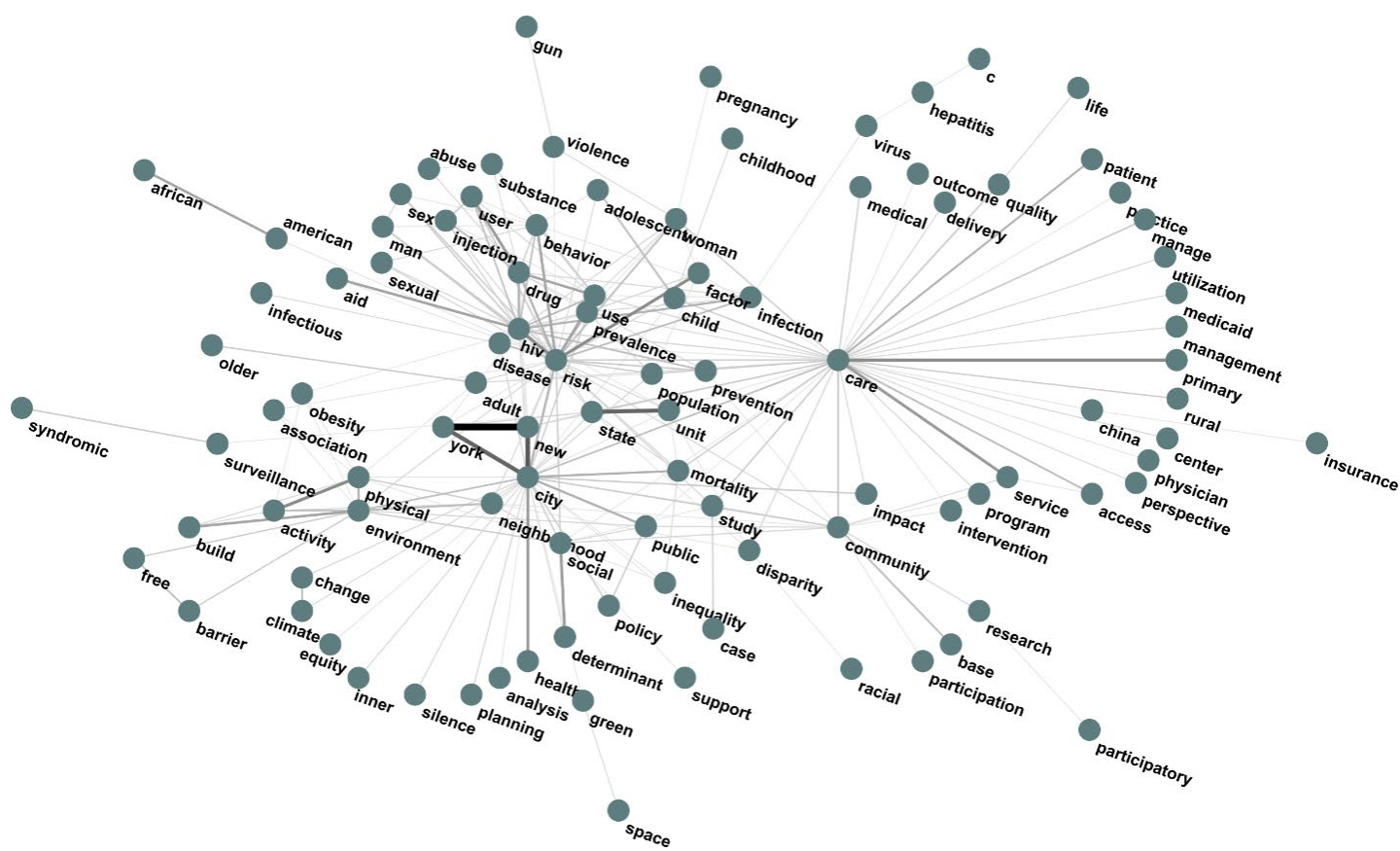


Рис. 9. Сеть совстречаемости ключевых слов ККп по всем годам

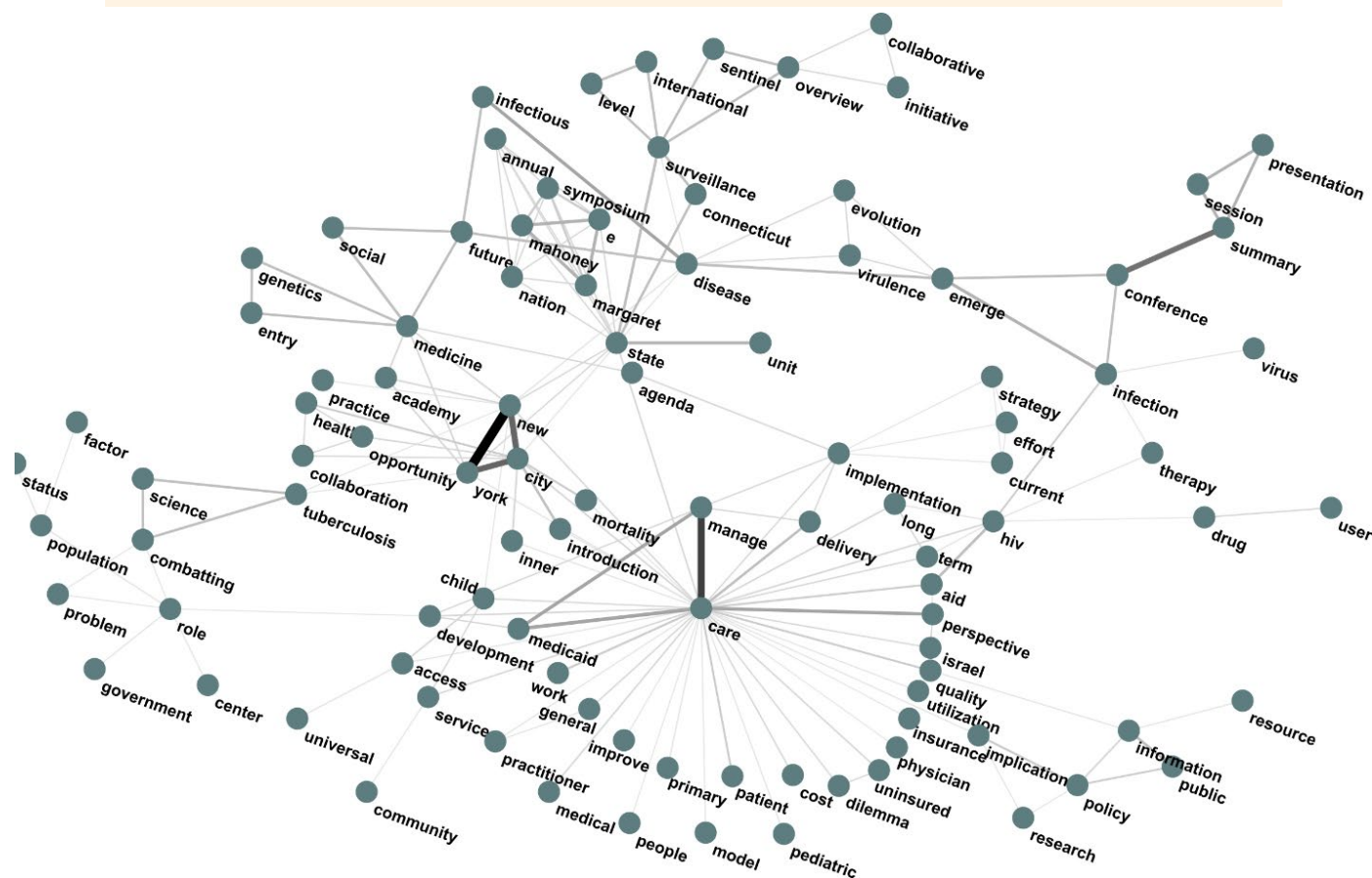


Из ключевых направлений проблемных областей мы выделяем внутренне связанный кластер слов, сопresentствующих с «hiv» — про ВИЧ и сопутствующие демографические и социальные аспекты. Второй субкластер — организация здравоохранения (блок слов, сопresentствующих с «care»). Третий кластер завязан на «community»: партиципаторные практики сообществ, связанных с медицинским уходом. Еще один значимый кластер посвящен ЗОЖ — «physical environment», «activity» и «build». Здесь видим крепкую связь медицины с городом.

