

## РЕДАКЦИОННАЯ ПОЛИТИКА / EDITORIAL POLICY



Оригинальная статья / Original paper

<https://doi.org/10.24069/SEP-22-11>

## Применение BI-аналитики для управления научным журналом\*

Д. Ю. Большаков  Акционерное общество «Концерн воздушно-космической обороны «Алмаз – Антей»,  
г. Москва, Российская Федерация [press@almaz-antey.ru](mailto:press@almaz-antey.ru)

**Резюме.** Приведены результаты использования бесплатного приложения анализа данных *Power BI* для управления научным журналом. Показано, что построение даже простого *dashboard* позволяет сделать ряд управленческих выводов для корректировки деятельности редакции. Использование данных о потоке статей за прошлые годы позволяет прогнозировать в системе *Power BI* поток статей на будущий период. Проанализированы все доступные бесплатные способы проведения прогноза в системе *Power BI* и проведен их сравнительный анализ по точности прогноза. Проведен анализ использования плагина системы *Power BI* для построения графа связей рецензентов в редакционной коллегии. На основании анализа графа связей рецензентов даны рекомендации научным редакторам по выбору рецензентов и развитию журнала. Упор в исследовании делается на простые решения, которые могут быть реализованы сотрудниками редакций научных журналов без специфических знаний в области *Data Science*.

**Ключевые слова:** бизнес-аналитика, научный журнал, поток статей, граф связей рецензентов, предсказание потока статей

**Для цитирования:** Большаков Д. Ю. Применение BI-аналитики для управления научным журналом. *Научный редактор и издатель*. 2022;7(1 Suppl):S23–S28. <https://doi.org/10.24069/SEP-22-11>

## Application of BI-analytics in the management of a scientific journal

D. Yu. Bolshakov  

Joint Stock Company “Almaz – Antey” Air and Space Defence Corporation, Moscow, Russian Federation

 [press@almaz-antey.ru](mailto:press@almaz-antey.ru)

**Abstract.** The results of applying the *Power BI* free software in the management of a scientific journal are presented. It is shown that even the creation of the most basic *dashboard* allows management solutions aimed at improving the editorial work to be obtained. Data on the manuscript flow for the past several years can be used as input to *Power BI* to make informed decisions about future trends. All available free forecasting techniques in *Power BI* are compared in terms of forecast accuracy. The *Power BI* force-directed graph plugin for establishing connections between reviewers is analysed. On this basis, recommendations to scientific editors concerning the selection of reviewers and the development of a scientific journal are provided. This research is focused on the search for simple solutions that can be applied by the editorial team without acquiring specific knowledge in the field of *Data Science*.

**Keywords:** business intelligence, scientific journals, manuscript flow, graph of reviewers' connections, manuscript flow forecast.

**For citation:** Bolshakov D. Yu. Application of BI-analytics in the management of a scientific journal. *Science Editor and Publisher*. 2022;7(1 Suppl):S23–S28. (In Russ.) <https://doi.org/10.24069/SEP-22-11>

\* Статья написана по материалам доклада, представленного на 10-й Международной научно-практической конференции «Научное издание международного уровня – 2022: от настоящего к будущему», которая прошла 26–29 апреля 2022 г. в Московском государственном университете им. О.Е. Кутафина (МГЮА) (см. [https://rassep.ru/upload/iblock/35c/Bolshakov\\_2022.03.14-Doklad-na-konferentsii-ANRI-NIMU-2022.pdf](https://rassep.ru/upload/iblock/35c/Bolshakov_2022.03.14-Doklad-na-konferentsii-ANRI-NIMU-2022.pdf) [презентация]).

## Введение

Термин *Business Intelligence* (сокращенно – *BI*), в русском переводе – «бизнес-аналитика» или «*BI-аналитика*», появился в 1958 г. благодаря американскому ученому Х.П. Луну (*H.P. Luhn*). *BI* – компьютерная система сбора информации о реальном мире и преобразование ее в понятные человеку данные через машинную обработку, пригодную для анализа и принятия управленческих решений [1]. В настоящее время существуют множество инструментов для решения задач *BI*. Наиболее известные – это *QlikView*, *Tableau* и *Power BI*. Толчком для разработки и внедрения систем *BI-аналитики* послужило развитие в конце 80-х – начале 90-х гг. XX в. компьютерной техники.

В сознании большинства руководителей *BI* ассоциируется не с принятием решений, а только с построением так называемых «дашбордов» (от англ. *dashboard* – приборная доска). В настоящее время особенно в русском языке термин «дашборд» используется для графического представление информации, меняющееся во времени и показывающее изменение основных показателей деятельности. Простым примером *dashboard* является простой механический будильник, который показывает данные по изменению информации во времени.

