**Отчет по работе с Gephi**

1. **Общие сведения, и выбор данных для анализа**

Gephi — самый известный инструмент для визуализации графов и сетевого анализа. С его помощью можно исследовать и население Викторианской Англии, и отношения героев «Войны и мира», и странички друзей Вконтакте. Рассказываем максимально просто, как пользоваться Gephi.

Для начала нам надо найти и скачать некоторые данные для анализа. Вот несколько ссылок:

<https://github.com/gephi/gephi/wiki/Datasets>

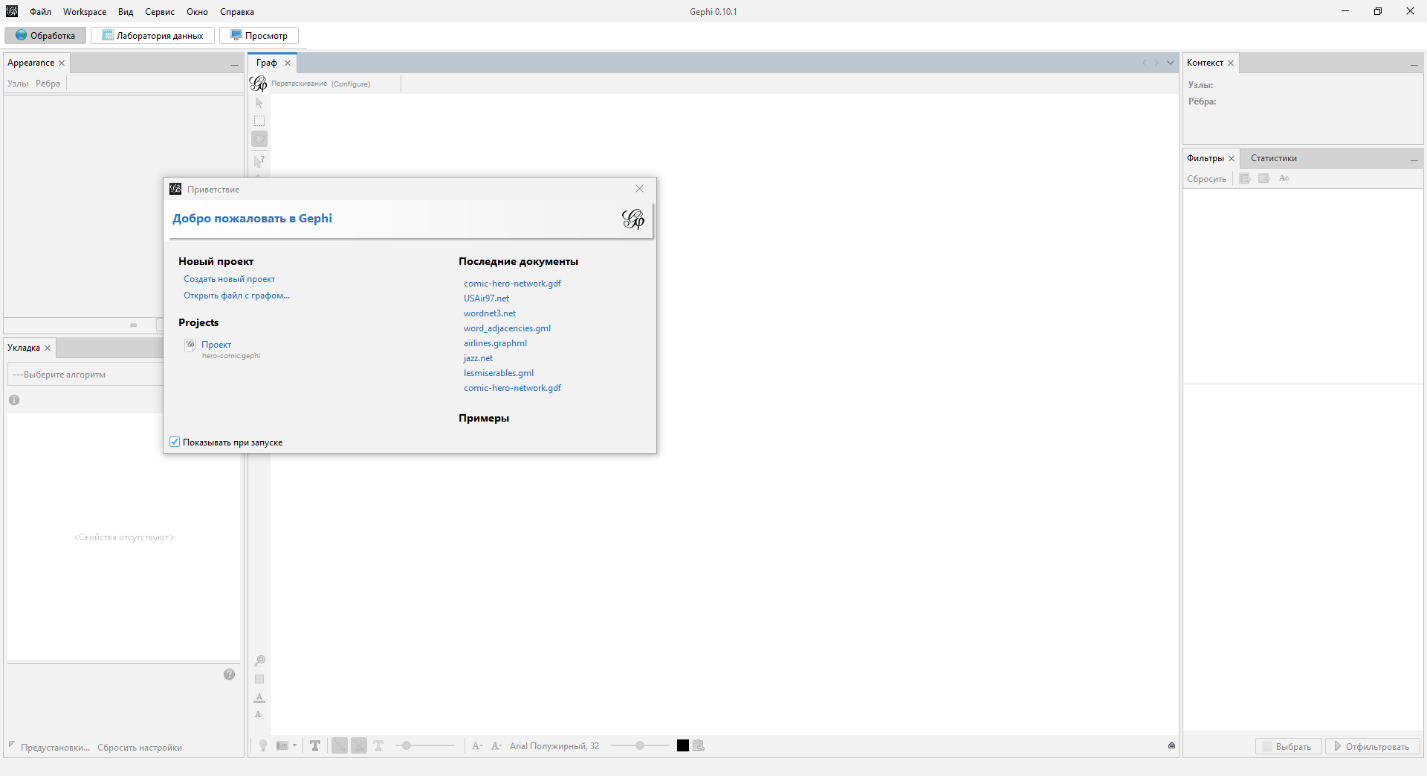
([GDF file](https://gephi.org/datasets/comic-hero-network.gdf.zip). **Comic and Hero Network** Same data as above, but this includes the comics the heroes appear in.)