Кластеры связаны друг с другом: кластер дискуссий о ВИЧ связан с кластером перспективы организации здравоохранения, а лечение предполагает академический и социальный контекст, предоставляющий соучастие для эффективного достижения целей.

В редуцированной сети работ и ключевых слов 1990—2000-х годов число работ составило 26 519, ключевых слов — 1603. В сети следующего десятилетия показатели были соответственно 51 316 и 3778, в сети 2010—2020 гг. — 55 535 и 6195, за 2021—2022 годы — 2659 работ и 3034 ключевых слова. По редуцированным сетям выстроены сети совстречаемости ключевых слов по указанным периодам, причем число узлов в каждой сети соответствует числу ключевых слов в периоде. Через подход островов мы выделили основные кластеры связанных ключевых слов (рис. 10—13).

Рис. 10. Сеть совстречаемости ключевых слов ККн1 за 1990–2000-е годы



В сети первого периода (рис. 10) видим управленческий академический дискурс. Описываются новые методы и результаты исследований либо организационные вопросы лечения. Эти два направления соединяются только в точках, отражающих массовые болезни (см. «tuberculosis», «hiv»).

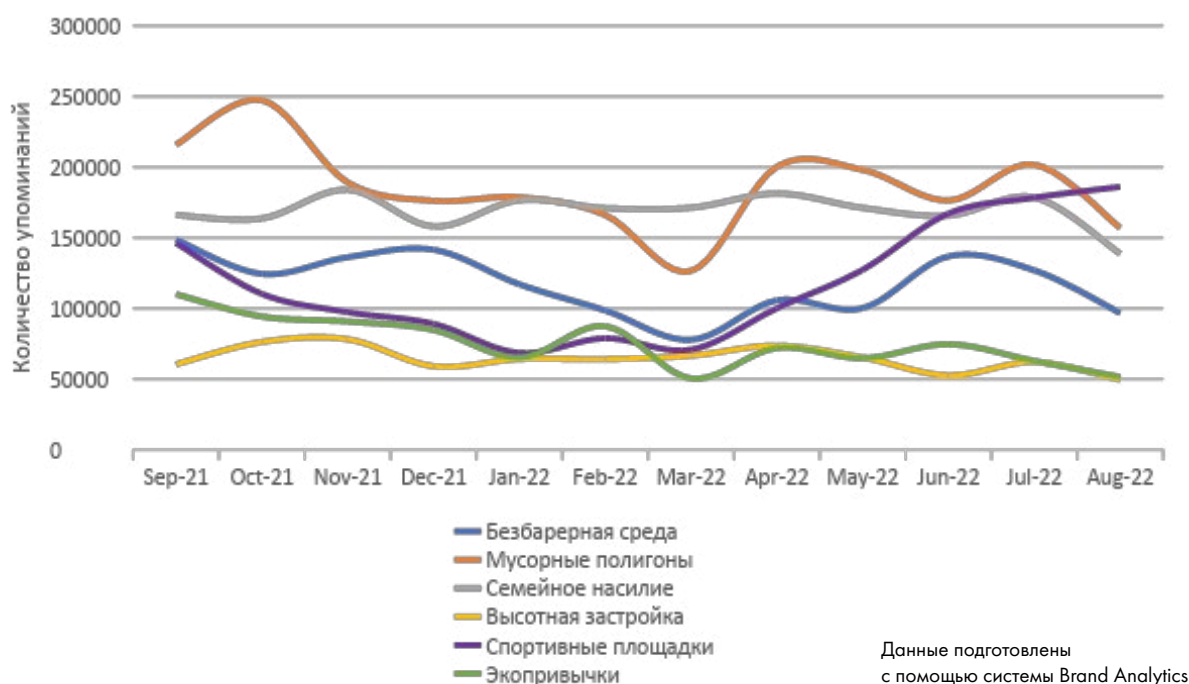
В 2010-е годы (рис. 11) дискурс о «здоровой» городской среде центрируется вокруг ВИЧ-инфекций. Сеть показывает «сексуально-наркотическую» гипотезу об их распространении. На этой почве вырос дискурс о социальном неравенстве и роли сообществ в медицине. Выходит, сеть имеет тематические «отростки», связывающие сообщества с академическими исследованиями, которые включают расовую и гендерную справедливость в систему координат организации медицины.