## 1. Построение *dashboard*

Особенностью *dashboard* в *Power BI* в отличие от аналогичного графика в *Excel* является возможность создания интерактивной информации, которая меняется от времени и от действий пользователя автоматически перестраивая визуализацию. Например, интерактивный элемент системы *Power BI* на рис. 1 в левой нижней части *dashboard* позволяет строить произвольные разрезы по годам.

При кажущейся простоте построение даже такого *dashboard* как на рис. 1 позволяет сделать несколько выводов:

1. Конверсия статей, т.е. отношение количества опубликованных статей к общему количеству, поступивших за последние 5 лет в журнале практически постоянна и равна в среднем 40% (публикуется 2 из 5 поступивших статей). Таким образом, следует ожидать, что в ближайшие годы вряд ли конверсия сменится радикально, что подтверждает анализ предыдущих лет – нет резких взлетов и падений, например, от 10 до 100%.

2. Поток статей в журнал с 2012 по 2016 г. имел экспоненциальный рост, что связано с применением в этих годах определенных маркетинговых коммуникаций, но в связи с достижением предела по использованию каналов безбюджетного маркетинга вышел на плато.



**Рис. 1.** Интерактивная аналитическая панель – *dashboard* – научно-технического журнала «Вестник Концерна ВКО «Алмаз – Антей», построенная в программе *Power BI*.

**Fig. 1.** Interactive analytical dashboard of the scientific and technical Journal of "Almaz – Antey" Air and Space Defence corporation, built in the *Power BI*

3. Поток публикуемых статей хоть и вырос с 2011 г. почти вдвое, но в среднем с 2015 г. остается постоянным, и, следовательно, можно предположить, что останется постоянным, а значит задача редакции – привлекать в журнал  $40/0,4 = 100$  статей в год.

В статье [2] показано, что для потока входящих статей в целом можно найти оценку снизу (так называемую, «миноранту») в виде пуассоновского потока. Знание характеристик пуассоновского потока дает нам информацию о событиях, что в журнал не придет ни одной статьи за любой промежуток времени. Например, для научно-технического журнала «Вестник Концерна ВКО «Алмаз – Антей» вероятность того, что за три месяца не придет ни одной статьи составляет  $10^{-5}$  (один случай из 10 000), а за год –  $10^{-16}$ . Миноранта в виде пуассоновского потока характерна для всего потока статей с выхода первого номера с 2011 г. А это значит, что пуассоновский поток статей в журнал не прекратится в будущем и останется пуассоновским, если редакция сознательно не будет предпринимать действий, чтобы остановить этот поток.

## 2. Прогнозирование потока статей в журнал

Прогноз по количеству статей, которые поступят в журнал, позволяет принять заблаговременные управленческие решения по подготовке базы данных рецензентов в случае большого потока статей и запланировать маркетинговые коммуникации для привлечения авторов в случае малого потока статей [3].

У всех научных журналов есть поток входящих статей, а значит есть и даты поступления статей. Сбор и систематизация этой информации за несколько прошлых лет позволяет построить прогноз по поступлению статей на будущий период. Это связано с тем, что поток статей имеет сезонную зависимость [4]. На основании полученных данных автором выявлено, что для научно-технического журнала «Вестник Концерна ВКО «Алмаз – Антей» эта зависимость имеет периодичность в 1 год [4; 5]. Автором лично проверен алгоритм предсказания потока статей за 2020 г. и показано, что в целом алгоритм довольно точно аппроксимирует поток [4; 5]. Однако для построения прогноза из статьи [4] требуются определенные знания по математике и расширенное знание возможностей *Excel*.

В настоящем материале предлагается использовать возможности системы *Power BI* для построения прогноза. Для этого необходимо собрать данные о датах поступающих статей в таблицу

и минимально их обработать с использованием функции «Частота» программы *Excel* как показано в табл. 1.

**Таблица 1.** Данные для построения прогноза  
**Table 1.** Data for a forecast

| Дата поступления статьи | Интервал дат | Количество статей |
|-------------------------|--------------|-------------------|
| 02.01.2022              | 01.01.2022   | 0                 |
| 15.01.2022              | 01.02.2022   | 2                 |
| 03.02.2022              | 01.03.2022   | 1                 |

Для построения прогноза используются данные двух последних столбцов табл. 1. Столбец «Интервал» нужно вставить в поле «Date» алгоритма предсказания, а значение «Количество» в поле «Value» алгоритма. Для построения минимально корректного прогноза по месяцам данных должно быть не менее 36, т.е. должны быть данные по потоку поступающих статей как минимум за три года. В табл. 2 собраны данные использования алгоритмов системы *Power BI* при использовании данных о потоке статей в научно-технический журнал «Вестник Концерна ВКО «Алмаз – Антей» за 10 лет (2011–2020).

