

# Профилирование подписчиков пабликов ВКонтакте

“Хунта Экспресс”, “Балканский  
Экспресс” и “Она развалилась”

---

Мария Казакова

# Задача

Определить примерные интересы, пол, возраст и место жительства (город, страна) типичного подписчика всех трёх пабликов

- “ХУНТА ЭКСПРЕСС” – про страны Третьего мира
- “БАЛКАНСКИЙ ЭКСПРЕСС” – про Балканы
- “ОНА РАЗВАЛИЛАСЬ” – про СССР в 1985-1999 годах

# Почему эти паблики?

- Изначально создавались и админились одними и теми же людьми
- Содержат перекрёстные ссылки друг на друга
- Особая эстетика упадничества и абсурда исторических процессов в регионах, к которым не приковано мейнстримное внимание

# Этапы реализации

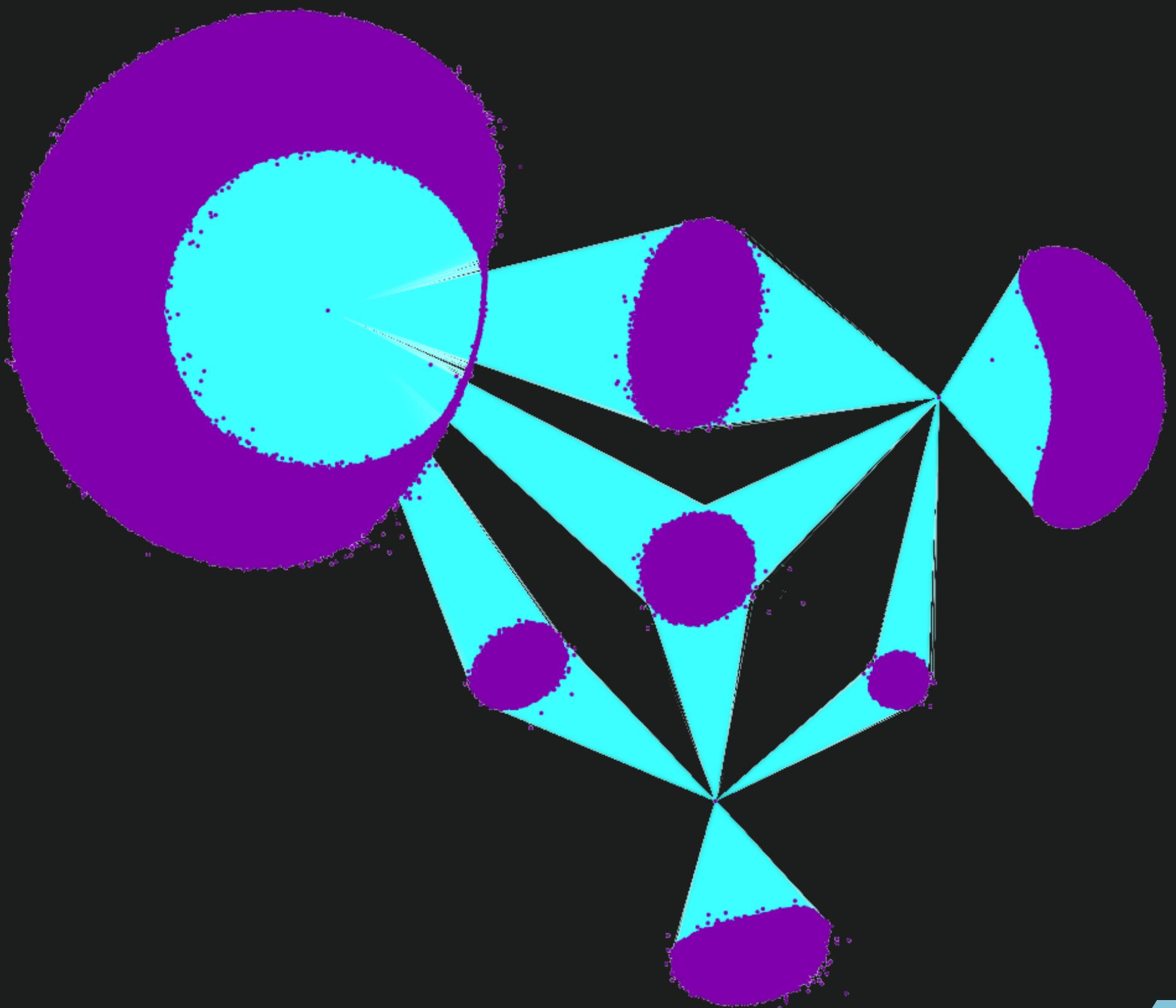
- Сбор подписчиков пабликов – Github 1
- Выявление пересечения (ядра, центрального скопления) аудиторий – Github 2
- Описание пола, возраста, стран ядра – Github 3
- Выявление в ядре самого большого сообщества и анализ его интересов по пабликам-подпискам –  
Github 4

# Визуализации. Шаг 1

Когда мы только собрали аудитории всех пабликов и нашли их пересечение-ядро, у нас получилась такая картина ->

Центр слева – паблик "Она развалилась",  
центр справа – "Хунта Экспресс",  
центр снизу – "Балканский Экспресс".

Шапочки вокруг центров – это подписчики, которые подписаны исключительно на этот паблик, но не подписаны на другие.



# Визуализации. Шаг 2

Затем мы оставили только узлы ядра, центрального скопления.

Собрали друзей этих id и попытались визуализировать.