Во втором десятилетии XXI века ВИЧ отошел на второй план, оставив центр внимания «риску», что переводит внимание с уже больных на тех, кто мог бы потенциально заболеть.

Расширяется географический спектр городского, к которому обращаются исследователи: если в предшествующие периоды слово «город» («city») как имя собственное проявляло себя лишь в сочетании «New York City», то на рисунке 12 уже видны отсылки к европейским городам или Гонконгу.

Выделяются три смежные области, сопоставимые с наблюдаемыми в общей сети (рис. 9): факторы риска, городская среда и организация здравоохранения.

Рис. 14. Динамика обсуждений вопросов, касающихся здоровья города, в социальных медиа с сентября 2021 по август 2022 года



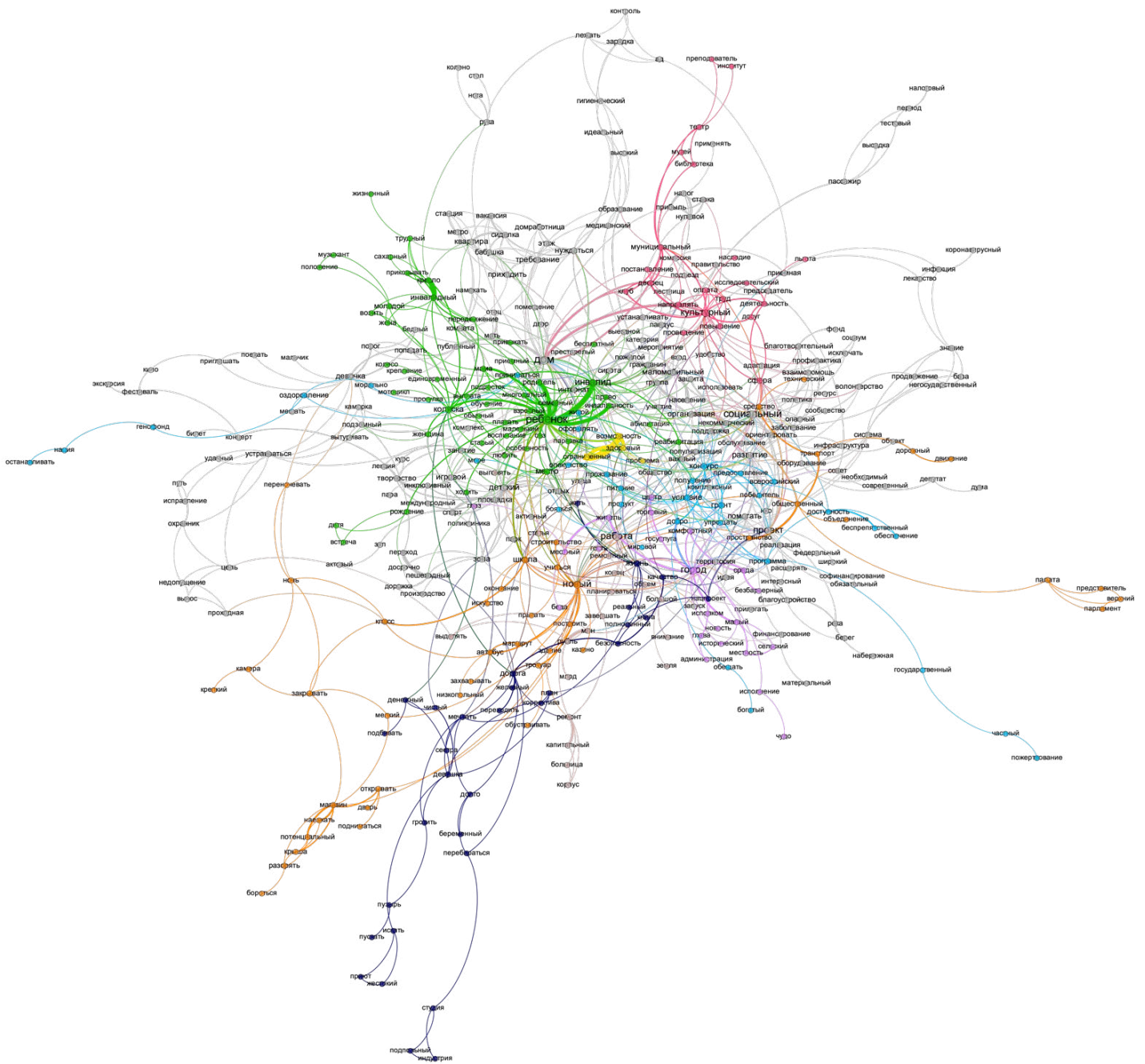
Безбарьерная среда воспринимается пользователями как социальная интеграция инвалидов в городскую жизнь через мероприятия, предоставление условий для жизни, учебы, работы и пр.