Как видно из табл. 2 все методы дают примерно одинаковый коэффициент корреляции с данными в 80%, однако различаются по точности. Например, наименьшее значение среднеквадратического отклонения от реальных данных имеет алгоритм *Forecasting TBATS* с настройками. Следует отметить, что без настроек этот алгоритм выдавал пустой результат. Следует также отметить, что все вышеприведенные алгоритмы позволяют дополнительно настроить прогноз, но автором статьи не ставилась задача исследовать коэффициенты корреляции и «подгон» под данные, а ставилась задача, чтобы алгоритмы работали без «тонкой» настройки. Единственная настройка, которая сделана для всех алгоритмов – это «*Forecasting length*», изменена на 12 (по умолчанию 10) потому что прогноз строится на год, то есть на 12 месяцев.

Для более точного прогноза можно варьировать параметры работы алгоритма на тестовых данных. Например, можно построить прогноз на прошлый год по данным за позапрошлый и предшествующие года с целью добиться варьированием параметров алгоритма достижения максимального возможного коэффициента корреляции и минимально возможного среднеквадратического отклонения. У автора статьи получилось добиться значения коэффициента корреляции в 98%, а среднеквадратического отклонения в 2,74.

**Таблица 2.** Результаты обработки потока входящих статей алгоритмами предсказания с прогнозом на 2021 г. по месяцам и реальный поток статей, а также данные о точности прогноза в виде коэффициента корреляции и данных о среднеквадратическом отклонении реального потока от прогнозного (данные прогноза округлены до целых чисел)

**Table 2.** Results of processing the flow of incoming manuscripts by prediction algorithms with a forecast for 2021 by months and the real flow of manuscripts, as well as data on the forecast accuracy in the form of a correlation coefficient and the standard deviation of the real flow from the forecasted flow (forecast data rounded to integers)

| Метод прогноза и реальный поток статей  | Количество статей по месяцам |         |      |        |     |      |      |        |          |         |        |         | Коэффициент корреляции | Среднеквадратическое отклонение |
|---|------------------------------|---------|------|--------|-----|------|------|--------|----------|---------|--------|---------|------------------------|---------------------------------|
|   | Январь                       | Февраль | Март | Апрель | Май | Июнь | Июль | Август | Сентябрь | Октябрь | Ноябрь | Декабрь |                        |                                 |
| Forecast Using Multiple Models by MAQ Software  | 10                           | 13      | 12   | 9      | 8   | 12   | 15   | 8      | 9        | 14      | 35     | 10      | 0,77                   | 6,73                            |
| Forecast Using Neural Network by MAQ Software   | 3                            | 6       | 16   | 1      | 5   | 6    | 13   | 4      | 2        | 14      | 35     | 8       | 0,80                   | 6,31                            |
| Time Series Forecasting Chart   | 7                            | 6       | 10   | 2      | 4   | 9    | 8    | 4      | 3        | 14      | 35     | 5       | 0,81                   | 5,59                            |
| Forecasting with ARIMA  | 8                            | 7       | 10   | 4      | 6   | 10   | 9    | 6      | 5        | 13      | 24     | 7       | 0,80                   | 3,28                            |
| Forecasting TBATS (с настройками Seasonal factor #1 – year, Seasonal factor #1 – month) | 6                            | 6       | 8    | 5      | 5   | 9    | 10   | 5      | 5        | 14      | 20     | 13      | 0,83                   | 2,84                            |
| Реальный поток статей в 2021 г.   | 6                            | 4       | 10   | 6      | 3   | 8    | 4    | 6      | 6        | 8       | 20     | 13      | 1                      | 0                               |

### 3. Построение графа связей между рецензентами

Подробно о построении графа рецензентов и показателях журнала, в том числе наукометрических, описано в статье [6]. Построение графа связей между рецензентами с использованием функции *Force-Directed Graph* системы *Power BI* позволяет визуализировать информацию о том, в каких направлениях рецензенты наиболее сильны, а там, где будет малое количество связей либо очень узкая и редкая тематика, будет точка для роста и поиска новых рецензентов в направлении. Визуализация позволяет находить рецензентов, которые имеют множество связей с коллегами, например, рецензента № 16 (см. рис. 2). Такие рецензенты характеризуются широким кругозором, потому что рецензируют статьи по различным направлениям. Следует отметить, что анализ текстовых связей между рецензентами

возможен, но так как связи представляется в виде 16<->68, 44<->68, 16<->44, 59<->79, 16<->79, 16<->59, 80<->90, 79<->90, их очень сложно оценивать без визуализации.