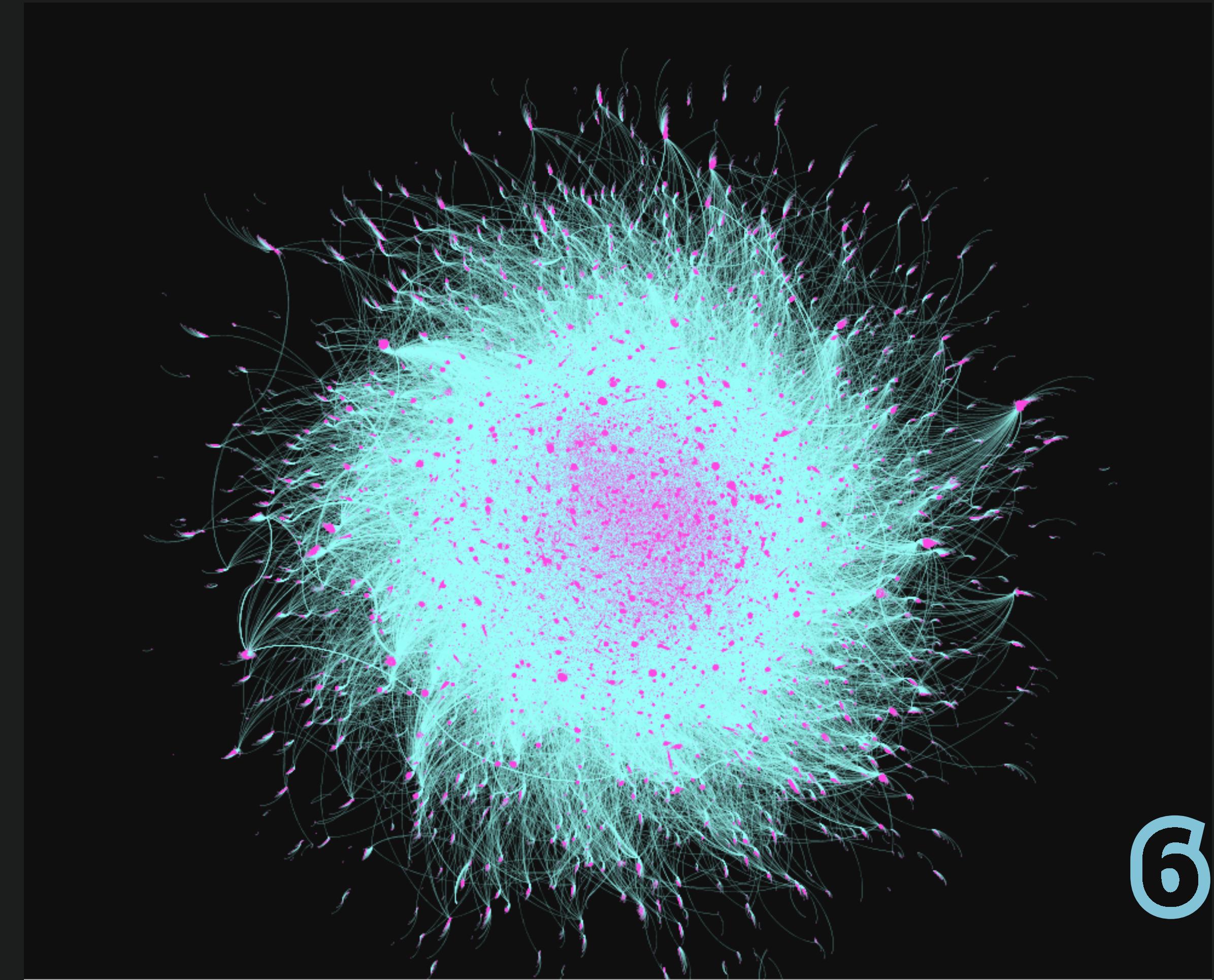
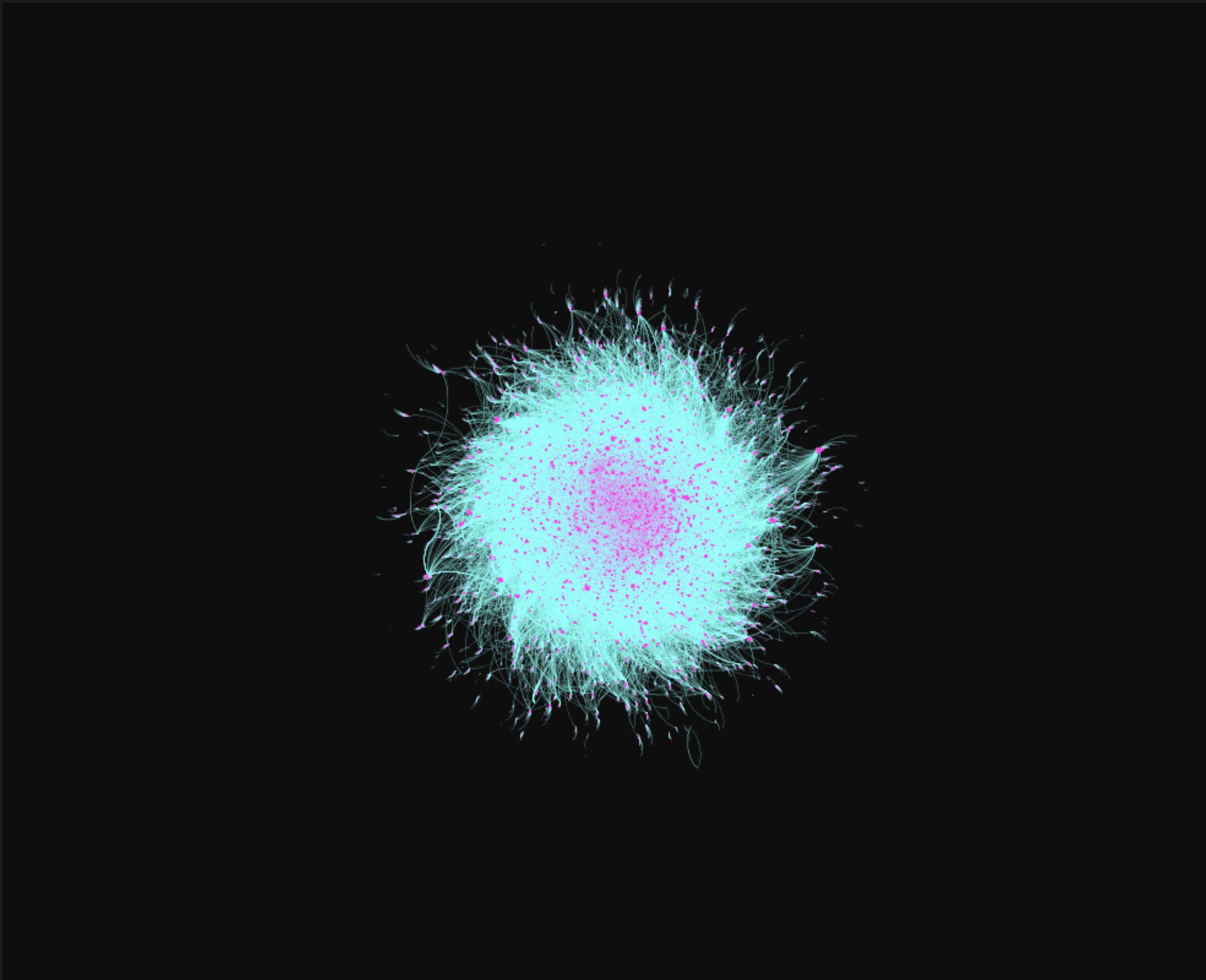
Jupyter Notebook не потянул такую нагрузку, и мы перешли в Gephi.

Gephi тоже несколько раз ломался.