В центре обсуждения здесь — инклюзивная среда для детей-инвалидов и условия для передвижения родителей с колясками. Также видно благоустройство городской среды и проекты по созданию среды для маломобильных граждан.

К социальному компоненту обсуждения относятся мероприятия для инвалидов и научно-просветительские лекции; налоговые льготы для культурных учреждений, проводящих выездные культурные мероприятия для людей с инвалидностью, сирот и пожилых граждан. Также видна финансовая поддержка инвалидов.

Рис. 15. Структура обсуждений вопросов безбарьерной среды в социальных медиа с сентября 2021 по август 2022 года

Цвета репрезентируют разные блоки обсуждений, выделенные на основе автоматической кластеризации сети. Связь между парой слов — их семантическая близость как минимум в 200 упоминаниях



Обсуждение полученных результатов

Тематика здоровых городов в последние десятилетия центрирована на вопросах о важности эмпирического изучения детерминант здоровья, социальной справедливости в отношении здравоохранения и решения проблем с доступом к медицине через медицинское обслуживание в городских сообществах. Эта область чувствительна к социальным проблемам в области здравоохранения, поэтому сегодня фокус в обсуждениях на COVID-19. Появление программных работ и обсуждение методологии исследований по тематике здоровых городов указывают на зрелость направления: складывается самостоятельная предметная область.

С точки зрения используемой методологии работа с различными источниками данных (поиск по ключевым словам в WoS; по статьям в конкретном журнале в WoS; сбор упоминаний в социальных медиа) и использование анализа ключевых слов, публикаций, тематического моделирования и пр. позволили получить более надежные и разнообразные результаты. В частности, выводы на базе анализа сетей цитирования нашли дополнительное подтверждение в анализе ключевых слов, а анализ данных социальных медиа дал возможность увидеть, как «здоровье» города обсуждается в публичном дискурсе.