<https://snap.stanford.edu/data/>

Я выбрал данные про - **Сеть комиксов и героев** супергероев Marvel в которые входят комиксы, в которых появляются герои.

1. Визуализация

После открытия Gephi нас встречает такое окно



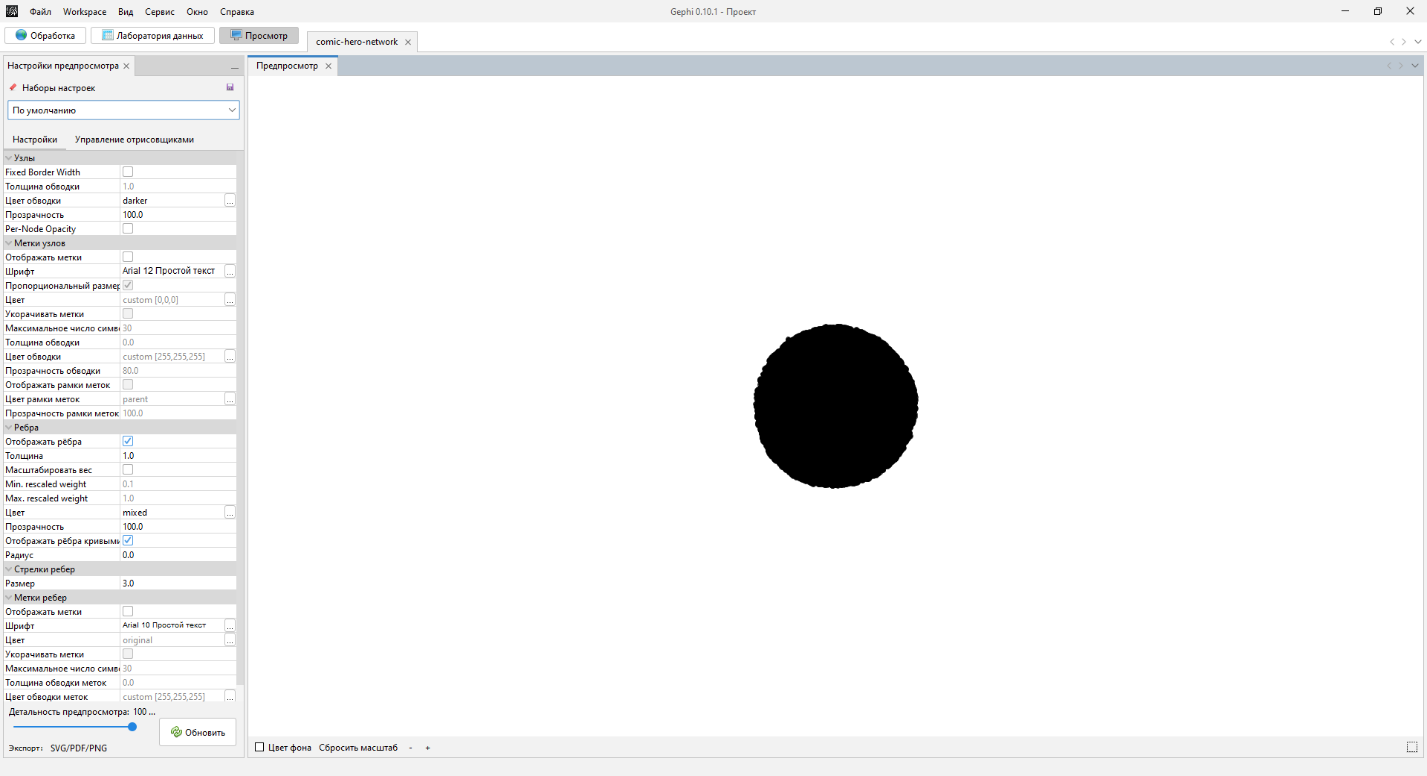
После этого надо выбрать скачанный файл для анализа.

Мы сразу увидим, сколько узлов и рёбер в нашем графе. выбирать здесь ничего специально не нужно. А в будущем ещё будет важно проверять, что наш граф добавляется в новое рабочее пространство (New workspace). Теперь просто жмем ОК.

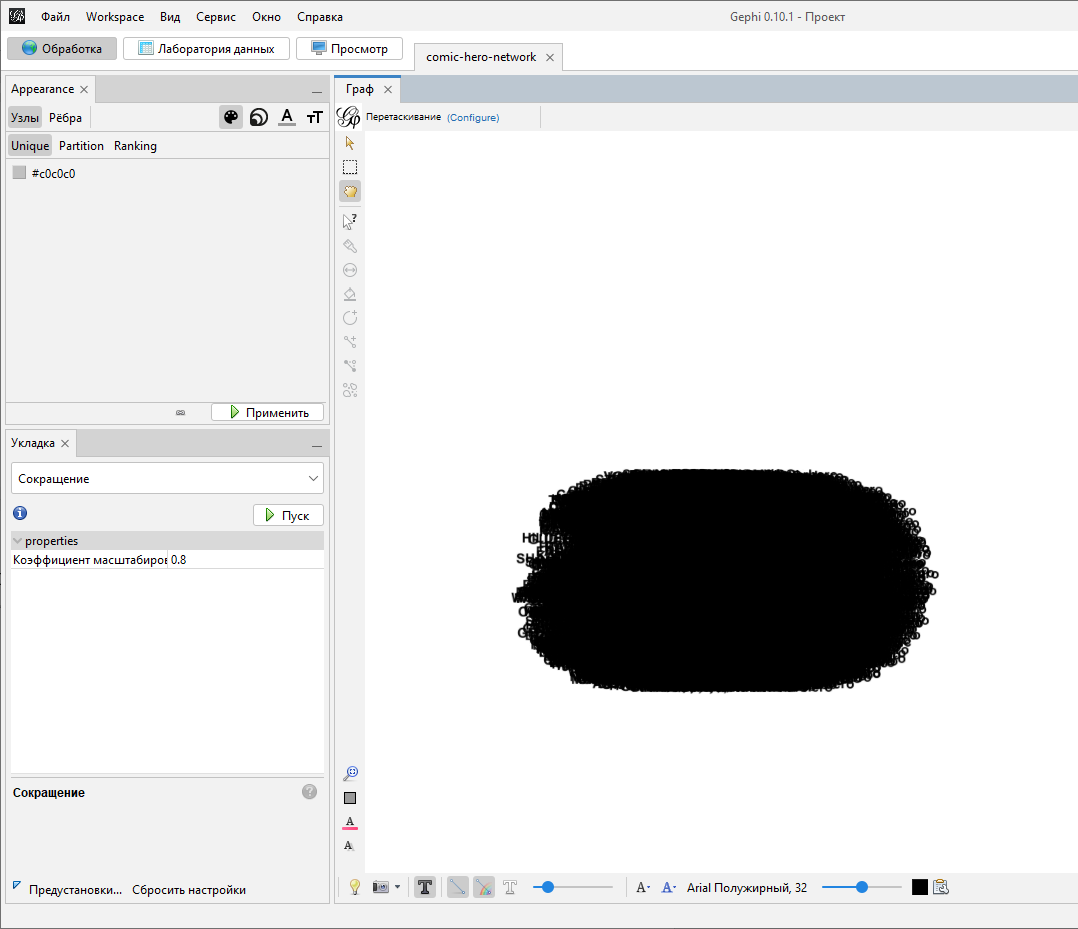
Обратимся к вкладке Просмотр: мы либо окажемся на ней автоматически после загрузки файла, либо перейдем сами. Все вкладки подписаны в самом верху.



Если данных много, то изначально будет просто черная масса. Как в моем случае.