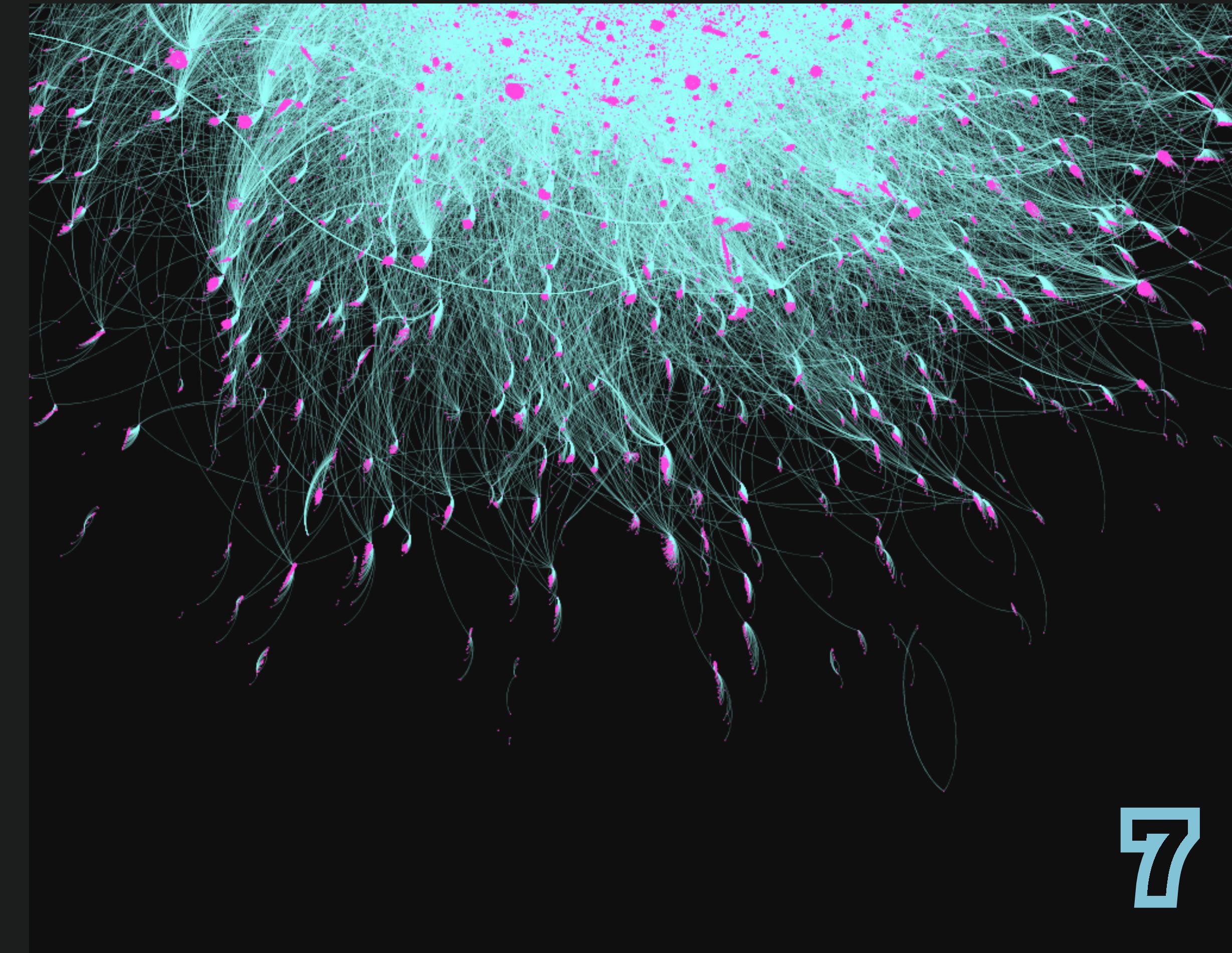
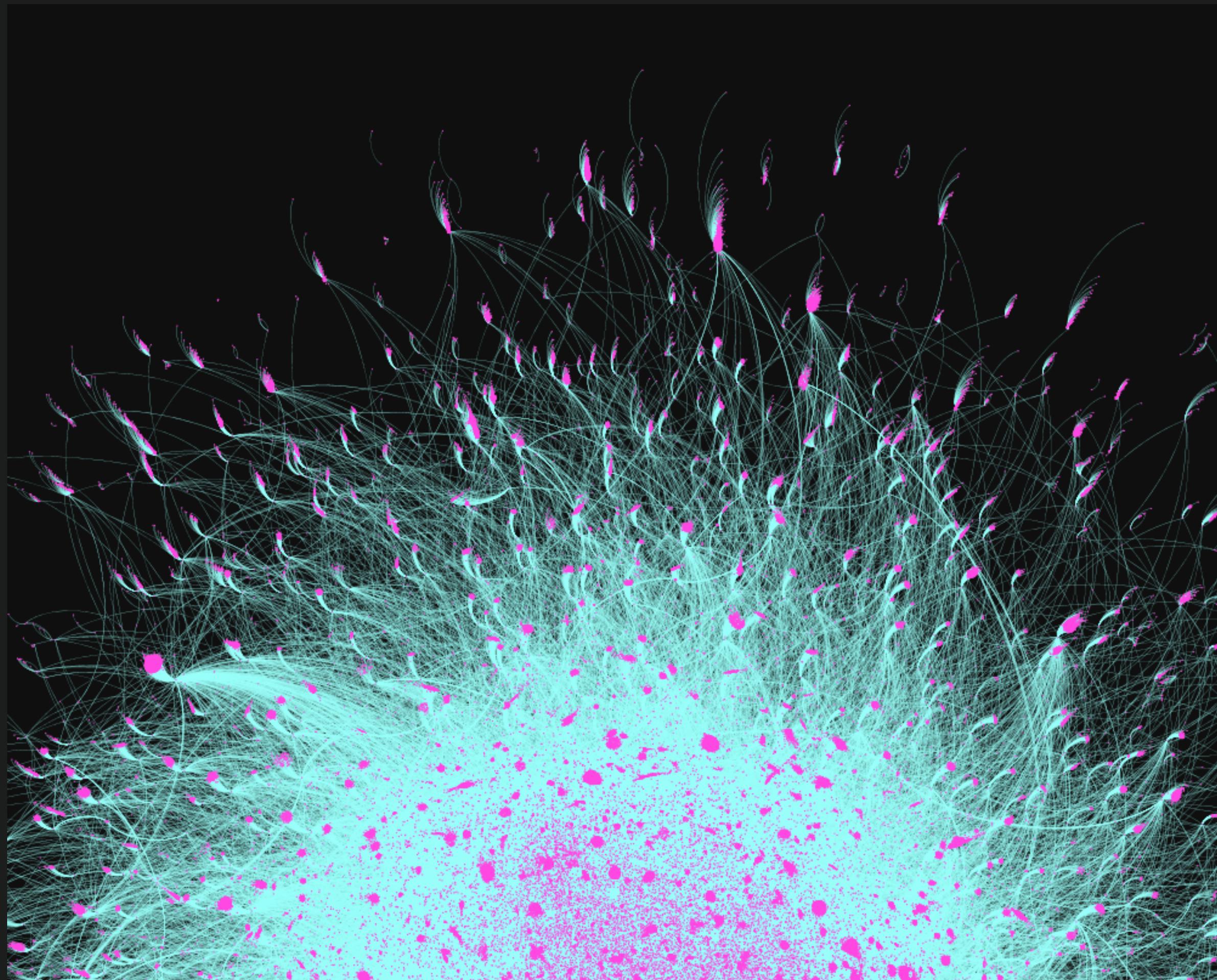
Всё же нам удалось заскриншотить в неполном качестве некий итог и сделать выводы.

Результат – на следующих слайдах.

# Визуализации. Шаг 2



# Визуализации. Шаг 2

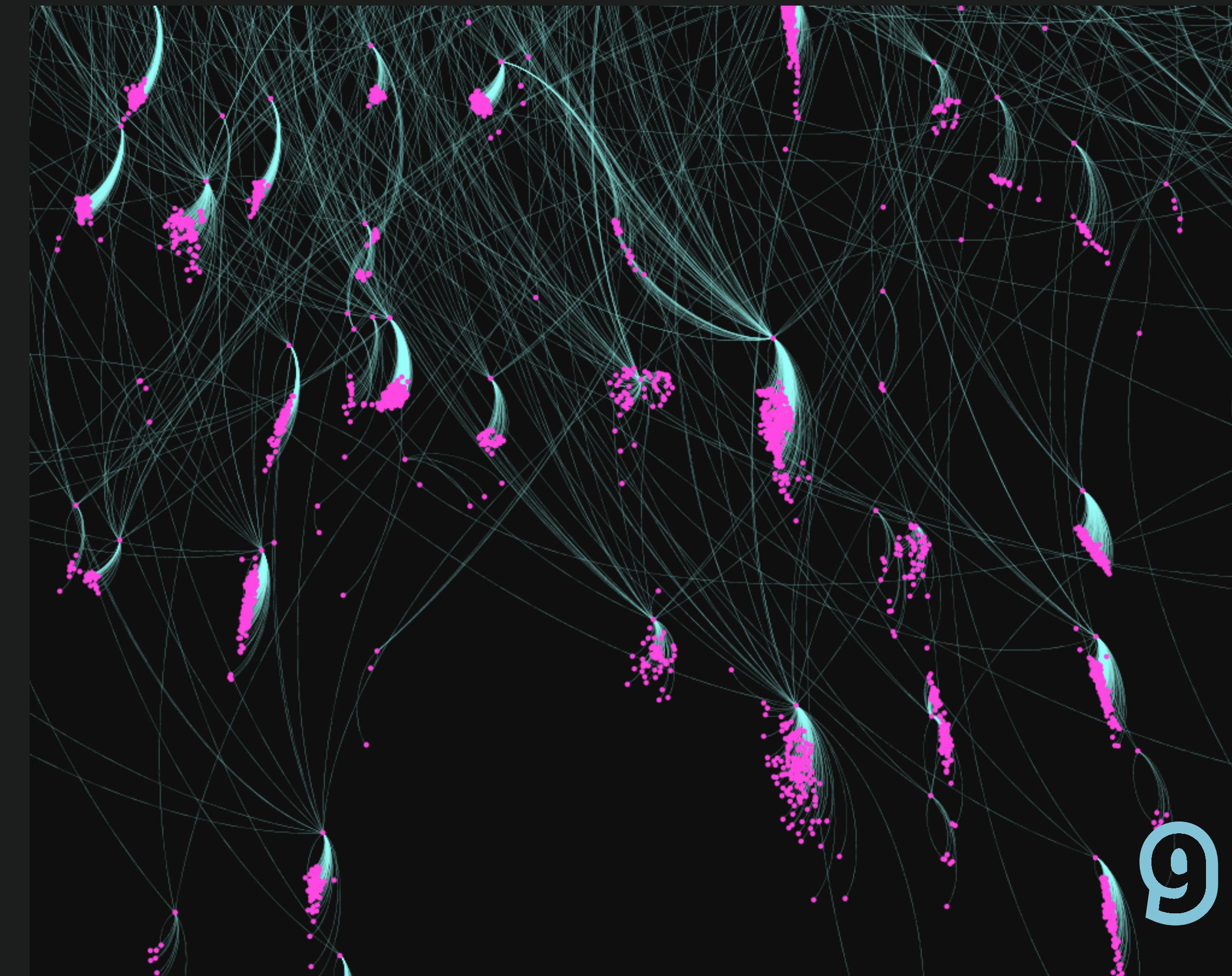
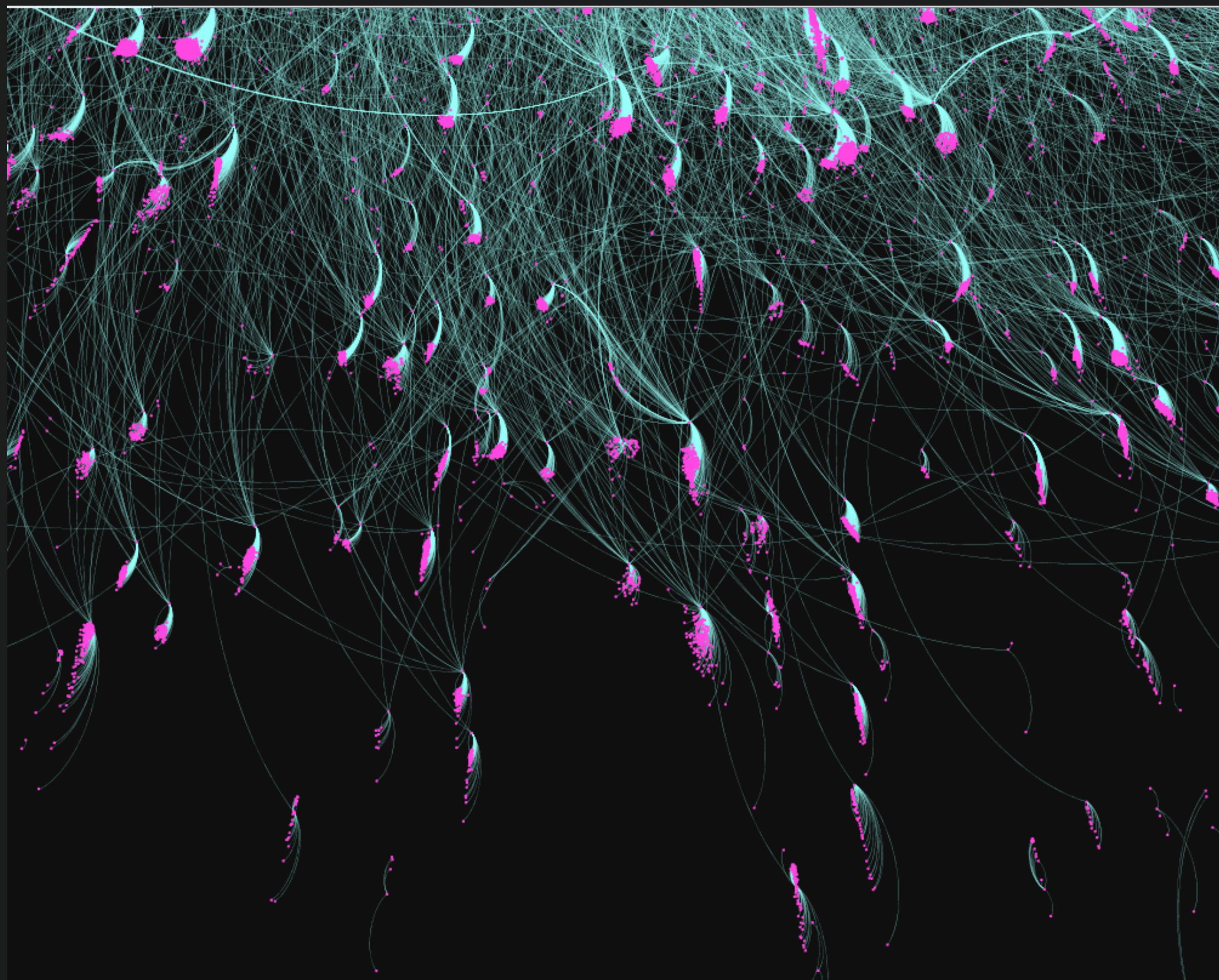


# Визуализации. Шаг 2. Цитата

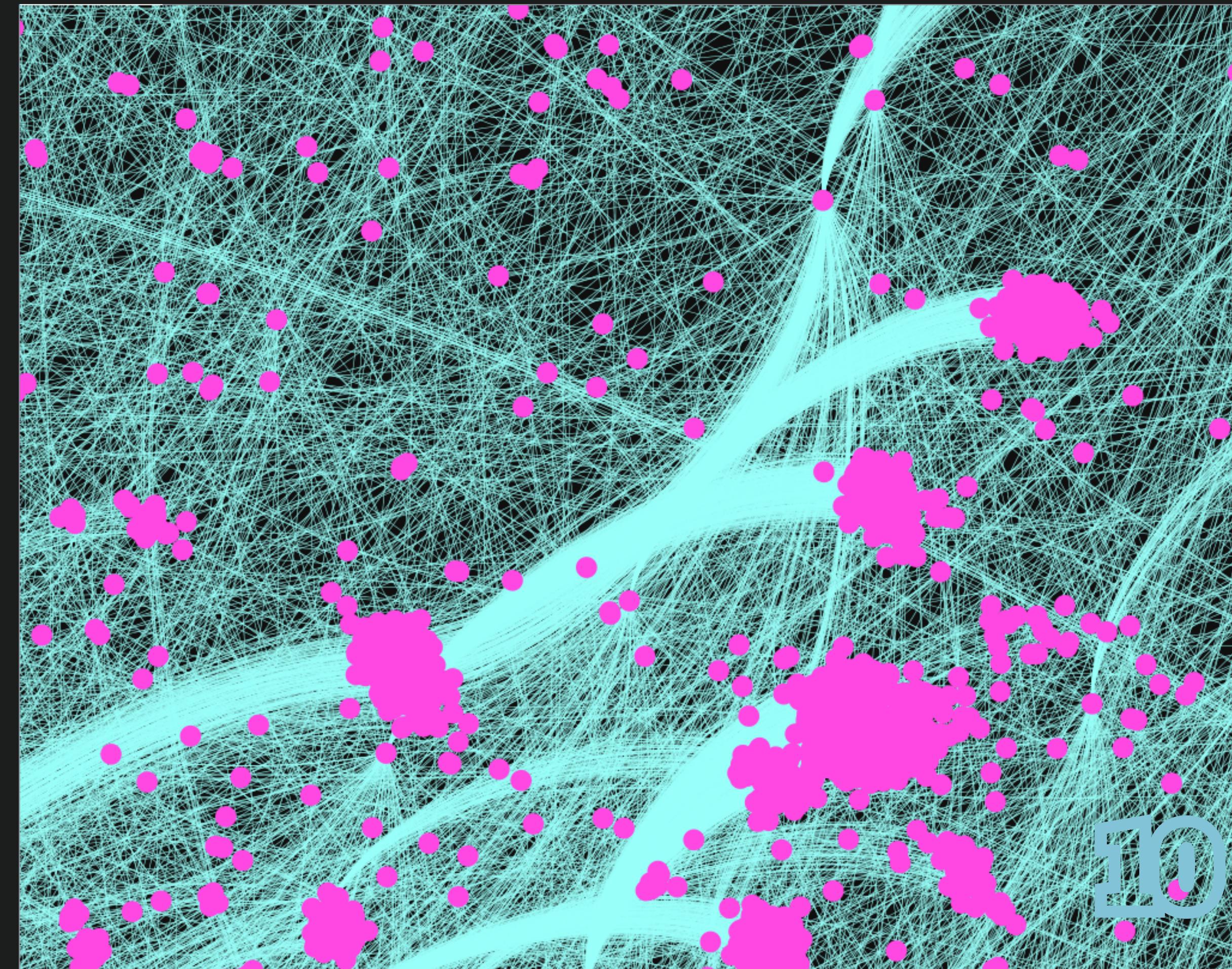
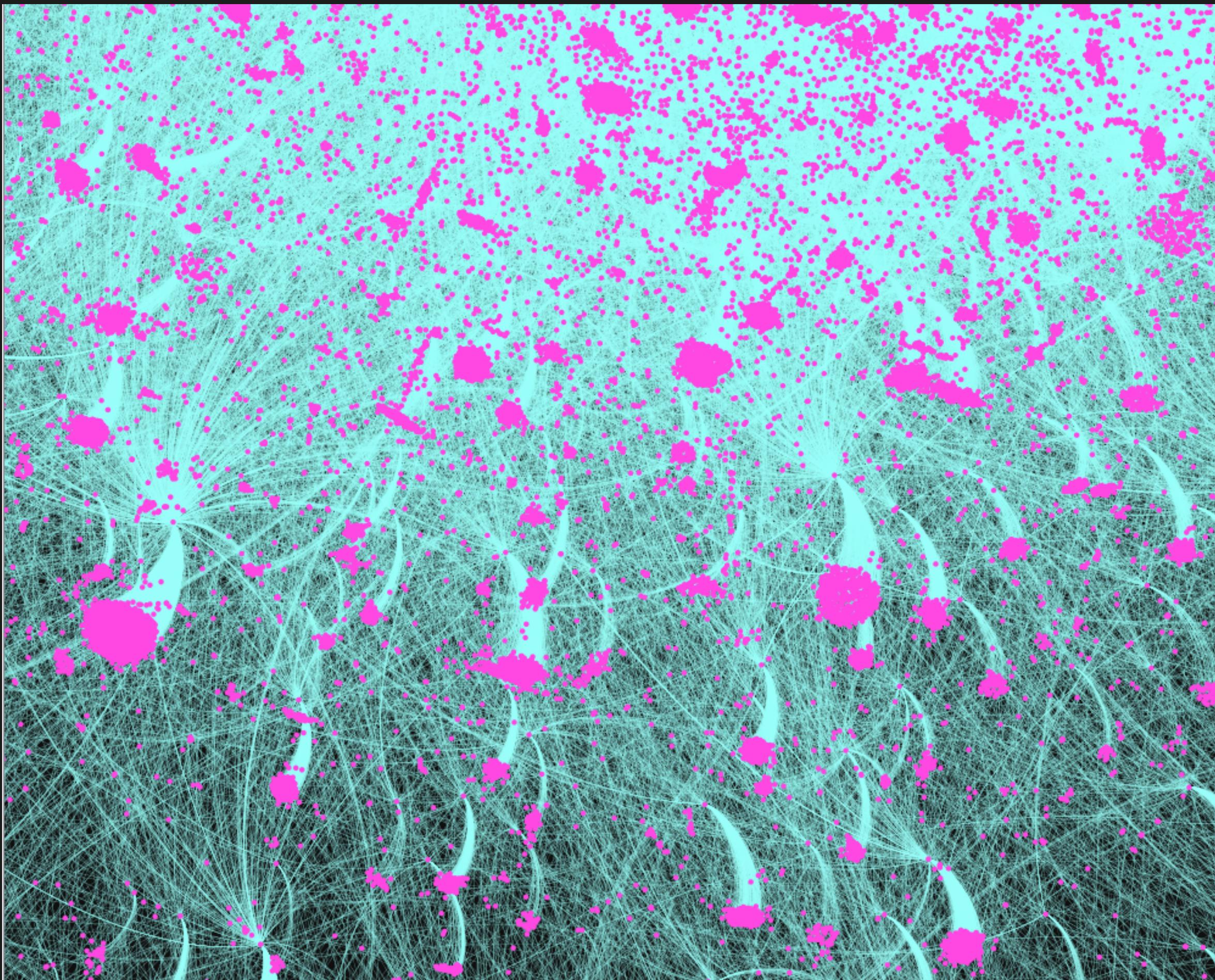
“Мне этот мир абсолютно понятен, и я здесь ищу только одного: покоя, умиротворения и вот этой гармонии от слияния с бесконечно вечным. От созерцания этого великого фрактального подобия и от вот этого вот замечательного всеединства существа. Бесконечно вечного, куда ни посмотри: хоть вглубь – в бесконечно малое, хоть ввысь – в бесконечное большое, понимаешь?”

Идущий к реке

# Визуализации. Шаг 2



# Визуализации. Шаг 2



# Визуализации. Шаг 2

Мы были настолько поражены сложностью и красотой этого графа, что совсем упустили из глаз его очевидность.

Издалека нам казалось, что розовые скопления узлов – это и есть сообщества. На деле же “конские хвостики” – это случаи, когда у одного человека есть друзья, которые больше ни с кем не пересекаются, вот они и выталкиваются алгоритмом Force Atlas 2 к краю холста.

Продолжаем анализ.

