АНАЛИТИКА БОЛЬШИХ ДАННЫХ: ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

Кафедра «Анализ конкурентных систем»

ИНСТИТУТ МЕЖДУНАРОДНЫХ ОТНОШЕНИЙ МИФИ

Артамонов Алексей Анатольевич, к.т.н., доц.,
заведующий кафедрой «Анализ конкурентных систем» МИФИ

ТИПЫ ДАННЫХ

Структурированные – это данные, которые имеют формализованную структуру, например, база данных, где в искомой таблице строка представляет некий объект с характеристиками, представленными в соответствующих полях и заранее установленными типами данных.

Слабоструктурированные – это данные, которые имеют слабо формализованную структуру, например, текстовая информация, размеченная при помощи языков разметки, таких как HTML. Из-за вариативности визуального представления некого объекта, исходный HTML код страницы информационного ресурса может кардинально разниться.

Неструктурированные – это данные, которые имеют неформализованную структуру, например, текст письма электронной почты, сообщение в социальной сети и т.д.

РАБОТА С BIG DATA

V7

VOLUME
(объем)

VELOCITY
(скорость)

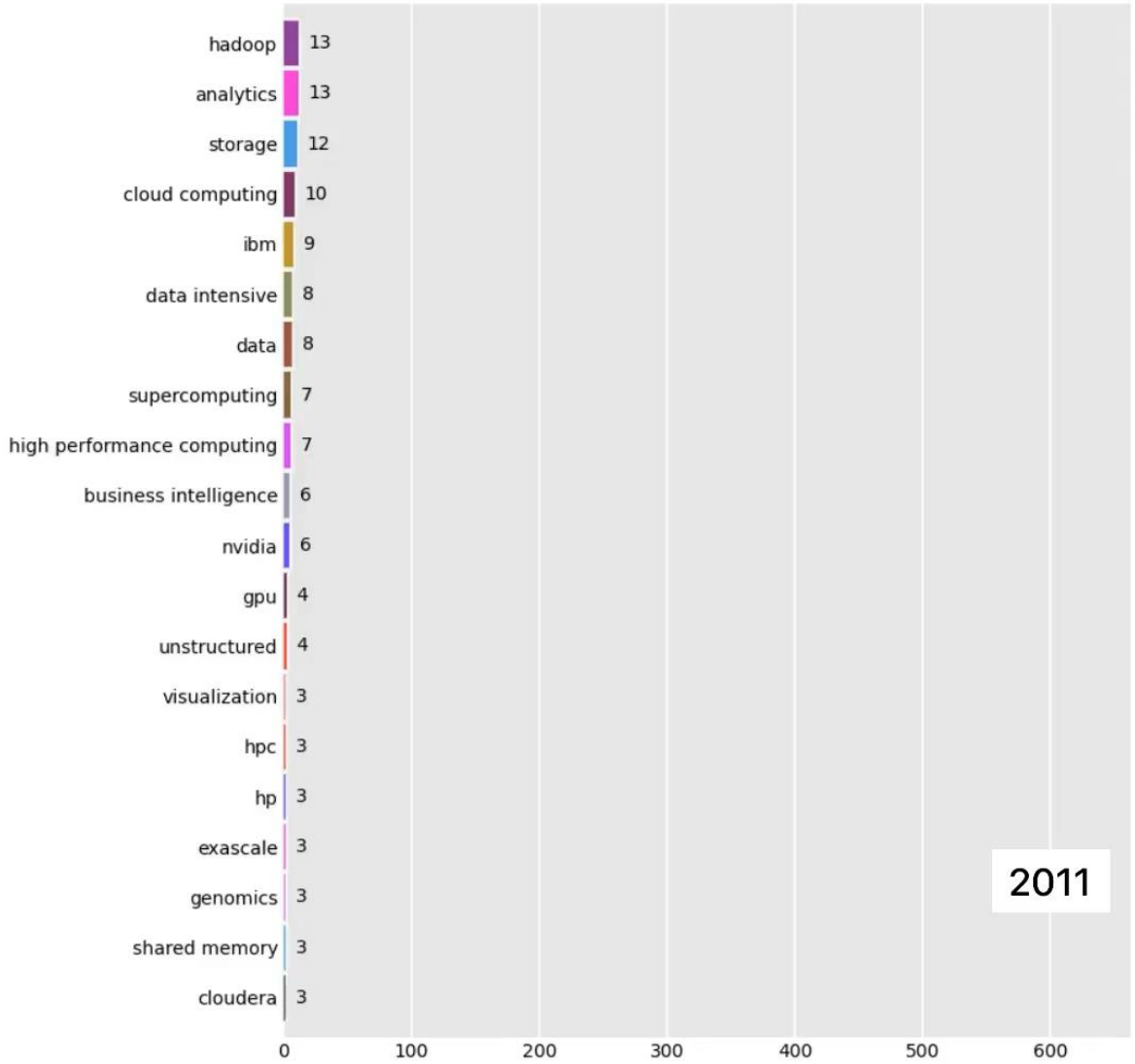
VARIETY
(разнообразие)

VARIABILITY
(изменчивость)

