

## Информация

УДК 004.912

**Перспективы использования семантического анализа для стратегической аналитики отрасли***И.Ф. Кузьминов, И.В. Логинова, А.А. Беликов, П.А. Лобанова*

Национальный исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва

E-mail: ikuzminov@hse.ru, abelikov@hse.ru

В современном мире аквакультура является одной из наиболее стремительно растущих отраслей экономики. В силу ряда факторов масштабы аквакультурного производства в России на сегодняшний день невелики. Вместе с тем, в ключевых нормативных документах федерального уровня рыбохозяйственной отрасли уделяется большое внимание, перед ней ставятся амбициозные задачи увеличения объёмов производства и экспорта. Реализация программ развития аквакультуры должна базироваться, помимо прочего, на внедрении новых технологий, развитии научно-технологического потенциала и адаптации опыта зарубежных стран. Планирование развития любой отрасли в современном мире должно носить комплексный и системный характер, а для определения перспективных технологических решений и управленческих практик необходим современный и точный аналитический инструментарий. В настоящей статье предлагается использование стратегической аналитики как концептуального подхода к анализу данных в условиях их современного объёма и уровня диверсификации. В качестве инструмента анализа перспективных направлений развития отрасли аквакультуры предлагается использовать разработанную НИУ ВШЭ систему интеллектуального анализа больших данных iFORA. В качестве основного метода в статье предлагается использование семантического анализа больших объёмов текстовой информации. Предложенную методологию планируется использовать в анализе отрасли аквакультуры в последующих публикациях.

**Ключевые слова:** большие данные, семантический анализ, стратегическая аналитика, мониторинг трендов, агропромышленный комплекс, аквакультура.

**ВВЕДЕНИЕ**

В современном мире всё более отчётливо проявляется тенденция к резкому увеличению объёмов информации, производимой человечеством. В качестве аргументации данному тезису достаточно привести несколько фактов:

1) По прогнозам экспертов к 2025 году объём цифровых данных в мире увеличится

более чем в 5 раз (с 33 зеттабайт в 2018 году до 175 зеттабайт) [Reinsel et al., 2018].

2) В период с 2003 по 2016 годы число научных исследований, публикуемых в течение года, выросло с 1,2 млн до 2,3 млн т. е. почти вдвое [National Science Board ..., 2018].

3) Примерная численность исследователей в США, Европе, России, Китае, Японии и Кореи суммарно увеличилась с 4 млн

в 2000 г. до 6,3 млн человек в 2015 г. Растёт международная образовательная мобильность, число научных организаций, численность докторов наук [National Science Board ..., 2018].

Таким образом, тенденция к увеличению объёмов производимой информации обуславливается, с одной стороны, развитием инфраструктуры сбора, документирования и хранения данных, а с другой — диверсификацией сферы науки, ростом числа мировых центров компетенций, параллельных исследований и разработок, реализуемых большим количеством научных коллективов по всему миру.

В контексте постоянного роста объёмов информации и увеличения степени её диверсификации повышаются также и требования к аналитическим инструментам, используемым в качестве фундамента для принятия управленческих решений как в государственном управлении в целом, так и в отдельных отраслях экономики.

Перед отраслью аквакультуры, как частью агропромышленного комплекса России, поставлены амбициозные задачи значительного роста объёмов производства и экспорта [Постановление Правительства ..., 2014; Приказ Минсельхоза ..., 2015; Проект Стратегии развития ..., 2017]. Эта цель обоснована, поскольку масштабы аквакультурного производства в России на фоне других стран невелики: доля России в мировом объёме аквакультурного производства составляет лишь 0,2%. При этом аквакультура является одной из наиболее активно развивающихся подотраслей АПК и одной из самых быстрорастущих отраслей рыбохозяйственного комплекса в мире по причине высокой рыночной востребованности конечных продуктов. Это подтверждается статистикой мирового аквакультурного производства: по данным ФАО в период с 2000 по 2016 годы рост составил 147% (с 32,4 млн тонн до 80 млн тонн), а доля продукции аквакультуры в общем объёме рыбной продукции в 2016 году составила 46,8% (в 2000 г. — 25,7%).

В связи с этим, по экспертным оценкам авторов, повышение годовых темпов ро-

ста аквакультуры в России вполне возможно при правильном регулировании, минимизации барьеров для бизнеса, поддержке экспорта, благоприятствовании импорту средств производства на уровне таможенных пошлин. Увеличение темпов роста кажется ещё более перспективным с учётом возможности развития отдельных направлений аквакультуры и повышения её устойчивости. На фоне нарастающих признаков снижения темпов роста мировой (прежде всего морской) аквакультуры для России существует окно возможностей для импортозамещения широкой номенклатуры продукции, повышения потребления отечественного недорогого и качественного белка населением, захвата ряда ценных экспортных ниш.

Вместе с тем для грамотного и обоснованного принятия решений и управления аквакультурным производством необходимы понимание существующих мировых и российских тенденций, анализ перспективных технологических разработок, не применяемых или применяемых недостаточно в современной российской практике.

Концепция стратегической аналитики, которую предлагают авторы данной статьи, направлена на поддержку принятия управленческих решений, основанных на данных, и может стать опорой для формирования приоритетов стабильного развития аквакультуры.

## СТРАТЕГИЧЕСКАЯ АНАЛИТИКА

В существующей на сегодняшний день российской практике термин стратегической аналитики используется как аналог понятий бизнес- или стратегического консалтинга и означает, прежде всего, оценку деятельности конкретной компании, анализ конъюнктуры рынка и использования существующих активов для разработки стратегии развития компании [PwC, 2019; Маринко, 2005]. В таком же ключе термин стратегической аналитики может использоваться и в образовательной сфере [Coursera, 2019].

В зарубежной литературе встречается более широкое толкование понятия стратегической аналитики, включающее анализ

перспективных и альтернативных путей развития для экономики страны [Papadimitriou et al., 2013, 2014]. Подобное употребление находится значительно ближе к пониманию стратегической аналитики авторами настоящей статьи; тем не менее, в понимании авторов, стратегическая аналитика имеет ряд специфических черт, в том числе методологического и технического характера. В настоящей статье понятие стратегической аналитики будет употребляться в следующем значении: стратегическая аналитика — вид научно-практической деятельности, направленный на проведение стратегических исследований и получение ценных для лиц, принимающих решения, рекомендаций о направлениях политики. Она охватывает ряд направлений, включая разработку стратегий социально-экономического, научно-технологического и инновационного развития стран, отраслей и центров компетенций, стратегические маркетинговые исследования (в том числе исследования рынков, их объёмов, структуры и динамики), аналитику по отраслям экономики, направлениям научно-технологического развития. В основе стратегической аналитики лежат автоматизированный тренд-споттинг и тренд-анализ как наиболее востребованные и перспективные аналитические инструменты, используемые в целях производства тех объективных и релевантных данных, которые помогают выстроить систему принятия стратегических решений на основе чёткого видения процессов внешней микро- и макросреды, понимания того, какие аналитические категории являются приоритетными в вопросах отраслевого стратегического планирования [Кузьминов, Логинова, Лобанова, 2019].

С понятием «стратегическая аналитика» в значительной степени связано понятие «форсайта». Поскольку в настоящей российской аналитической и исследовательской практике этот термин представлен нешироко, он требует дополнительных пояснений. Форсайт (от англ. foresight — «предвидение») — систематические действия, направленные на оценку долгосрочных перспектив развития науки, технологий, экономики и общества, для того чтобы определить стра-

тегические направления исследований и идентифицировать новые технологии, способные принести наибольшие социально-экономические блага [Соколов, 2007; Miles et al., 2016]. Иными словами, форсайт представляет собой систему методов экспертной оценки стратегических направлений социально-экономического и инновационного развития, выявления технологических прорывов зарождающихся и уже оформившихся трендов, способных оказать воздействие на экономику и общество в средне- и долгосрочной перспективе.

В контексте роста объёмов информации (и неизбежного роста её противоречивости) сфера стратегической аналитики сталкивается с необходимостью использования принципиально новых методов обработки данных, основанных на автоматизации и использовании больших данных.

Среди типов больших данных, которые можно использовать для целей стратегической аналитики, ключевыми являются текстовые данные. Текстовые данные обладают рядом особенностей, которые обуславливают эффективность их применения именно в стратегической аналитике. Операционные локальные задачи возможно решать с применением количественных данных, например, официальной статистики. Однако количественные данные зачастую не могут обобщить и конкретизировать исчерпывающую специфику рассматриваемого явления, выявить существующие и только зарождающиеся риски или возможности. Текст, как полагают авторы данной статьи, способен передавать более сложные концепции, идеи, рекомендации, обоснования, экспертные мнения о перспективных направлениях развития и многие другие аналитические нарративы. Если в операционной аналитике важно выявить из данных тот или иной тренд и произвести короткий аналитический вывод, то в стратегической аналитике необходимо из большого числа полифонии мнений экспертов, гипотез о потенциальных путях развития, с учётом латентных процессов, не фиксируемых статистикой, построить определённый генерализированный мета-нарратив, который сможет стать основой

для разработки рекомендаций для лиц, принимающих решения.

В международной исследовательской практике на сегодняшний день существует немало примеров использования больших объёмов текстовой информации и их семантического анализа. Не являются исключением аквакультура: научные публикации можно найти как по общетеоретическим вопросам использования больших данных для анализа отрасли [Costa, Rihtar, 2016], так и по более локальным отраслевым проблемам. В частности, с помощью анализа больших данных можно агрегировать симптоматику и причины заболеваний рыбы [Froehlich et al., 2017], предсказать в долгосрочной перспективе темпы её роста при различных внешних условиях [Bar, Radde, 2009] или выявить отношение общества и СМИ к тем или иным технологическим решениям и инновациям [Rao et al., 2017].

Кроме того, технологию семантического анализа больших данных авторы настоящей статьи применяли также и к ряду других отраслей агропромышленного комплекса. В частности, с помощью анализа больших данных были выявлены виды насекомых-вредителей, в перспективе угрожающее свеклосахарной отрасли. Некоторые из этих вредителей традиционно не являются угрозой для сахарной свёклы, однако под воздействием внешних факторов (к примеру, климатических условий) постепенно меняют привычный рацион и становятся все более опасными для этой сельскохозяйственной культуры [Кузьминов и др., 2018].

Ещё одним примером использования анализа больших данных для стратегической аналитики отраслей экономики является анализ совместного упоминания компонентов, входящих в состав комбикормов для домашней птицы и распространённых заболеваний птицы и их симптомов. Ряд компонентов комбикормов (витамины, микроэлементы и др.) рассматриваются в научной и практико-ориентированной литературе в контексте профилактики или лечения заболеваний, другие — в качестве причин развития болезней [Кузьминов, Логинова, Лобанова, 2019].

Алгоритмы текстовой аналитики (естественно-языковая обработка, сетевой анализ, векторные представления и др.) легли в основу созданной в НИУ ВШЭ системы интеллектуального анализа данных iFORA. Особенность системы iFORA заключается в использовании математического аппарата, который традиционно предназначен для оптимизации операционных задач бизнеса, в целях автоматизированного выявления принципиальных для стратегической аналитики содержательных категорий (глобальные тренды, возникающие рынки и продукты, зарождающиеся технологии). Система предназначена для усиления экспертной аналитики в рамках практически любого тематического направления. Актуальная база данных iFORA содержит порядка 350 миллионов документов, в том числе материалы отраслевой периодики, тексты научных публикаций и патентные заявки.

Система интеллектуального анализа данных iFORA может служить инструментом для анализа глобальных тенденций в сфере науки и технологий, применяемых в тех или иных отраслях экономики. Не является исключением и аквакультура, особенно с учётом описанного ранее имеющегося международного опыта применения технологий анализа больших данных в данной отрасли.

## ЗАКЛЮЧЕНИЕ

В условиях тенденции на увеличение объёмов информации, роста её структурной сложности и диверсификации мировых центров компетенций и генерации нового знания, перед исследователями, аналитиками и лицами, принимающими управленческие решения, стоит принципиальная задача поиска новых методов и инструментов анализа данных. В последние годы в науке начинают активно применяться методы работы с большими данными, в том числе семантический анализ неструктурированной текстовой информации.

Разработанная НИУ ВШЭ система интеллектуального анализа данных iFORA может служить основой для применения на практике описанной в данной статье концепции

стратегической аналитики. В следующем выпуске «Трудов ВНИРО» авторы настоящей статьи планируют опубликовать результаты исследования с использованием семантического анализа больших данных применительно к отрасли аквакультуры, которые, как мы надеемся, позволят наглядно проиллюстрировать перспективные возможности этого нового для российской науки метода работы с информацией.

#### ЛИТЕРАТУРА

- Кузьминов И. Ф., Логинова И. В., Лобанова П. А. 2018. Перспективы использования технологий анализа больших данных для стратегической аналитики агропромышленного комплекса // Сахарная свекла. № 9. С. 2–7.
- Кузьминов И. Ф., Логинова И. В., Лобанова П. А. 2019. Технология анализа больших данных для стратегической аналитики отрасли // Комбикорма. 2019. № 4. С. 46–52.
- Маринко Г. 2005. Управленческий консалтинг. М: ИНФРА-М. 381 с.
- Постановление Правительства РФ от 15.04.2014 г. N314 «Об утверждении государственной программы Российской Федерации «Развитие рыбохозяйственного комплекса». Доступно через: <http://base.garant.ru/70644222> 15.08.2019.
- Приказ Министерства сельского хозяйства РФ от 16.01.2015 г. N10 «Об утверждении отраслевой программы «Развитие товарной аквакультуры (товарного рыбоводства) в Российской Федерации на 2015–2020 годы». Доступно через: <https://www.garant.ru/products/ipo/prime/doc/70751534/> 15.08.2019.
- Проект Стратегии развития рыбохозяйственного комплекса Российской Федерации на период до 2030 года. Доступно через: <http://fish.gov.ru/files/documents/files/proekt-strategiya-2030.pdf> 15.08.2019.
- Соколов А. В. 2007. Форсайт: взгляд в будущее // Форсайт. № 1. С. 8–15.
- Bar N. S., Radde N. 2009. Long-term prediction of fish growth under varying ambient temperature using a multiscale dynamic model // BMC Systems Biology, 3(1). P. 107–125.
- Costa J., Rihtar M. 2016. Data Analytics in Aquaculture. SIKDD2016. Accessible via: [https://www.researchgate.net/publication/312024207\\_Data\\_Analytics\\_in\\_Aquaculture](https://www.researchgate.net/publication/312024207_Data_Analytics_in_Aquaculture) 07.10.2019.
- Coursera. Accessible via: <https://www.coursera.org/specializations/strategic-analytics> 16.08.2019.
- Froehlich H. E., Gentry R. R., Rust M. B., Grimm D., Halpern B. S. 2017. Public Perceptions of Aquaculture: Evaluating Spatiotemporal Patterns of Sentiment around the World. PLoS ONE12(1): e0169281. Accessible via: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169281> 07.10.2019.
- Miles I., Saritas O., Sokolov A. 2016. Foresight for Science, Technology and Innovation. Switzerland: Springer Nature. 270 p.
- National Science Board. 2018. Science and Engineering Indicators 2018. Alexandria, VA: National Science Foundation. 1060 p.
- Papadimitriou D., Nikiforos M., Zezza G. 2013. The Greek Economic Crisis and the Experience of Austerity: A Strategic Analysis. Accessible via: <http://www.levyinstitute.org/publications/summary-fall-2013> 07.10.2019
- Papadimitriou D., Nikiforos M., Zezza G. 2014. Prospects and Policies for the Greek Economy. Accessible via: <http://www.levyinstitute.org/publications/prospects-and-policies-for-the-greek-economy> 07.10.2019
- PwC Россия. Accessible via: <https://www.pwc.ru/ru/industries/real-estate/strategy-consulting-services.html> 16.08.2019.
- Reinsel D., Gantz J., Rydning J. 2018. The Digitization of the World. From Edge to Core. // IDC White Paper I Doc# US44413318 I November 2018. Accessible via: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataaage-whitepaper.pdf> 16.08.2019
- Venkateswara Rao P., Ramamohan Reddy A., Sucharita V. 2017. Big Data Analytics in Aquaculture Using Hive and Hadoop Platform // Exploring the Convergence of Big Data and the Internet of Things. India, IGI Global. P. 29–36.

Поступила в редакцию 31.07.2019 г.

## Information

Prospects for the use of semantic analysis  
for strategic industry analytics

I.F. Kuzminov, I.V. Loginova, A.A. Belikov, P.A. Lobanova

Higher School of Economics — National Research University, Moscow, Russia

E-mail: ikuzminov@hse.ru, abelikov@hse.ru

Aquaculture is nowadays one of the fastest growing sectors of the economy. In Russia, however, the volume of aquaculture production is low due to several factors. At the same time, the key regulatory documents of the Federal level pay great attention to the fishery industry, ambitious goals to increase production and exports are set. The implementation of aquaculture development programs should be based, among other things, on the introduction of new technologies, the development of scientific and technological potential and the adaptation of the experience of foreign countries. Planning the development of any industry in the modern world should be comprehensive and systematic, and determining promising technological solutions and management practices requires contemporary and accurate analytical tools. This article proposes the use of strategic analytics as a conceptual approach to the data analysis in terms of their current volume and diversification level. It is proposed to use the iFora big data mining system developed by the HSE as a tool for analysing promising trends in the development of the aquaculture industry. The article proposes the use of semantic analysis of large amounts of textual information as the main method. The proposed methodology will be used in the analysis of the aquaculture industry in subsequent publications.

**Keywords:** Big data, semantic analysis, strategic analytics, trend monitoring, agriculture, fisheries, aquaculture.

## REFERENCES

- Kuzminov I.F., Loginova I.V., Lobanova P.A. 2018. Perspektivy ispol'zovaniya tekhnologij analiza bol'shikh dannykh dlya strategicheskoy analitiki agropromyshlennogo kompleksa [Prospects for using big data technologies for strategic analysis of agro-industrial complex] // *Sakharnaya svekla* № 9. S. 2–7.
- Kuzminov I.F., Loginova I.V., Lobanova P.A. 2019. Tekhnologiya analiza bol'shikh dannykh dlya strategicheskoy analitiki otrasli [Big data analytics technology for strategic industry analytics] // *Kombikorma* № 4. S. 46–52.
- Marinko G. 2005. Upravlencheskij konsalting [Management consultancy]. M: INFRA-M. 381 p.
- Postanovlenie Pravitel'stva RF ot 15.04.2014 g. N314 «Ob utverzhdenii gosudarstvennoj programmy Rossijskoj Federatsii “Razvitie rybokhozyajstvennogo kompleksa» [Order of the Government of the Russian Federation of 15.04.2014 N314 «About the approval of the state program of the Russian Federation development of fishery complex»]. Accessible via: <http://base.garant.ru/70644222/> 15.08.2019.
- Prikaz Ministerstva sel'skogo khozyajstva RF ot 16.01.2015 g. N10 «Ob utverzhdenii otraslevoj programmy «Razvitie tovarnoj akvakul'tury (tovarnogo rybovodstva) v Rossijskoj Federatsii na 2015–2020 gody» [Order of the Ministry of agriculture of the Russian Federation of 16.01.2015 N10 «about the approval of the industry program» development of commodity aquaculture (commodity fish farming) in the Russian Federation for 2015–2020»]. Accessible via: <https://www.garant.ru/products/ipo/prime/doc/70751534/> 15.08.2019.
- Proekt Strategii razvitiya rybokhozyajstvennogo kompleksa Rossijskoj Federatsii na period do 2030 goda [Draft strategy for the development of the

- fisheries sector of the Russian Federation for the period up to 2030]. Accessible via: <http://fish.gov.ru/files/documents/files/proekt-strategiya-2030.pdf> 15.08.2019
- Sokolov A. V.* 2007. Forsajt: vzglyad v budushchee [Foresight: looking to the future] // Foresight. № 1. S. 8–15.
- Bar N.S., Radde, N.* 2009. Long-term prediction of fish growth under varying ambient temperature using a multiscale dynamic model // BMC Systems Biology, 3(1). P. 107–125.
- Costa J., Rihtar M.* 2016. Data Analytics in Aquaculture. SIKDD2016. Accessible via: [https://www.researchgate.net/publication/312024207\\_Data\\_Analytics\\_in\\_Aquaculture](https://www.researchgate.net/publication/312024207_Data_Analytics_in_Aquaculture) 07.10.2019.
- Coursera.* Accessible via: <https://www.coursera.org/specializations/strategic-analytics> 16.08.2019.
- Froehlich H.E., Gentry R.R., Rust M.B., Grimm D., Halpern B.S.* 2017. Public Perceptions of Aquaculture: Evaluating Spatiotemporal Patterns of Sentiment around the World. PLoS ONE12(1): e0169281. Accessible via: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169281> 07.10.2019.
- Miles I., Saritas O., Sokolov A. V.* 2016. Foresight for Science, Technology and Innovation. Switzerland: Springer Nature. 270 p.
- National Science Board.* 2018. Science and Engineering Indicators 2018. Alexandria, VA: National Science Foundation. 1060 p.
- Papadimitriou D., Nikiforos M., Zezza G.* 2013. The Greek Economic Crisis and the Experience of Austerity: A Strategic Analysis. Accessible via: <http://www.levyinstitute.org/publications/summary-fall-2013> 07.10.2019.
- Papadimitriou D., Nikiforos M., Zezza G.* 2014. Prospects and Policies for the Greek Economy. Accessible via: <http://www.levyinstitute.org/publications/prospects-and-policies-for-the-greek-economy> 07.10.2019
- Pw C. Russia.* Accessible via: <https://www.pwc.ru/industries/real-estate/strategy-consulting-services.html> 16.08.2019.
- Reinsel D., Gantz J., Rydning J.* 2018. The Digitization of the World. From Edge to Core. // IDC White Paper I Doc# US44413318 I November 2018. Accessible via: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> 16.08.2019
- Venkateswara Rao P., Ramamohan Reddy A., Sucharita V.* 2017. Big Data Analytics in Aquaculture Using Hive and Hadoop Platform // Exploring the Convergence of Big Data and the Internet of Things. India, IGI Global. P. 29–36.