# Визуализации. Шаг 3

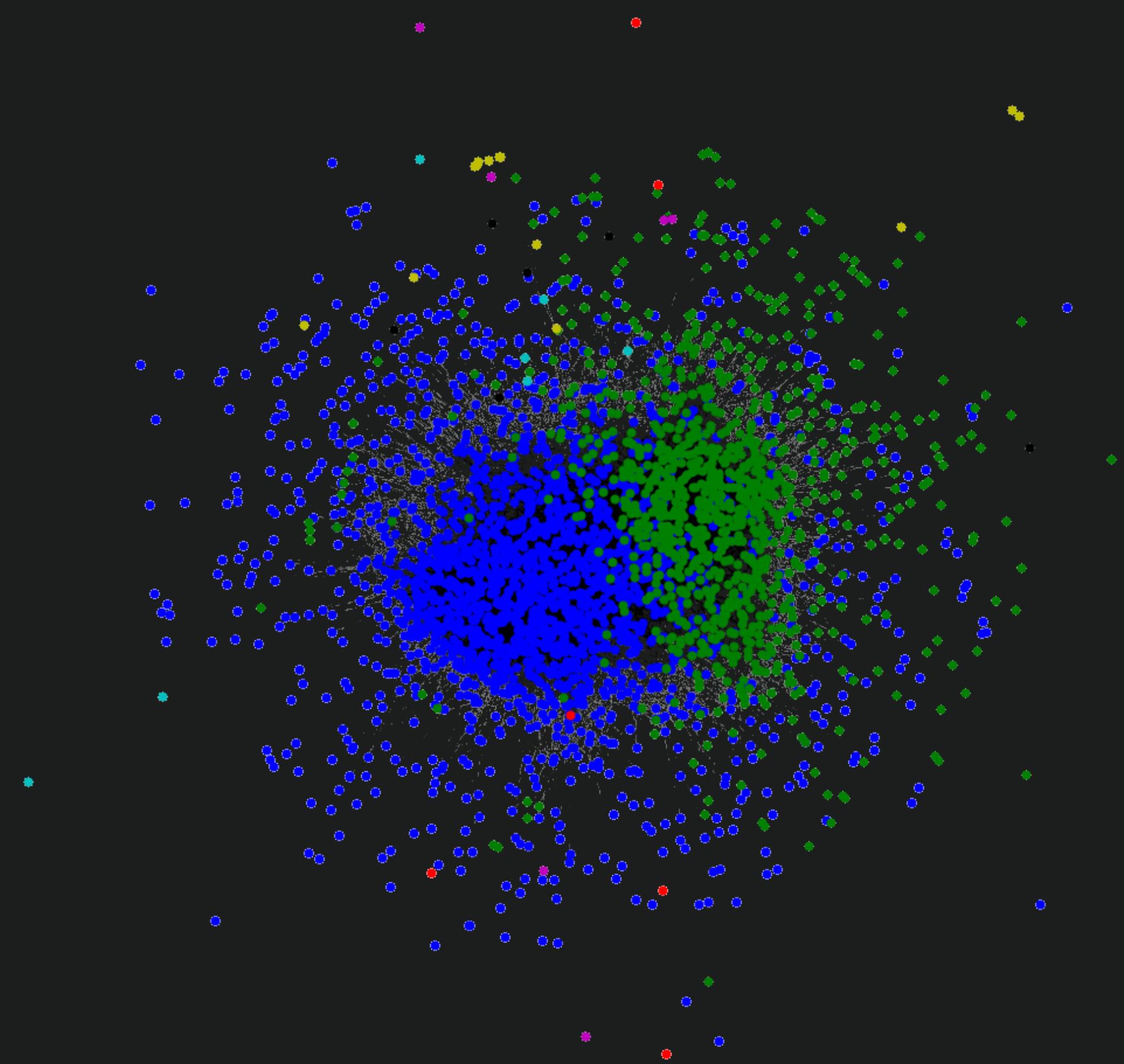
Далее мы отсортировали граф так, чтобы у узлов было degree не менее 10, то есть чтобы у каждого человека в графе было не менее десяти друзей.

Затем пропустили ядро через алгоритмы communities detection.

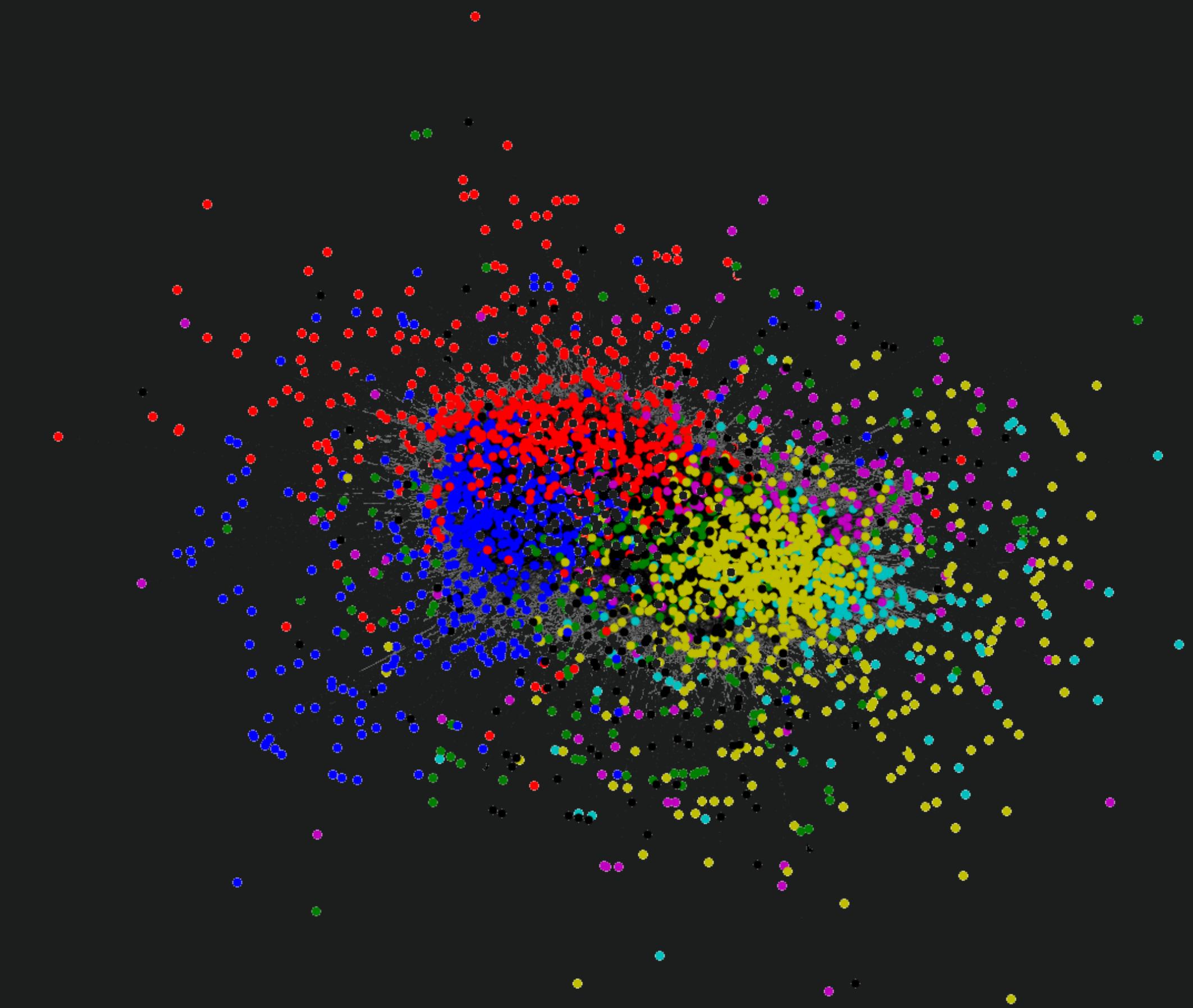
Выявили для каждого алгоритма самое большое сообщество.

Нашли пересечение результатов выполнения алгоритмов Label Propagation, Fast Greedy и Modularity и оформили в единое сообщество из 979 членов.

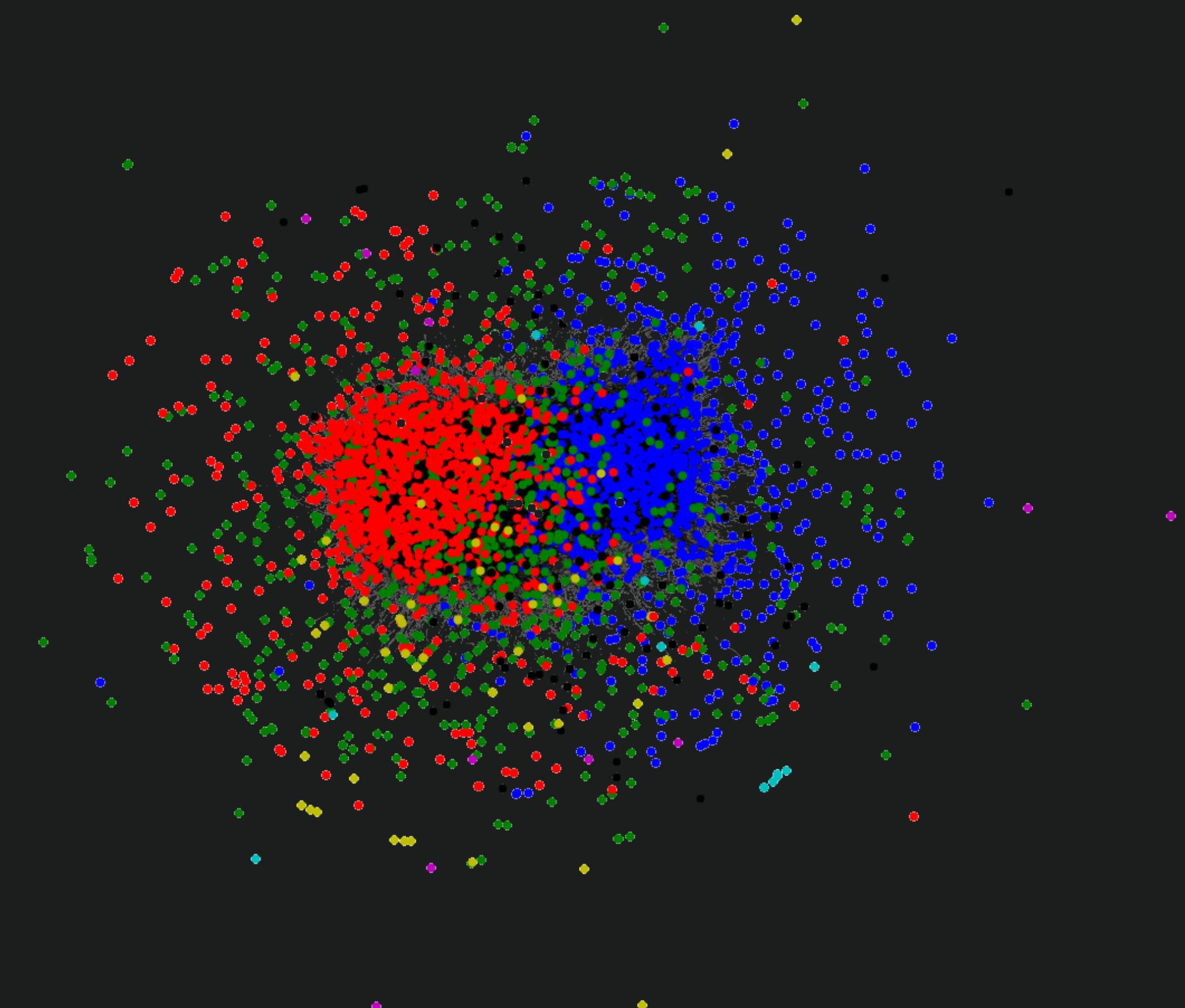
# Визуализации. Label Propagation



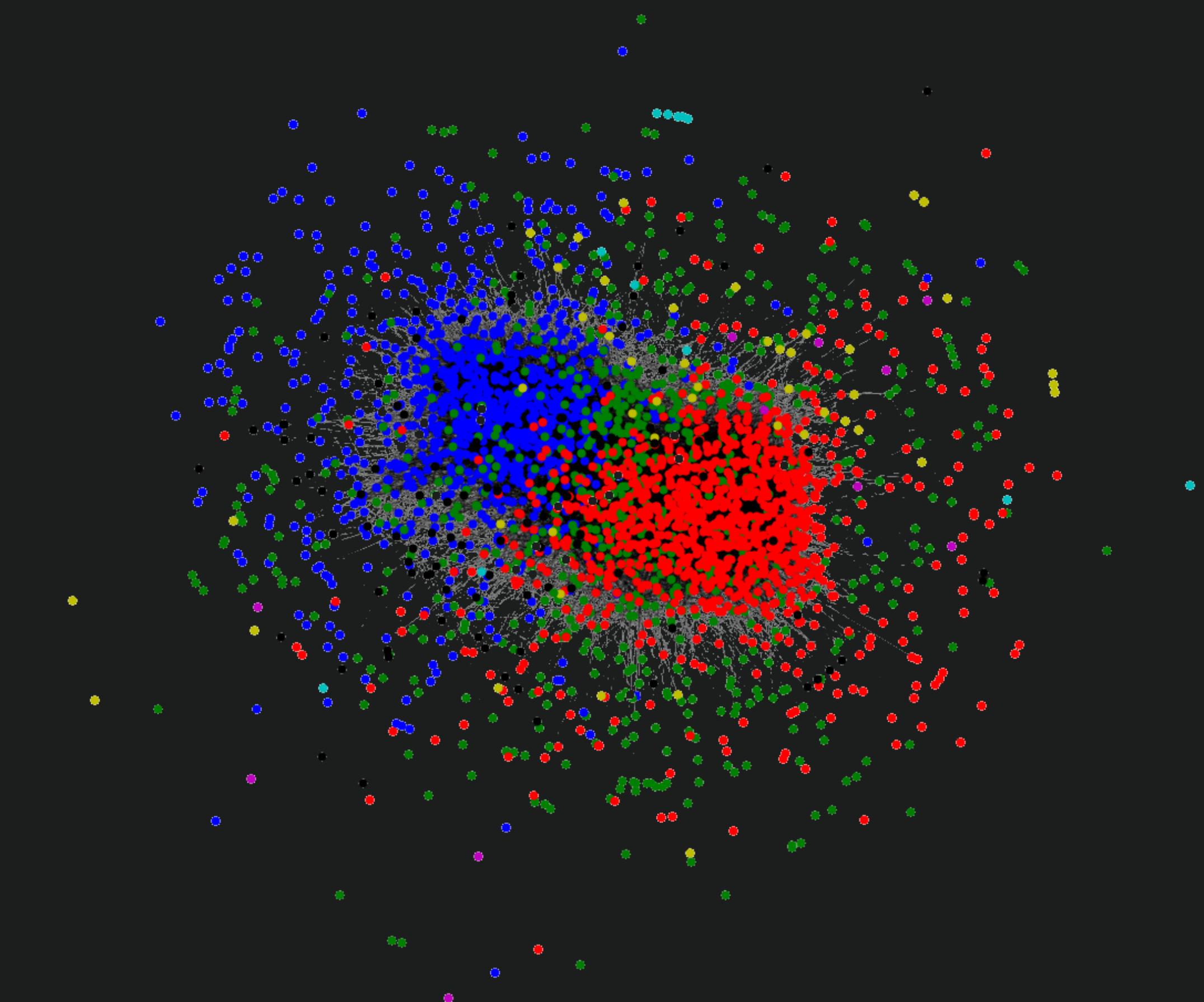
# Визуализации. Louvain



# Визуализации. Fast Greedy

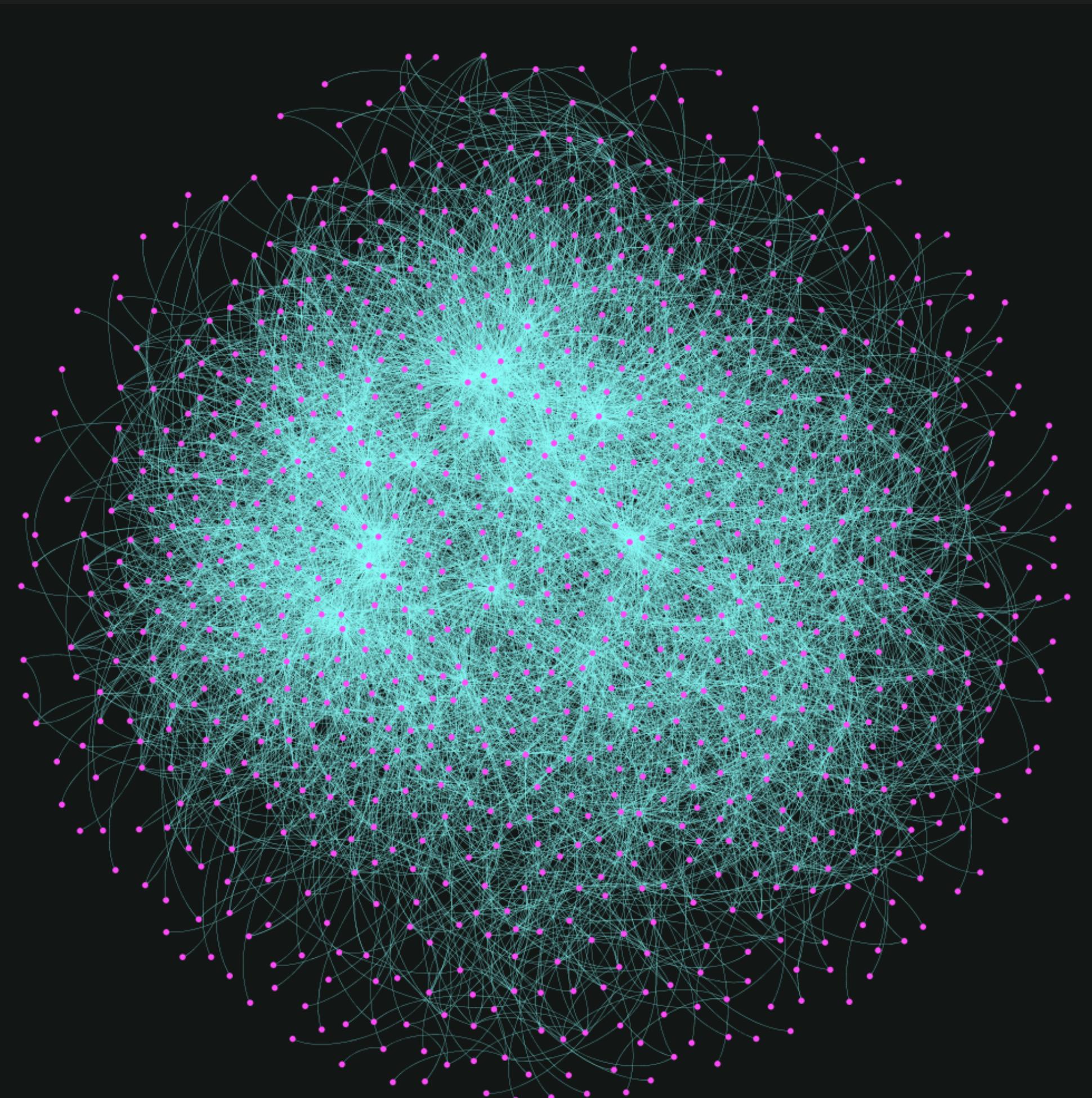


# Визуализации. Modularity



# Визуализации. Шаг 3. Итог

979



человек

## Выводы. Пол

Мужчин 86%, а женщин 14%.

Согласно данным Mediascope, в среднем в аудитории ВКонтакте 54% женщин и 46% мужчин.

В исследуемом ядре очевидно превалируют мужчины (на 32% больше мужчин, чем в среднем ВКонтакте), в то время как женщин подобный контент почти не интересует (на 40% меньше, чем в среднем ВКонтакте).

# Выводы. Возраст

## От 20 до 34 лет

В ядре подписчиков преобладает аудитория 25-34 (50% подписчиков).

В сравнении с общей аудиторией ВКонтакте, показатель больше средних значений на 22-24%.

Также больше средних значений аудитория 12-24. Для детального рассмотрения делим на три возрастные группы: 12-15, 16-19, 20-24. В 15-16 лет человек обычно заканчивает школу и поступает в университет. В 19-20 лет обычно заканчивает бакалавриат. Процент аудитории до 20 лет – 5%, а от 20 до 24 – 24%.

Получается, что наибольший интерес к контенту группы проявляют совершеннолетние молодые люди от 20 до 34 лет.

# Выводы. Регион

Россия, Казахстан, США, Восточная и Центральная Европа + бывшие колониальные острова у берегов Новой Зеландии и в Карибском море

Хотя нам стоило бы выбрать не заливку, а кружочки для визуализации, так как данные не нормированы на площадь или на население, на этой визуализации мы хотели скорее выявить регионы, в которых возникает интерес к контенту паблика. И с этой задачей наша визуализация справляется.