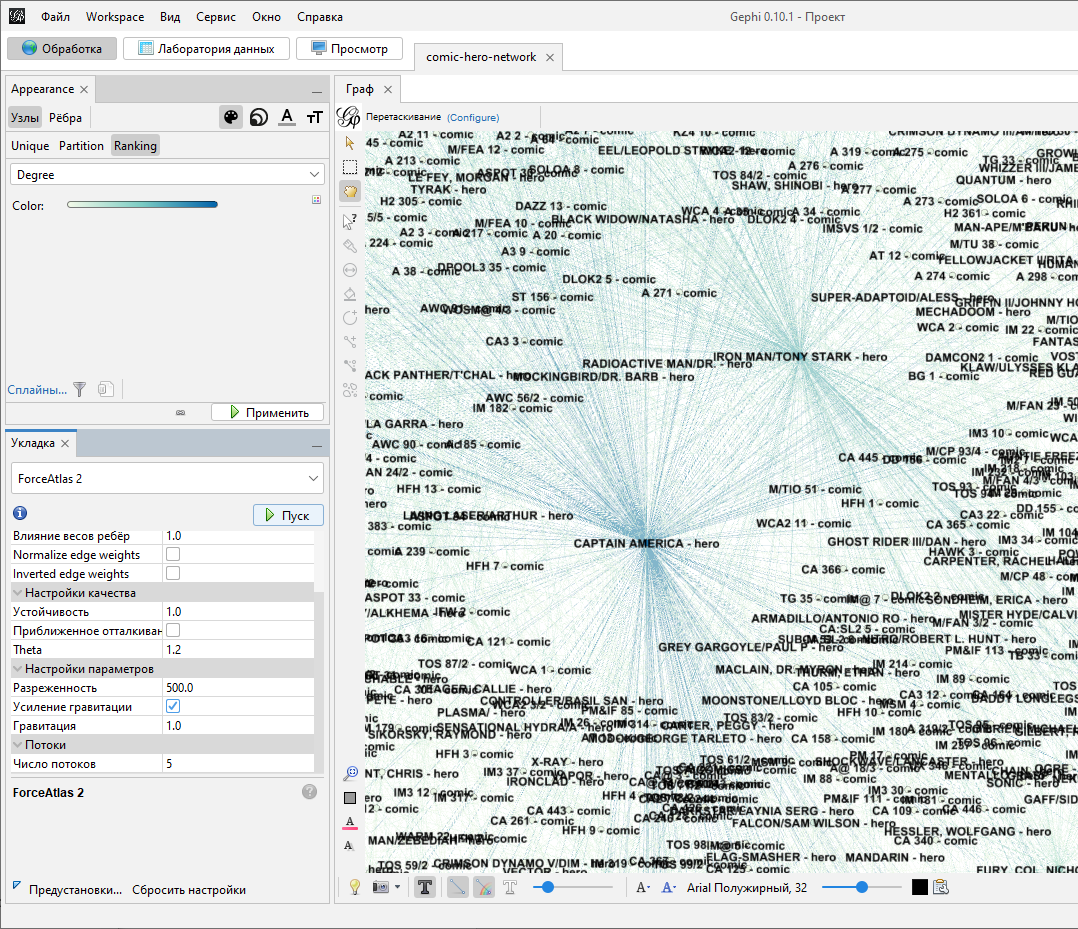
Чтобы сделать наш граф красивым и наглядным, а самое главное — понятным, нужно будет настроить цвет и размер узлов, их расположение, длину ребер и т. д. Для этого заходим во вкладку Обработка:



**Добавляем цвет**

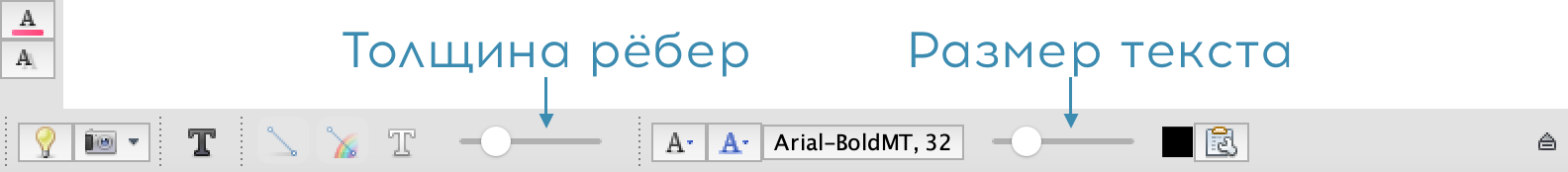
Первое, что мы сделаем, раскрасим наш граф. Это делается через вкладку Appearance — выбираем Nodes и значок палитры. По умолчанию все узлы и ребра раскрашены одним цветом (Unique). Помимо монохрома у нас есть два варианта раскраски:

* разными цветами по одному из признаков (Partition; в наш файл зашито несколько таких признаков: пол; принадлежность к группе, например «Купцы»; количество произнесенных за пьесу слов);
* одним цветом, но будет отличаться его интенсивность и яркость: по признакам «количество произнесенных слов» и «степень» (количество связей этого узла) (Ranking > Degree), чем больше значение признака, тем ярче цвет.



Укладка графа(упаковка)

Следующий наш шаг — добавить графу ясности и включить подписи узлов. Это делается по нажатию на букву Т на нижней панели под графом:



Сейчас имена могут налезать друг на друга и сбиваться в кучу. Ничего страшного: всё это решается укладкой. Изначально, весь граф уложен случайным образом, все элементы расположены хаотично. Но у Gephi есть несколько встроенных способов укладки. Сейчас мы порекомендуем вам ForceAtlas2, вы же можете опробовать все другие и выбрать понравившийся. Чтобы укладка состоялась, необходимо нажать «Пуск», а после этого «Стоп», когда граф примет нужный вид, чтобы он больше не двигался. В этом же меню мы можем выбрать «Укладка меток», чтобы наши названия узлов (метки) точно не наслаивались друг на друга.



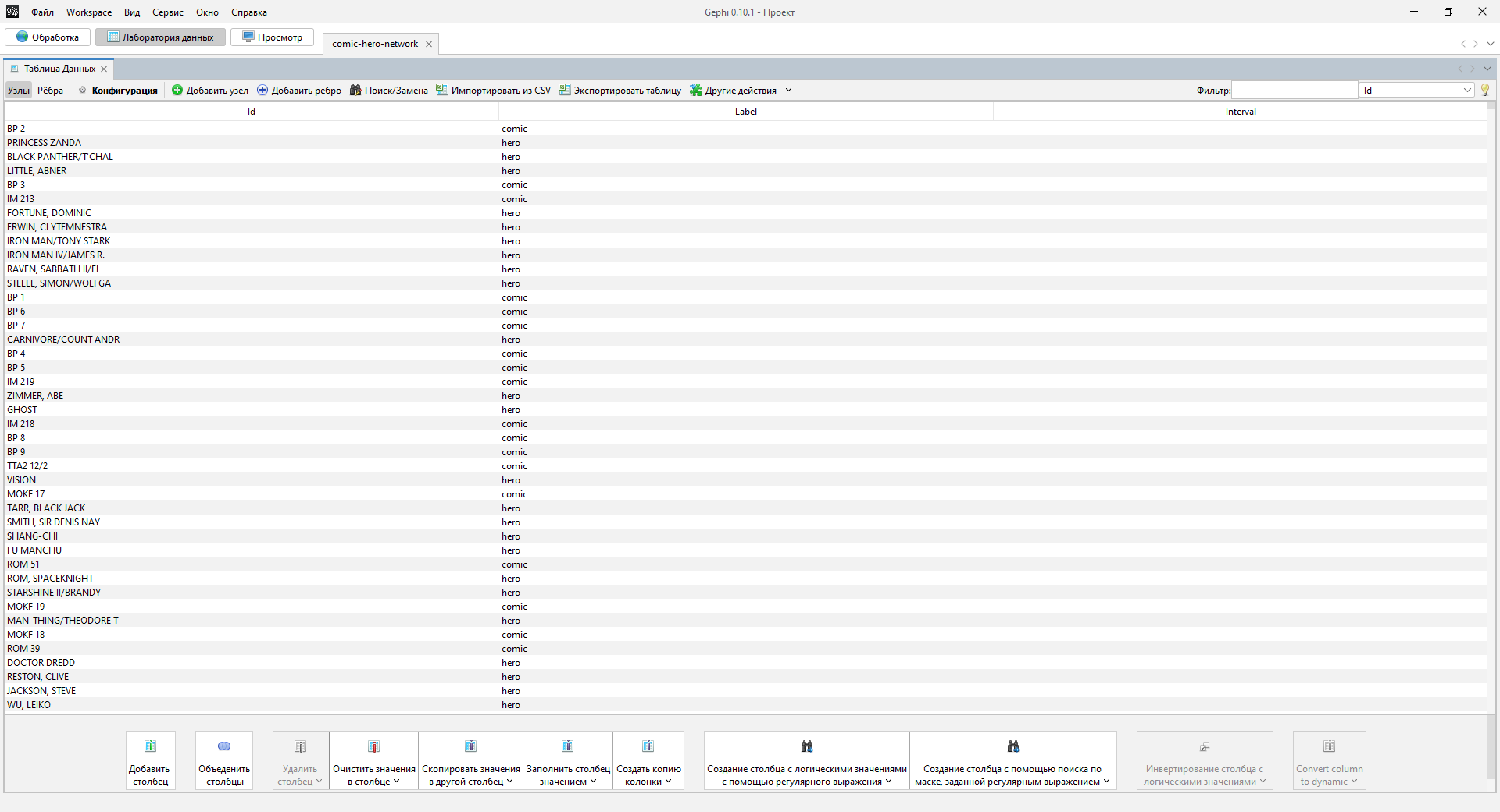
Вот так это будет выглядеть после укладки и с добавлением подписей.

Размер узлов

Последнее, о чём мы поговорим из базовых функций — это размер узлов. Это делается также во вкладке Appearance — Nodes, но уже под значком совмещенных кругов. Мы также можем задать единый размер для всех или проранжировать по какому-то признаку. 

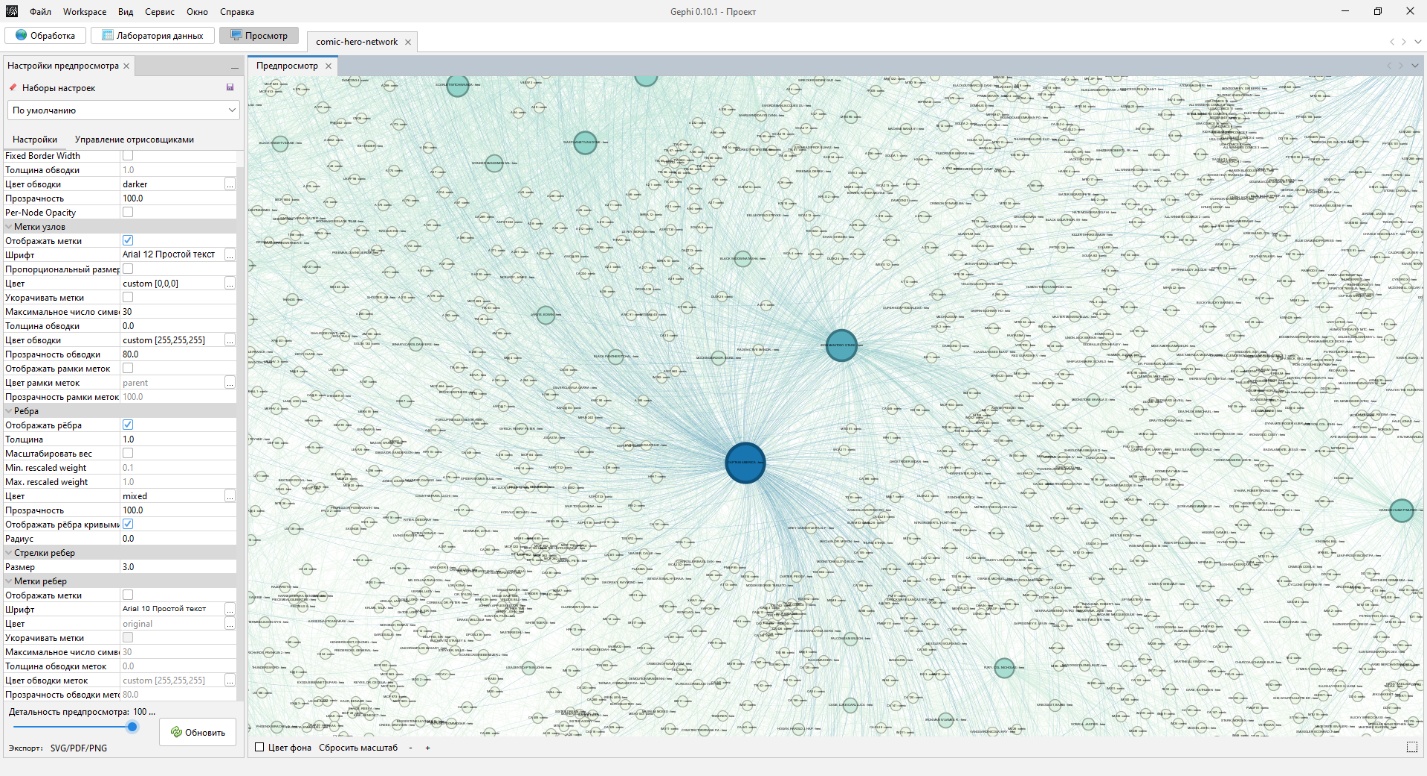
Лаборатория данных

Следующий этап знакомства с Gephi — вкладка «Лаборатория данных». Здесь мы можем посмотреть на все наши данные в табличном виде, увидеть ту самую цифру под визуалом. Мы можем рассмотреть каждую строчку, удалить что-то или добавить новое, можем добавить новый столбец — это будет новый признак, который будет характеризовать тех или иных героев в нашем случае.



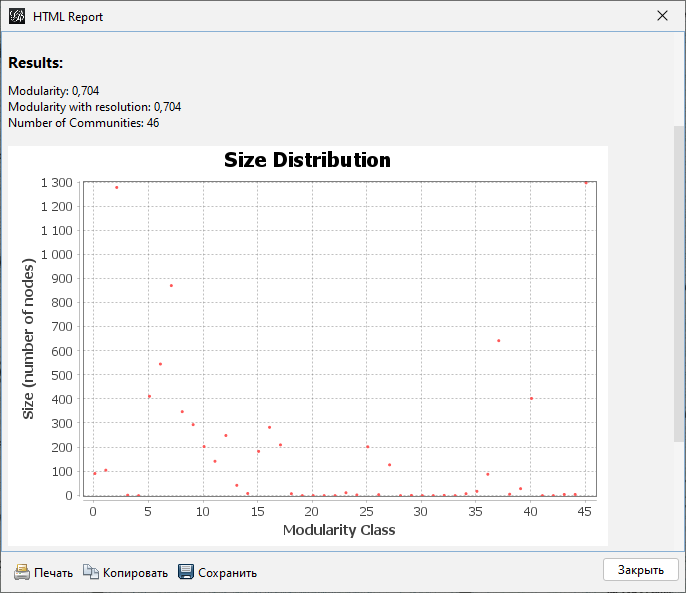
Просмотр

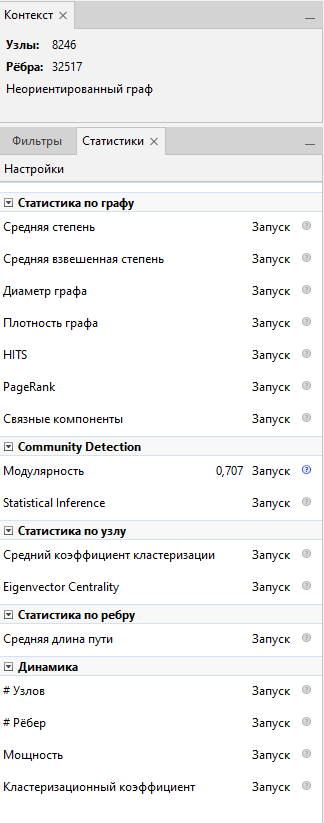
Финальный этап — мы снова возвращаемся во вкладку «Просмотр». Чтобы все наши наработки появились здесь, мы должны нажать «Обновить» (здесь это такое же правило как с «Применить» в «Обработке»: всегда жмём «Обновить» после любого изменения). Кроме того, нам потребуется заново включить отображение меток. Если они наслаиваются друг на друга, то рекомендуем снять галочку с «Пропорционального размера». Остальные же настройки оставим на ваш вкус, тут снова важно экспериментировать и не бояться пробовать. Здесь уже ничего нельзя «испортить», а любое действие можно поправить.



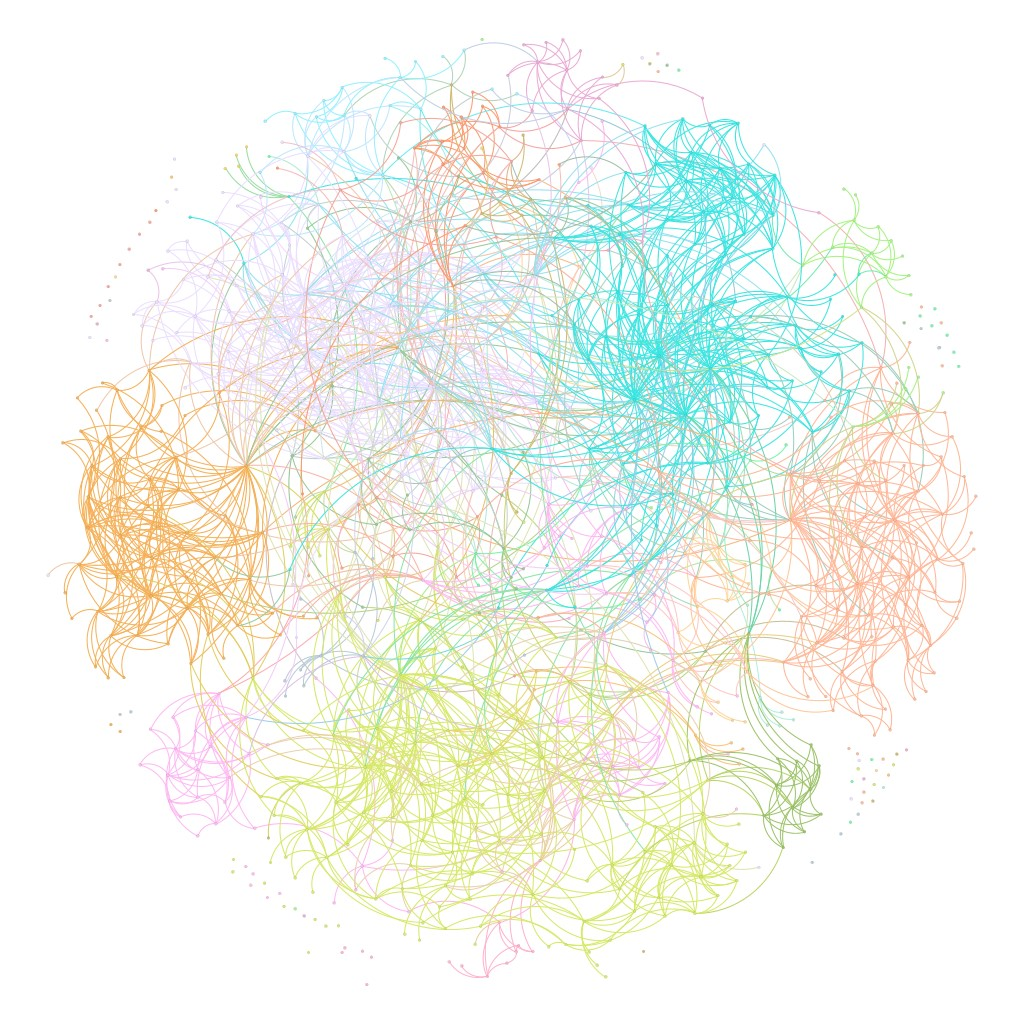
Теперь добавим статистику.

Сначала просчитаем модулярность:





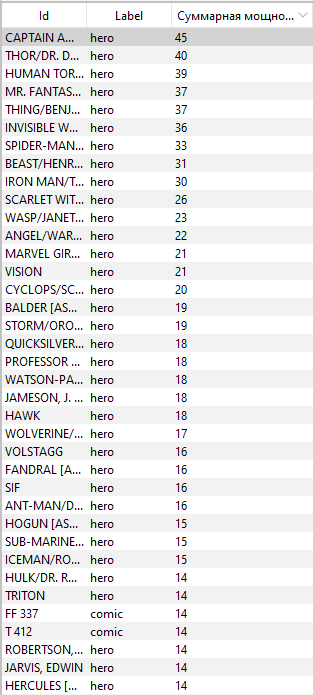
Теперь мы можем их раскрасить по модулярности.



Добавим Degree Centrality(кол-во связей)

в gephi это называется суммарная мощность

Капитан-Америка связывает большинство узлов.

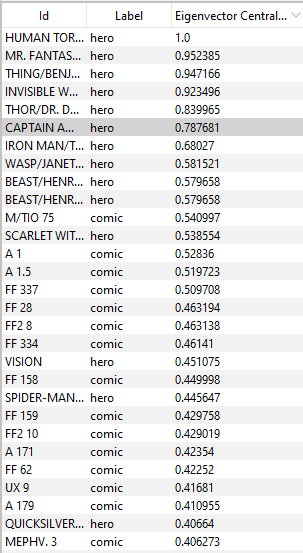


Посчитаем betweenness centrality (Кто является мостом между другими узлами)

В итоге Капитан-Америка является основным мостом между всеми узлами



eigenvector centrality (близость к важным узлам)



closeness centrality (Близость к всем)