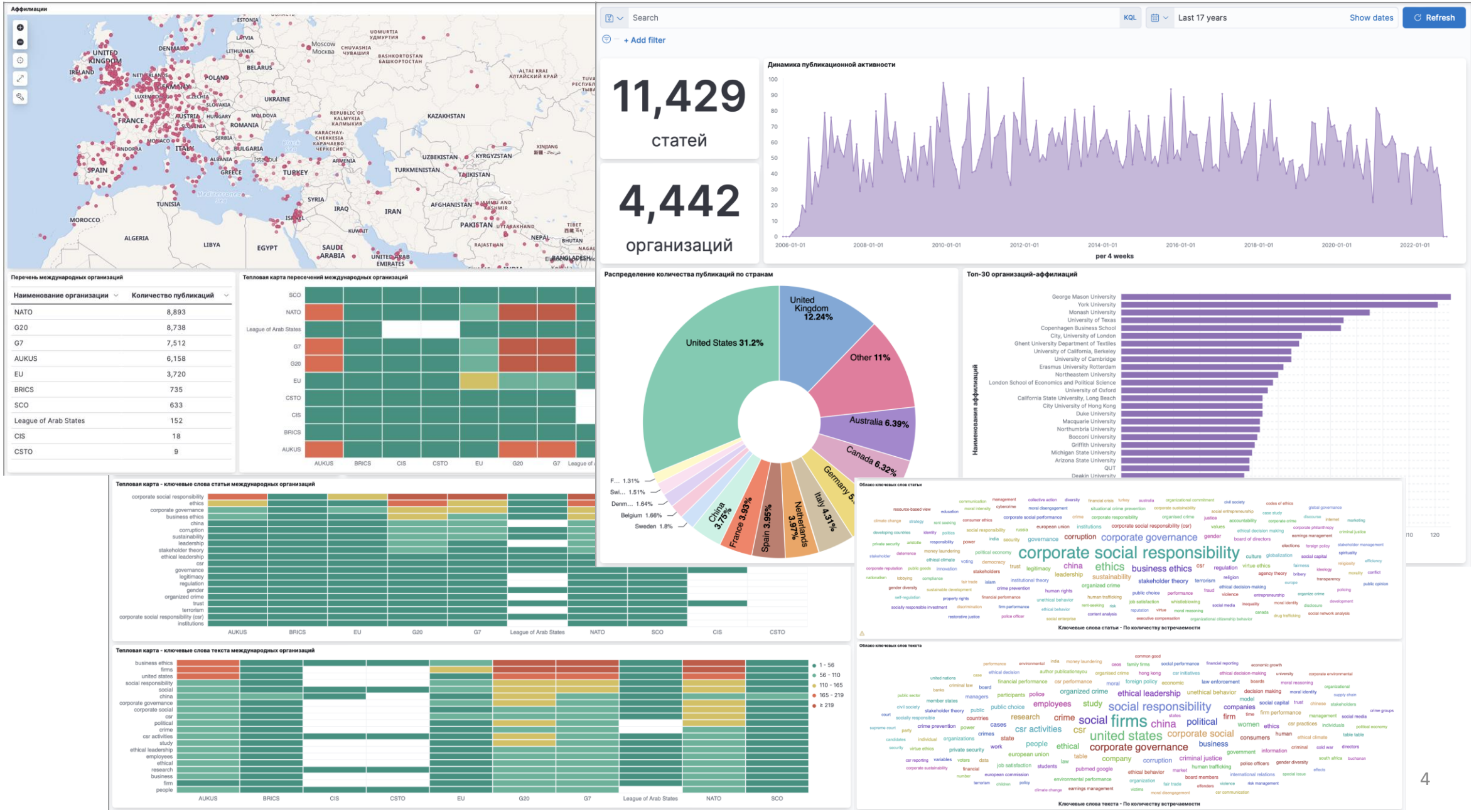
VALUE
(ценность)

VERACITY
(достоверность)

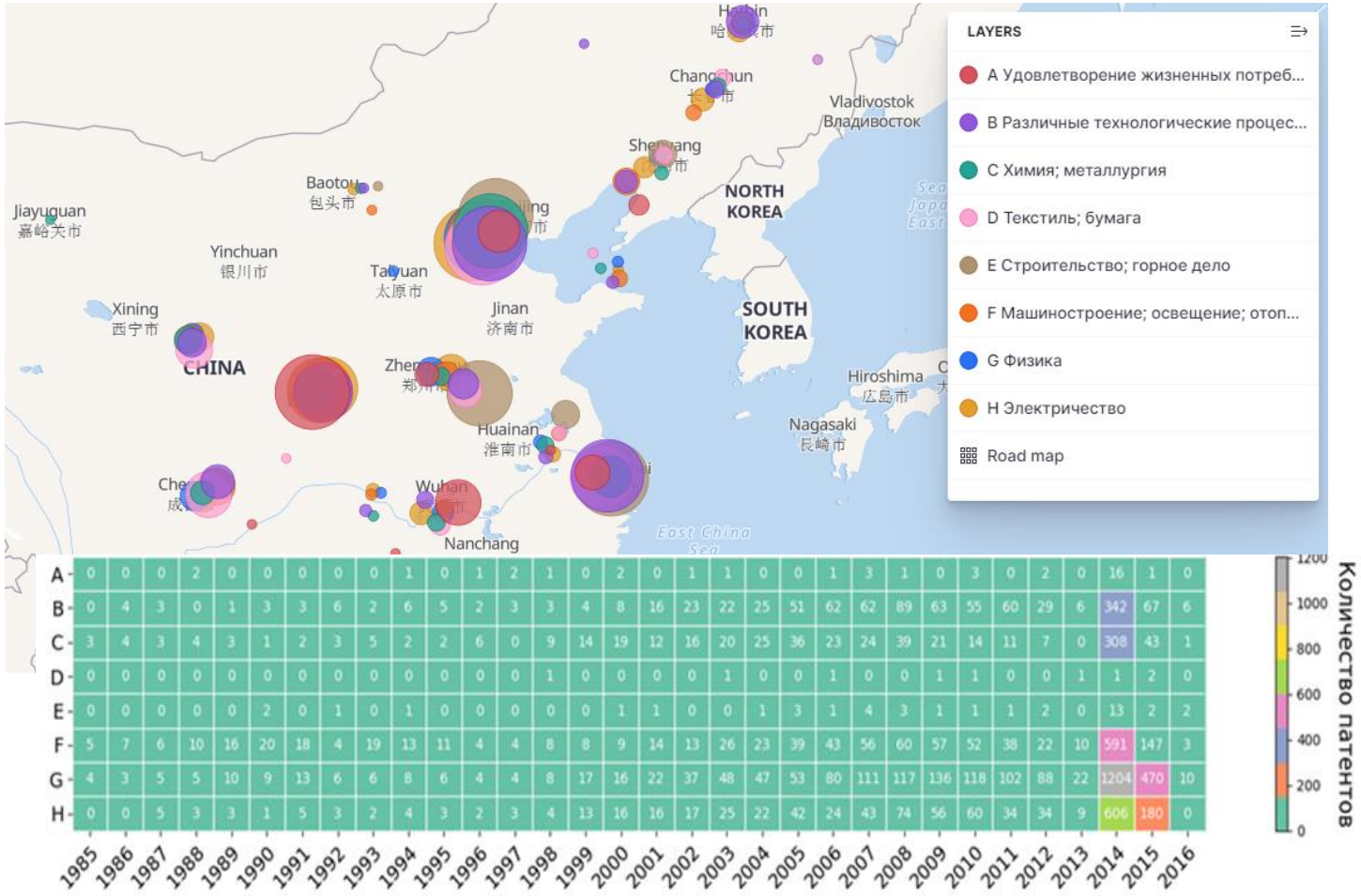
VISUALIZATION
(визуализация)



Озеро данных по публикациям в области Финансовой безопасности

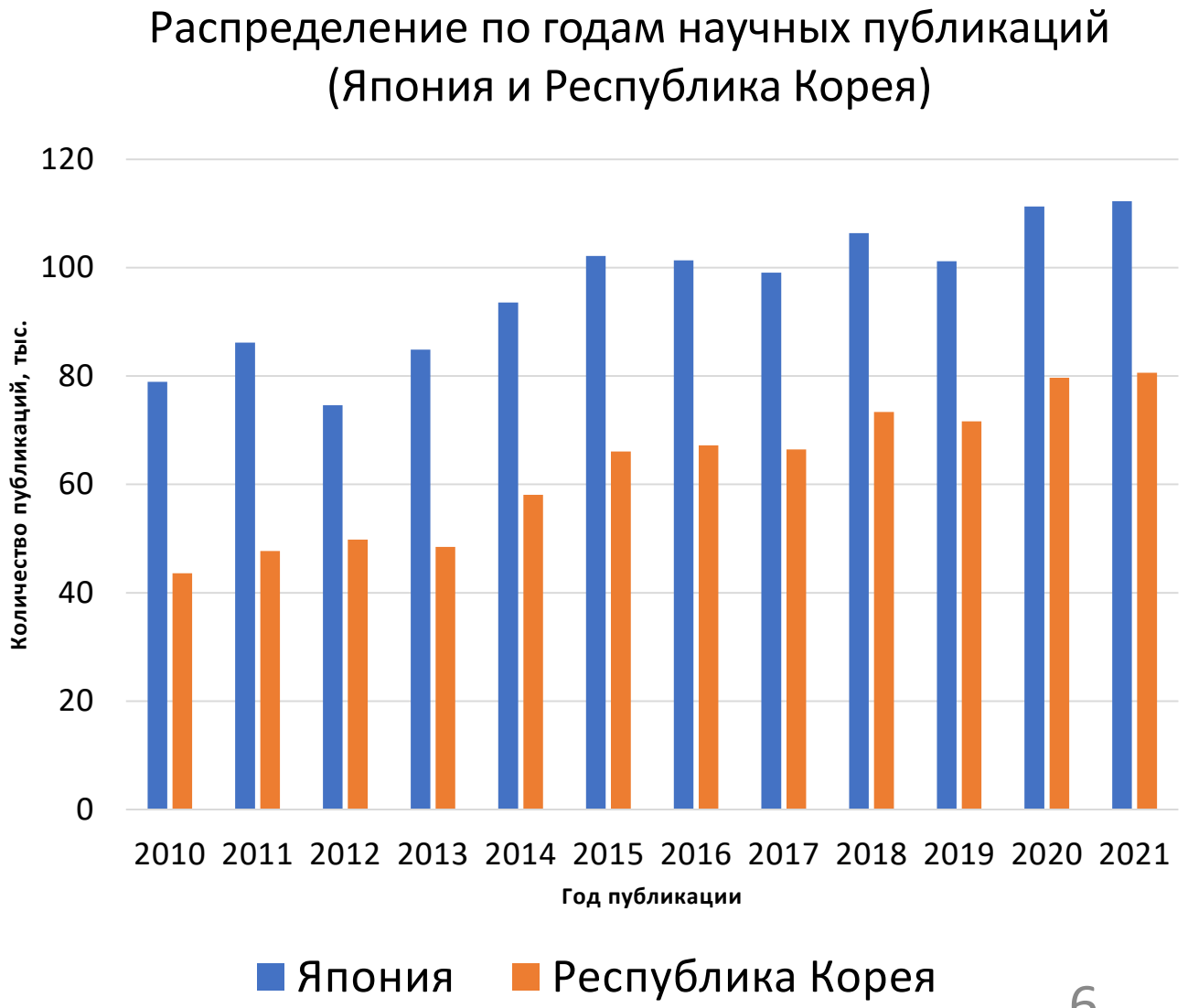


Исследование научно-технических направлений развития КНР по патентной информации (на китайском языке)



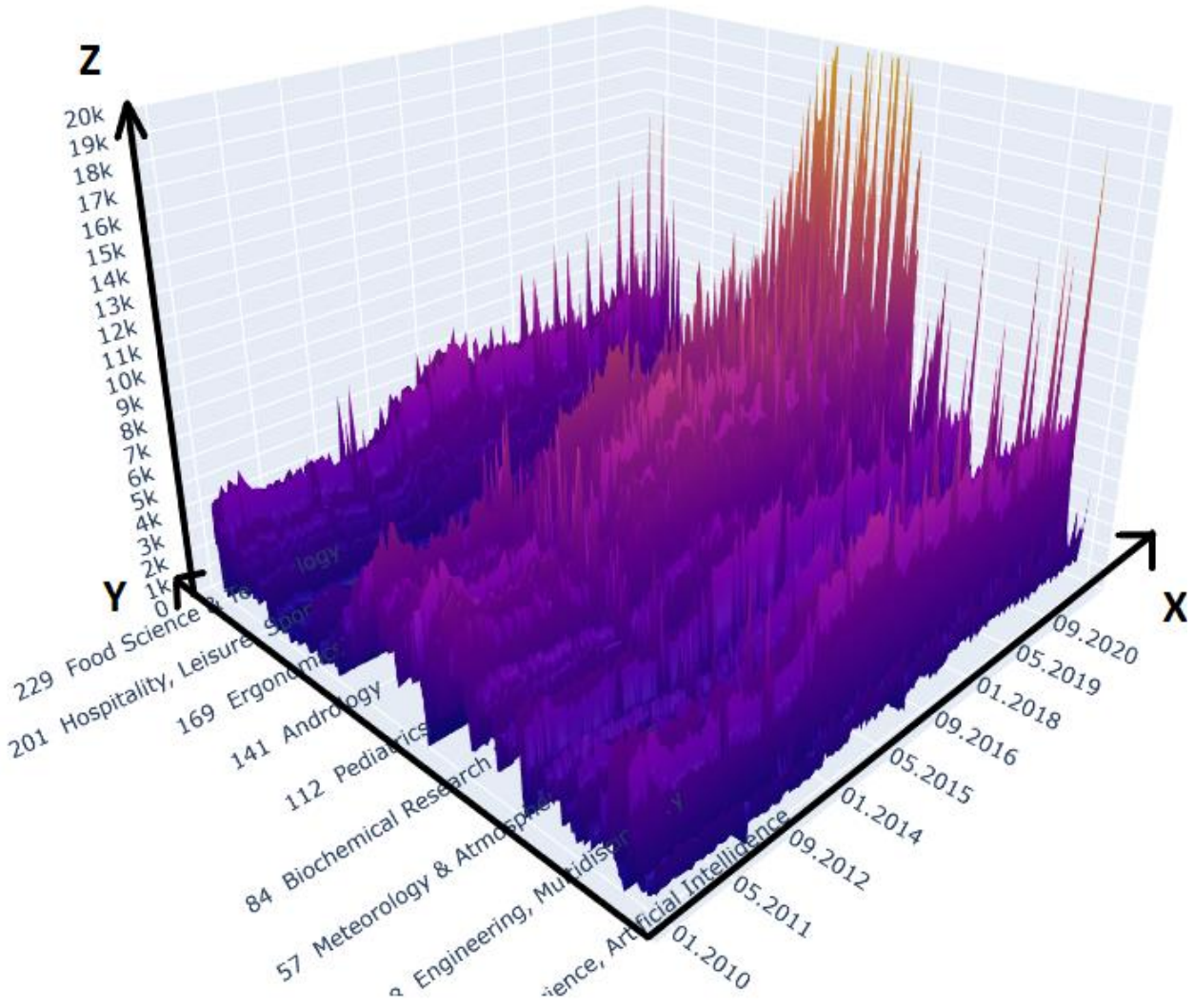
Формирование набора данных

	Япония	Республика Корея
Временной период	2010 – 2021 гг.	2010 – 2021 гг.
Количество научных публикаций	1 050 000	753 000
Объем данных	24.3 Гб	16.3 Гб

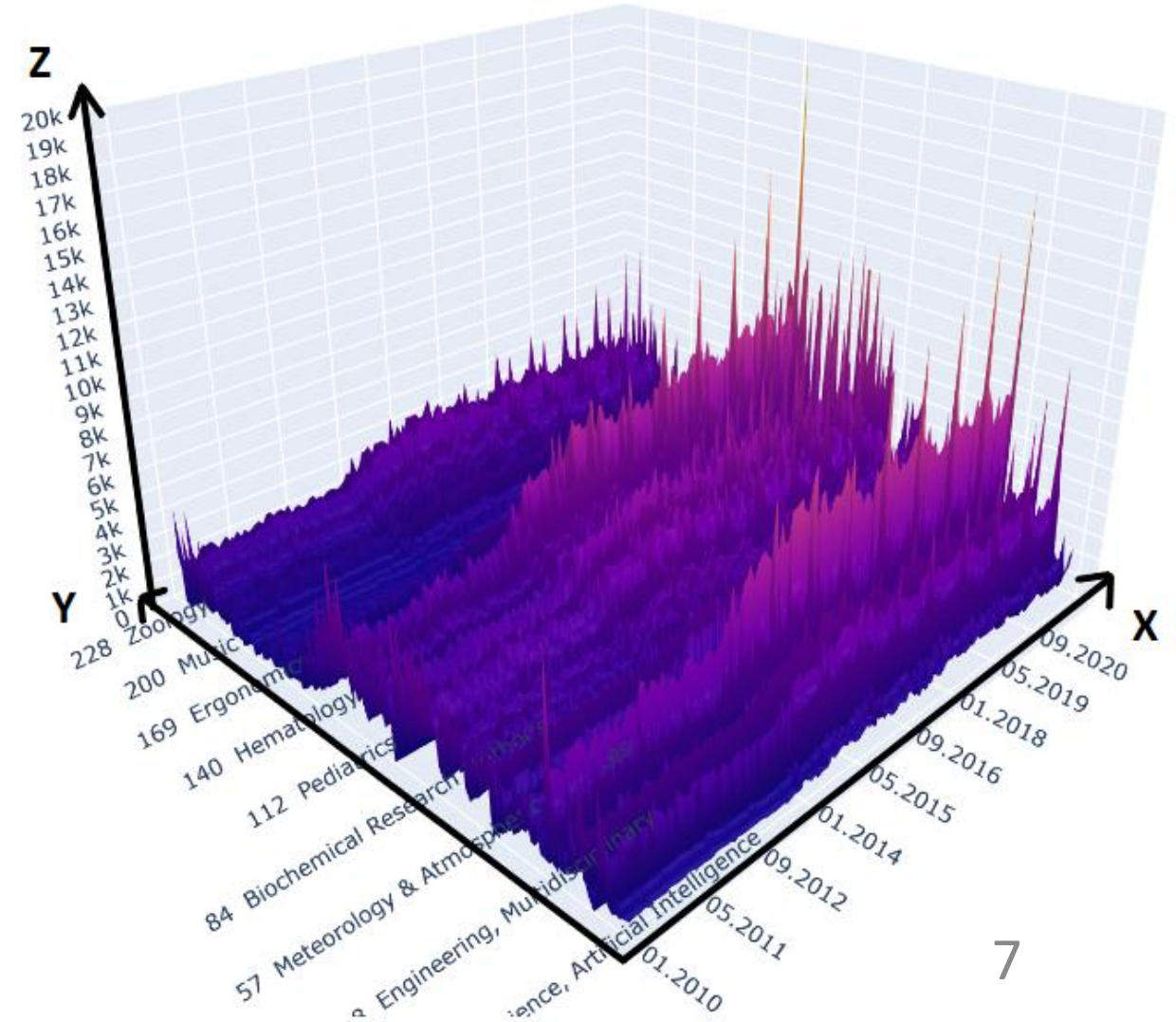


Научно-технологический ландшафт страны

НТЛ страны (Япония)