Благодарности

Мы благодарим за помощь наших коллег и партнеров: Владимира Батагеля (МЛ прикладного сетевого анализа НИУ ВШЭ, Университет Любляны, Университет Копера) — за сбор библиометрических данных, Василия Черного и Валерию Вебер (Brand Analytics) — за сбор данных соцмедиа и возможность использовать систему анализа Brand Analytics, Ирину Павлову и Аню Пищикову (МЛ прикладного сетевого анализа НИУ ВШЭ) — за оформление результатов работы. Исследование выполнено в рамках Программы фундаментальных исследований НИУ ВШЭ.

Библиография

Булычева Е. Е., Мальцева Д. В. (2020) Выделение актуальных тематик в социологии: взгляд сквозь призму анализа сети цитирований // Мониторинг общественного мнения: экономические и социальные перемены. No. 6. С. 113—140. <https://doi.org/10.14515/monitoring.2020.6.971>.

Мальцева Д. В., Моисеев С. П. (2018) (<https://www.hse.ru/org/persons/65832822>). Отбор источников для систематического обзора литературы: сравнение экспертного и алгоритмического подходов (<https://publications.hse.ru/view/226445733>) // Социология: методология, методы, математическое моделирование. Т. 47. С. 7—43.

- Bastian M., Heymann S., Jacomy M. (2009) Gephi: An Open-Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 3. No. 1. P. 361–362. <https://ojs.aaai.org/index.php/ICWSM/article/view/13937>.
- Batagelj V. (2020) On Fractional Approach to Analysis of Linked Network. *Scientometrics*. Vol. 123. No. 2. P. 621–633.
- Batagelj V., Doreian P., Ferligoj A., Kejžar N. (2014) Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution. Chichester: Wiley. <https://doi.org/10.1002/9781118915370>.
- Batagelj V., Ferligoj A., Squazzoni F. (2017) The Emergence of a Field: A Network Analysis of Research on Peer Review. *Scientometrics*. Vol. 113. No. 1. P. 503–532.
- Blei D. M., Ng A. Y., Jordan M. I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*. Vol. 3. P. 993–1022.
- Carley K. M. (2020) Network Text Analysis: The Network Position of Concepts. *Text Analysis for the Social Sciences*. Routledge. P. 79–100.
- De Nooy W., Mrvar A., Batagelj V. (2018) Exploratory Social Network Analysis with Pajek: Revised and Expanded Edition for Updated Software. Cambridge: Cambridge University Press.
- Freeman L. (2005) The Development of Social Network Analysis: A Study in the Sociology of Science. *Social Networks*. Vol. 27. No. 4. P. 377–384.
- Hummon N. P., Carley K. (1993) Social Networks as Normal Science. *Social Networks*. Vol. 15. No. 1. P. 71–106. [https://doi.org/10.1016/03788733\(93\)90022d](https://doi.org/10.1016/03788733(93)90022d).
- Maltseva D., Batagelj V. (2019) Social Network Analysis as a Field of Invasions: Bibliographic Approach to Study SNA Development. *Scientometrics*. Vol. 121. No. 2. P. 1085–1128. <https://doi.org/10.1007/s1119201903193x>.
- Maltseva D., Batagelj V. (2020) Towards a Systematic Description of the Field Using Keywords Analysis: Main Topics in Social Networks. *Scientometrics*. Vol. 123. No. 1. P. 357–382. <https://doi.org/10.1007/s11192-020-03365-0>.
- NetworkX. *Software Library*. By Aric Hagberg Pieter Swart Dan Schu and others. <https://networkx.org/>.
- NLTK Project. NLTK – Developers and contributors. By Steven Bird, Edward Loper, Ewan Klein and others. <https://www.nltk.org>.
- R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Zachlod C., Samuel O., Ochsner A., Werthmüller S. (2022) Analytics of Social Media Data—State of Characteristics and Application. *Journal of Business Research*. Vol. 144. P. 1064–1076.

 Приложения:

