

О.Ю. Колесниченко¹, Г.Н. Смородин², Д.А. Яковлева³,
Л.С. Мазелис⁴, С.И. Баландин⁵, Ю.Ю. Колесниченко⁶

Текстовая аналитика Big Data

Второй этап исследования

Dell EMC External Research and Academic Alliances

¹Главный редактор информационного Бюллетеня «Анализ безопасности», PhD, Москва, Россия;

E-mail: oykolesnichenko@list.ru

²Руководитель Академического Партнерства Dell EMC в России и СНГ (Dell EMC External Research and Academic Alliances), PhD, MBA, Санкт-Петербург, Россия;

E-mail: gennady.smorodin@dell.com

³Ассистент кафедры информационных технологий и систем Владивостокского государственного университета экономики и сервиса, Владивосток, Россия;

E-mail: darya.yakovleva15@vvsu.ru

⁴Заведующий кафедрой математики и моделирования Владивостокского государственного университета экономики и сервиса, доктор экономических наук, профессор, Владивосток, Россия;

E-mail: lev.mazelis@vvsu.ru

⁵Президент Ассоциации Открытых Инноваций «FRUCT» (Finnish-Russian University Cooperation in Telecommunications), профессор Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики (Университет ИТМО) и Технологического университета Тампере (Tampere University of Technology, Finland), PhD, MBA, Санкт-Петербург, Россия;

E-mail: Sergey.Balandin@fruct.org

⁶Директор информационного портала «Бюллетень «Анализ безопасности», www.SecurityAnalysisBulletin.com, Москва, Россия.

E-mail: green-apple_2000@mtu-net.ru

© О.Ю. Колесниченко, Г.Н. Смородин, Д.А. Яковлева, Л.С. Мазелис, С.И. Баландин, Ю.Ю. Колесниченко, 2016

© Dell EMC External Research and Academic Alliances, 2016

© Национальный Суперкомпьютерный Форум, 2016

Аннотация. Представлен второй этап исследования по текстовой аналитике Больших данных (Big Data), проводимого Академическим Партнерством Dell EMC в России (Dell EMC External Research and Academic Alliances). Исследование названо «Третья волна» в честь концептуальной работы американского философа Элвина Тоффлера. В анализ вошли 49 стран из разных регионов мира. Выполнялась текстовая аналитика Big Data из открытых ресурсов Интернета, при выполнении Data Mining использовался доступ через API к текстам дата-центров Google и Яндекс, что методически рассматривается как использование Google и Яндекс в качестве неклассических суперкомпьютеров в режиме «as-a-Service» Cloud model. Авторская методика по технике Data Mining через API-Google и API-Яндекс разработана Д.А. Яковлевой. Собрана морфологическая матрица метаданных из 20 ключевых словосочетаний (Morphological Matrix). Анализ слов осуществлялся с использованием модифицированной карты «Цепочка ценности» (Value Chain Map) Саймона Водлей (Simon Wardley), которая позволяет проводить стратегический анализ развития рынка. Выполнен кластерный анализ k-means (Д.А. Яковлева). Страны распределены на три типа кластеров:

высокой, средней или низкой активности в Интернете (High, Middle, Low level of Internet activity and citing). Кластерный анализ морфологической матрицы ключевых слов позволил выявить важный индикатор смены технологического уклада – повышение цитируемости словосочетания «облачные технологии» с 2008 года к 2015 году со сменой кластера. Для трактовки результатов исследования проведен краткий обзор научно-технического развития в разрезе Кондратьевских волн, показывая роль ключевых слов из морфологической матрицы для каждой из Кондратьевских волн. Страны сравнивались по регионам. Анализируя соотношение паттернов ключевых слов, выявлялись прогнозируемые лидеры экономического развития для каждого из регионов.

Ключевые слова и фразы: API-доступ, текстовая аналитика Больших данных, морфологического анализ, Google, Яндекс, неклассические суперкомпьютеры, Data Mining, кластерный анализ К-средних, карта «Цепочка ценности», Кондратьевские волны.

Порядок разделов статьи:

Введение (3).

Методика (4).

Основные результаты и обсуждение (7):

Кластерный анализ стран (8);

Кластерный анализ морфологической матрицы (8);

Модифицированная карта «Цепочка ценности» (Value Chain Map) (14);

Анализ модифицированной карты Value Chain Map по регионам (28).

Заключение и выводы (36).

Список литературы (37).

Приложение 1. *Географический захват начала Шестой Кондратьевской волны и воздействие неравенства и стресс-нагрузки (41);*

Приложение 2. *Пространственное двухмерное соотношение ключевых слов морфологической матрицы (49).*

Введение

Представлен второй этап исследования по текстовой аналитике Больших данных (Big Data), проводимого Академическим Партнерством Dell EMC в России (Dell EMC External Research and Academic Alliances). Первый этап исследования был доложен на Национальном Суперкомпьютерном Форуме в 2015 году [3]. Часть результатов второго этапа ранее была доложена на VI Международной социологической Грушинской конференции ВЦИОМ [4].

Исследование названо «Третья волна» в честь концептуальной работы американского философа Элвина Тоффлера [8]. Философ определил в истории цивилизации три крупных и последовательных периода, которые он называет «волнами перемен». Первая волна по Тоффлеру была сельскохозяйственной, Вторая волна – индустриальной, а нынешняя Третья волна – информационная цивилизация, скачок в инфосферу. Тоффлер описал свою теорию в 1980 году, когда еще не было Интернета и не возникла концепция Интернета вещей (Internet of Things, IoT), не появились Облачные технологии, 3D печать, мобильные телефоны, Большие данные. Однако Тоффлер предвидел главный революционный поворот в рыночных процессах, который он назвал «prosumer» (термин образован из двух английских слов «producer» – производитель, и «consumer» – потребитель). Экономика «Третьей волны» – это новая экономика, когда стирается исторически сложившийся разрыв между производителем и потребителем. IoT, Облачные технологии, мобильные приложения, 3D печать – все эти прямые, а не косвенные, черты информационной эры и стали сегодня причиной смены технологического уклада, причиной принципиальной трансформации рынка в направлении, которое описал Тоффлер, со стиранием различий между производителем и потребителем. Именно этот трансформационный процесс и изучает данное исследование при помощи текстовой аналитики открытых ресурсов Интернета.

На первом этапе исследования «Третья волна» [3] целью морфологического анализа (подсчет числа ключевых слов) был ответ на вопрос – что представляет собой Интернет с точки зрения его хаотичного заполнения текстами глобальной аудитории, есть ли связь при подсчете ключевых слов с реальными событиями в глобальном социуме? Подсчетом ключевых слов занимаются в США в компании Google, это известные и популярные проекты Google Flu Trends, Google Trends, Google Books Ngram. Также ключевые слова регистрируют в рамках проектов US National Security Agency (NSA), например программа PRISM. В исследовании Dell EMC External Research and Academic Alliances «Третья волна» на первом этапе при сборе ключевых слов через API-доступ (Application Programming Interface) к хранилищам данных Google и Яндекс были получены обнадеживающие результаты: выявлены сильные корреляционные связи с некоторыми статистическими данными международных баз метаданных, определены группы для проанализированных стран (позитивная и негативная по имиджу группы и группы с указанием на смену технологического уклада и без указаний), сделана попытка построить рейтинг стран и найдена ковариация с рейтингом недееспособности государств Fragile States Index.

На втором этапе исследования, результаты которого представлены в этой статье, подходы к анализу были усложнены – проведен кластерный анализ. Помимо продолжения работы по построению рейтинга стран (в данной статье эта часть исследования не рассматривается), был применен новый дизайн анализа ключевых слов с использованием модифицированной карты «Цепочка ценности» (Value Chain Map) Саймона Водлей (Simon Wardley), которая позволяет проводить стратегический анализ развития рынка [22, 30, 31]. При подсчете слов Интернет рассматривается как глобальная система «People-to-IT», которая обладает гибкостью с точки зрения реакции на события, происходящие в глобальном социуме, в том числе в экономике. Численность тех или иных ключевых слов в Интернете отражает массовую реакцию людей в разных аспектах: рыночном, рекламном, социальном, политическом, психологическом, стрессовом. В ходе построения модифицированной карты Value Chain Map выбранные ключевые слова рассматривались как признаки экономической активности стран в определенных областях. Если в привязке к той или иной стране больше пишут о конкретной технологии, то значит, для этого есть информационные поводы,

которые в свою очередь, появляются в связи с реальными событиями в экономической жизни страны.

Методика

В анализ вошли 49 стран из разных регионов мира. Выполнялась текстовая аналитика Big Data из открытых ресурсов Интернета, при выполнении Data Mining использовался доступ через API к текстам дата-центров Google и Яндекс, что методически рассматривается как использование Google и Яндекс в качестве неклассических суперкомпьютеров в режиме «as-a-Service» Cloud model. Карты распределения дата-центров Google и Яндекс представлены на рисунке 1.

Авторская методика по технике Data Mining через API-Google и API-Яндекс разработана Д.А. Яковлевой. Кластерный анализ k-means выполнен Д.А. Яковлевой. Графическая обработка табличных метаданных выполнена в Microsoft Excel, использовался линейный корреляционный анализ по Пирсону в Microsoft Excel.

Собрана морфологическая матрица метаданных из 20 ключевых словосочетаний (Morphological Matrix), которые подсчитывалась в миллионах в привязке к названиям стран и годам – 2008 и 2015 гг. Годы были выбраны исходя из следующих соображений: 2008 год – острый период мирового финансового кризиса, 2015 год – завершившийся последний год на момент набора данных. Языки API-поиска ключевых словосочетаний – русский и английский, данные суммировались. Морфологическая матрица была разделена на четыре набора слов (units). Ключевые слова перечислены в таблице 1. Многоуровневая гибридная схема организации процесса Data Mining представлена в таблице 2.

Таблица 1

Морфологическая матрица из 20 ключевых слов по Data Mining

Морфологическая матрица / Morphological Matrix	
Unit «Информация» / «Information»	компьютер / computer; облачные технологии / cloud computing; мобильный телефон / mobile phone; курс доллара / dollar exchange rate; курс евро / euro exchange rate
Unit «Энергия & Материалы» / «Energy & Material»	электромобиль / electric cars; солнечные батареи / solar panel; капельное орошение / drip irrigation; поставки газа / gas supplies; нефть / oil; цена на нефть / oil price; атомная электростанция / nuclear power plant
Unit «Социальное неравенство» / «Social inequality»	кризис / crisis; инфляция / inflation; рост цен / price increase; безработица / unemployment; увольнения / dismissal
Unit «Стрессовая нагрузка» / «Stress load»	мигранты / migrants; беженцы / refugees; мусульмане / Muslims

Таблица 2

**Многоуровневая гибридная схема организации Data Mining
в многоцентровом исследовании «Третья волна»**



Страны были подвергнуты кластерному анализу по всему списку ключевых слов на основе таблиц с метаданными (две таблицы-матрицы из подсчитанных ключевых слов за 2008 год и 2015 год). Пример исходной таблицы Microsoft Excel, подготовленной для переноса ее в программу кластеризации методом k-means (K-средних) представлен на рисунке 2. Кластерный анализ был выбран для того, чтобы разбить страны на группы с учетом всей матрицы подсчитанных слов и определить их принадлежность к высокой, средней или низкой активности в Интернете. Метод k-means позволяет разбить выборку на группы с минимальным разбросом значений.



Рис. 1. Карты распределения дата-центров Google и Яндекс.

Активность стран в Интернете оценивалась в данном исследовании в зависимости от числа ключевых слов, привязанных к этим странам. Эта активность не одинакова и зависит от многих факторов: число компьютеров у населения страны; уровень компьютерной грамотности населения; число IP-соединений и доступность Интернета в стране в

морфологической матрицы ключевых слов (II); сравнение данных по модифицированной карте «Цепочка ценности» (Value Chain Map) (III) с распределением стран по регионам (IV).

I. Кластерный анализ стран

Кластерное распределение по 2008 году показало наличие четырех кластеров, при этом по активности и интенсивности цитирования в Интернете уровню высокой активности High level of Internet activity and citing соответствует Cluster № 1, и в него вошли только США. К уровню средней активности Middle level of Internet activity and citing отнесены Cluster № 2 и Cluster № 3. А уровню низкой активности Low level of Internet activity and citing соответствует Cluster № 4. На рисунке 3 показано кластерное распределение стран по результатам Data Mining ключевых слов.

Кластерное распределение стран по 2015 году (определено пять кластеров) в зависимости от активности и интенсивности цитирования в Интернете позволило разделить страны также на три уровня: высокая активность High level of Internet activity and citing – Cluster № 1 и Cluster № 3; средняя активность Middle level of Internet activity and citing – Cluster № 2 и Cluster № 4; низкая активность Low level of Internet activity and citing – Cluster № 5. На рисунке 4 показано кластерное распределение стран по результатам Data Mining ключевых слов.

Можно отметить динамику в распределении стран на уровне высокой интернет-активности и цитируемости за годы с 2008 по 2015 год. Если в 2008 году лидером были только США, то к 2015 году США догнали Германия, Канада и Китай. А вот Великобритания оказалась на уровень ниже, вместе с Австралией, Японией и Россией. Эти данные косвенно указывают на глобальное экономическое соревнование стран-лидеров и, в частности, на соотношение сил Германии и Великобритании. В политико-экономической жизни 23 июня 2016 года в Великобритании был проведен референдум, результатом которого был объявлен выход страны из Евросоюза, то есть отделение Великобритании от общего рынка с Германией. Возможно, именно соревнование двух стран за топ-позиции на мировом рынке и явились причиной выхода Великобритании из Евросоюза, где в более выгодном положении оказалась Германия.

При анализе графиков, построенных по карте «Цепочка ценности» (Value Chain Map), которые будут представлены далее, следует делать поправки на различия стран, попавших в разные кластеры. Однако принципиально на трактовку результатов сравнения стран по картам Value Chain Map различия по кластерам не повлияли. На частоту встречаемости ключевых слов при региональном распределении стран можно ориентироваться, даже если страны локализованы в разных кластерах, так как в одном регионе зачастую страны находятся в близких кластерах по активности в Интернете. Также выявлены случаи, когда страны из одного кластера на карте Value Chain Map показывают значительно различающуюся частоту встречаемости ключевых слов.

II. Кластерный анализ морфологической матрицы

Кластерное распределение по 2008 году показало четыре кластера ключевых слов, подсчитанных для всех проанализированных 49 стран мира. На рисунке 5 представлено распределение морфологической матрицы по кластерам. В кластер максимального цитирования попали два словосочетания – «компьютер» и «мобильный телефон». На уровень ниже, но также большим цитированием в Интернете отличились ключевые слова «нефть», «кризис» и «рост цен».

Кластерное распределение по 2015 году также характеризуется четырьмя кластерами. Лидером цитирования стало слово «компьютер». А частота цитирования словосочетания «мобильный телефон» сильно уменьшилась, перескочив через кластер вниз с уровня Max на уровень Middle-2 (среднее цитирование второго уровня). Данные представлены на рисунке 6.

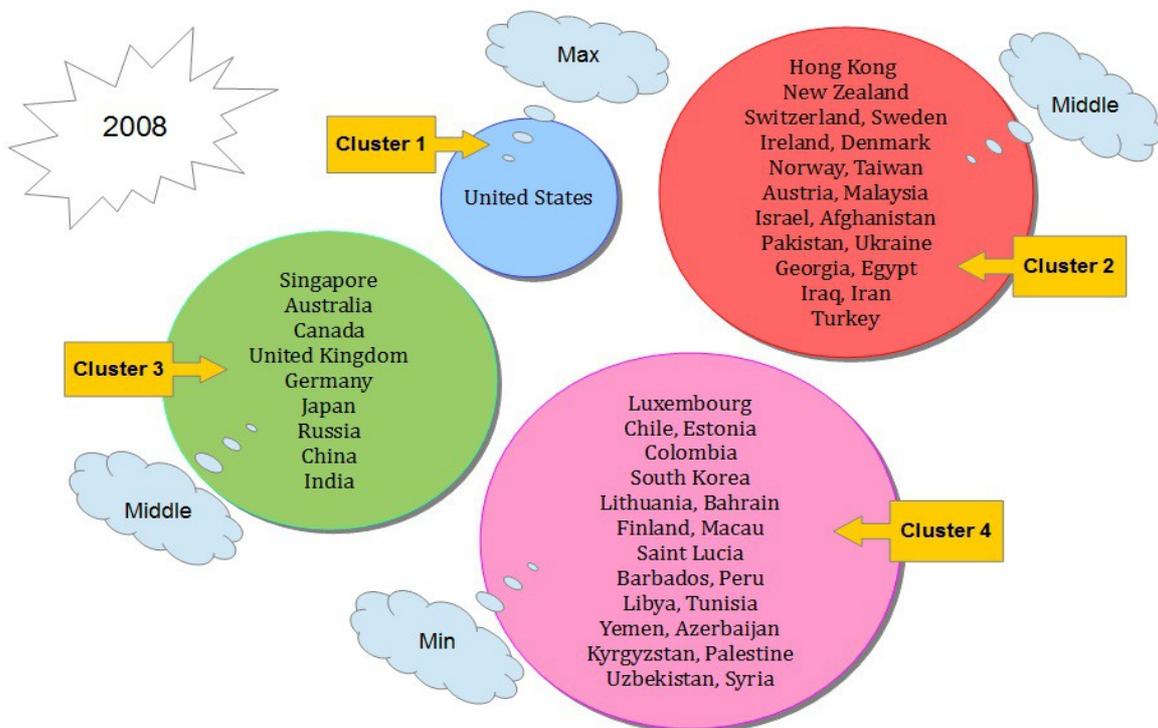
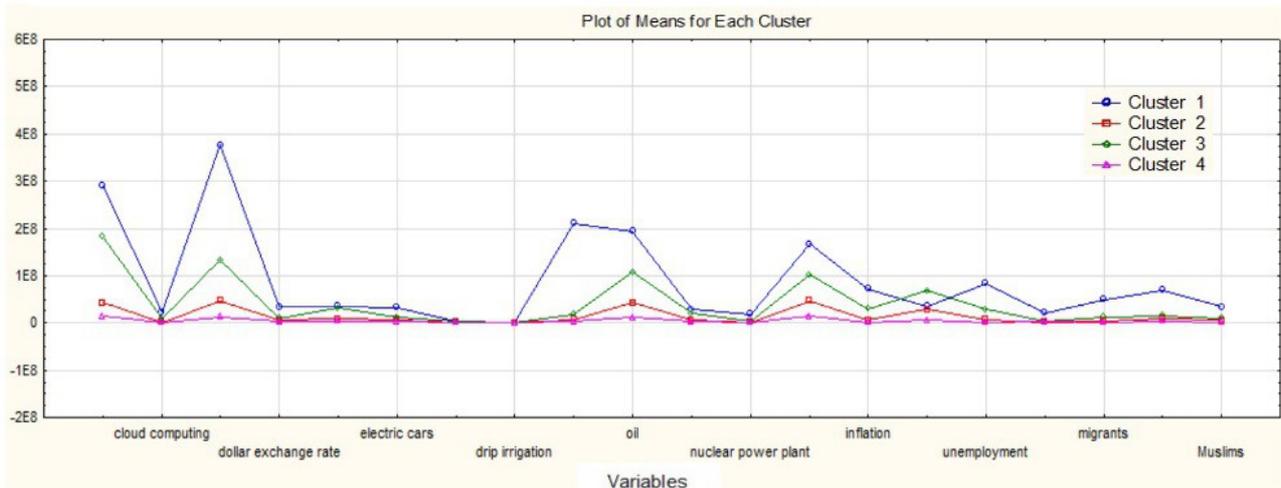


Рис. 3. Распределение стран по кластерам, 2008 год. Variables – ключевые слова (всего 20 словосочетаний, список см. в таблице 1).

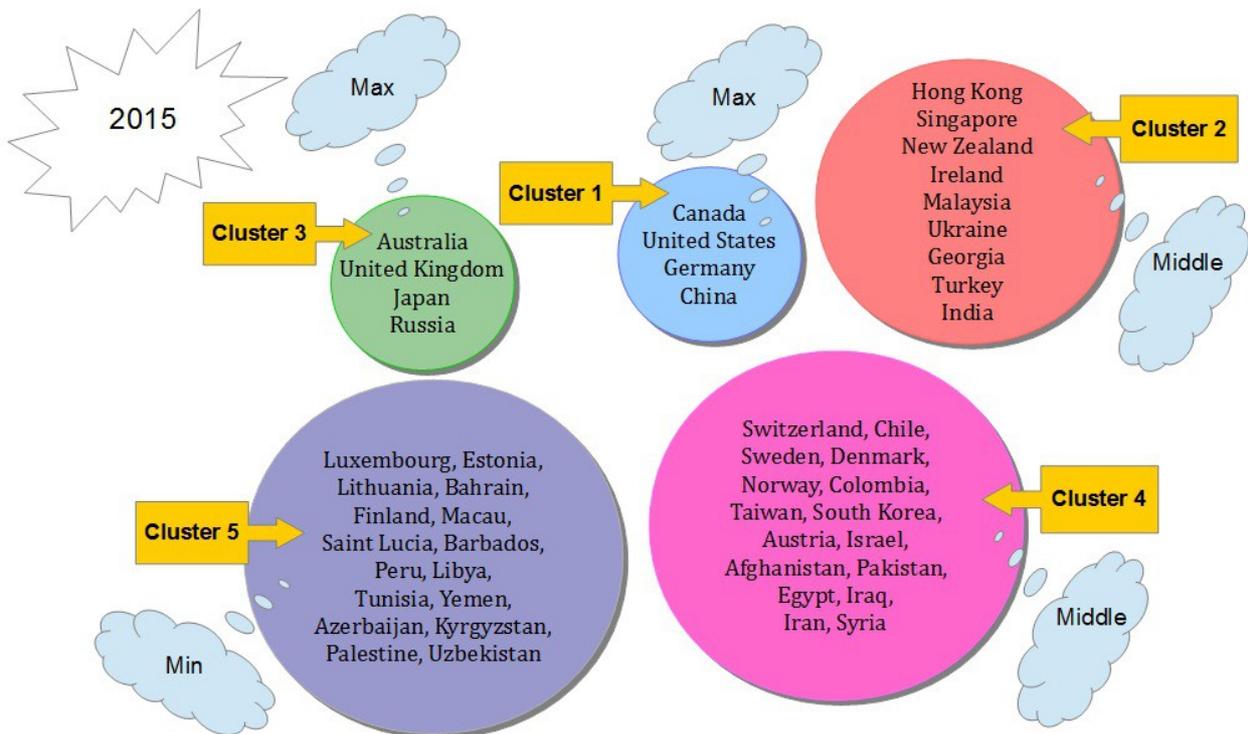
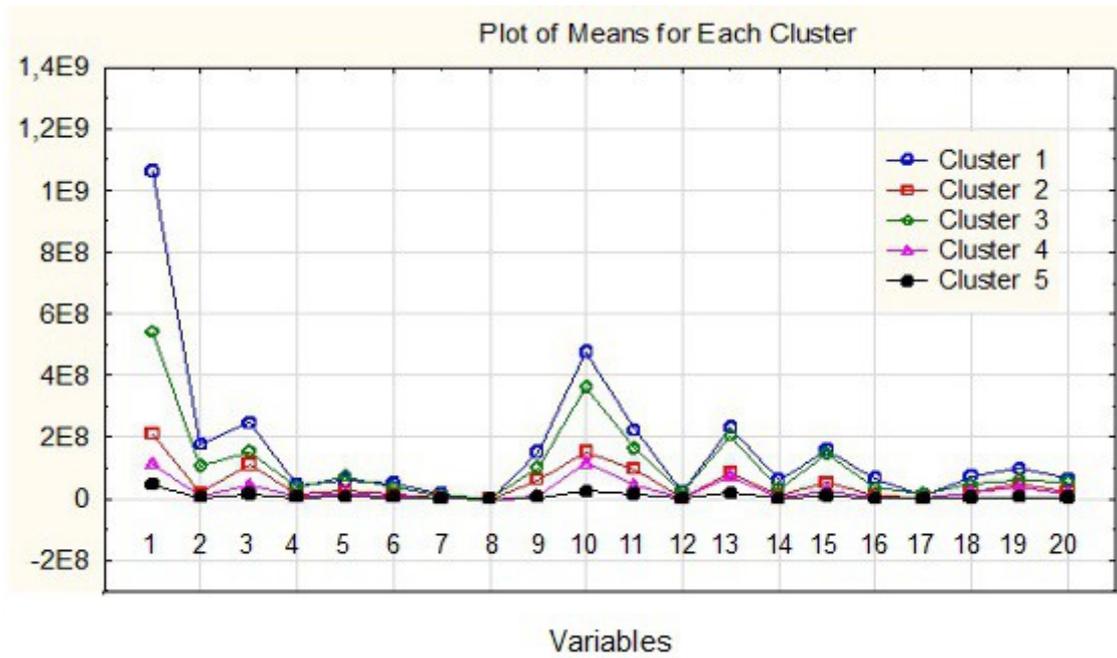


Рис. 4. Распределение стран по кластерам, 2015 год. Variables – ключевые слова (всего 20 словосочетаний, список см. в таблице 1).

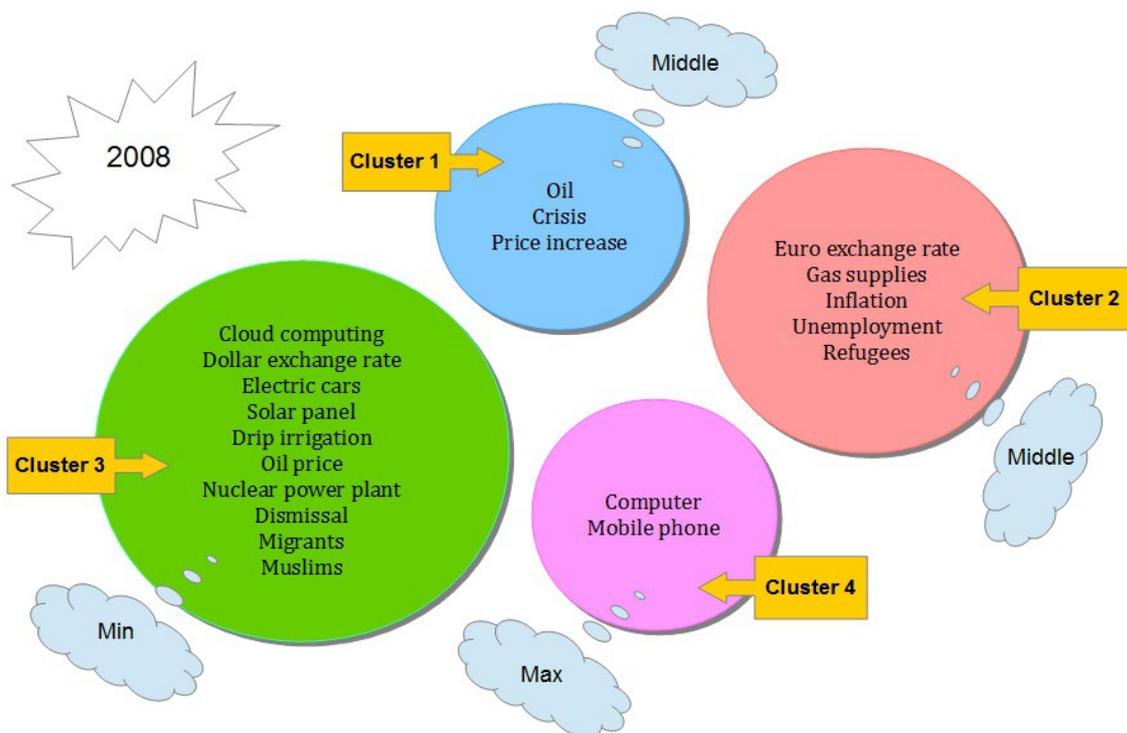
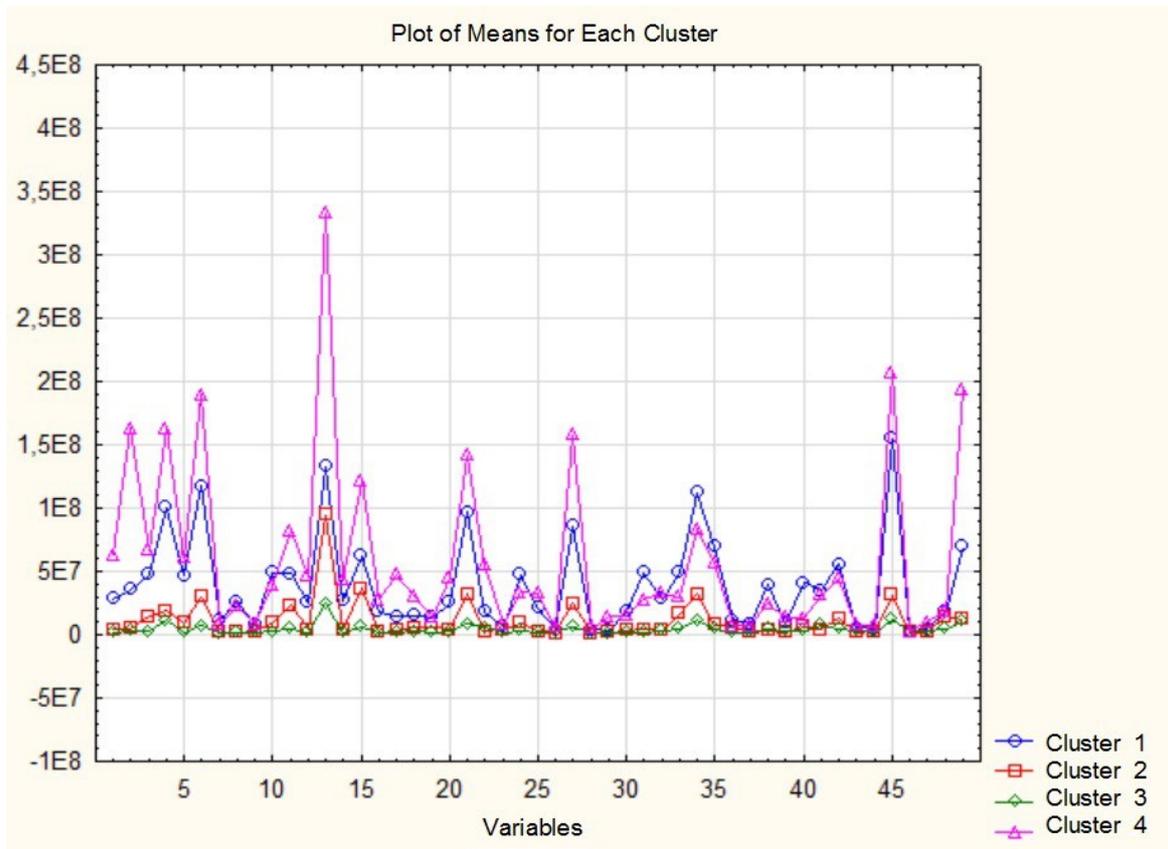


Рис. 5. Распределение морфологической матрицы по кластерам, 2008 год. Variables – страны.

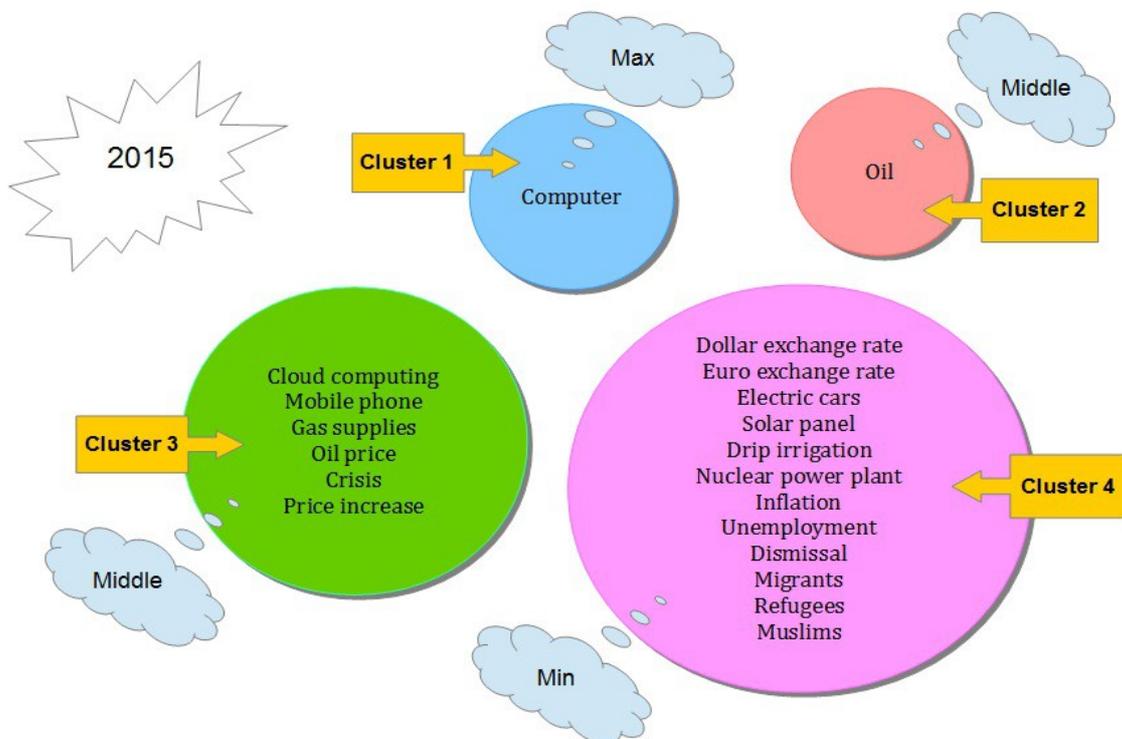
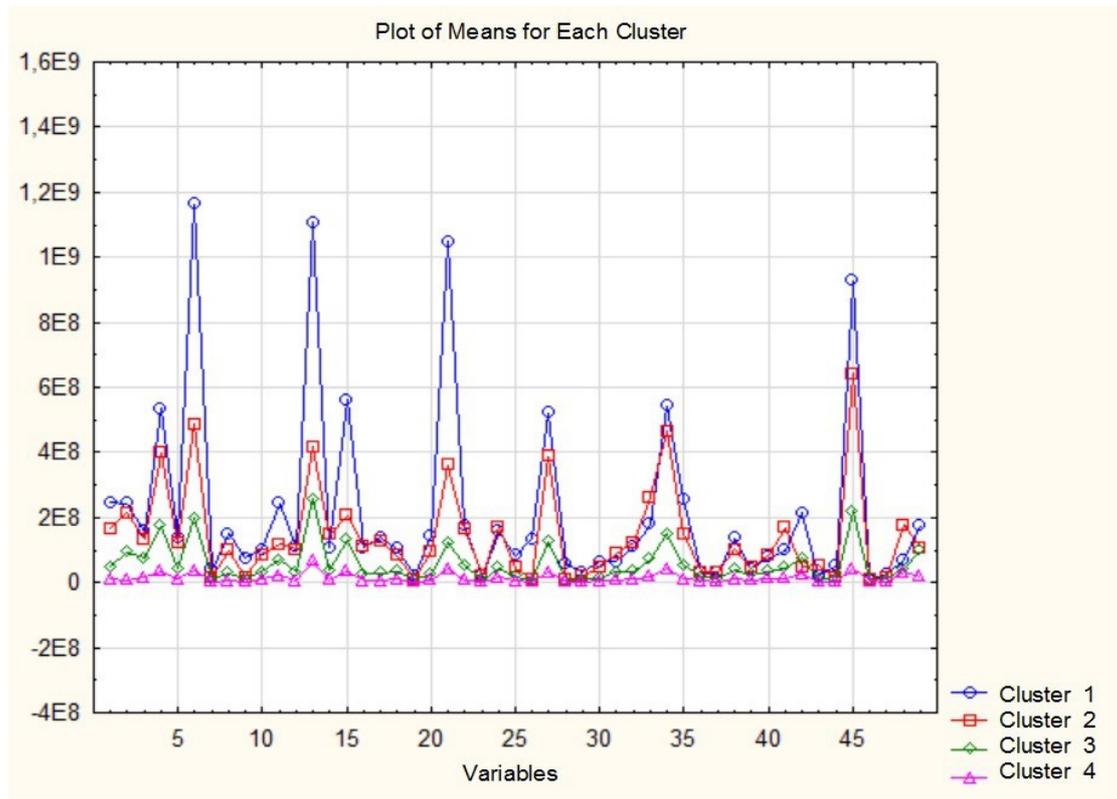


Рис. 6. Распределение морфологической матрицы по кластерам, 2015 год. Variables – страны.

На рисунке 7 представлена сводная информация по кластерному распределению ключевых слов из морфологической матрицы. Можно отметить, что слово «компьютер» стабильно является самым цитируемым из представленных ключевых слов. Без изменений в одних и тех же кластерах остаются за прошедшие годы слова и словосочетания «нефть», «поставки газа», «электромобиль», «солнечные батареи», «капельное орошение», «атомная электростанция» и еще несколько слов. Поменяли кластер с уменьшением частоты цитируемости такие слова и словосочетания, как «мобильный телефон», «кризис», «рост цен» и еще несколько ключевых слов.

Только два словосочетания поднялись по цитируемости на кластер выше. Это «облачные технологии» и «цена на нефть». На мировом рынке в 2015 году цена на нефть падала, что бурно обсуждалось мировым сообществом. Падение цен на нефть обусловлено сменой технологического уклада, что будет рассмотрено дальше на примере Кондратьевских волн. А вот Облачные технологии – это то, что приходит на смену сырьевой экономике. И кластерный анализ показал этот индикатор – повышение цитируемости словосочетания «облачные технологии» со сменой кластера. Интересным результатом является то, что повышение цитируемости словосочетаний «облачные технологии» и «цена на нефть» произошло синхронно, что как раз и отражает процесс смены приоритетов в экономическом развитии вместе с уходом старой экономики.

Keywords	2008	2015	Dynamics
Computer	Max	Max	☰
Mobile phone	Max	Middle 2	↓
Oil	Middle 1	Middle 1	☰
Crisis	Middle 1	Middle 2	↓
Price increase	Middle 1	Middle 2	↓
Euro exchange rate	Middle 2	Min	↓
Gas supplies	Middle 2	Middle 2	☰
Inflation	Middle 2	Min	↓
Unemployment	Middle 2	Min	↓
Refugees	Middle 2	Min	↓
Cloud computing	Min	Middle 2	↑
Dollar exchange rate	Min	Min	☰
Electric cars	Min	Min	☰
Solar panel	Min	Min	☰
Drip irrigation	Min	Min	☰
Oil price	Min	Middle 2	↑
Nuclear power plant	Min	Min	☰
Dismissal	Min	Min	☰
Migrants	Min	Min	☰
Muslims	Min	Min	☰

Рис. 7. Итоговая таблица по сравнению кластерного анализа морфологической матрицы за 2008 и 2015 годы.

III. Модифицированная карта «Цепочка ценности» (Value Chain Map)

Для сравнения паттернов ключевых слов была применена карта «Цепочка ценности», которую в контексте развития ИТ-технологий представил Саймон Водлей ([Simon Wardley](#)) на ведущем международном форуме по ИТ O'Reilly Conference «Open Source Convention» / OSCON (конференция по открытым ИТ-ресурсам) [30, 31]. Саймон Водлей является главным научным сотрудником лондонской компании Leading Edge Forum of Computer Sciences Corporation, а также сопредседателем форума O'Reilly Conference OSCON и членом оргкомитетов форумов Open Forum Europe, Strata + Hadoop World, Cloud Camp London. Подход Саймона Водлей к анализу развития рынка ИТ в динамике, с определением критических моментов для конкурентной борьбы и точек опережающего роста, был модифицирован под результаты данного исследования по текстовой аналитике Больших данных.

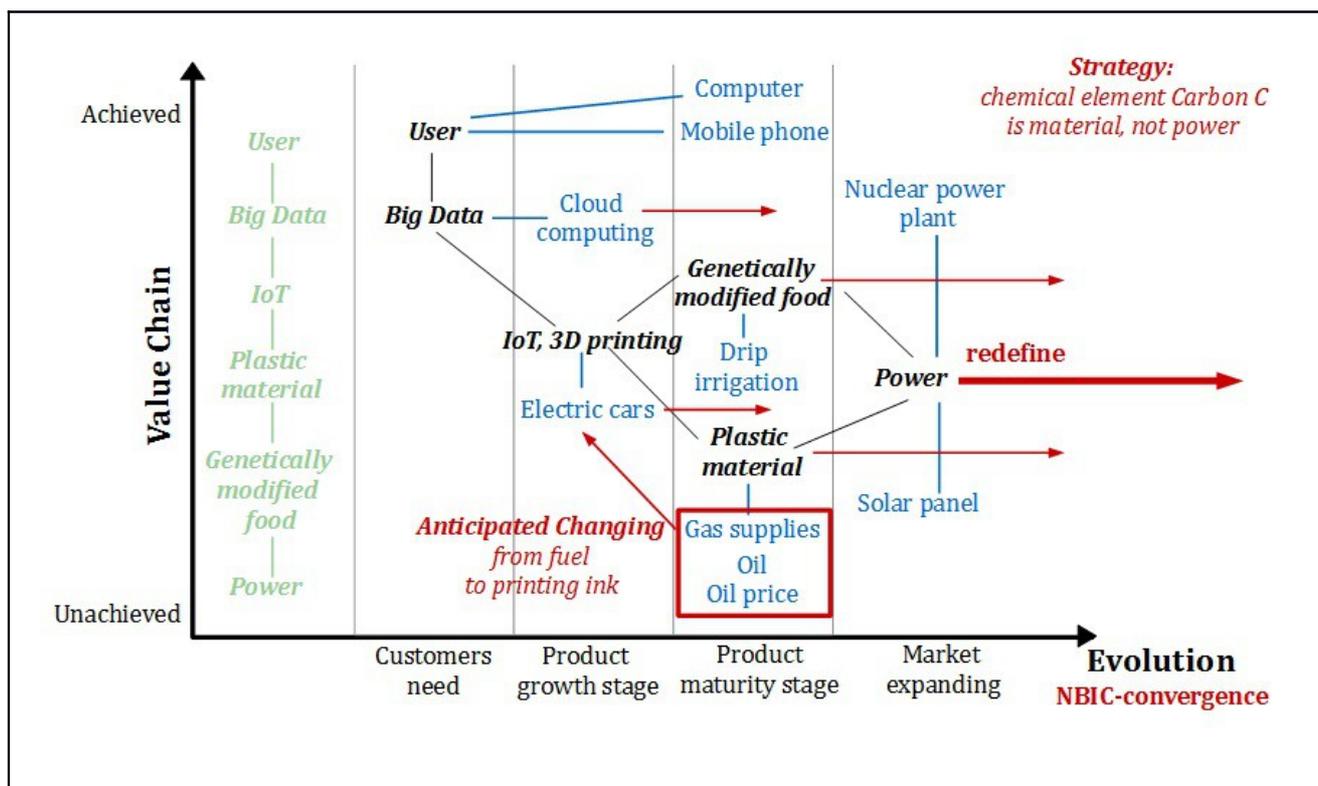


Рис. 8. Модифицированная карта «Цепочка ценности» (Value Chain Map) с ключевыми словами из морфологической матрицы (синий шрифт). Все компоненты цепочки показывают карту в развитии слева направо, в интересах конкурентной борьбы на рынке.

На рисунке 8 представлена полученная карта, для ее построения использовано десять ключевых словосочетаний (синий шрифт на графике). По оси Y пространство карты распределено от точки достигнутого до точки пока еще не достигнутого места на рынке. Вертикально вписана (зеленый шрифт) сама цепочка развития рынка (цепочка ценности). Пользователи получили доступ к ИТ-технологиям, они производят Большие данные, что закладывается в идеологию Интернета вещей, который несет в себе настоящую промышленную революцию благодаря распространению технологии 3D печати (включая машиностроение). Именно технология 3D печати и есть тот революционный поворот, который Элвин Тоффлер предвидел в 80-х годах прошлого века, описывая новый тип экономики «prosumer» (об этом уже было сказано в данной статье в разделе «Введение»).

Облачные технологии, мобильные приложения и 3D печать являются в настоящее время точкой опережающего роста. На шаг вперед, пока еще менее внедренные в жизнь, но уже вставшие в ряд ожидания на ближайшую перспективу, два направления – это новые

материалы, имеющие уникальные свойства, совместимые с ИТ-технологиями и являющиеся одновременно «чернилами» для 3D печати (включая продукты нефтегазохимии), и генно-модифицированные сельскохозяйственные культуры (рис, пшеница, овощи, фрукты и т.п.), которые выращиваются на «умных полях» (сельскохозяйственный Интернет вещей / Agriculture's IoT). Эти два направления как следующий шаг на ближайшую перспективу также отражены на карте Value Chain Map.

Критическим моментом для конкурентной борьбы в данной цепочке является электроэнергия. Объем электроэнергии, которая будет распределяться по умным сетям электроснабжения (Smart Grid), определяет потолок роста всех перечисленных технологий экономики «prosumer» (дата-центры, подключенные объекты к IoT, компьютерный парк, 3D-принтерный парк). Соответственно, энергетика (возобновляемая и атомная) является опорной платформой для пирамиды новой экономики Третьей волны, то есть информационной эры. На карте (рис. 8) все упомянутые компоненты цепочки развернуты в динамике, в развитии слева направо, в интересах конкурентной борьбы на рынке. К звеньям цепочки прикреплены ключевые слова, которые подсчитывались в данном исследовании (синий шрифт), и именно они и служили индикаторами для анализа ситуации для каждой из стран. Красным маркером отмечены критические моменты для конкурентной борьбы и точки опережающего роста.

По оси X представленной карты на рисунке 8 точки опережающего роста находятся в зоне Product growth stage, в описанной цепочке это Облачные технологии и электромобили. В зоне рынка, достигшего плато развития, Product maturity stage, находятся нефть и газ, но предстоит обновление подходов с изменением вектора рынка углеводородов в сторону нефтегазохимии и получения новых материалов для 3D печати. Ожидаемые изменения будут сопровождаться переделом рынка с появлением новых пионеров и лидеров.

В качестве критического момента для конкурентной борьбы по оси X указана энергетика, роль которой предстоит переосмыслить в ближайшее время. Лидерство странам в новых экономических условиях обеспечит возобновляемая энергетика (особенно солнечная), не имеющая лимита, и атомная энергетика, относящаяся к ископаемой лимитированной энергетике, но еще далекой от исчерпания ресурсов и открытой для новых технологических возможностей.

В логике описанной карты Value Chain Map представлены данные по подсчету ключевых слов по 49 странам, сгруппированным на графиках по своим географическим регионам (рис. 21-26). Эти графики будут рассмотрены дальше в статье в разделе IV.

Для трактовки результатов исследования представляется необходимым провести краткий обзор динамики научно-технического развития в разрезе Кондратьевских волн [1, 2, 5, 6], показывая роль ключевых слов из морфологической матрицы для каждой из Кондратьевских волн, но не выходя за рамки темы статьи и не углубляясь в экономическую тематику обсуждения.

К-волны или Кондратьевские волны (Кондратьевские циклы, Большие циклы конъюнктуры) описаны русским ученым Николаем Дмитриевичем Кондратьевым (1892-1938 гг.) в 1920-е годы, когда он возглавлял Конъюнктурный институт при Наркомате финансов Союза ССР, который был организован по инициативе самого Кондратьева. Труды и идеи ученого о теории экономических циклов были запрещены Сталиным с момента отстранения Кондратьева от должности в 1928 году, а в 1938 году ученый был расстрелян как враг народа, и вплоть до времен перестройки 1987 года теория Кондратьева была предана забвению. Идеи ученого называли «кондратьевщина», термин был символом вредительства народному хозяйству СССР. Экономический ущерб от запрета трудов Н.Д. Кондратьева ощутим особенно сегодня, так как знания о времени прихода новых технологических циклов позволяют подготовиться заранее и действовать на опережение, а не догонять в условиях экономического кризиса. Уже на современном этапе теория Н.Д. Кондратьева была развита и дополнена его последователями (рис. 9) [1, 2, 5, 6].

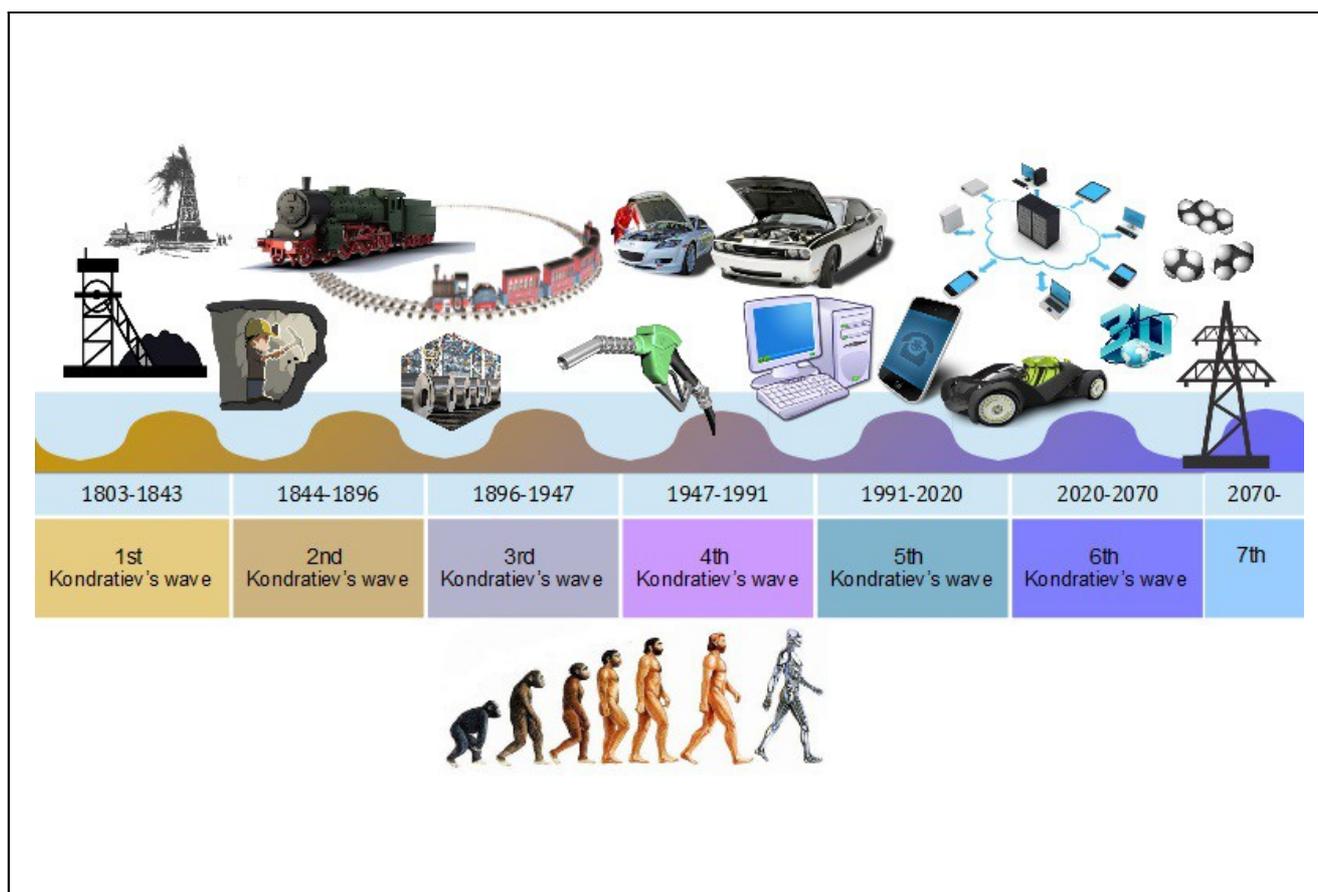


Рис. 9. Схема смены Кондратьевских волн в разрезе научно-технического прогресса.

В данном исследовании выбранные ключевые слова для Data Mining были тематически распределены по Кондратьевским волнам, описанным современными последователями ученого. Это позволило понять, как построить карту «Цепочка ценности» (Value Chain Map), представленную выше (рис. 8), и оценить ландшафт конкурентной борьбы стран.

1) **«Computer / Компьютер».** Это ключевое слово из всей собранной матрицы ключевых слов имеет самую большую встречаемость с отрывом от остальных слов, что указывает на произошедшую смену технологического уклада с преобладанием интереса к ИТ-технологиям. Компьютерная эра началась не сегодня, еще в прошлом веке. Начало Кибернетической революции относят к периоду с 1940 года до 1980 года. Наблюдая за изменениями, происходящими в тот период, Элвин Тоффлер в 1980 году издал свою концептуальную работу «Третья волна», где описаны черты информационного общества и принципиальная смена экономики с рыночной на экономику «производства для себя», с совсем иными экономическими законами и правилами [8]. Первая волна развития цивилизации по Тоффлеру – сельскохозяйственная, главное в экономике было владение землей. Вторая волна – индустриальная, когда в экономике лидируют промышленное производство, углеводороды в качестве сжигаемого топлива для двигателя внутреннего сгорания, и товары массового потребления. Производство массовых товаров и владение сырьем – главное для экономики Второй волны. Сегодня страны с отстающими в инновациях экономиками все еще находятся на уровне Второй волны, живут в условиях индустриализации. Третья волна, уже охватившая страны-лидеры мирового рынка, характеризуется приоритетом информационных технологий в экономике и процессами

глобализации. В результате для Третьей волны главное – это участие в процессах глобального управления через информационные технологии [8, 9].

Кондратьевские волны являются параллельным направлением изучения динамики мировых экономических процессов, волны описывают более короткие периоды времени – 30-50 лет, они не столь обобщенные, как волны Тоффлера, а более детальные. Кондратьевские волны имеют две равные по времени фазы – повышательную (рост экономики) и понижательную (спад экономики) [1, 2, 5, 6]. В данном аспекте прогнозирование характеристик следующей Кондратьевские волны должно помочь сконцентрировать усилия в ходе понижательной фазы текущей Кондратьевские волны не на удержании устаревающих тенденций, а на опережающем развитии новых технологий, которые будут свойственны новой волне, что даст больший экономический эффект в период повышательной фазы новой волны.

Старт Кибернетической революции совпал с Четверной Кондратьевской волной (период с 1947 года по 1991 год). Ключевое слово «computer / компьютер» можно отнести к признаку Четверной Кондратьевской волны. Сегодня эта экономическая волна уже завершила свой цикл. В морфологической матрице ключевое слово «Computer / Компьютер» доминирует по численности (по данным за 2008 и 2015 годы), что свидетельствует о широком распространении этого товара (Product maturity stage). Это можно наблюдать по странам и регионам на графиках, представленных на рисунках 21-26.

2) **«Mobile phone / Мобильный телефон».** Это ключевое словосочетание также характеризует информационные технологии. Ключевое слово «mobile phone / мобильный телефон» можно отнести к признаку Пятой Кондратьевской волны (период с 1991 года по 2020 год). Это текущая на данный момент экономическая волна. Распространенным продуктом на рынке (Product maturity stage) мобильный телефон стал лет 10-15 назад. Пятая Кондратьевская волна находится на данный момент в понижательной фазе. Сама по себе мобильная связь уже не является чем-то инновационным, акценты сместились в наполнение мобильности смыслом и содержанием. Мобильные телефоны способствовали формированию нескольких базовых черт информационной эры: бурный рост объема информационных данных (Больших данных); постоянная минимизация размеров носимых устройств, связанных с информационными технологиями; массовое применение пластмасс для изготовления устройств, связанных с информационными технологиями. Большие данные, малогабаритные носимые ИТ-устройства и массовое использование пластмасс для ИТ-производства – это то, что появилось и укрепилось в период Пятой Кондратьевской волны. Разброс подсчитанного количества ключевых словосочетаний «mobile phone / мобильный телефон» в привязке к разным странам намного больше, чем для слова «computer / компьютер». Пятая Кондратьевская волна еще не завершена, также следует учитывать различия в цитируемости между странами (разную кластерную локализацию стран) и разный уровень захвата Пятой Кондратьевской волной стран мира (то есть разный уровень развития экономик стран). При кластерном анализе морфологической матрицы ключевых слов выявлено, что словосочетание «mobile phone / мобильный телефон» поменяло кластерное позиционирование в период с 2008 года по 2015 год на более низкий уровень цитируемости. Это косвенно свидетельствует в пользу того, что мобильный телефон служит знаком уходящей Пятой Кондратьевской волны.

3) **«Cloud computing / Облачные технологии».** Это ключевое словосочетание можно считать символом наступающей Шестой Кондратьевской волны (период с 2020 года по 2070 год). Шестая Кондратьевская волна должна завершить долгий период Кибернетической революции. Сама концепция Облачных технологий появилась в 2006 году, что сразу вызвало взрывной рост объема данных. Через два года, в 2008 году большой объем данных стал настолько острой проблемой, что редактор журнала Nature Клиффорд Линч написал статью и подготовил отдельный выпуск журнала на эту тему. Клиффорд Линч впервые предложил

термин «Большие данные». Концепция Интернета вещей (Internet of Things, IoT) появилась впервые в 1999 году, ее описал исследователь Кевин Эштон из Массачусетского технологического института. Но настоящий импульс для развития Интернет вещей получил только с появлением Облачных технологий [27, 32]. Облачные технологии пока не достигли стадии распространенного продукта на рынке, Облачные технологии находятся на стадии поиска своей ниши, жесткой конкурентной борьбы за раздел рынка, появления новых стартапов и новых идей (Product growth stage). Сегодня на фоне экономического спада понижательной фазы Пятой Кондратьевской волны, которая закончится в 2020 году, Облачные технологии являются тем самым направлением, которое нужно интенсивно развивать, чтобы получить «окно возможностей» в период экономического роста на фоне повышательной фазы Шестой Кондратьевской волны.

В Приложении 1 представлены диаграммы по всем проанализированным в исследовании странам, на которых наглядно показана морфологическая матрица ключевых слов. По виду диаграмм страны можно разделить на три группы: с преобладанием ключевых слов Unit «Информация» / «Information» (группа стран, которые вошли в географический захват нового технологического уклада информационной волны); с преобладанием ключевых слов Unit «Энергия & Материалы» / «Energy & Material» (группа стран, которые пока не захвачены сменой технологического уклада информационной волны); с преобладанием ключевых слов Unit «Социальное неравенство» / «Social inequality» и Unit «Стрессовая нагрузка» / «Stress load» (группа стран, для которых более актуальны проблемы социальной нестабильности, чем экономическое развитие). На рисунке 11 представлена сводная картина по анализу выше упомянутых диаграмм – зона географического захвата нового технологического уклада информационной волны, зоны старого сырьевого уклада и эпицентр нестабильности и социальной напряженности.

После завершения Шестой Кондратьевской волны предполагается переход человечества в некие новые социальные условия существования, что называют трансгуманизмом (Transhumanism, кратко Н+). Трансгуманисты ожидают этот переход раньше 2070 года, после 2045 года. Трансгуманизм предполагает «совершенствование» человека – его интеллекта, физических возможностей, психологических характеристик, социальных аспектов – при помощи ИТ-технологий, электроники, дополнительных устройств и чипов, гаджетов, биотехнологий. Сейчас это направление называют «дополненная реальность» или «расширенная реальность» (Augmented Reality, AR), введение в поле восприятия сенсорных данных с целью дополнения и улучшения восприятия информации (Magic Leap, HoloLens, Smart Glasses, умная одежда, умные лобовые стекла самолета и т.д.).

Дорога к трансгуманизму сегодня пролегает через Облачные технологии, Большие данные, Интернет вещей, NBIC-конвергенцию, робототехнику, искусственный интеллект. Экосистема Интернета вещей, состоящая из разнообразных носимых на теле и портативных устройств (гаджетов), включая мобильные телефоны, и приборов, машин, сенсоров и других окружающих человека предметов, базируется на Облачных технологиях [11, 14, 15, 21, 24, 29]. Скоро многие из этих объектов Интернета вещей можно будет создавать, используя 3D печать и пластмассы в качестве «чернил» для 3D-принтеров. Согласно докладу McKinsey Global Institute «Digital America: A tale of the haves and have-mores» [24], одним из трендов, меняющих экономику, будет массовое использование в ближайшем будущем потребителями технологии 3D печати для изготовления нужных товаров для себя (custom-ordered 3D products). Это и есть описанный интуитивно Тоффлером в прошлом веке принцип «prosumer» («producer» / производитель и «consumer» / потребитель в одном) со стиранием разрыва между производителем и потребителем.

В Приложении 2 представлены графики, с помощью которых проанализировано соотношение ключевых словосочетаний: «электромобиль / electric cars», «поставки газа / gas supplies», «облачные технологии / cloud computing», «цена на нефть / oil price», «кризис / crisis». Среди всех проанализированных стран только три страны в 2015 году попали на

графике в зону, обозначенную как инновационная зона «Modern zone», в которой сочетается интерес к Облачным технологиям и поставкам газа – США, Великобритания и Германия. Причем условия построения графика таковы, что чтобы попасть в «Modern zone», нужно, чтобы у страны интерес к Облачным технологиям превысил интерес к цене на нефть. Остальные страны в основном находятся в ресурсной зоне «Old industrial zone», некоторые перешли туда в 2015 году из транзитной зоны «Transition zone», где они находились в 2008 году. Этот переход обусловлен выявленным общим интересом в 2015 году к поставкам газа, что в данном исследовании трактуется как интерес к газонефтехимии в связи с повышением производства пластмасс для создания различных ИТ-устройств и приборов (в том числе создаваемых путем 3D печати). Можно констатировать, что данные, представленные на графиках Приложения 2, указывают на прогнозируемое лидерство США, Великобритании и Германии на мировом рынке в период предстоящей Шестой Кондратьевской волны.

Интересен график с позиционированием Сирии в 2008 году (до военной дестабилизации страны) в зоне экономики будущего «Future oriented zone», где интерес к Облачным технологиям и электромобилям превышает интерес к нефти и газу. Анализировать этот результат сложно, так как в связи с войной в Сирии найти в Интернете достаточное количество действующих сирийских сайтов с текстами прошлых лет сложно, многие из них закрыты, и уточнить происхождение подсчитанных ключевых слов через API из дата-центров не представляется возможным. Но найдены тексты, в которых описаны проекты компании Toyota по поставкам на рынок Сирии первых гибридных машин Prius (с дополнительным электромотором к бензиновому мотору). Также с 2007 года в Сирии тестировались электромобили собственного производства (рис. 10), причем электродвигатели работали на солнечной энергии [10].

Гражданская война в Сирии началась в 2011 году. Было ли заинтересовано нефтяное лобби региона Ближнего Востока в том, чтобы воспрепятствовать инновационному пути Сирии, ориентированному на электромобили и солнечную энергетику, и было ли это одной из причин расшатывания ситуации в Сирии, сказать по результатам данного исследования нельзя. Но предположить, что это могло быть одной из причин, толкнувшей внешние региональные силы на дестабилизацию Сирии, определенно можно. Углубляться в данный вопрос в рамках заданной темы статьи не представляется целесообразным.



Solar electric car prototypes in Syria. Damascus University model (left), Privately constructed (right).

Рис. 10. Сирийские электромобили на солнечных батареях. Ресурс фото [10].

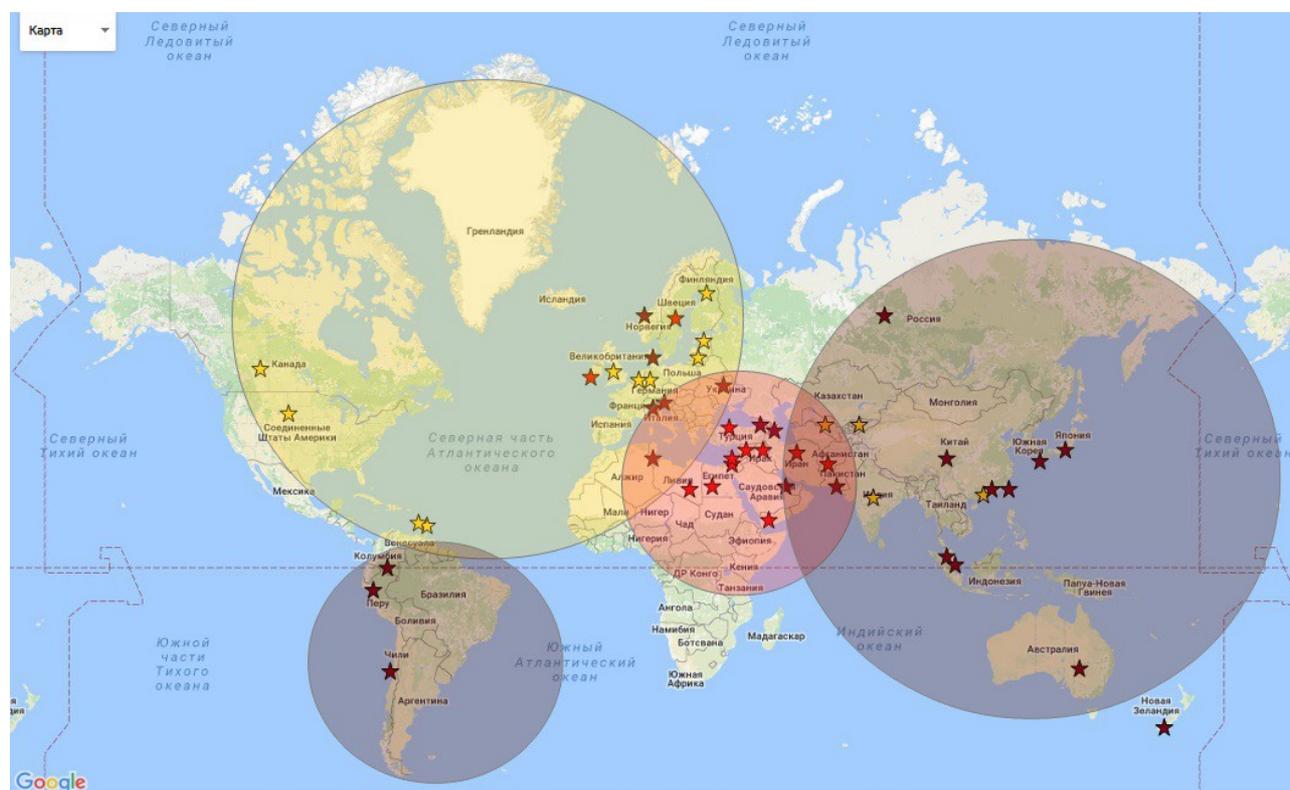


Рис. 11. Географический захват нового технологического уклада информационной волны (желтые маркеры и круг); зоны старого сырьевого уклада (коричневые маркеры и круги); эпицентр нестабильности и социальной напряженности (красные маркеры и круг).

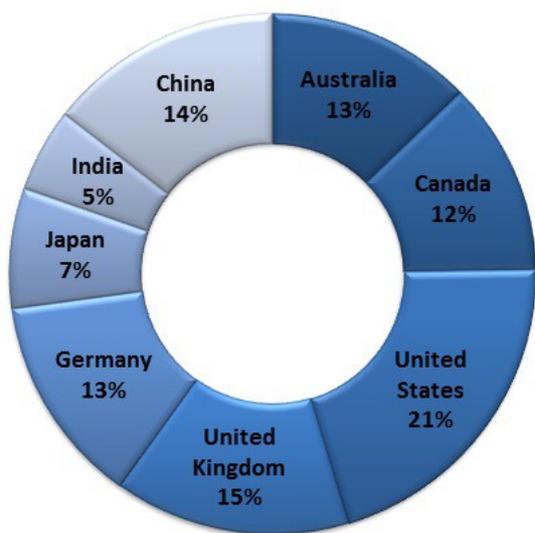


Рис. 12. Распределение стран, имеющих наибольшее число ключевых словосочетаний «cloud computing / облачные технологии».

Если рассматривать весь ряд стран по числу ключевых словосочетаний «облачные технологии / cloud computing», то только 8 стран характеризуются выпадающим, большим количеством этих слов, относящихся к 2015 году. Все они из кластера максимальной активности High level of Internet activity and citing (Cluster № 1 и Cluster № 3), и Индия из кластера средней активности Middle level of Internet activity and citing – Cluster № 2 (рис. 12). Стоит особо отметить, что хотя Россия и входит в Cluster № 3 (кластер максимальной активности), но в лидеры по числу упоминания ключевых словосочетаний «облачные технологии / cloud computing» Россия не попала. При рассмотрении долевого соотношения стран-лидеров по упоминанию Облачных технологий выявлено, что на первом месте – США.

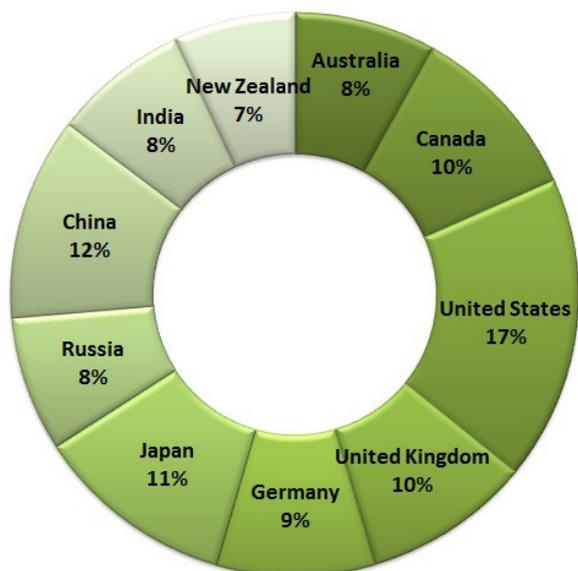


Рис. 13. Распределение стран, имеющих наибольшее число ключевых словосочетаний «electric cars / электромобиль».

4) «Electric cars / Электромобиль».

Это словосочетание связано с очень важным технологическим трендом современности, который уже опрокинул нефтяной рынок. Электромобили являются серьезным фактором, меняющим технологический уклад и экономику стран мира. Если рассматривать весь ряд стран по числу ключевых словосочетаний «electric cars / электромобиль», то 10 стран характеризуются выпадающим, большим количеством этих слов, относящихся к 2015 году (имеют повышенный интерес к электромобилям). Это все страны из двух кластеров максимальной активности High level of Internet activity and citing, Cluster № 1 и Cluster № 3, а также Индия и Новая Зеландия из кластера средней активности Middle level of Internet activity and citing, Cluster № 2 (рис. 13).

По рейтингу продаж ведущего аналитического портала www.HybridCars.com США, Китай, Япония, Германия, Великобритания, Канада входят в первую десятку пользователей электромобилями. Норвегия также в лидерской десятке, но в лидеры по ключевым словам она не попала (Норвегия вошла в Cluster № 4, средней активности). Австралия, Индия, Новая Зеландия и Россия имеют меньшее число электромобилей в сравнении с лидерами рынка, в убывающем порядке.

Если вернуться к прошлому и вспомнить характеристики описанных волн Кондратьева, то Первая Кондратьевская волна (период с 1803 года по 1843 год) характеризовалась промышленной добычей и переработкой угля; Вторая Кондратьевская волна (период с 1844 года по 1896 год) – развитием черной металлургии и строительством железных дорог; Третья Кондратьевская волна (период с 1896 года по 1947 год) – развитием тяжелого машиностроения, электрификацией, производством стали, производством электродвигателей для промышленности. Электрические двигатели появились давно, в 1821 году английский физик Майкл Фарадей создал первую модель электродвигателя. Российско-немецкий физик Мориц Герман Якоби в 1834 году разработал более мощный электродвигатель, а сербский физик Никола Тесла в США в 1888 году продал свои патенты на двухфазный электродвигатель переменного тока американскому промышленнику Джорджу Вестингаузу, основателю компании «Вестингауз Электрик» (сегодня американская Westinghouse Electric Company контролируется японской Toshiba Group и лидирует на рынке ядерной энергетики).

Несмотря на такое активное развитие электротехники, следующая Четвертая Кондратьевская волна (период с 1947 года по 1991 год) была не только стартом Кибернетической революции, но и периодом активного промышленного использования нефти и газа, а также временем двигателя внутреннего сгорания (который в истории впервые появился в 1807 году как поршневой водородный двигатель, его изобрел Франсуа Исаак де Риваз; а первый бензиновый двигатель сделал русский изобретатель Огнеслав Костович в 1880-х годах). Но стоит напомнить высказывание русского ученого с мировым именем Дмитрия Менделеева в его докладе «Бакинское нефтяное дело» от 1886 года, что сжигать нефть – это все равно, что топить печку ассигнациями. Менделеев настаивал на строительстве нефтеперерабатывающих заводов, и хотя его высказывание касалось сжигания тяжелых фракций нефти, остающихся после получения керосина, все же эта фраза универсальная. Сегодня ее можно трактовать по-другому: тратить нефть на двигатели внутреннего сгорания – это все равно, что отнимать у своей экономики будущее. Новый тренд – это нефтехимия, получение пластмасс для всевозможных устройств Интернета вещей, 3D печати и производства роботов.

Сегодня не только автомобили меняют двигатели на электрические, но и самолеты. Так, Airbus Group вместе с британской компанией Rolls-Royce Group plc разрабатывают новый гибридный авиалайнер E-Thrust с шестью электромоторами. Airbus Group уже имеет легкий электросамолет E-Fan. Но если захват авиации электромоторами прогнозируют в Airbus Group только к 2050 году [18], то компания Bloomberg уже в настоящее время прогнозирует дальнейшее падение цен на нефть из-за роста потребительского рынка электромобилей – Tesla Motors, Chevrolet Ford, Volkswagen, Nissan, BMW, Apple, Google, АвтоВАЗ, Local Motors. (рис. 14).



Рис. 14. Российский электромобиль Lada Ellada, ОАО «АвтоВАЗ» (слева) и американский электромобиль Strati, Local Motors, созданный на 3D-принтере.

В инновационном секторе автомобилестроения наиболее интересным трендом является разработка технологий 3D печати для производства самих машин [25]. Американская компания Local Motors создала электромобиль *Strati* – это первый в мире электромобиль, сделанный с применением технологии 3D печати (full-body 3D-printed car) из армированного углеродным волокном термопластика (carbon-fiber reinforced plastic, CFRP). Одновременно американцы создали другой вариант автомобиля (газовый двигатель), получаемого методом 3D печати. Компания Divergent Microfactories представила электромобиль *Blade*, материал для 3D печати – алюминий и углеродное волокно. Обратная цепочка промышленного производства углеродного волокна упрощенно следующая: полиакрилонитрил – акрилонитрил – акриловая кислота – пропилен (газ, получаемый пиролизом нефтепродуктов и дегидрогенизацией пропана). Таким образом, нефть и природный газ переходят из одного сектора рынка в другой – из сектора топлива для двигателей в сектор материалов для промышленной 3D печати машин («чернила» для 3D-принтера).

Завершая данный подраздел статьи, посвященный электромобилям, необходимо подчеркнуть, что процесс 3D печати тесно связан с Облачными технологиями [15, 29], благодаря которым создаются целые производственные роботизированные линии на заводах или удаленные рабочие места в малом бизнесе. Соответственно, чем больше будут развиваться Облачные технологии, тем больше на мировом рынке будут производиться электромобили методом 3D печати.

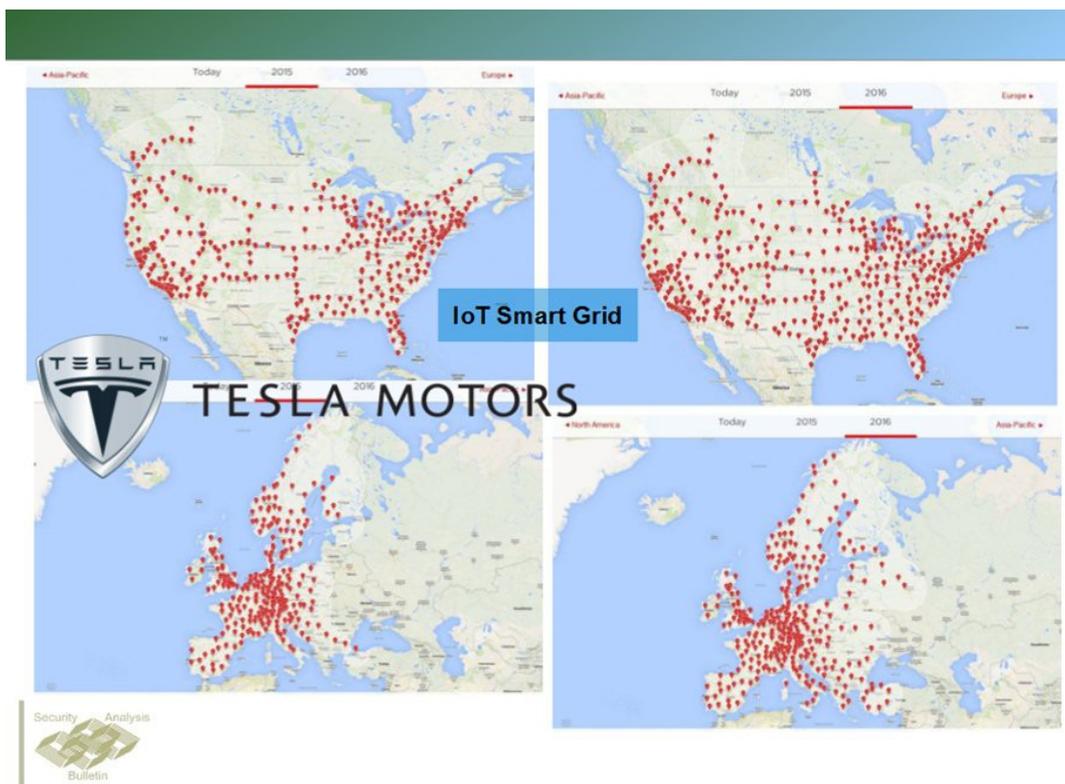


Рис. 15. Локализация в 2015 году (слева) и прирост в 2016 году (справа) зарядных станций для электромобилей.

Рынок электромобилей потребует бурного роста инфраструктуры электрозаправок (Electricity grid / Supercharger network), которая уже активно развивается в США и Европе (рис. 15). Массовое использование электрозаправок и конкуренция на рынке заставят компании проводить аналитику распределения электричества – аналитика Больших данных,

а также искать решения в рамках концепции Интернета вещей, совмещая стратегии эффективности расхода электричества как машинами, так и в домах / квартирах. В итоге этот процесс станет замкнутым кругом, так как возможность экономить благодаря «умным электросетям» (IoT Smart Grid) будет привлекать потребителей и способствовать массовому переходу на электромобили. IoT Smart Grid, как и 3D печать, связаны с Облачными технологиями.

В результате получается новая, взаимозависимая, самоподдерживающаяся и постоянно растущая в объеме цепочка:

- больше Облачных технологий,
- больше масштаб «умных электросетей»,
- больше электромашин,
- больше пользователей 3D печати,
- больше рынок материалов для 3D печати из пластмасс,
- больше потребность в нефтегазохимии (особенно в газе).

Стоит напомнить, что именно эта новая зависимость «Облачные технологии – поставки газа» и была обнаружена в данном исследовании (см. Приложение 2) только для трех стран-лидеров – США, Великобритании и Германии, что описано выше.

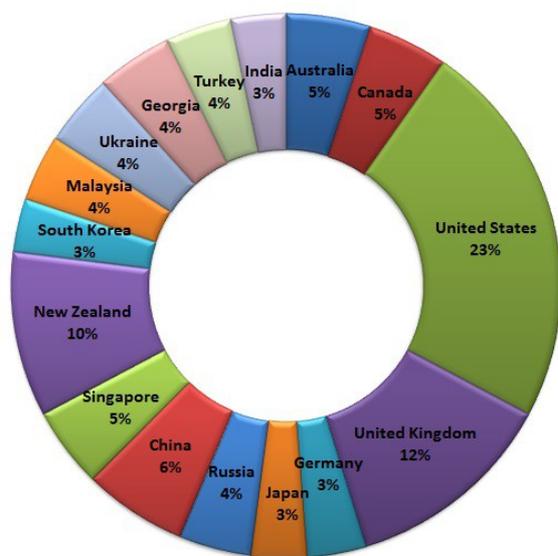


Рис. 16. Распределение стран, имеющих наибольшее число ключевых словосочетаний «gas supplies / поставки газа».

5) «Gas supplies / Поставки газа»; «Oil / Нефть»; «Oil price / Цена на нефть».

Эти три словосочетания с одной стороны отражают уходящий сырьевой технологический уклад. А с другой стороны, как уже было сказано выше, в период Шестой Кондратьевской волны (период с 2020 года по 2070 год) углеводороды поменяют свою функцию с топлива на материалы. Нефть и газ не будут сжигать, а каждый баррель и кубометр будут на счету. Из нефти и газа будут производить пластмассы, которые займут основное место в промышленности в качестве «чернил» для 3D-принтеров. По прогнозам McKinsey Global Institute [24] новые тенденции промышленного производства включают шесть основных черт: ИТ-платформы, Интернет вещей, Большие данные, искусственный интеллект, робототехника и 3D печать.

Углеводороды (уголь, нефть и газ) или каустобиолиты сформированы за миллионы лет в результате накопления углерода (C) в планктоне и водорослях мирового океана за счет фотосинтеза. Использование этих природных ископаемых в мировой экономике можно отнести к использованию продуктов фотосинтеза. Сегодня инновационные технологии позволяют получать кислород и водород, а также энергию из воды с использованием искусственного фотосинтеза. Научной задачей остается изобретение способа извлечения углерода из атмосферы для промышленных нужд. В грядущую эру роботов и трансгуманизма искусственный биохимический способ добычи углерода из атмосферы будет критически востребованной технологией. Пока же появились способы добычи углерода из атмосферного воздуха с использованием гибридных бактерий, которые при помощи

фотосинтеза продуцируют н-бутанол и полигидроксibuтират – эту технологию изобрели в Университете Калифорнии в Беркли [23].

Синтетические материалы, в основе которых углерод, станут самыми востребованными на рынке в эпоху трансгуманизма, из них будут делать роботов и всевозможные устройства, дополняющие человека. Запасы нефти, газа и угля, уровень развития нефтегазохимии и инновационные технологии искусственного фотосинтеза к концу Шестой Кондратьевской волны будут определять экономическое лидерство стран на мировом рынке. На рис. 16 представлено 16 стран, для которых выявлено выпадающее, большое количество ключевых словосочетаний «gas supplies / поставки газа», относящихся к 2015 году. Можно предположить, что эти страны будут активно осваивать новый рынок пластмасс для 3D печати в качестве пионеров.

На текущем начальном этапе уже разработана целая линейка материалов для 3D печати: полилактид (PLA, ПЛА), акрилонитрилбутадиенстирол (ABS, АБС), поливиниловый спирт (PVA, ПВА), нейлон (Nylon), поликарбонат (PC, ПК), полиэтилен высокой плотности (HDPE, ПНД), полипропилен (PP, ПП), поликапролактон (PCL), полифенилсульфон (PPSU), полиметилметакрилат (Acrylic, оргстекло, акрил, ПММА), полиэтилентерефталат (PET, ПЭТ), ударопрочный полистирол (HIPS), древесные имитаторы (LAYWOO-D3, BambooFill), имитаторы песчаника (Laybrick), имитаторы металлов (BronzeFill). Материалы будущего будут производиться с использованием нанотехнологий.

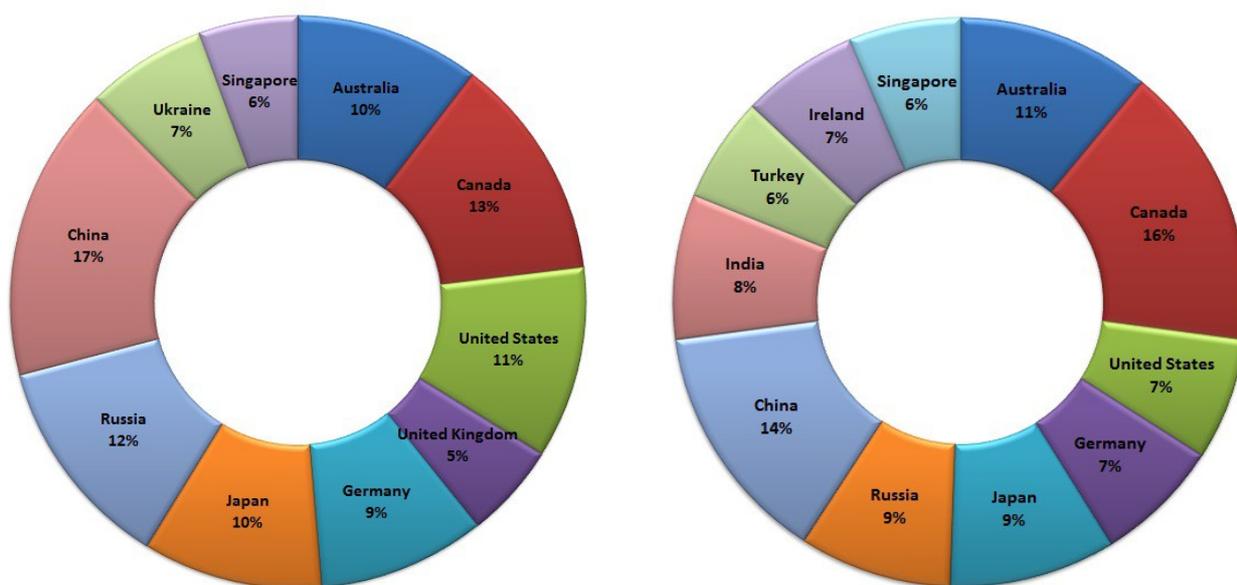


Рис. 17. Распределение стран, имеющих наибольшее число ключевых словосочетаний «Oil / Нефть» (слева) и «Oil price / Цена на нефть» (справа).

Интересным представляется рассмотреть внимание стран к нефти и цене на нефть (рис. 17) в контексте реакции на нефтяной биржевой кризис последних лет. Можно отметить, что США за период 2015 года реагировали меньше, чем Россия, Германия тоже, а Великобритания даже не попала в число стран, активно реагировавших на изменения цен на нефть. Эти три страны – США, Великобритания и Германия – уже перестроились на новый технологический уклад. Зато Китай и Канада особо выделяются повышенной реакцией на тему нефти. Нефтяной кризис задел экономики этих стран, и, например, в Канаде в 2016 году был подготовлен доклад «Canada’s New Trade and Technology Paradigm» (The Globe and Mail, июль 2016), в котором были поставлены новые приоритеты, в целом обозначенные как

цифровая Data-driven экономика: новые материалы; хранилища и устройства, аккумулирующие энергию; возобновляемая энергетика; робототехника; беспилотные машины; 3D печать; Интернет вещей; Облачные технологии; программное обеспечение уровня искусственного интеллекта [11, 20].

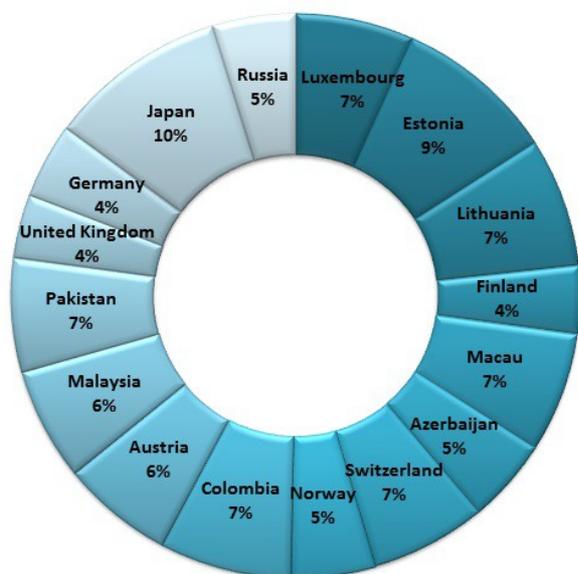
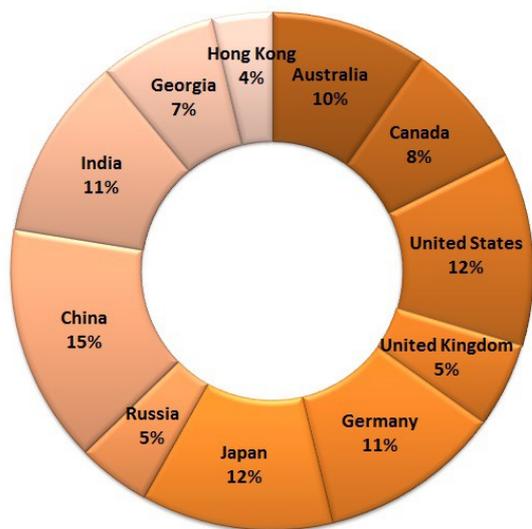


Рис. 18. Распределение стран, имеющих наибольшее число ключевых словосочетаний «drip irrigation / капельное орошение».

Общий уровень количества этого ключевого словосочетания в морфологической матрице не высок – тема обсуждается не так активно, как, например, нефть. Более высокий уровень количества словосочетаний «drip irrigation / капельное орошение», относящихся к 2015 году, определен для 16 стран из разных регионов, что косвенно указывает на интерес этих стран к данной технологии (рис. 18).

7) «Solar panel / Солнечные батареи». Это ключевое словосочетание отражает интерес стран к солнечной энергетике. Возобновляемые источники энергии являются одним из прогнозируемых важных трендов Шестой Кондратьевской волны [13] и экономики будущего (например, Venus Project Жака Фреско). В исследовании «Третья волна» ранее были выявлены сильные положительные корреляционные связи между уровнем ВВП стран (GDP PPP /purchasing power parity/, \$) и встречаемостью ключевого словосочетания «solar panel / солнечные батареи» в привязке к названию страны. А также суммарная мощность вырабатываемой электроэнергии в год на душу населения в странах (Electricity production, kWh/capita) имела сильную положительную корреляционную связь с ключевым словосочетанием «solar panel / солнечные батареи» [3]. Эта связь логически очень понятна, и она показывает, что солнечная энергетика стала реальным сектором экономики, она включена в основной тренд суммарной выработки электроэнергии в странах.

6) «Drip irrigation / Капельное орошение». Это ключевое словосочетание является индикатором очень важных трендов. Системы капельного орошения создают условия для сбора Больших данных по показателям полива полей, что формирует концепцию «электронных полей» и «умного сельского хозяйства» (Интернет вещей для сельского хозяйства). Такой подход обеспечивает новый уровень экономии и урожайности. Капельное орошение служит знаком внедрения новых технологий в сельское хозяйство, включая выращивание ГМО-культур. Есть еще один аспект – глобальное потепление, преодоление последствий которого в южных регионах тесно связано со стабильностью национальных экономик.



Из всего проанализированного ряда стран 11 стран характеризуются выпадающим, большим количеством ключевых словосочетаний «solar panel / солнечные батареи», относящихся к 2015 году. Это все страны из двух кластеров максимальной активности High level of Internet activity and citing, Cluster № 1 и Cluster № 3, а также Индия, Грузия и Гонконг из кластера средней активности Middle level of Internet activity and citing, Cluster № 2 (рис. 19).

Рис. 19. Распределение стран, имеющих наибольшее число ключевых словосочетаний «solar panel / солнечные батареи».

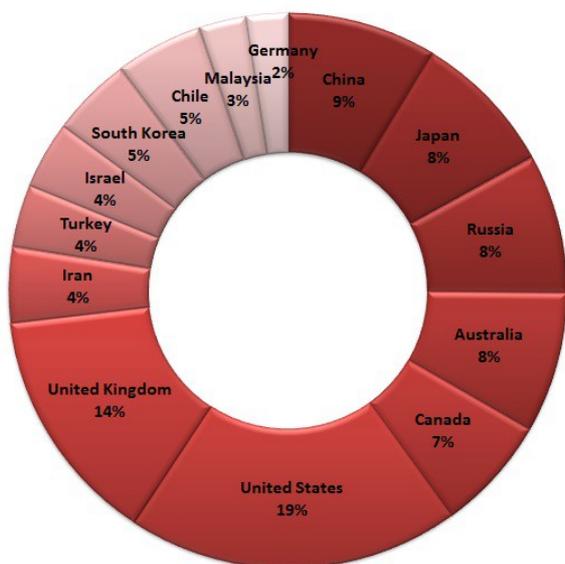


Рис. 20. Распределение стран, имеющих наибольшее число ключевых словосочетаний «nuclear power plant / атомная электростанция».

8) «Nuclear power plant / Атомная электростанция». Это ключевое словосочетание отражает интерес стран к тематике ядерной энергетики. На рис. 20 представлено 14 стран, для которых выявлено выпадающее, большое количество ключевых словосочетаний «nuclear power plant / атомная электростанция». В данном случае можно разделить страны на те, которые имеют атомные станции и те, которые вовлечены в активную дискуссию о возможности строительства атомной станции в стране (Австралия, Турция, Израиль, Чили, Малайзия). Эти пять стран входят в разные кластеры по активности стран в Интернете, Австралия – в кластер максимальной активности High level of Internet activity and citing, Cluster № 3. Остальные четыре страны – в кластеры средней активности Middle level of Internet activity and citing, Cluster № 2 и Cluster № 4.

IV. *Анализ модифицированной карты Value Chain Map по регионам*

На рисунках 21-26 представлено сравнение стран по выбранным ключевым словосочетаниям. Паттерны, отражающие количество подсчитанных слов, размещены по оси X в соответствии с картой «Цепочка ценности» (см. рис. 8). Анализируя соотношение паттернов, необходимо делать поправку на различие стран по локализации в кластерах, характеризующих активность в Интернете. С другой стороны, более активное присутствие страны в Интернете соотносится больше с уровнем ее экономического развития, чем с численностью населения страны.

На рис. 21 представлено сравнение количества ключевых слов для стран постсоветского пространства, из числа стран, взятых в исследование. Можно отметить, что из представленных стран повышенный интерес к Облачным технологиям, электромобилям и солнечной энергетике проявляют Россия и Грузия. В отношении ядерной энергетике наибольший интерес проявляет Россия. Хотя Россия и Грузия входят в более активные кластеры, все же абсолютной зависимости от принадлежности к кластерам на данном графике не видно. Например, Украина не показывает тех же позитивных инновационных паттернов, как Грузия и Россия. А для России выявлен слабый интерес к капельному орошению, в сравнении с прибалтийскими странами, входящими в кластер минимальной активности. Азербайджан вместе с Россией, Украиной и Грузией проявляет повышенный интерес к углеводородам и входит в перечень потенциальных лидеров будущего рынка новых пластмасс для 3D печати.

На рис. 22 представлено сравнение количества ключевых слов для выбранных стран Балтийского региона. Для Германии отмечен выпадающий, большой паттерн, указывающий на интерес к Облачным технологиям. Несмотря на то, что Россия входит в кластер максимальной активности, как и Германия, разница в пользу Германии очевидна. То же самое можно отметить и в отношении интереса к солнечной энергетике. А вот интерес к электромобилям почти сопоставим для России и Германии. Россия показывает наибольший интерес к ядерной энергетике.

На рис. 23 представлено сравнение количества ключевых слов для выбранных стран регионов MENA и Азии. Представленные страны входят либо в кластер средней активности, либо в кластер минимальной активности. Можно отметить повышенный интерес Турции к Облачным технологиям, незначительно повышенный интерес Пакистана к капельному орошению. Турция, Израиль и Иран выделяются повышенным интересом к теме ядерной энергетике (при этом Турция и Израиль не имеют пока АЭС, а только вовлечены в активную дискуссию по этому вопросу).

На рис. 24 представлено сравнение количества ключевых слов для выбранных стран Трансатлантического пространства. В отношении интереса к Облачным технологиям можно отметить четыре лидера – США, Великобритания, Германия и Канада. Среди паттернов, указывающих на интерес к электромобилям, отличаются повышенным количеством ключевых слов США, Канада, Великобритания, Германия и Россия. В отношении солнечной энергетике можно выделить те же пять стран, но все же лидерами показали себя США и Германия. Интерес к ядерной энергетике показали США, Великобритания, Канада и Россия. Несмотря на то, что большие паттерны показывают страны, входящие в кластер максимальной активности, разница между ними в разных паттернах очевидна. В отношении стран из кластеров с меньшей активностью можно заметить некоторые моменты: Швеция проявляет повышенный интерес к Облачным технологиям и солнечной энергетике; Ирландия – к электромобилям.

На рис. 25 представлено сравнение количества ключевых слов для стран Азиатско-Тихоокеанского региона, из числа стран, взятых в исследование. Повышенный интерес к Облачным технологиям отмечен у Китая, Японии, Австралии и Индии. Эти же страны проявляют повышенный интерес к солнечной энергетике. Китай и Япония имеют повышенное количество ключевых слов, отражающих интерес к электромобилям. И хотя в данную группу стран вошли четыре страны из кластеров максимальной активности, только

две из них показали высокие паттерны ключевых слов, отражающих интерес к электромобилям. В отношении интереса к ядерной энергетике можно отметить три страны – Китай, Японию и Австралию, при этом Австралия не имеет АЭС, а только активно участвует в дискуссии по поводу возможного строительства АЭС. Паттерн России в отношении интереса к ядерной энергетике сопоставим с Китаем и Японией.

На рис. 26 представлено сравнение количества ключевых слов для выбранных стран Южной Америки и Карибского бассейна. Этот график очень показателен для сравнительного анализа – для выбранных стран региона не обнаружено никаких выделяющихся паттернов по обозначенным ключевым словам, хотя страны относятся к разным кластерам со средней и минимальной активностью в Интернете.

Делая обзор всех описанных графиков (рис. 21-26) суммарно, можно выделить страны, которые показывают повышенный интерес к основным трендам наступающей Шестой Кондратьевской волны – Облачным технологиям и солнечной энергетике. Учитывая то, что в настоящее время идет понижательная фаза Пятой Кондратьевской волны, страны, показывающие опережающий интерес к технологиям наступающей новой волны экономического развития, обеспечивают себе лидерство, особенно в период повышательной фазы новой волны. Из списка обсуждаемых в данном исследовании стран, лидерские позиции можно прогнозировать для России и Грузии на постсоветском пространстве; для Германии в Балтийском регионе; среди рассмотренных стран регионов MENA и Азии лидера не выявлено; для региона Трансатлантического пространства лидерами показывают себя США и Германия; в Азиатско-Тихоокеанском регионе лидерские позиции за Китаем, Японией, Австралией и Индией; среди выбранных стран Южной Америки и Карибского бассейна лидер не обнаружен.

Россия имеет хороший потенциал развития в направлении рынка электромобилей. Повышенный интерес к электромобилям на постсоветском пространстве показывает вместе с Россией Грузия. В регионах мира конкуренцию России на этом рынке составят Германия, США, Канада, Великобритания, Китай и Япония.

Выгодным преимуществом для России в связи с прогнозируемой возрастающей потребностью мирового рынка (IoT) в электроэнергии, является развитая атомная энергетика. По данным представленного исследования лидерство России в этом направлении прогнозируется не только на постсоветском пространстве, но и в Балтийском регионе. Лидерами также себя позиционируют США, Канада, Великобритания, Китай, Япония. Для регионов MENA и Азии из списка представленных стран лидером показывает себя Иран. Данные анализа ключевых слов перекликаются с реальным рынком мирного атома, который делят американско-японская компания «Westinghouse» и российский Росатом.



Страны постсоветского пространства

Кластеры:

- Россия – Cluster № 3, High level of Internet activity and citing;
- Эстония, Литва – Cluster № 5, Low level “_”;
- Украина, Грузия – Cluster № 2, Middle level “_”;
- Азербайджан – Cluster № 5, Low level “_”;
- Киргизия, Узбекистан – Cluster № 5, Low level “_”.

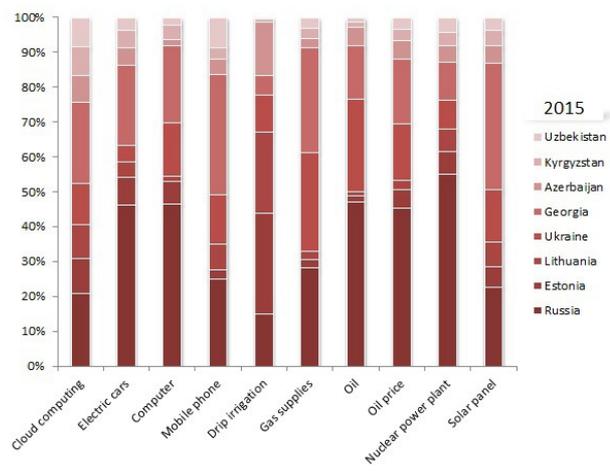
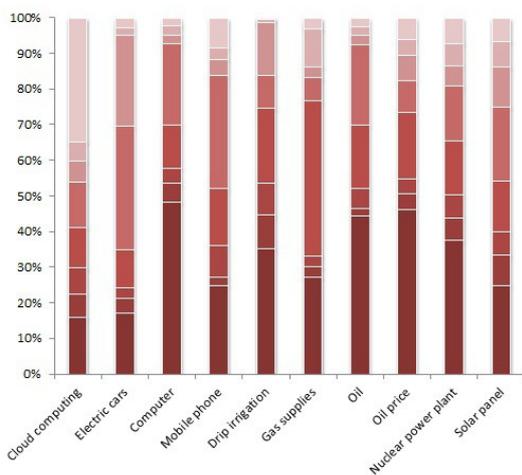
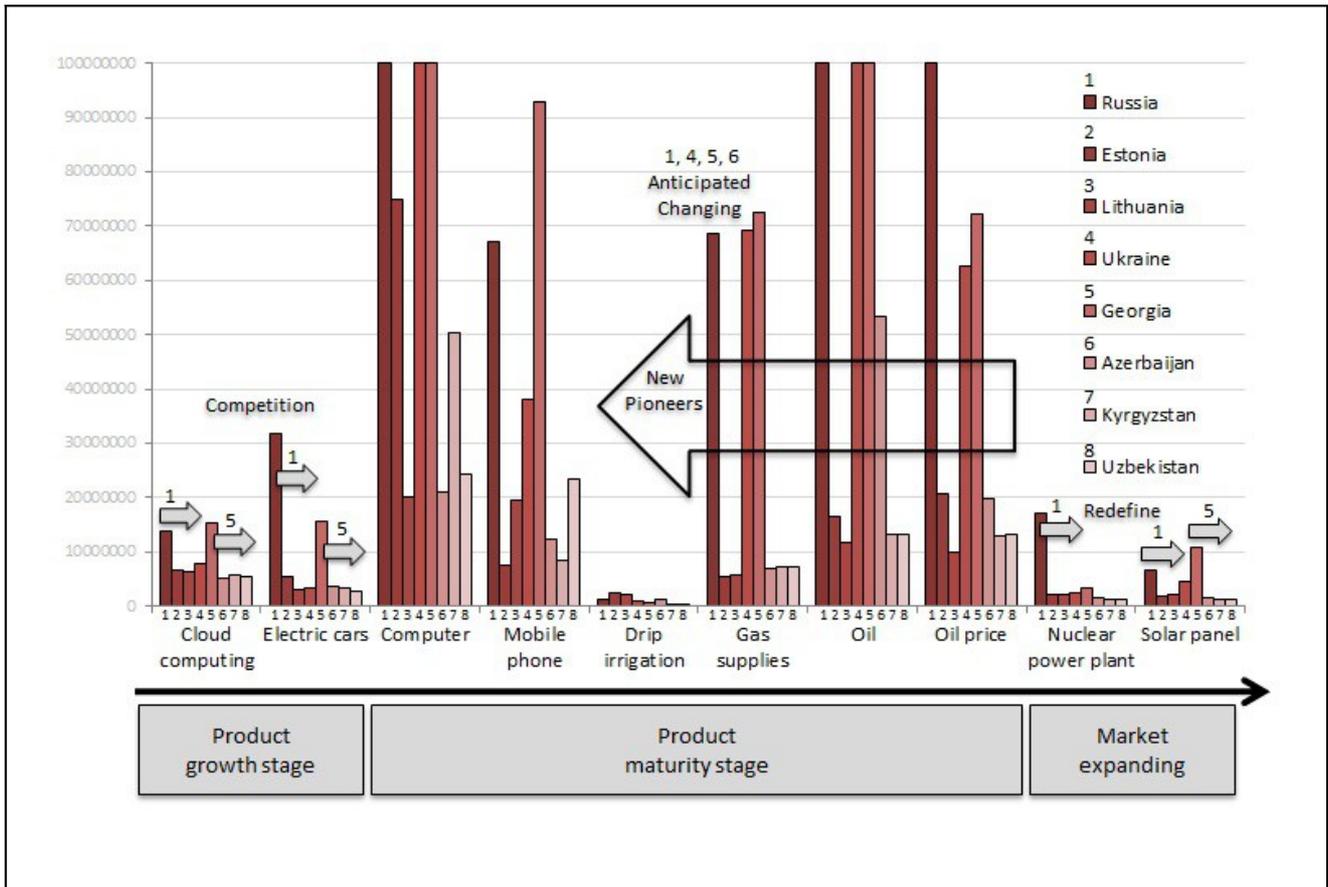


Рисунок 21. Сравнение количества ключевых слов для стран постсоветского пространства.



Страны Балтийского региона

Кластеры:

- Россия – Cluster № 3, High level of Internet activity and citing;
- Финляндия – Cluster № 5, Low level “_”;
- Дания – Cluster № 4, Middle level “_”;
- Германия – Cluster № 1, High level “_”;
- Швеция – Cluster № 4, Middle level “_”;
- Эстония, Литва – Cluster № 5, Low level “_”.

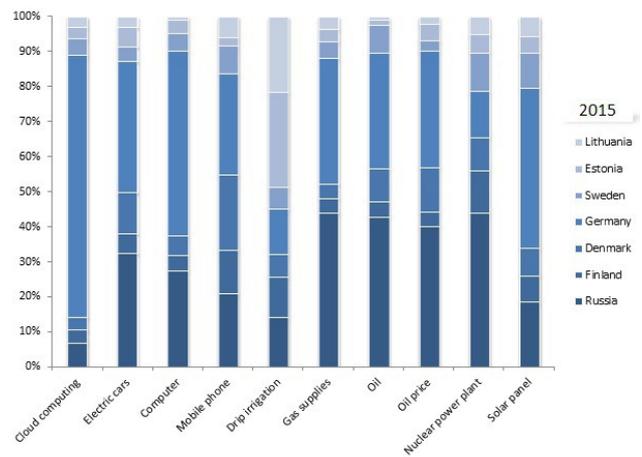
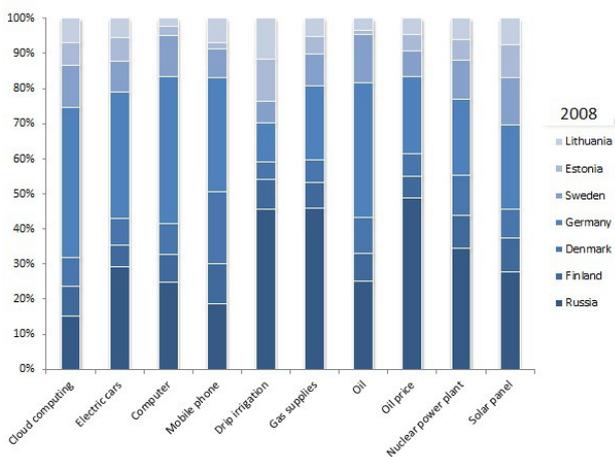
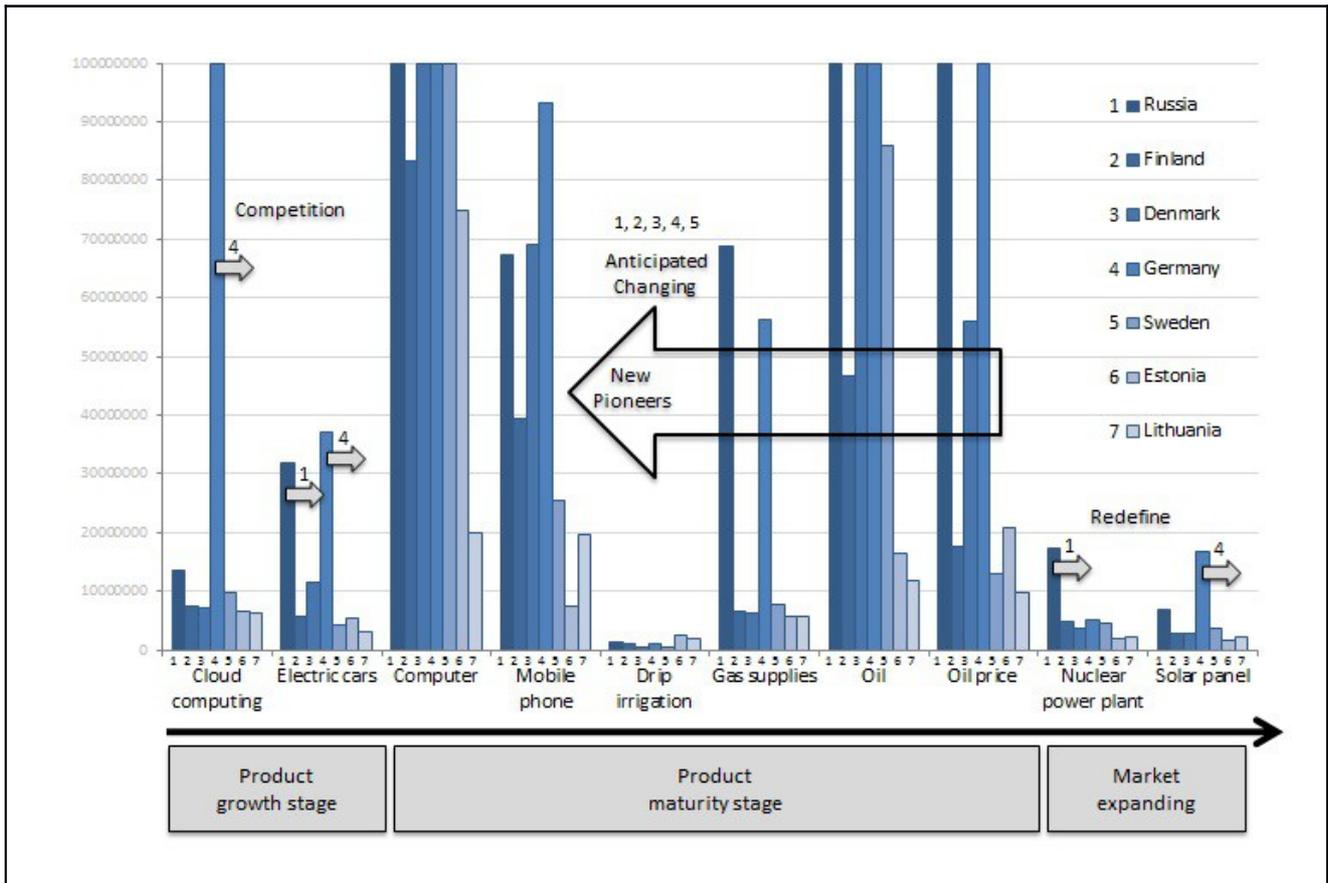


Рисунок 22. Сравнение количества ключевых слов для стран Балтийского региона.



Страны MENA (Middle East & North Africa) и Азии

Кластеры:

- Турция – Cluster № 2, Middle level of Internet activity and citing;
- Сирия, Израиль – Cluster № 4, Middle level “_”;
- Палестина – Cluster № 5, Low level “_”;
- Ирак, Иран, Афганистан, Пакистан – Cluster № 4, Middle level “_”;
- Тунис, Ливия – Cluster № 5, Low level “_”;
- Египет – Cluster № 4, Middle level “_”;
- Йемен, Бахрейн – Cluster № 5, Low level “_”.

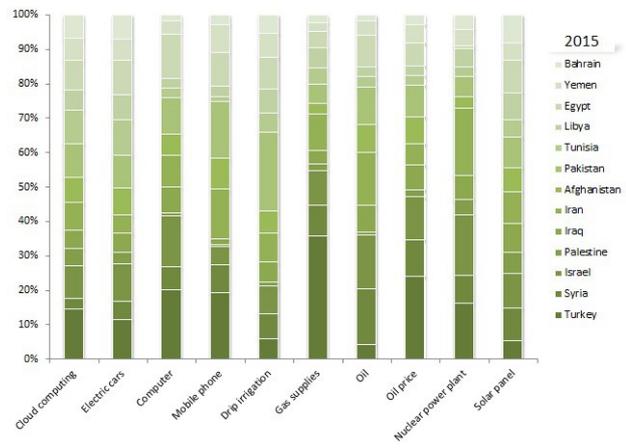
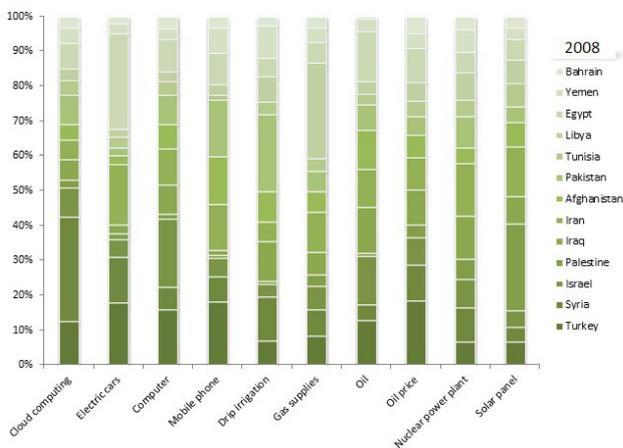
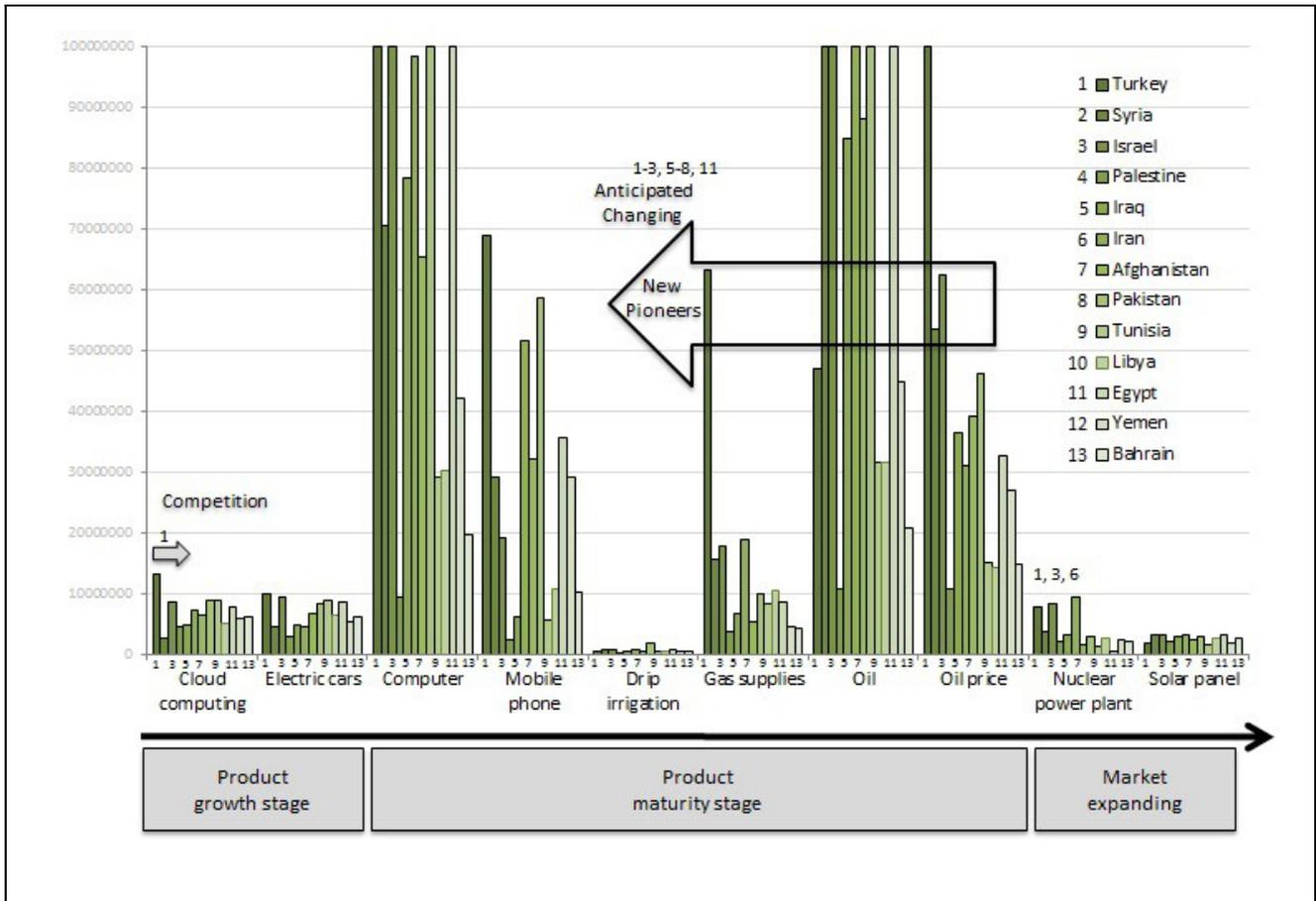


Рисунок 23. Сравнение количества ключевых слов для стран регионов MENA и Азии.



Страны Трансатлантического пространства

Кластеры:

- США, Канада – Cluster № 1, High level of Internet activity and citing;
- Великобритания – Cluster № 3, High level “_”;
- Ирландия – Cluster № 2, Middle level “_”;
- Германия – Cluster № 1, High level “_”;
- Дания, Норвегия, Швеция – Cluster № 4, Middle level “_”;
- Финляндия – Cluster № 5, Low level “_”;
- Австрия – Cluster № 4, Middle level “_”;
- Люксембург – Cluster № 5, Low level “_”;
- Швейцария – Cluster № 4, Middle level “_”;
- Россия – Cluster № 3, High level “_”.

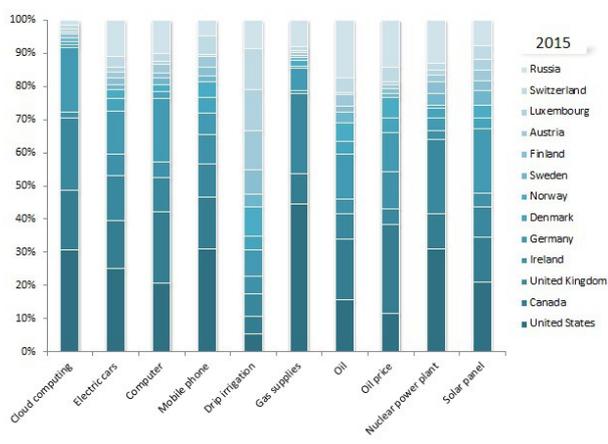
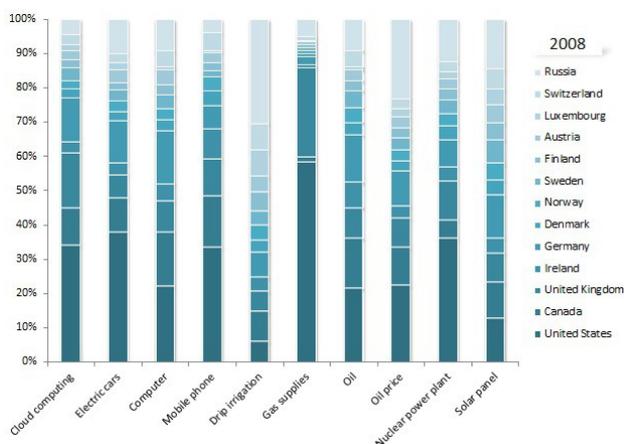
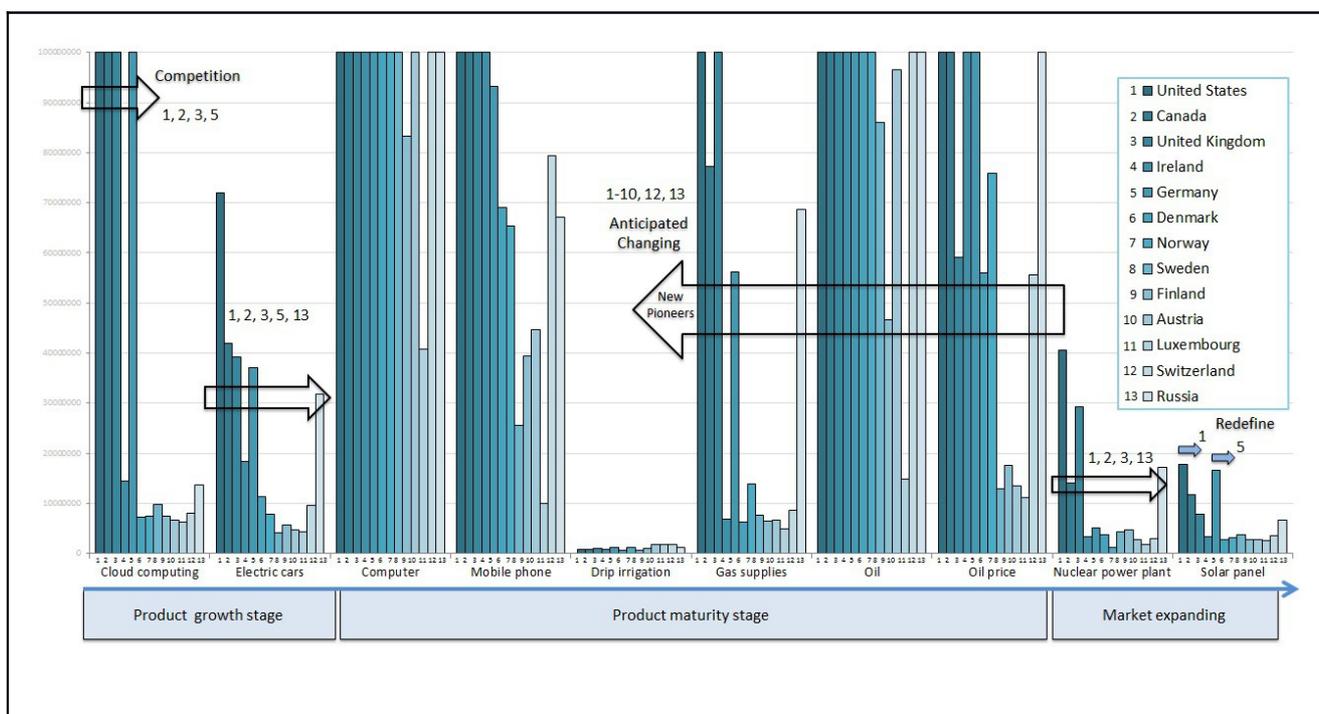


Рисунок 24. Сравнение количества ключевых слов для стран Трансатлантического пространства.



Страны Азиатско-Тихоокеанского региона

Кластеры:

- Китай – Cluster № 1, High level of Internet activity and citing;
- Гонконг – Cluster № 2, Middle level “_”;
- Макао – Cluster № 5, Low level “_”;
- Тайвань, Южная Корея – Cluster № 2, Middle level “_”;
- Япония – Cluster № 3, High level “_”;
- Сингапур, Малайзия, Новая Зеландия – Cluster № 2, Middle level “_”;
- Австралия – Cluster № 3, High level “_”;
- Индия – Cluster № 2, Middle level “_”;
- Россия – Cluster № 3, High level “_”.

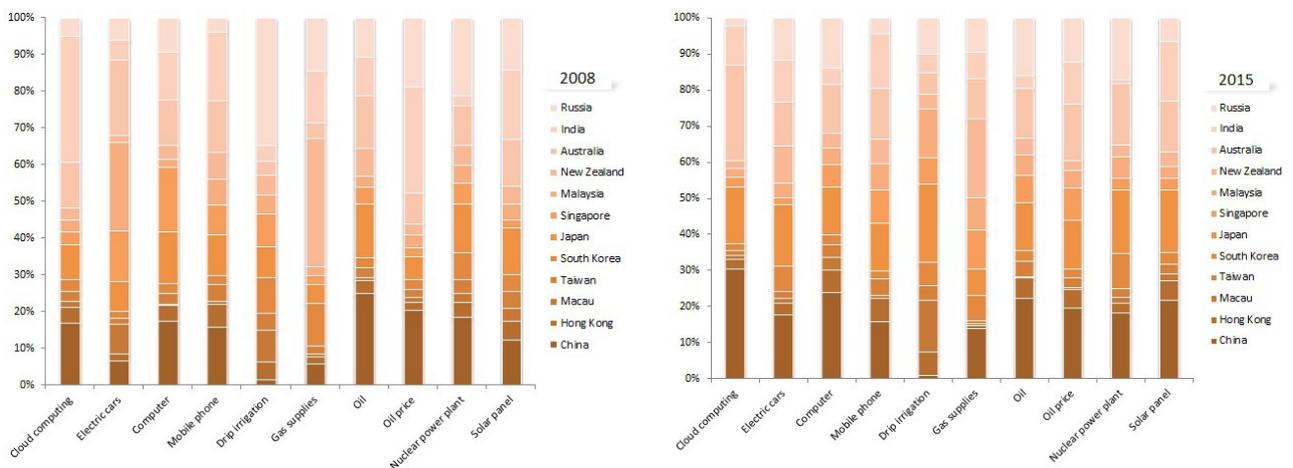
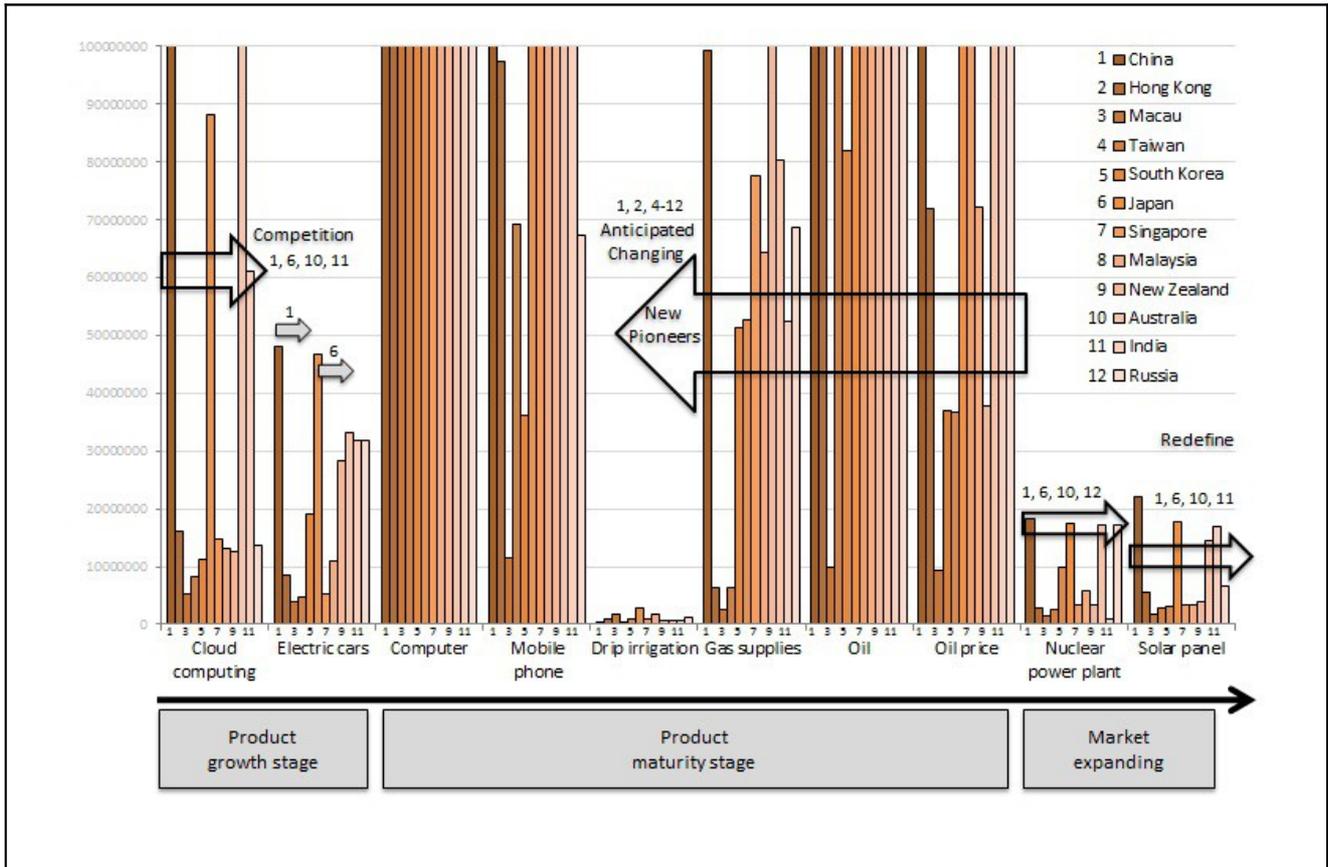


Рисунок 25. Сравнение количества ключевых слов для стран Азиатско-Тихоокеанского региона.



Страны Южной Америки и Карибского бассейна

Кластеры:

- Сент-Люсия – Cluster № 5, Low level “_”;
- Барбадос – Cluster № 5, Low level “_”;
- Колумбия – Cluster № 4, Middle level “_”;
- Перу – Cluster № 5, Low level “_”;
- Чили – Cluster № 4, Middle level “_”.

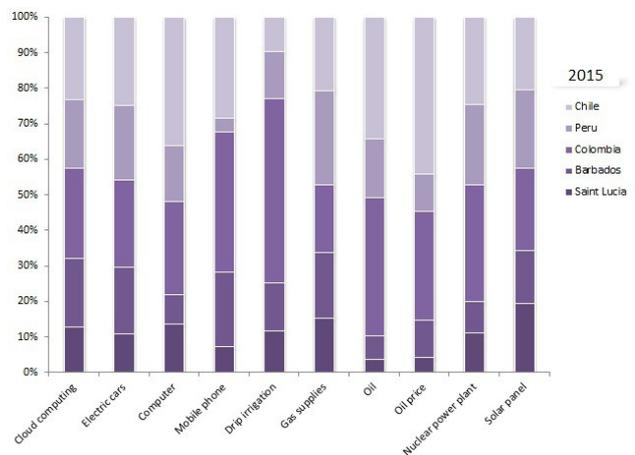
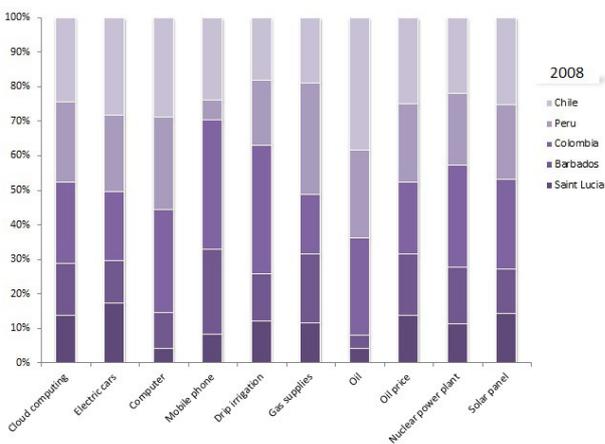
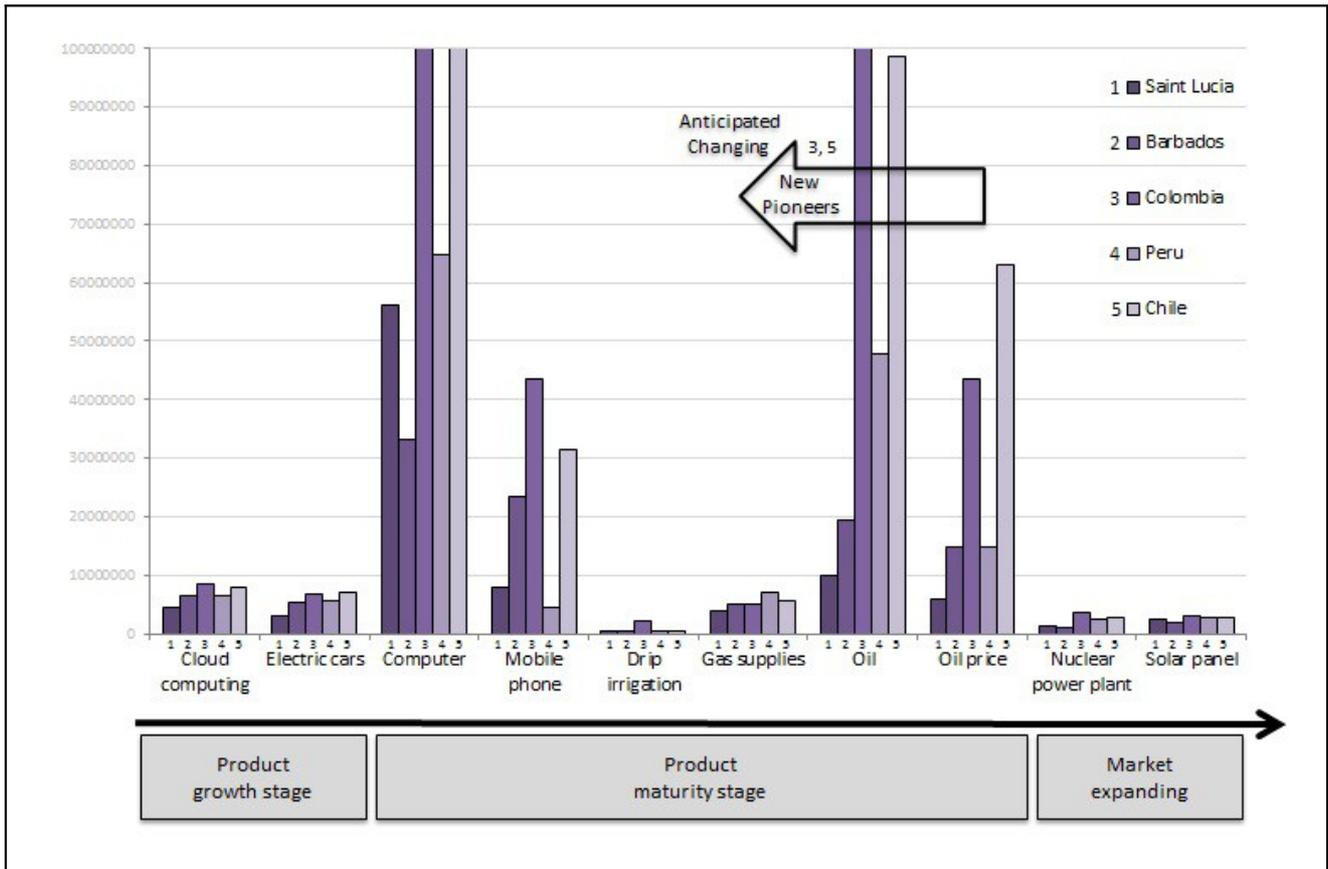


Рисунок 26. Сравнение количества ключевых слов для стран Южной Америки и Карибского бассейна.

Заключение и выводы

Среда Интернета открывает много возможностей для текстовой аналитики Больших данных. Даже морфологический подход, подразумевающий подсчет отдельных ключевых слов, при применении дополнительных методик анализа позволяет получать важную информацию из данных, которые по сути можно относить к Dark Data (данным, не имевшим ранее применения). Dark Data – «темные» или скрытые данные – являются сегодня самой главной темой в аналитике Big Data. По определению словаря Gartner IT Glossary (www.gartner.com) Dark Data относят к тем данным, которые накапливаются, но не имеют никакого использования для аналитики. Эксперты в области ИТ обычно оценивают объем Dark Data от 85 до 97% от всего объема оцифрованных Big Data, хотя это проверить невозможно, и объем скрытых и неучтенных данных может достигать 99,9%. Именно с Dark Data, то есть с креативным поиском новых параметров для аналитики, которые не имеют никакой классификации и не использовались ранее [26], и связывают будущее аналитики Big Data [16].

Морфологическая текстовая аналитика – это на сегодняшний день доступный и разрабатываемый пласт аналитики текстовых Big Data, бесконечно накапливающихся в Интернете. Поисковые системы и соцсети предоставляют сервисы доступа к данным с небольшой разницей в качестве и подходах, в двух видах: для уровня простого пользователя (сложно делать какие-либо исследования, так как выдается упрощенный, готовый продукт с заданными от разработчика параметрами) и для специалиста в области ИТ (можно составлять дизайн исследования и проводить научную работу). Поиск и подсчет, а иногда и графическая обработка, проводятся по словам, которые могут иметь разную функцию – ключевые слова в текстах (keywords), гиперссылки на слова, формирующие дерево обсуждения (#hashtags), слова как названия сайтов-источников (sources), слова-запросы из поисковой строки (web search queries) [17]. Это огромный объем информации, который пока что весь находится в тени Dark Data, мало что из него учитывается и мало что изучается. Если можно допустить такое сравнение, то замер числа слов можно сравнить с измерением общей температуры тела у пациента. Это не дает детальных, конкретных ответов на диагностические вопросы, но все же достаточно четко указывает в нужном направлении.

Для молодых специалистов в области ИТ, выпускающихся из вузов и выходящих на рынок труда, тема Big Data чрезвычайно важна. Например, согласно данным World Economic Forum [19] прогнозируемый рынок по аналитике Big Data только в США уже требует заполнения 1,5 млн рабочих мест, а в ближайшие годы в американской экономике возникнет до 4,5 млн новых рабочих мест в области аналитики Big Data [19]. Развитию этого рынка в США помогает федеральная программа The Federal Big Data Initiative (действующая до 2017 года), стимулирующая по всем штатам разработку инноваций в сфере Big Data [28]. Этот прогноз можно экстраполировать и на Россию, то есть как минимум 4 млн рабочих мест в разных секторах экономики в ближайшей перспективе будут связаны с аналитикой Big Data.

Текстовая аналитика Big Data уже сегодня в России имеет свое место на рынке, и выпускнику вуза важно иметь в своем резюме не только упоминание об обучении и сертификации по тематике Big Data, но и информацию об участии в проекте, который доведен до наглядных результатов. Зачастую именно это является проблемой, так как даже в рамках грантовых научных работ руководители исследовательских групп в итоге не могут показать наглядно вывод по заказанной у ИТ-специалистов аналитике Big Data. И специалисту, создавшему свой математический алгоритм или математическую модель, нечего предъявить в списке своих профессиональных достижений в качестве участия в успешных проектах по аналитике Big Data. Именно с этой проблемой борется Академическое партнерство Dell EMC в России, собирая команды для аналитики Big Data так, чтобы обязательно были показаны наглядные выводы по проведенному исследованию, что напрямую поддерживает процесс развития рынка труда в области аналитики Big Data. Это новая научная область, мало изученная и постоянно растущая и меняющаяся. Каждый вид аналитики фактически создается исполнителями впервые, а в качестве ориентации

исследователи сравнивают в большей степени не результаты, а сам дизайн аналитики Big Data в успешных и доведенных до выводов и практического применения проектах [7].

К основным выводам данного этапа исследования «Третья волна» можно отнести следующее:

- Изучая морфологическую матрицу ключевых слов необходимо проводить кластерный анализ, позволяющий определять место стран в ряду упоминания в Интернете и динамику цитируемости выбранных слов;
- Анализ морфологической матрицы должен обязательно быть интегрирован в методики логического анализа, как в данном случае была применена карта «Цепочка ценности» (Value Chain Map) и теория Кондратьевских волн;
- Исследование выявило признаки смены технологического уклада с растущим новым рынком Облачных технологий;
- Данное исследование посвящено важной задаче по изучению возможности для аналитики среды Интернета через базы данных международных поисковых систем, функционирующих как неклассические суперкомпьютеры. Разработка разных вариантов анализа больших массивов текстов вносит свой вклад в аналитику, проводимую на суперкомпьютерах с любыми текстовыми базами Больших данных. Более того, анализ текстовых массивов Интернета позволяет изучать не просто текстовые Большие данные, а формирование глобальной системы «People-to-IT», что имеет глубокие перспективы для многих отраслей науки.

Список литературы

- [1] Гринин Л.Е. Кондратьевские волны, технологические уклады и теория производственных революций // Кондратьевские волны. Аспекты и перспективы / Ред. А.А. Акаев, Р.С. Гринберг, Л.Е. Гринин, А.В. Коротаев, С.Ю. Малков. – Волгоград: Изд-во «Учитель», 2012. С. 222-262.
- [2] Гринин Л.Е., Гринин А.Л. Кибернетическая революция и шестой технологический уклад // Кондратьевские волны: наследие и современность. Ежегодник / Ред. Л.Е. Гринин, А.В. Коротаев, В.М. Бондаренко. – Волгоград: Изд-во «Учитель», 2015. С. 83-106.
- [3] Колесниченко О.Ю., Смородин Г.Н., Ильин И.В., Журенков О.В., Мазелис Л.С., Яковлева Д.А., Дашонок В.Л. Текстовая аналитика Big Data: перспективы для суперкомпьютеров // Национальный Суперкомпьютерный Форум, 2015. Электронная публикация докладов, статья 09-475, секция «Суперкомпьютерные технологии при работе с большими объемами данных (Big Data)». Институт программных систем имени А.К. Айламазяна РАН. Переславль-Залесский. 24-27 ноября 2015 г.
- [4] Колесниченко О.Ю., Смородин Г.Н., Яковлева Д.А. Мазелис Л.С., Колесниченко Ю.Ю. Исследование глобального образа кризиса методом API-социологии // Тезисы VI международной социологической Грушинской конференции «Жизнь исследования после исследования: как сделать результаты понятными и полезными», 16-17 марта 2016. Фонд «ВЦИОМ», ФГБОУ ВПО «РАНХиГС при Президенте РФ» / Отв. ред. А.В. Кулешова. – М.: ВЦИОМ, 2016. С. 982-987.
- [5] Кондратьев Н.Д., Яковец Ю.В., Абалкин Л.И. Большие циклы конъюнктуры и теория предвидения. Избранные труды. – М.: Экономика, 2002. 766 с.
- [6] Коротаев А.В., Цирель С.В. Кондратьевские волны в мировой экономической динамике // Системный мониторинг. Глобальное и региональное развитие / Ред. Д.А. Халтурина, А.В. Коротаев. – М.: Либроком / URSS, 2010. С. 189-229.

- [7] Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / Пер. с англ. Гайдюк И. – М.: Манн, Иванов и Фербер, 2014. 240 с.
- [8] Тоффлер Э. Третья волна / Пер. с англ. – М.: Издательство АСТ, 2004. 781 с.
- [9] Форрестер Д. Мировая динамика / Пер. с англ. – М.: Изд-во АСТ; СПб.: Terra Fantastica, 2003. 379 с.
- [10] Almasri R., Muneer T., Cullinane K. The effect of transport on air quality in urban areas of Syria. *Energy Policy*, 39 (2011), January 2011. pp. 3605–3611. DOI: 10.1016/j.enpol.2011.03.062.
- [11] Building Smarter Manufacturing With The Internet of Things (IoT). Lopez Research LLC, San Francisco, CA. 2014.
- [12] Canada's New Trade and Technology Paradigm. Report. The Conference Board of Canada, 2016.
- [13] Cuthbertson A. «Moth eye» graphene breakthrough could create indoor solar cells. *Newsweek*, 2.27.2016. [Электронный ресурс]. URL: Europe.newsweek.com/indoor-solar-cells-made-possible-motheye-graphene-breakthrough-430827 (дата обращения: 10.09.2016).
- [14] Cuthbertson A. How to turn your smartphone into a 3D printer. *Newsweek*, 3.24.2016. [Электронный ресурс]. URL: Europe.newsweek.com/3dprinting-olo-smartphone-printer-440033?rm=eu (дата обращения: 10.09.2016).
- [15] Davidson G. 3D printing and the Cloud: a natural fit for manufacturers. *Industry Week*, 2015. [Электронный ресурс]. URL: <http://www.industryweek.com/emerging-technologies/3-d-printing-and-cloud-natural-fit-manufacturers> (дата обращения: 10.09.2016).
- [16] Dennies P. Factories Of The Future: The Value Of Dark Data. *Forbes*. Feb 19, 2015. [Электронный ресурс]. URL: <http://www.forbes.com/sites/teradata/2015/02/19/factories-of-the-future-the-value-of-dark-data/#6c45a1d1fa6c> (дата обращения: 10.09.2016).
- [17] Dietrich D., Heller B., Yang B. *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. USA: Wiley, 2015.
- [18] E-Thrust. The Vision and Perspective of an Electrical Distributed Propulsion System. e-Brochure. Airbus Group & Rolls-Royce plc. [Электронный ресурс]. URL: <http://www.airbusgroup.com/int/en/news-media/media~item=be674db0-2b69-42f6-946a-03cf3b0eef32~.html> (дата обращения: 10.09.2016).
- [19] Gupta A. Making Big Data Something More than the «Next Big Thing» // *The Global Information Technology Report: Rewards and Risks of Big Data*. World Economic Forum, 2014.
- [20] Hodgson G., Goldfarb D. Canada's new thinking must focus on technology, talent and trade. *The Globe and Mail*, Jul. 27, 2016. [Электронный ресурс]. URL: <http://www.theglobeandmail.com/report-on-business/rob-commentary/canadas-new-thinking-must-focus-on-technology-talent-and-trade/article31119293/> (дата обращения: 10.09.2016).
- [21] Koeppe H.R., Laskowski N., Nickolaisen N. *3D Printing Is at a Tipping Point*, e-publication, TechTarget, 2016.
- [22] Kolesnichenko O., Smorodin G., Yakovleva D., Mazelis L., Kolesnichenko Yu. Kondratiev's Waves, Big Data and Value Chain Mapping. 18th Conference of Open Innovations Association FRUCT and Seminar on Information Security and Protection of Information Technology (The 18th FRUCT & ISPIT Conference). National Research University of Information Technologies, Mechanics and Optics (ITMO). Finnish-Russian University Cooperation in Telecommunications. St. Petersburg, Russia, 18-22 April 2016. pp. 507-514.

- [23] Liu C., Gallagher J.J., Sakimoto K.K., Nichols E.M., Chang C.J., Chang M.C.Y., Yang P. «Nanowire-Bacteria Hybrids for Unassisted Solar Carbon Dioxide Fixation to Value Added Chemicals», *Nano Lett.*, 2015, 15 (5), pp. 3634–3639. DOI: 10.1021/acs.nanolett.5b01254.
- [24] Manyika J., Ramaswamy S., Khanna S., Sarrazin H., Pinkus G., Sethupathy G., Yaffe A. *Digital America: A tale of the haves and have-mores*. McKinsey Global Institute Report. New York, United States. December 2015. [Электронный ресурс]. URL: <http://www.mckinsey.com/industries/high-tech/our-insights/digital-america-a-tale-of-the-haves-and-have-mores> (дата обращения: 10.09.2016).
- [25] Mearian L. Start-up plans to release 3D-printed, street-ready cars next year. *Computer World*, Jul 9, 2015. [Электронный ресурс]. URL: www.computerworld.com/article/2946022/3d-printing/start-upplans-to-release-3d-printed-street-ready-cars-next-year.html (дата обращения: 10.09.2016).
- [26] Miles D. Dark Data could halt Big Data's path to success. December 2013. [Электронный ресурс]. URL: <http://www.computerweekly.com/opinion/Dark-data-could-halt-big-datas-path-to-success> (дата обращения: 10.09.2016).
- [27] Pepper R., Biggs P., Garrity J., LaSalle C., Polomska A. *Harnessing the Internet of Things for Global Development*. IoT Technologies (ITU). Contribution to the ITU/UNESCO Broadband Commission for Sustainable Development, 2015.
- [28] *The Massachusetts Big Data Report: A Foundation For Global Leadership*. USA, 2014.
- [29] 3D printing, Cloud computing collide changing the face of small-scale manufacturing. *Global Manufacturing*. May 26, 2015. [Электронный ресурс]. URL: www.manufacturingglobal.com/technology/449/3d-printingcloud-computing-collide-changing-the-face-of-smallscalemanufacturing (дата обращения: 10.09.2016).
- [30] Wardley S. *Open Source Convention – OSCON 2014: Anticipating the Future: An Introduction to Value Chain Mapping*. Keynote lecture. [Электронный ресурс]. URL: www.youtube.com/watch?v=NnFeIt-uaEc (дата обращения: 10.09.2016).
- [31] Wardley S. *Open Source Convention – OSCON 2015: Situation Normal, Everything Must Change*. Keynote lecture. [Электронный ресурс]. URL: www.youtube.com/watch?v=Ty6pOVEc3bA (дата обращения: 10.09.2016).
- [32] *What is the future of Cloud computing? Dell official website*. [Электронный ресурс]. URL: Powermore.dell.com/micro-sites/future-of-cloud-computing (дата обращения: 10.09.2016).

KOLESNICHENKO Olga, Editor in Chief of Security Analysis Bulletin, PhD, Moscow; Russia;

E-mail: oykolesnichenko@list.ru

SMORODIN Gennady, Head of Dell EMC External Research and Academic Alliances Russia & CIS, PhD, MBA, St. Petersburg; Russia;

E-mail: gennady.smorodin@dell.com

YAKOVLEVA Dariya, Assistant of the Department of Information Technology and Systems, Vladivostok State University of Economics and Service, Vladivostok; Russia;

E-mail: darya.yakovleva15@vvsu.ru

MAZELIS Lev, Head of the Department of Mathematics and Modeling, Vladivostok State University of Economics and Service, Professor, Doctor of Economic Sciences, Vladivostok; Russia;

E-mail: lev.mazelis@vvsu.ru

BALANDIN Sergey, General Chair of Open Innovations Association FRUCT (Finnish-Russian University Cooperation in Telecommunications), Professor of Saint Petersburg National Research University of Information Technologies, Mechanics and Optics (ITMO University), Professor of Tampere University of Technology (Finland), PhD, MBA, St. Petersburg; Russia;

E-mail: Sergey.Balandin@fruct.org

KOLESNICHENKO Yuriy, Chief Technical Officer of Cybersecurity, Security Analysis Bulletin, Moscow; www.SecurityAnalysisBulletin.com; Moscow; Russia;

E-mail: green-apple_2000@mtu-net.ru

Text Big Data Analytics

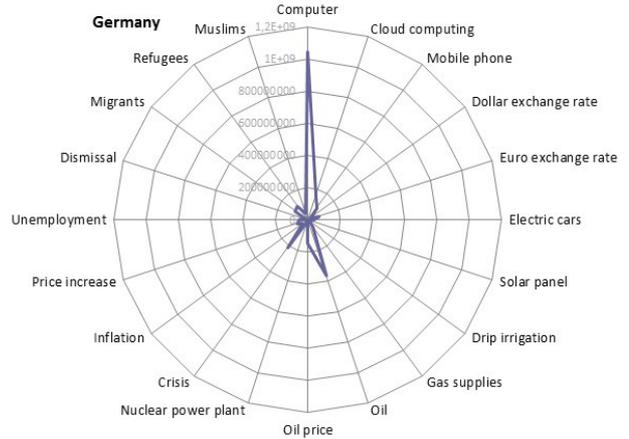
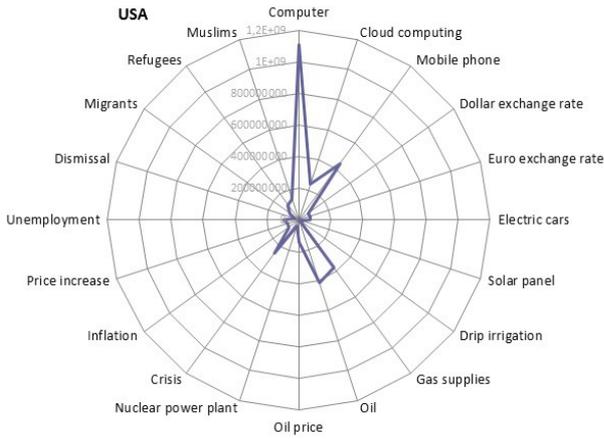
The second phase of the Study provided by Dell EMC External Research and Academic Alliances

Abstract. The second phase of Text Big Data Analytics Study «Third Wave» provided by Dell EMC External Research and Academic Alliances in Russia is presented. The study named after conceptual work of the American philosopher Alvin Toffler. The analysis has included 49 countries from different regions of the world. Data Mining was implemented by counting some keywords using API-Google and API-Yandex from open Internet resources. Both Google and Yandex are considered as non-classic supercomputers «as-a-Service» Cloud model. Dariya Yakovleva is the author of Data Mining technique using API-Google and API-Yandex. The Morphological Matrix of 20 key phrases was collected. The modified Value Chain Map of Simon Wardley was used for results analysis in terms of strategic market development. Cluster analysis k-means was implemented by Dariya Yakovleva. Countries were divided into different clusters of High, Middle and Low level of Internet activity and citing. Cluster analysis of Morphological Matrix allowed to find the important indicator of world technological order changing, that is increasing of number of keyword «Cloud computing» from 2008 to 2015. The theory of Kondratiev waves was used for interpretation of results analysis. Thorough keywords the features of Kondratiev waves in countries were investigated.

Key Words and Phrases: Text Big Data Analytics, Application Programming Interface, Data Mining, non-classic supercomputers, Google, Yandex, cluster analysis k-means, Value Chain Map, morphological analysis, Kondratiev waves.

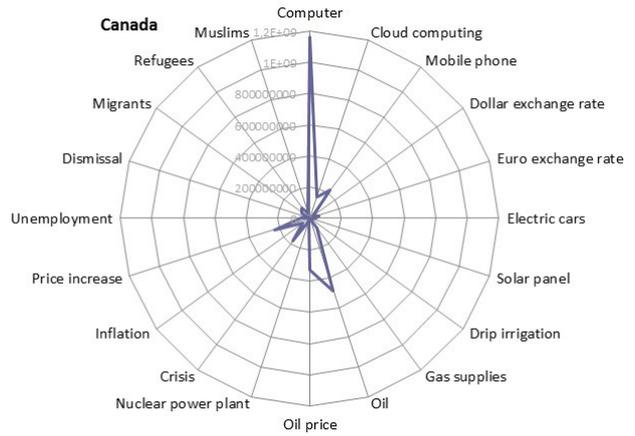
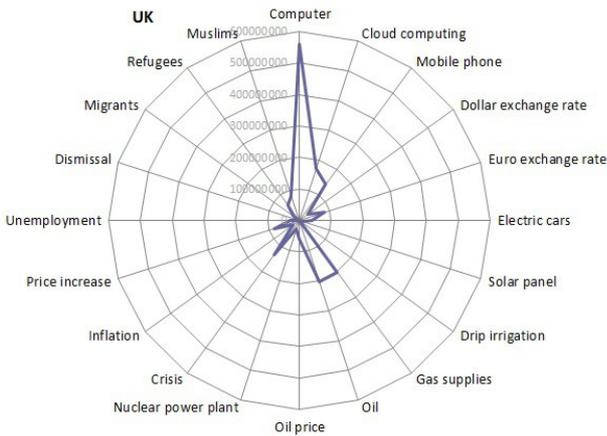
Приложение 1

**Географический захват начала Шестой Кондратьевской волны
и воздействие неравенства и стресс-нагрузки**



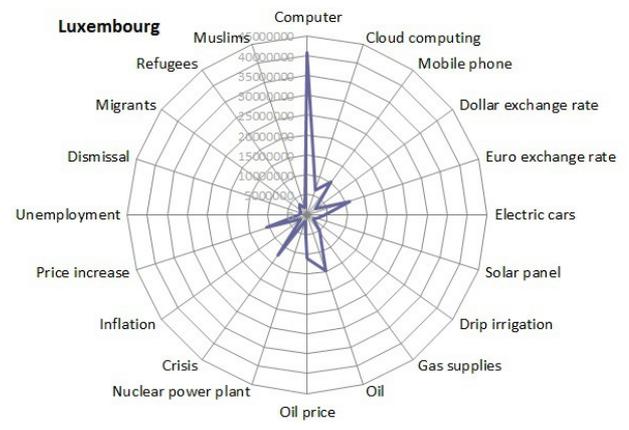
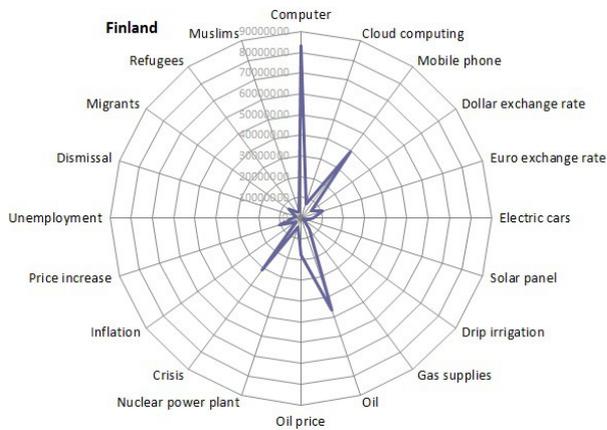
*Лидер географического захвата шестой К-волны.
Доминирование Unit «Информация» / «Information»*

Доминирование Unit «Информация» / «Information»



Доминирование Unit «Информация» / «Information»

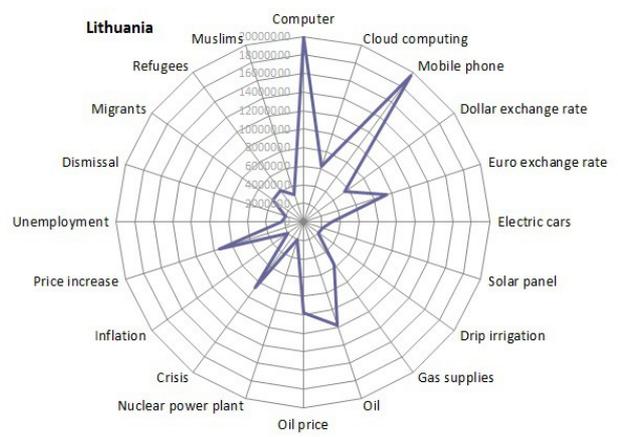
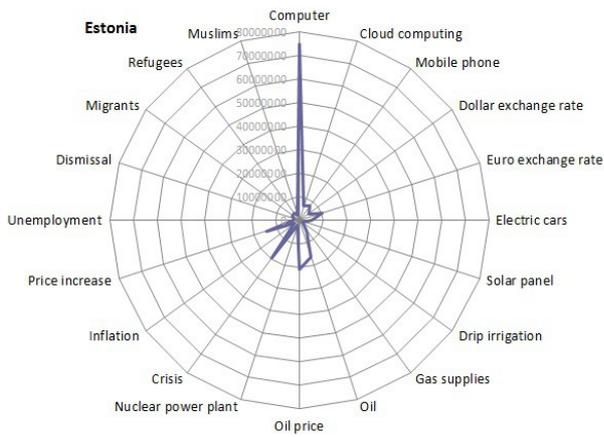
Доминирование Unit «Информация» / «Information»



Доминирование Unit «Информация» / «Information»

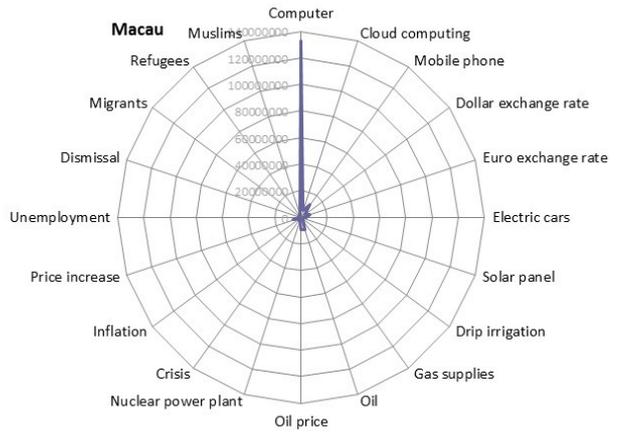
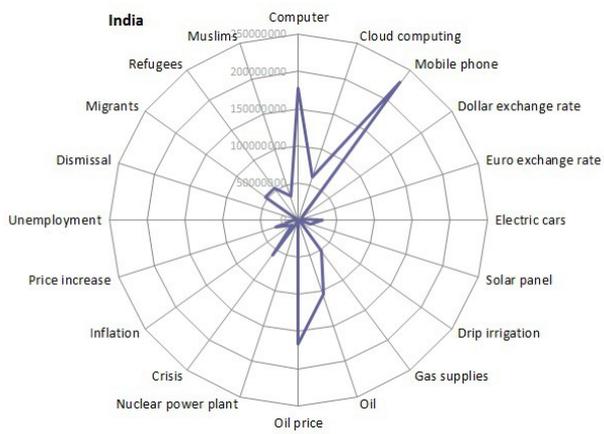
Доминирование Unit «Информация» / «Information»

Продолжение Приложения 1



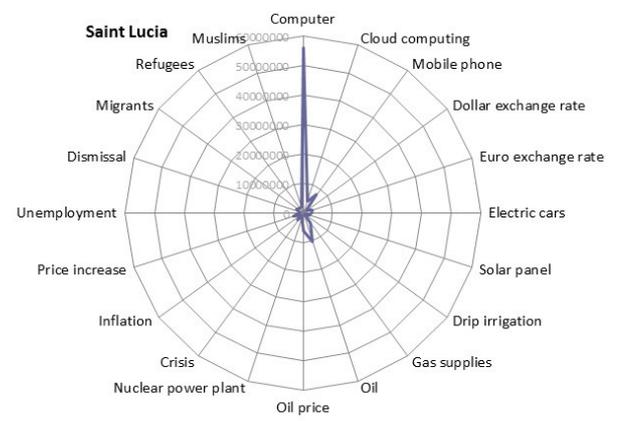
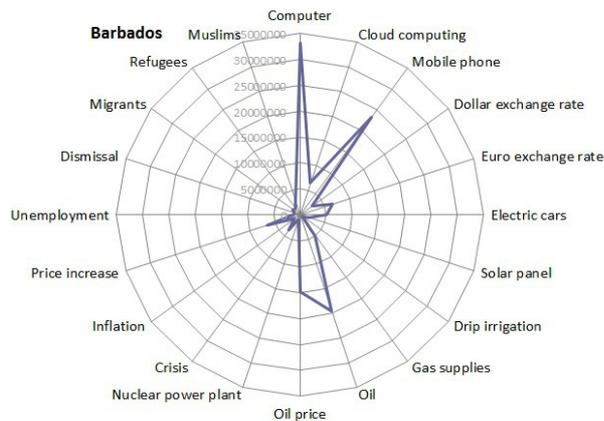
Доминирование Unit «Информация» / «Information»

Доминирование Unit «Информация» / «Information»



Доминирование Unit «Информация» / «Information»

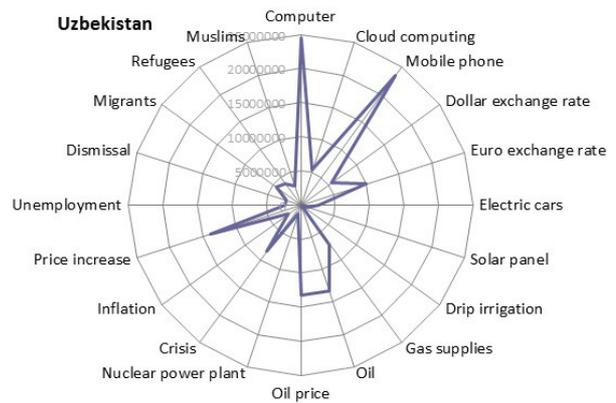
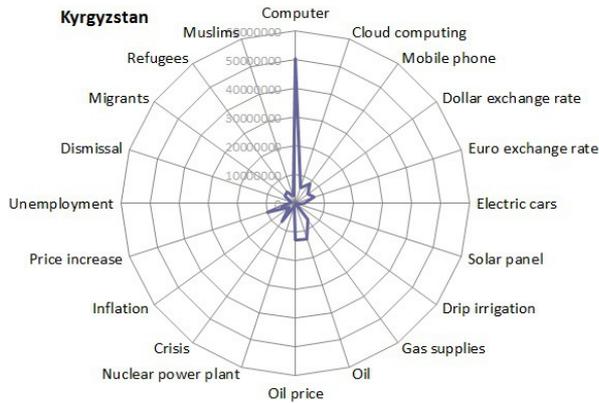
Доминирование Unit «Информация» / «Information»



Доминирование Unit «Информация» / «Information»

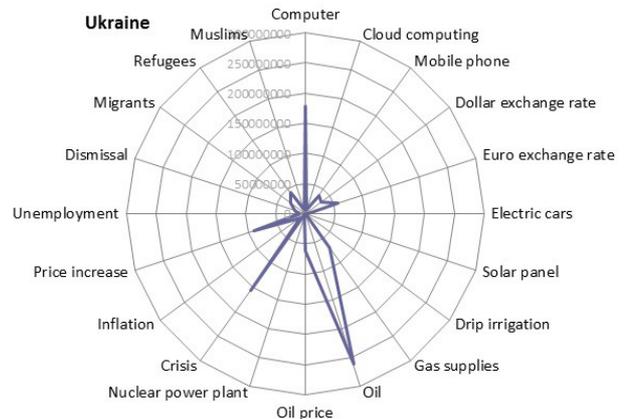
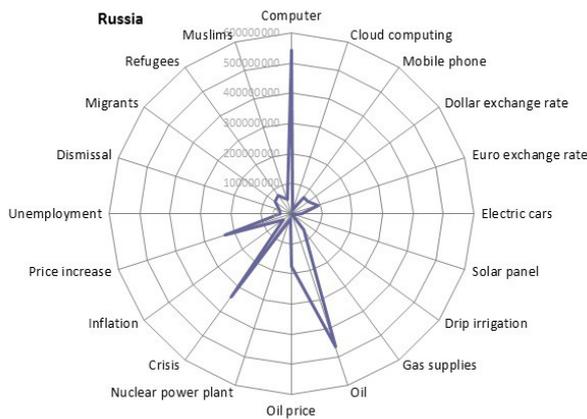
Доминирование Unit «Информация» / «Information»

Продолжение Приложения 1



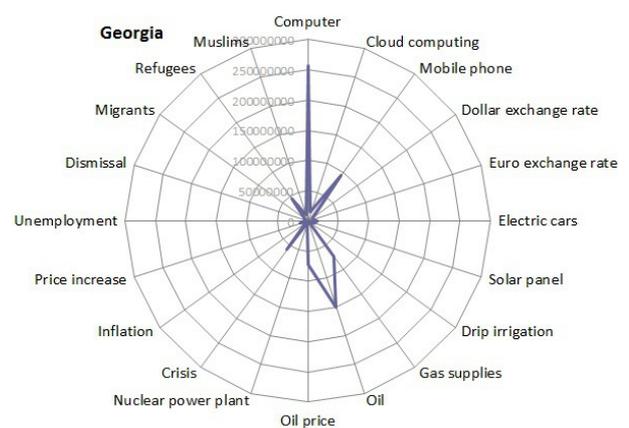
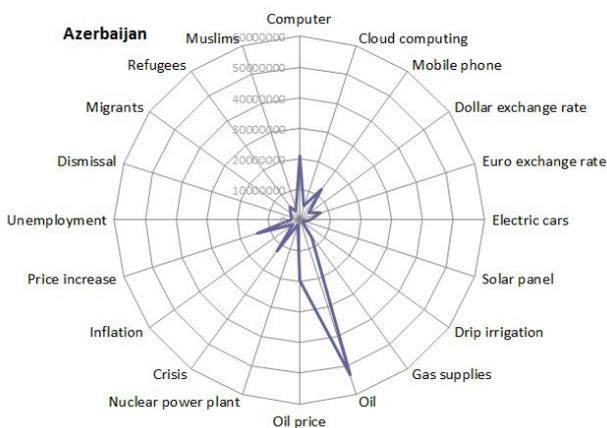
Доминирование Unit «Информация» / «Information»

Доминирование Unit «Информация» / «Information»



Доминирование Unit «Энергия & Материалы» / «Energy & Material»

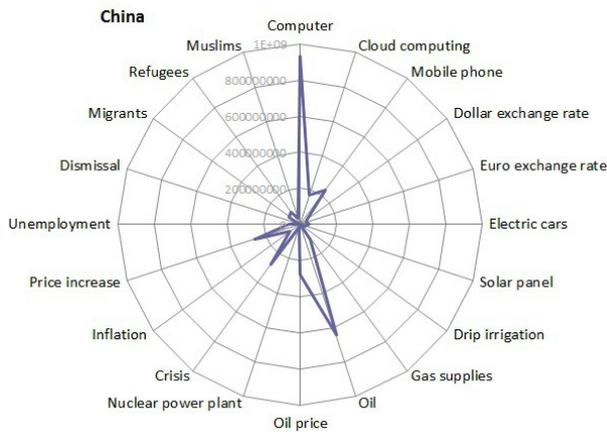
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



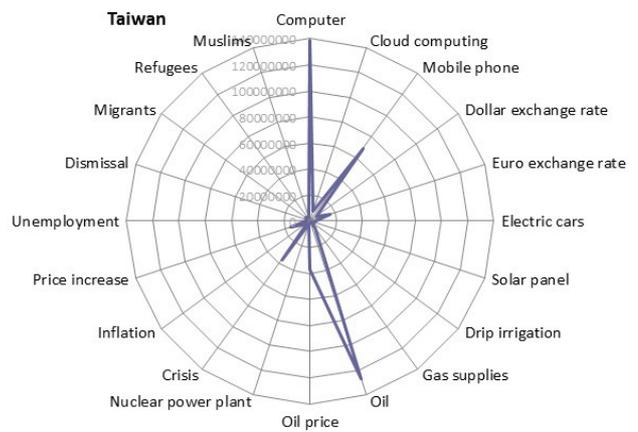
Доминирование Unit «Энергия & Материалы» / «Energy & Material»

Доминирование Unit «Энергия & Материалы» / «Energy & Material»

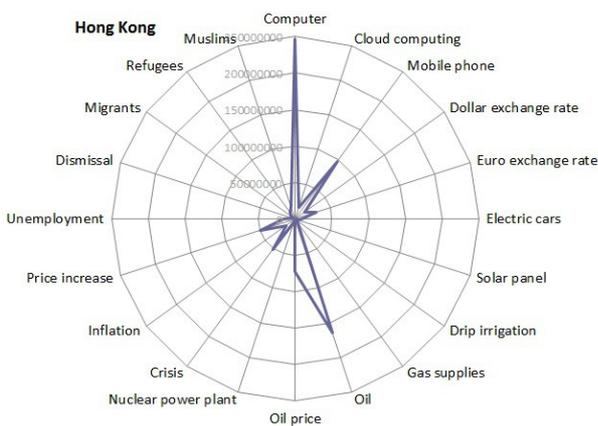
Продолжение Приложения 1



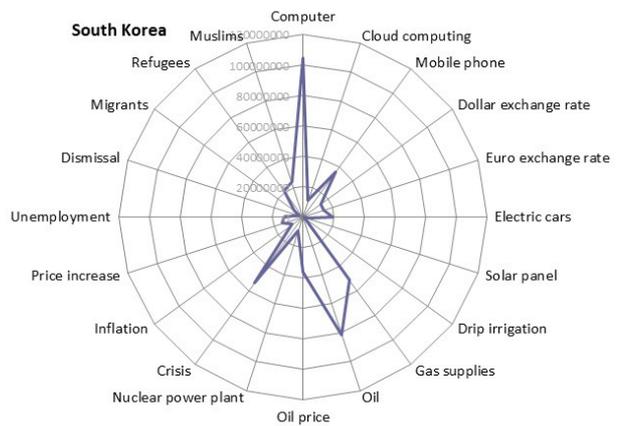
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



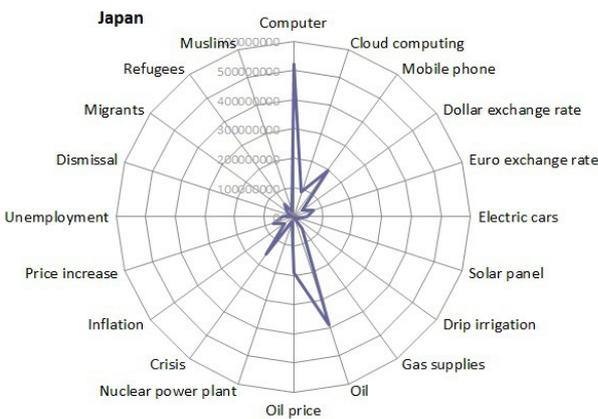
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



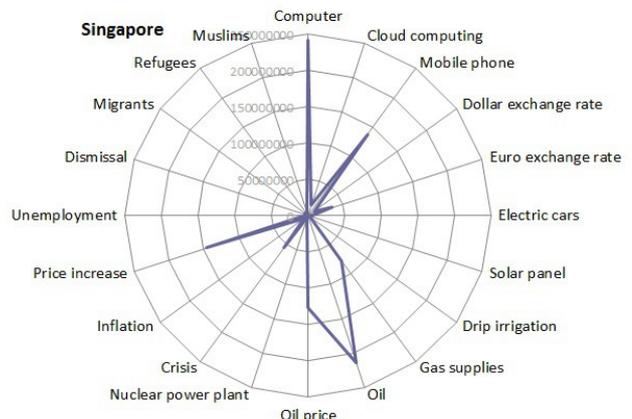
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



Доминирование Unit «Энергия & Материалы» / «Energy & Material»

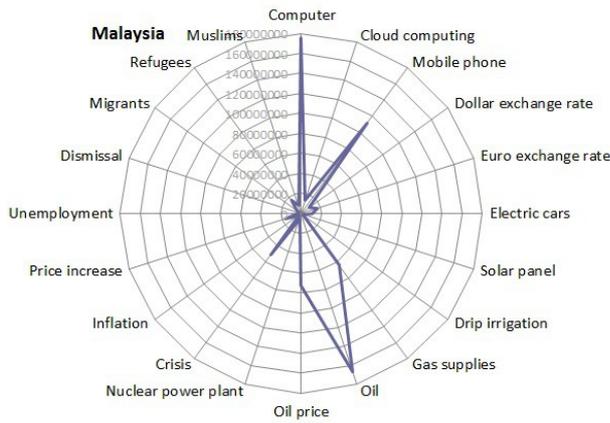


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

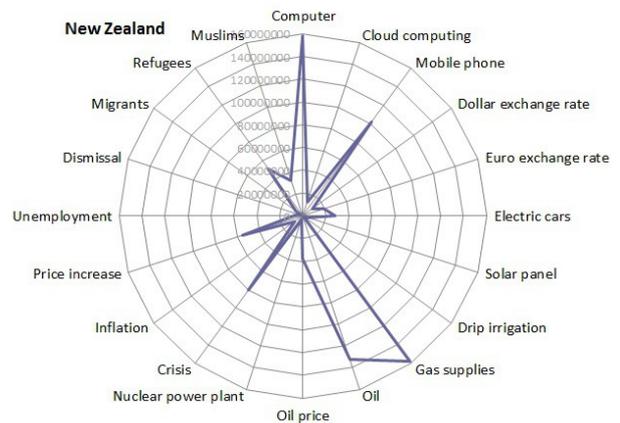


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

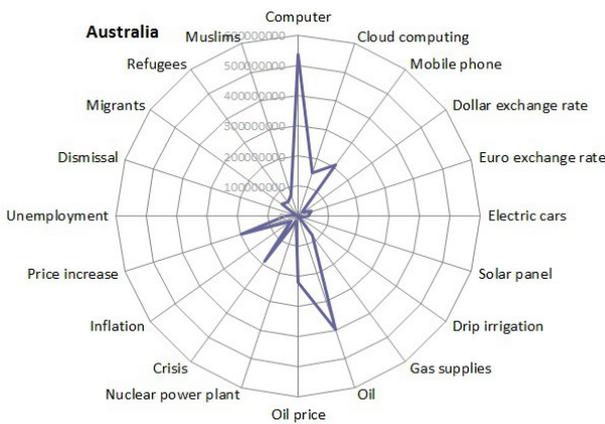
Продолжение Приложения 1



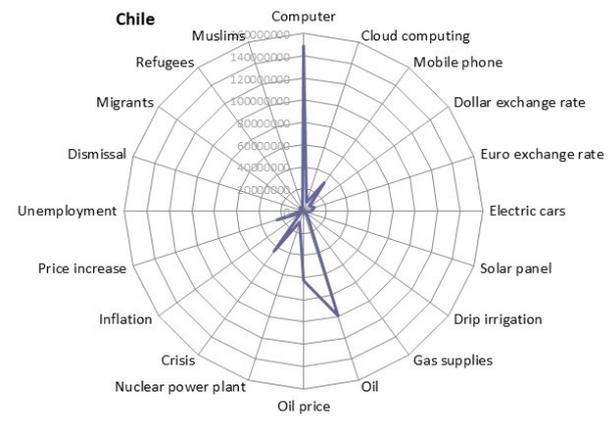
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



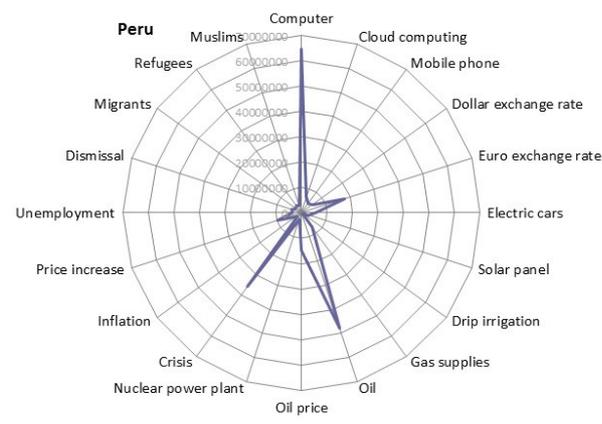
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



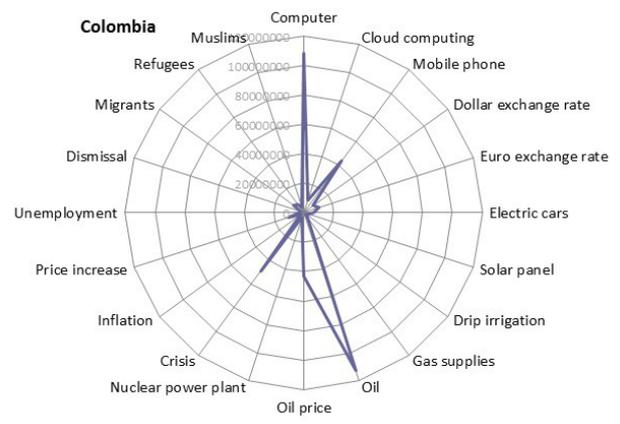
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



Доминирование Unit «Энергия & Материалы» / «Energy & Material»

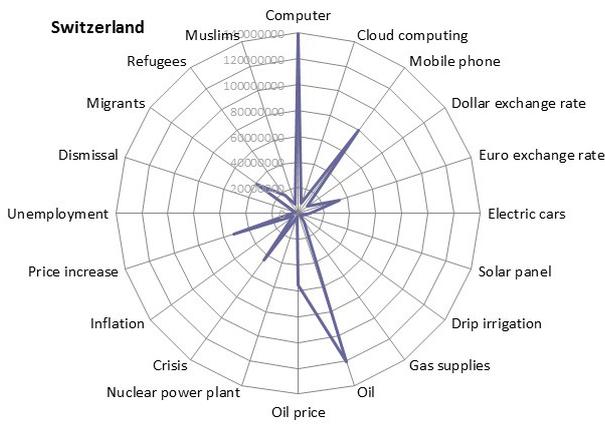


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

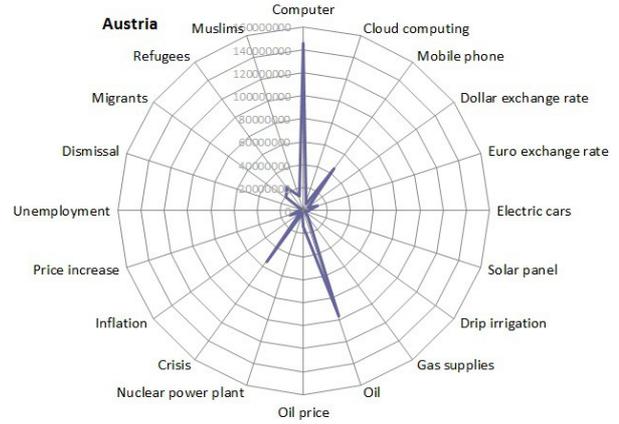


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

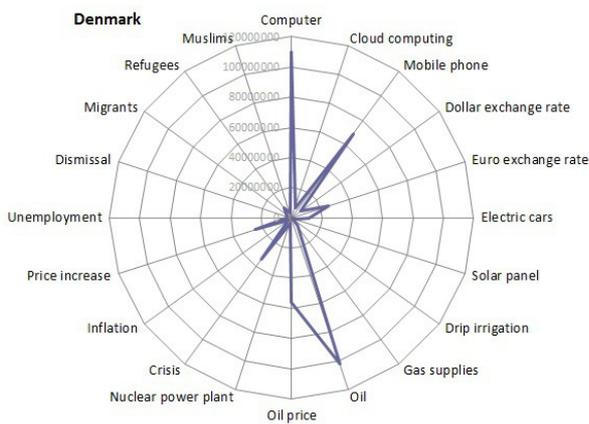
Продолжение Приложения 1



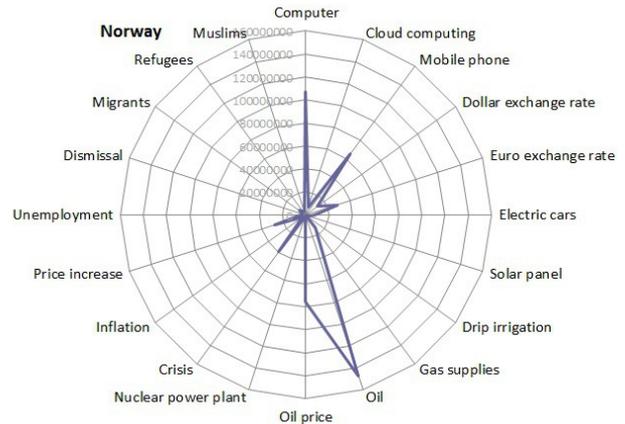
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



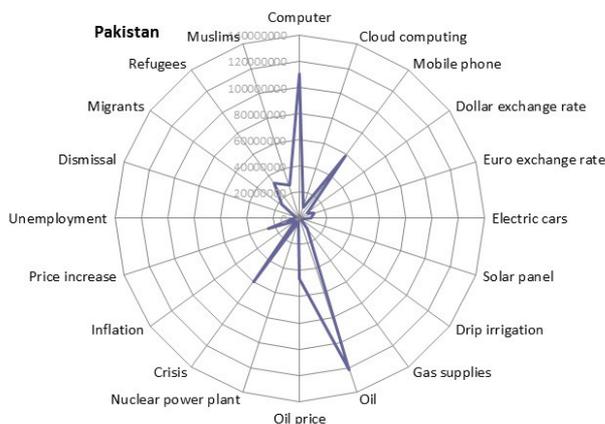
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



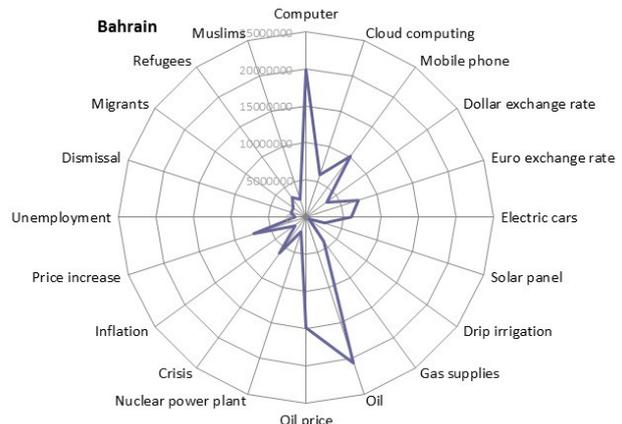
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



Доминирование Unit «Энергия & Материалы» / «Energy & Material»

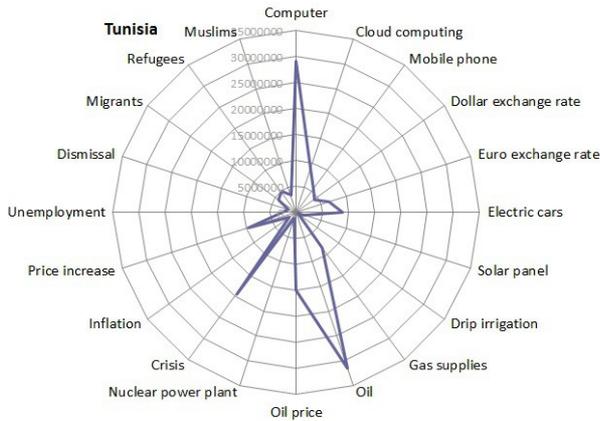


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

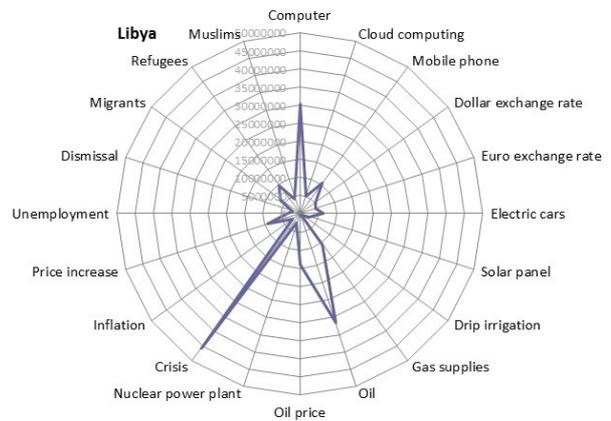


Доминирование Unit «Энергия & Материалы» / «Energy & Material»

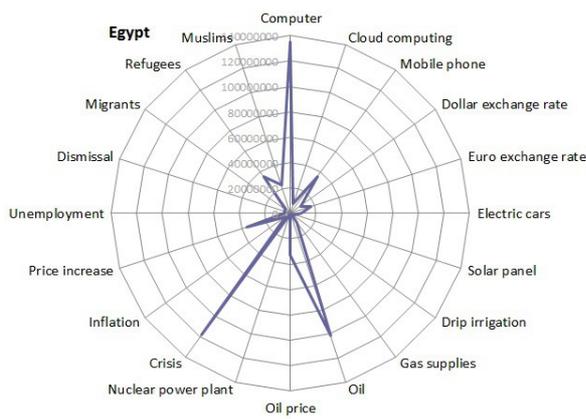
Продолжение Приложения 1



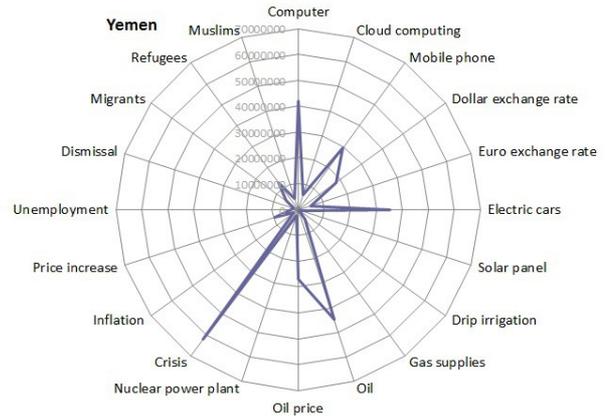
Доминирование Unit «Энергия & Материалы» / «Energy & Material»



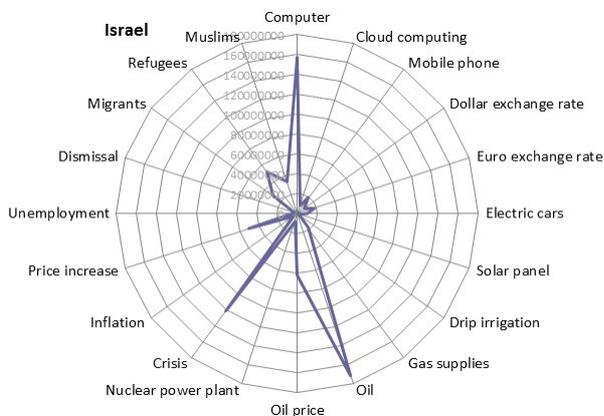
Доминирование Unit «Стрессовая нагрузка» / «Stress load»



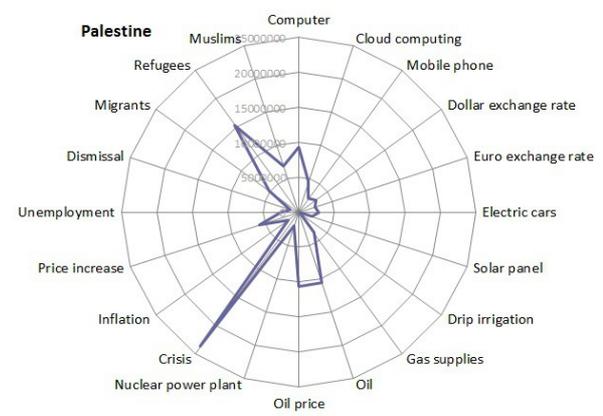
Доминирование Unit «Стрессовая нагрузка» / «Stress load»



Доминирование Unit «Стрессовая нагрузка» / «Stress load»

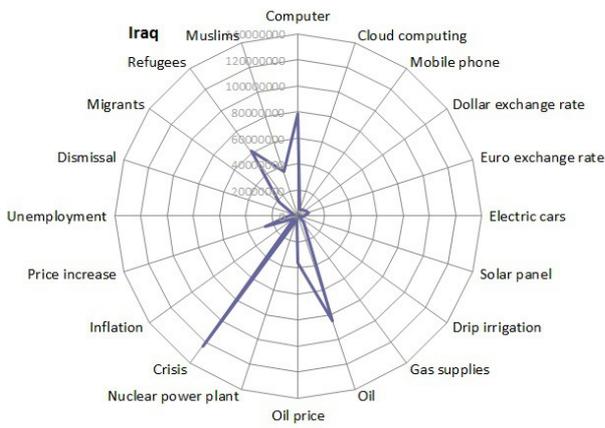


Доминирование Unit «Стрессовая нагрузка» / «Stress load»

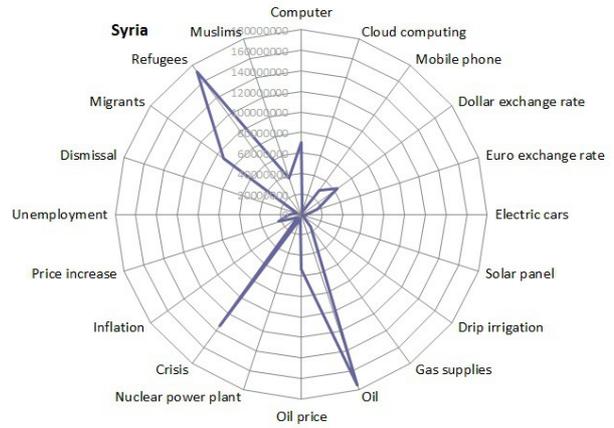


Доминирование Unit «Стрессовая нагрузка» / «Stress load»

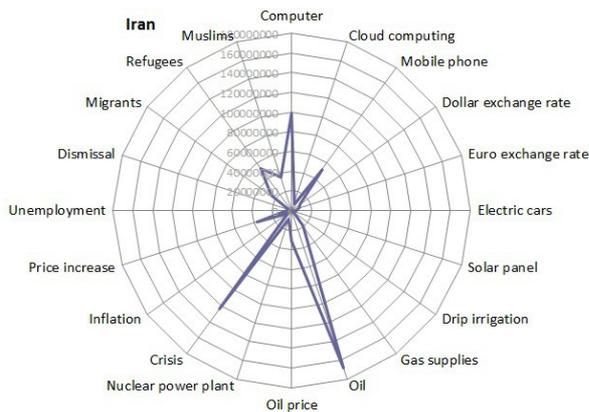
Продолжение Приложения 1



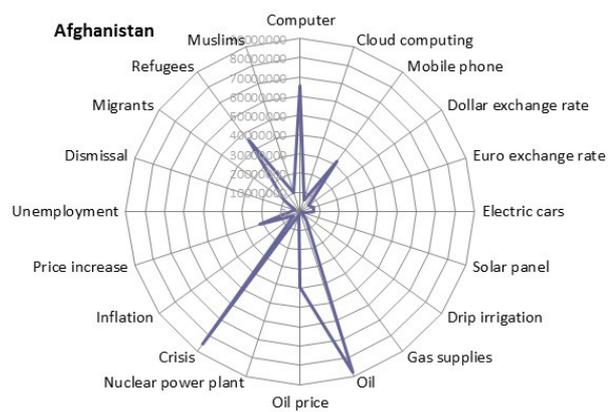
Доминирование Unit «Стрессовая нагрузка» / «Stress load»



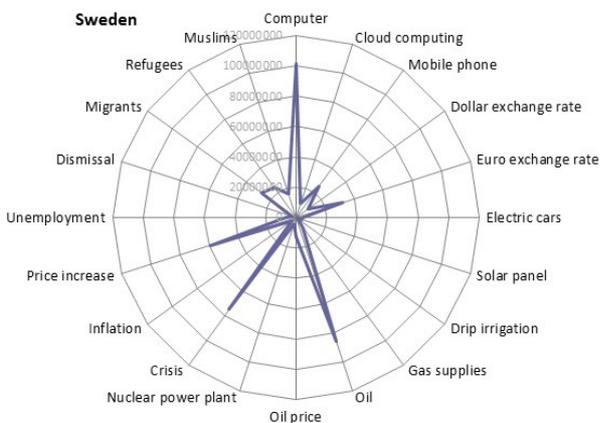
Доминирование Unit «Стрессовая нагрузка» / «Stress load»



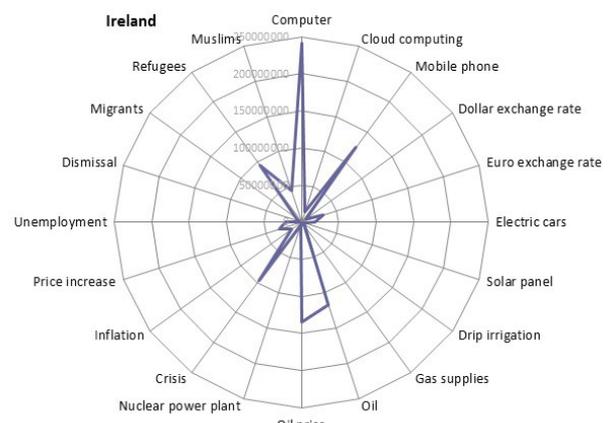
Доминирование Unit «Стрессовая нагрузка» / «Stress load»



Доминирование Unit «Стрессовая нагрузка» / «Stress load»

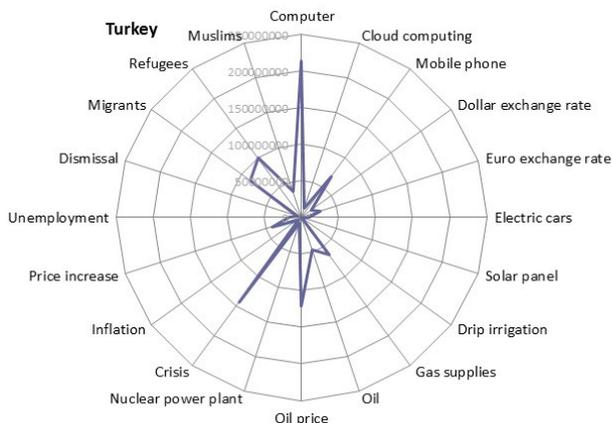


Доминирование Unit «Стрессовая нагрузка» / «Stress load»



Доминирование Unit «Стрессовая нагрузка» / «Stress load»

Продолжение Приложения 1



Доминирование Unit «Стрессовая нагрузка» / «Stress load»

Приложение 2

Пространственное двухмерное соотношение ключевых слов морфологической матрицы

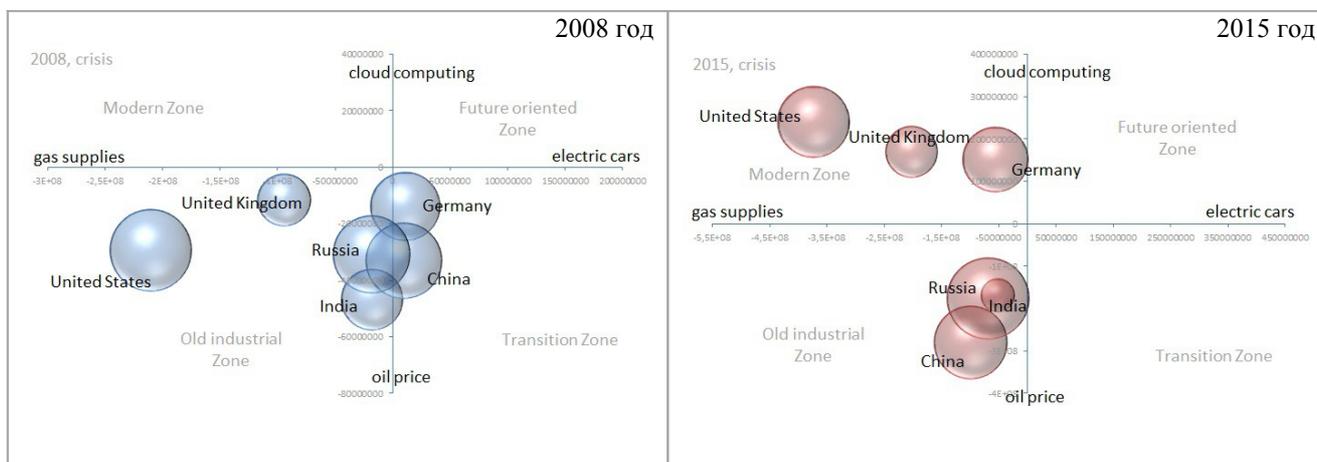
Ось X:

плюс – количество ключевых слов «электромобиль / electric cars» в млн,
минус – количество ключевых слов «поставки газа / gas supplies» в млн,
выбор исходя из доминирования в паре по числу ключевых слов.

Ось Y:

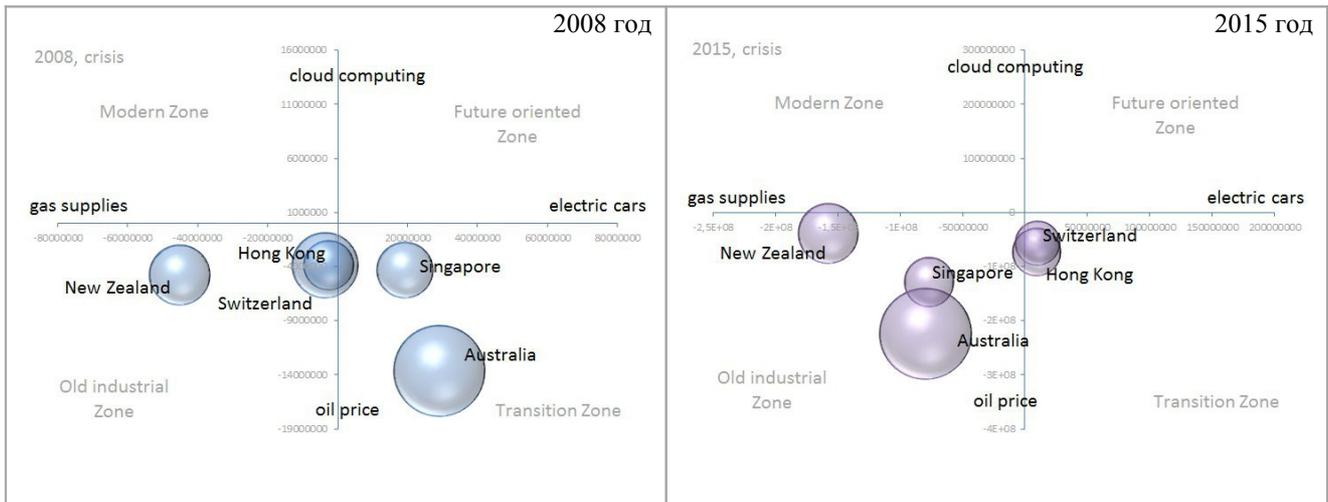
плюс – количество ключевых слов «облачные технологии / cloud computing» в млн,
минус – количество ключевых слов «цена на нефть / oil price» в млн,
выбор исходя из доминирования в паре по числу ключевых слов.

Размер круга: количество ключевых слов «кризис / crisis».

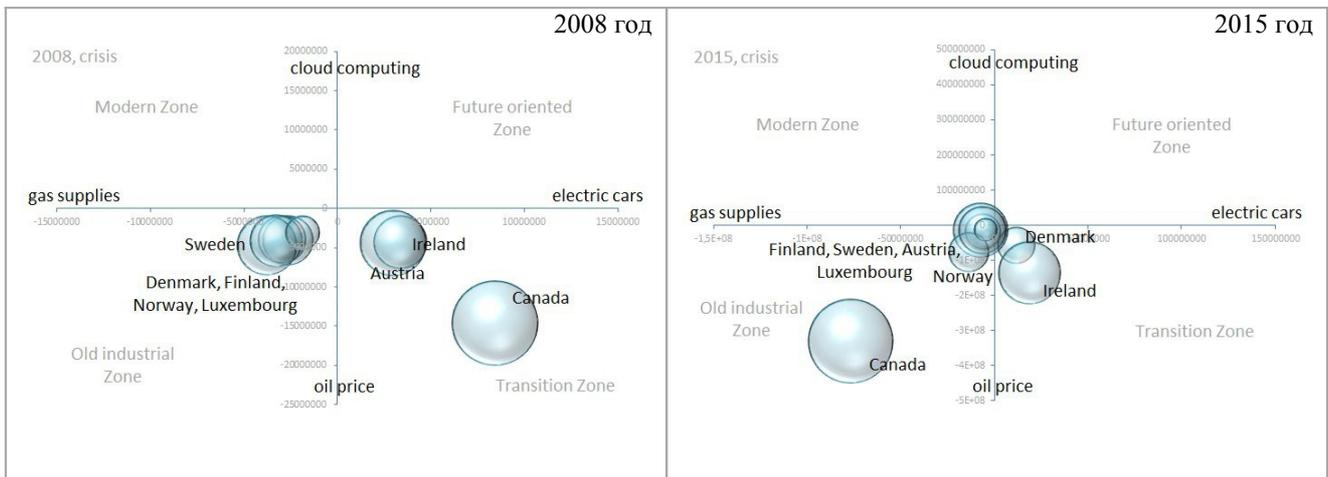


Группа стран с крупными экономиками

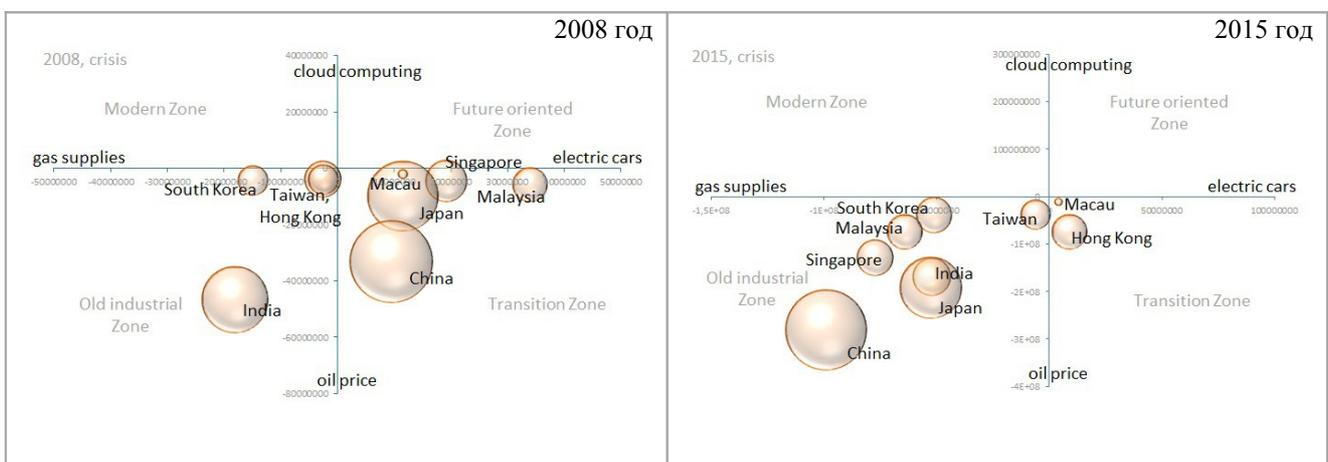
Продолжение Приложения 2



Группа стран Топ-5 рейтинга Index of Economic Freedom 2015

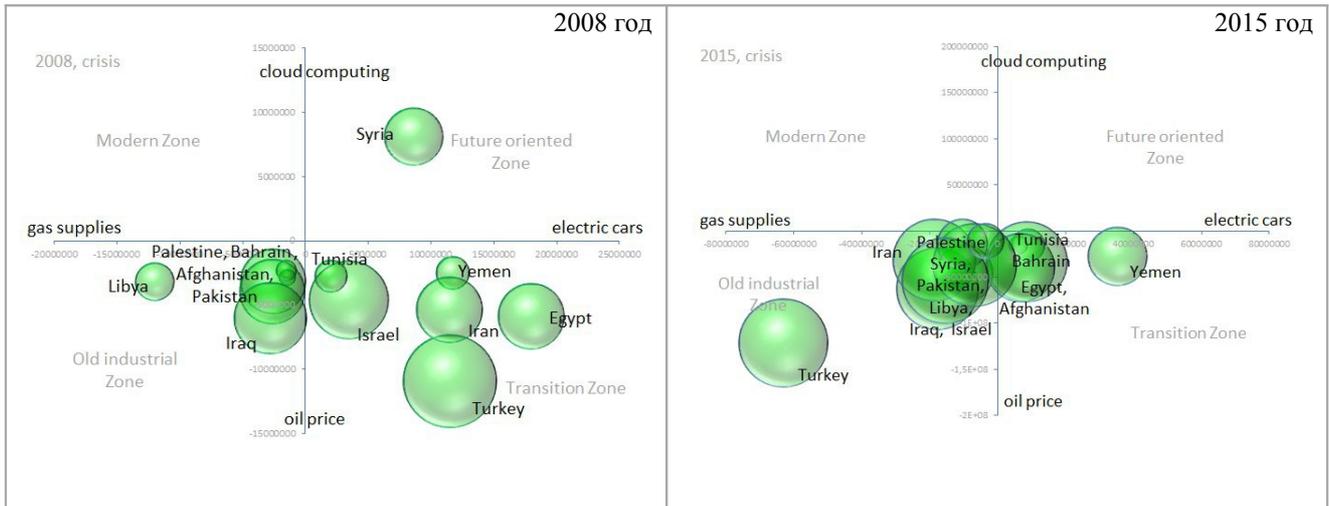


Группа западных стран

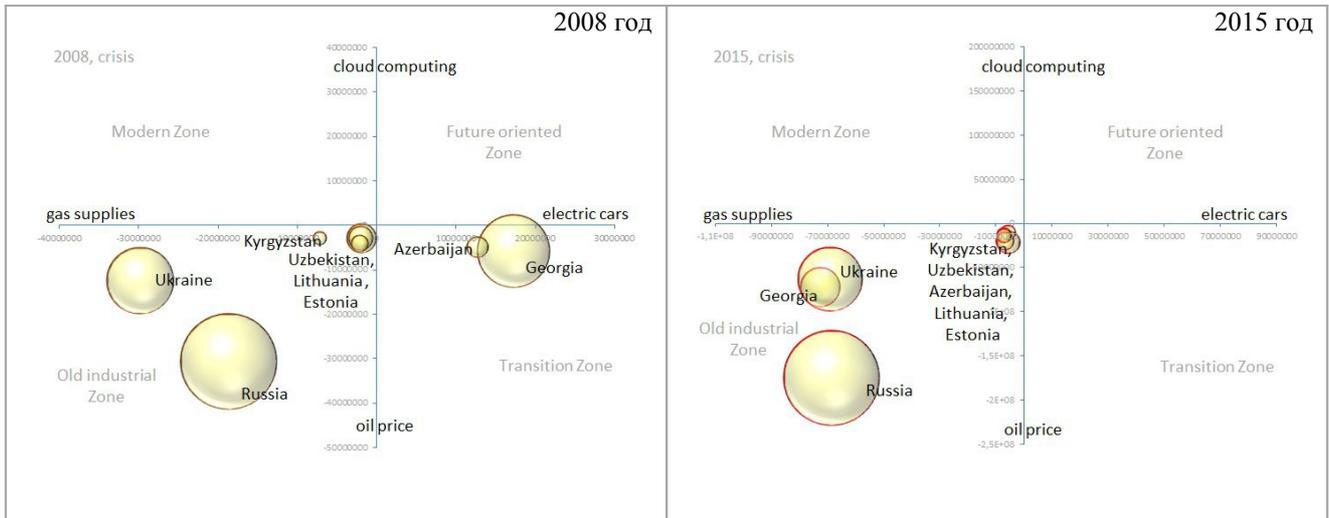


Группа стран АТР

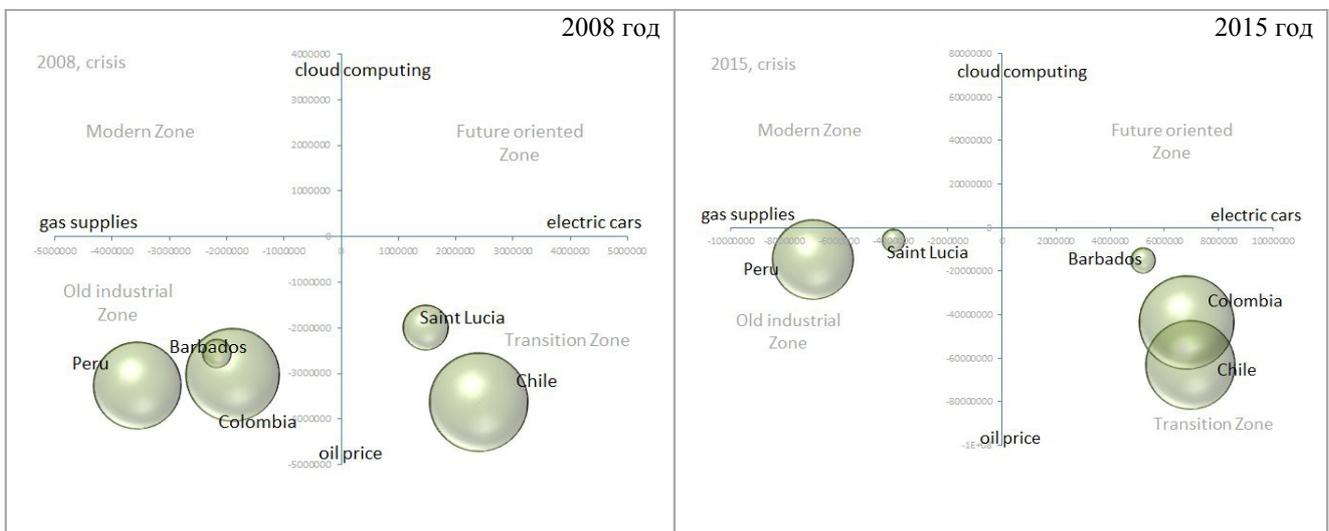
Продолжение Приложения 2



Группа стран MENA (Middle East & North Africa) и Азии



Группа постсоветских стран



Группа стран Южной Америки и Карибского бассейна