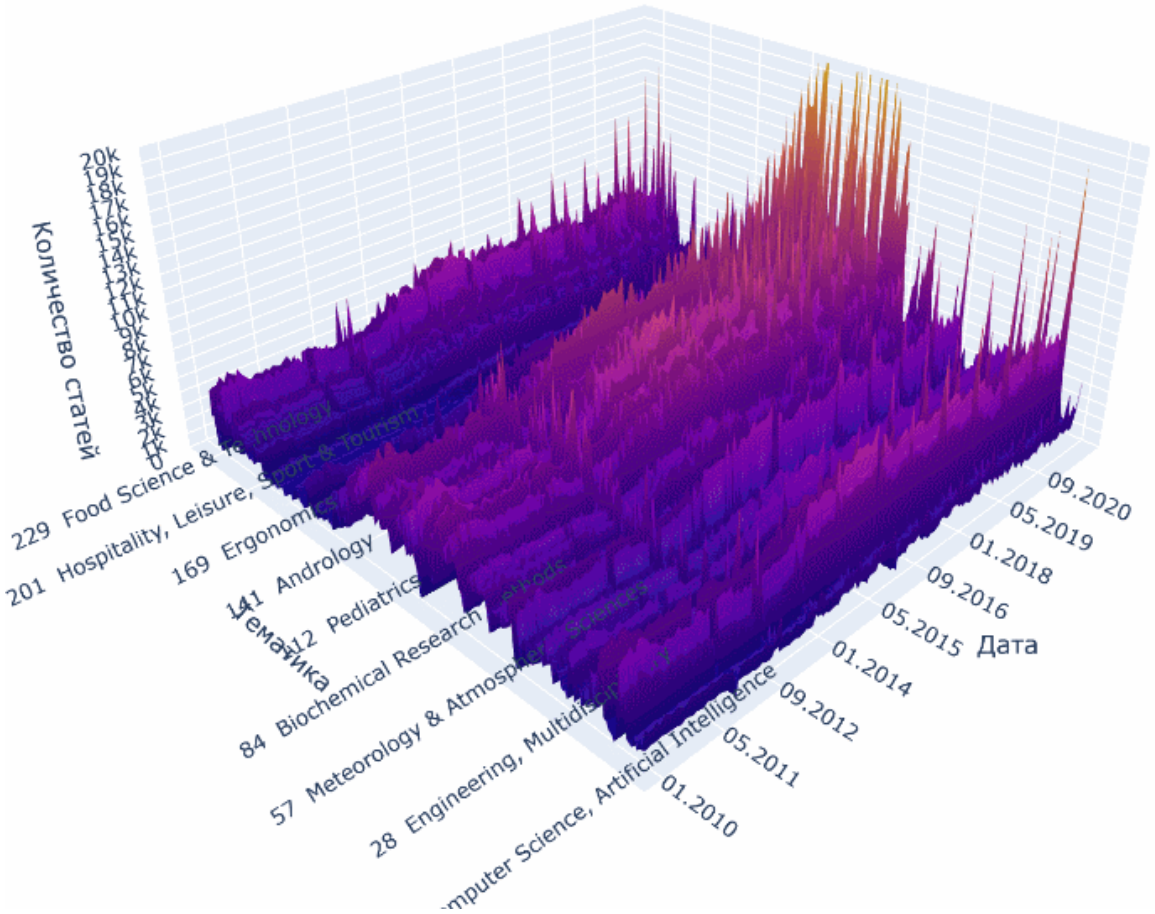


НТЛ страны (Республика Корея)

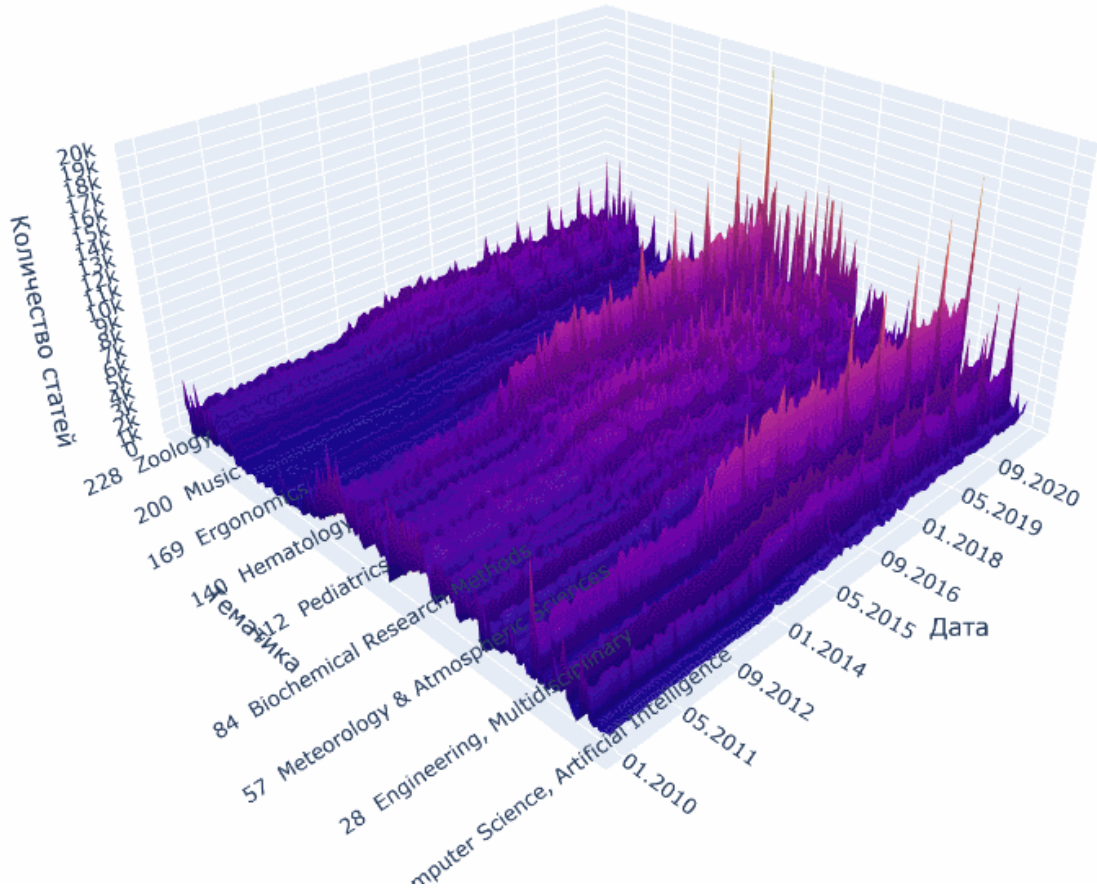


Научно-технологический ландшафт страны

НТЛ страны (Япония)

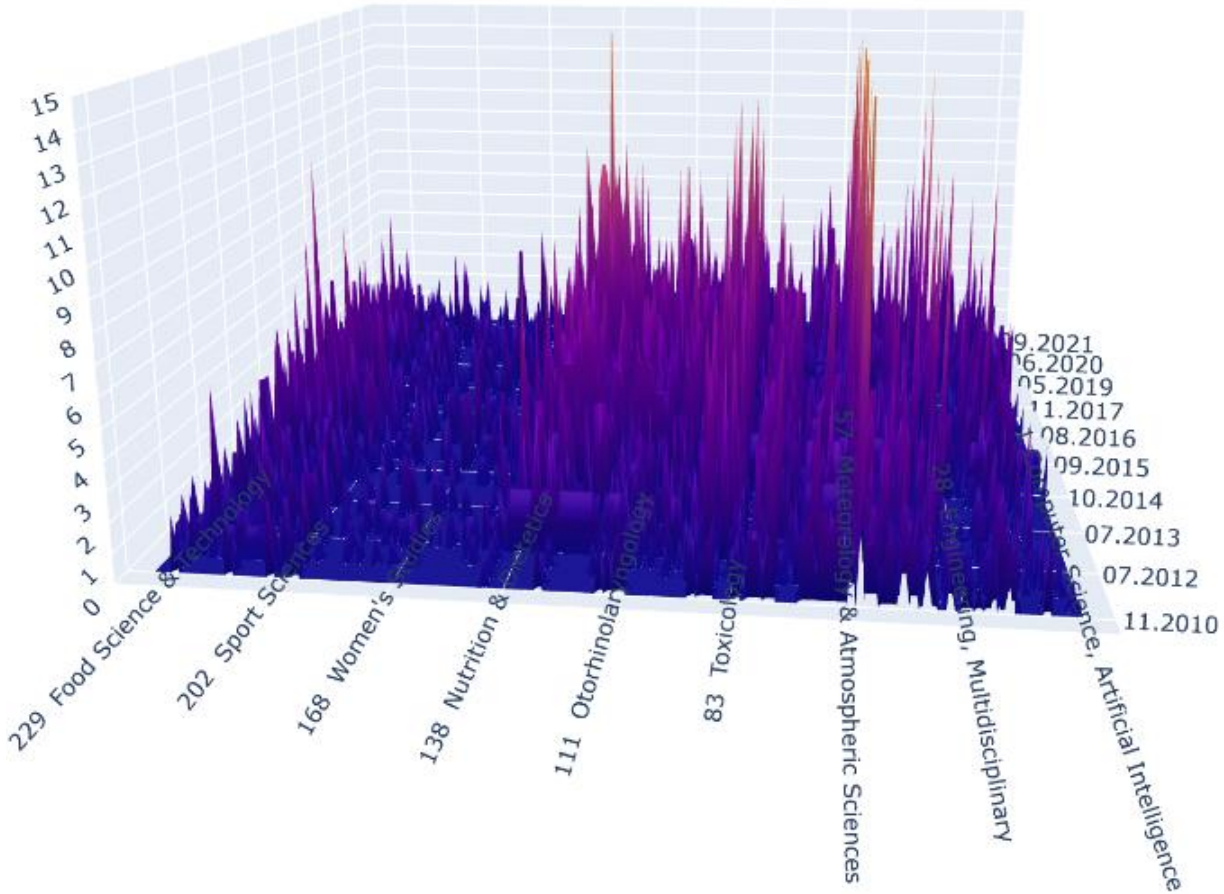


НТЛ страны (Республика Корея)

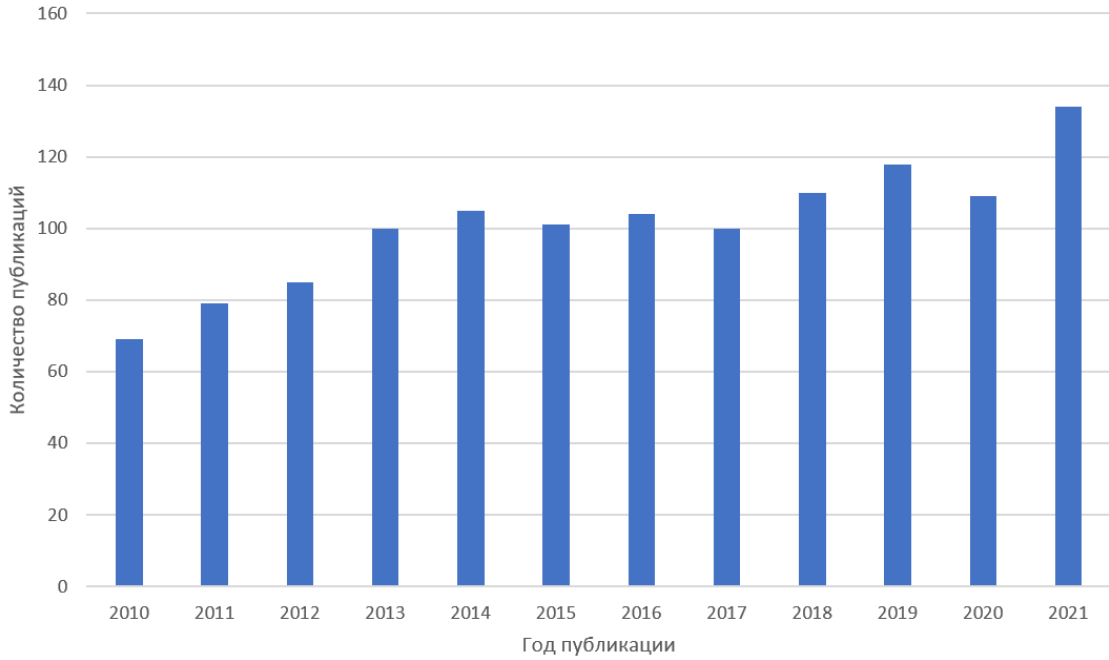


Научно-технологический ландшафт научной области

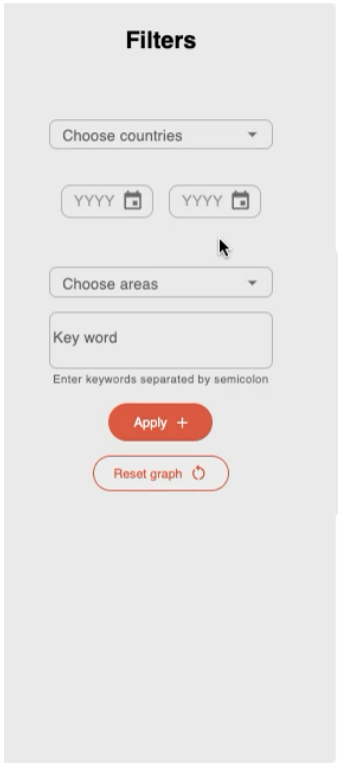
НТЛ области “FLASH терапия” (Япония)

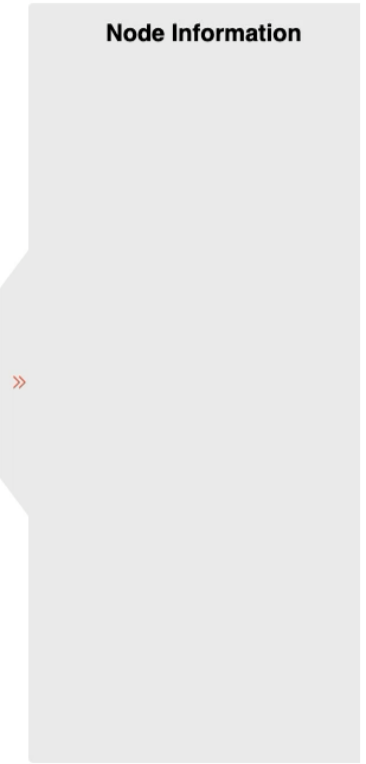


Распределение по годам научных публикаций по тематике "FLASH терапия"



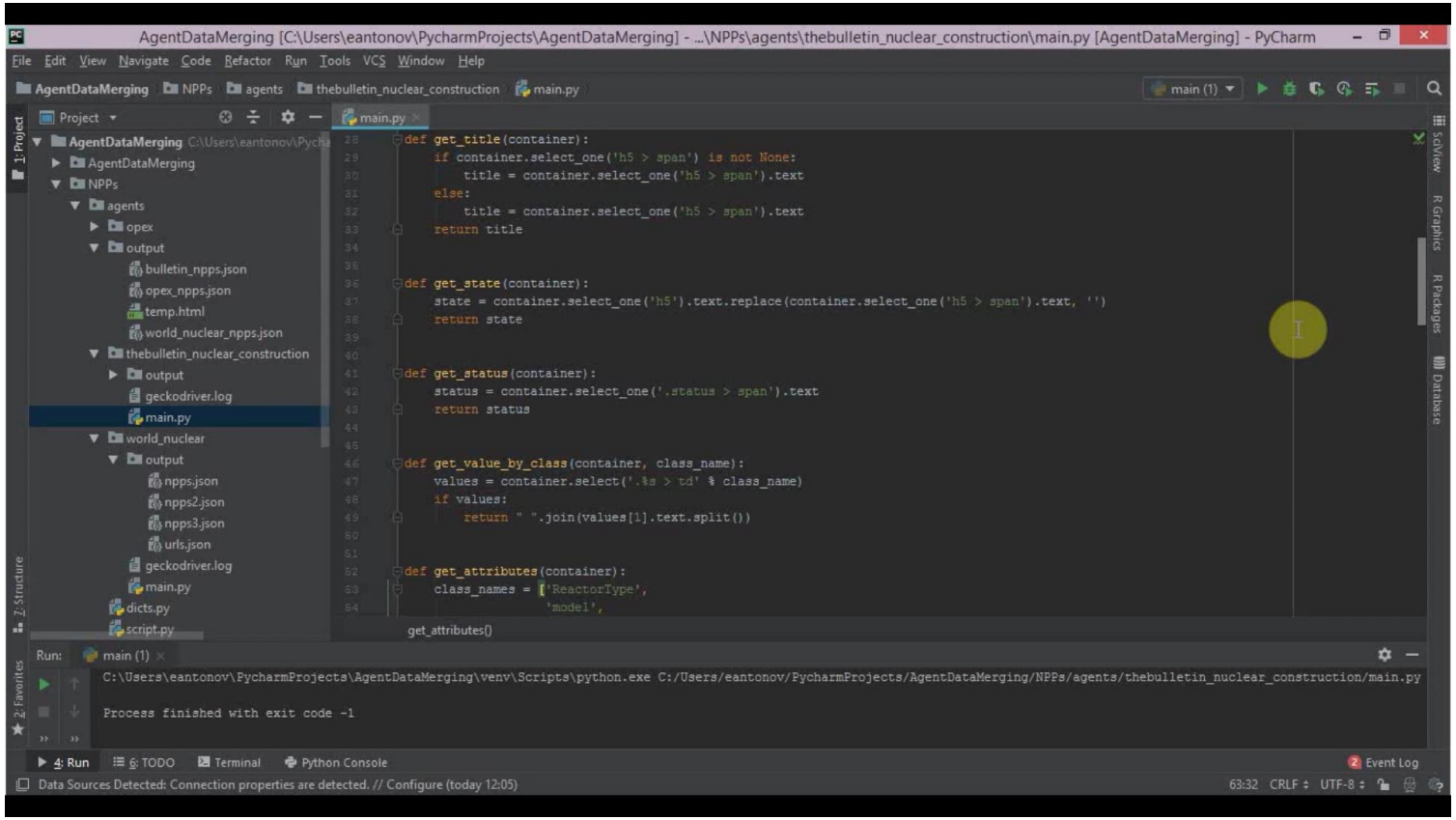
ВЫЯВЛЕНИЕ СВЯЗЕ МЕЖДУ ОБЪЕКТАМИ





АВТОМАТИЗАЦИЯ ДЕЯТЕЛЬНОСТИ

Разработка информационных агентов на языке программирования Phyton для формирования базы данных атомных станций



СПАСИБО ЗА ВНИМАНИЕ!