Визуализация выполнена в QGis.

# Выводы. Регион



# Выводы. Регион

Интересно, что у некоторых подписчиков указаны как родная страна непризнанные государства и зависимые территории.

Это коррелирует с интересами ядра.

# Выводы. Интересы

Интересы сообщества можно разделить на три группы: история, новости (оппозиционный взгляд), эстетика упадничества в немейнстримных регионах.

# Выводы. Интересы. История

Паблики:

“Застоялась”, “мир обещал быть”, “Конец истории”,  
“Когда боги смеялись” (все четыре – из того же  
семейства пабликов, что и анализируемые нами  
группы),

“HistoryPorn”, “Нечаевщина”, “Десятие”, “Нулевые”,  
“Память”, “Интересные факты”, и т.д.

# Выводы. Интересы. Новости

Паблики:

“Лентач”, “Медуза”, “РБК”, “Телеканал Дождь”,  
“Настоящий Лентач”, “Свободные новости”,  
“МБХ Медиа”, “Сталингулаг” и т.д.

# Выводы. Интересы. Эстетика

Паблики:

“Русская смерть”, “Эстетика ебеней”, “В Пхеньяне всегда солнечно”, “Неизвестная Россия” и т.д.

# Итого

Типичный подписчик пабликов “Хунта Экспресс”, “Балканский Экспресс” и “Она развалилась” – мужчина от 20 до 34 лет. Родом из/ живёт в России, США, Казахстане, Восточной или Центральной Европе, а также (возможно) в непризнанных государствах, бывших колониях и зависимых территориях.

# Спасибо большое за внимание!

Весь проект на Github –

<https://github.com/undine-su-menulio/Junta-Project>

Автор проекта: Мария Казакова, Москва, 2020 год.

Почта: marikasakowa@gmail.com

Instagram: [@undine\\_su\\_menu](https://www.instagram.com/undine_su_menu)