Анализируя связи между рецензентами, которые показаны на рис. 2, можно сделать несколько выводов:

1. Количество больших кружков в левой нижней части графа соответствуют тематике «Электроника. Радиотехника» – одному из основных направлений деятельности Концерна ВКО «Алмаз – Антей», и там существует много связей между рецензентами (размер кружков соответствует количеству связей).

2. Рецензент 16 имеет рекордно большое количество связей с коллегами – 38 и является глубоко эрудированным специалистом, который рецензирует статьи по довольно широкой тематике.



3. Малые кружки характеризуют рецензентов, которые недавно включились в работу, либо рецензируют статьи по крайне узкой тематике. Например, рецензент № 40 (справа) имеет очень узкую специализацию «Радиология», которая не часто встречается на страницах журнала.

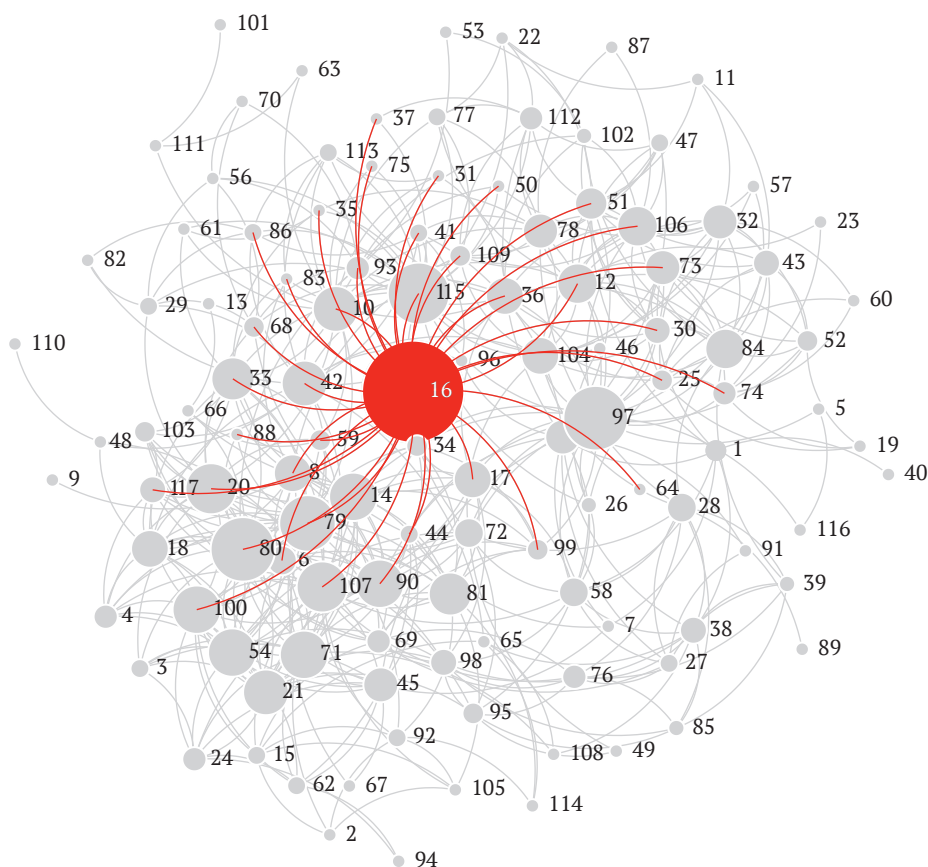
Использовать граф рецензентов в деятельности редакции можно следующим образом. Например, когда основной рецензент по тематике загружен, зная его связи, можно автоматически найти другого рецензента. Также граф рецензентов говорит нам об основной тематике журнала, но более ценная информация о новой тематике, где связей мало и, возможно, нужно искать рецензентов на слабо охваченную тематику, где специалистов и связей не так много. Например, в научно-техническом журнале «Вестник Концерна ВКО «Алмаз – Антей» так были найдены необходимые специалисты по нескольким научным направлениям (теплофизика, теплозащитные покрытия и др.).

Следует отметить, что для построения графа были использованы данные о рецензентах

и подготовленных рецензиях за 7 лет (с 2015 г.), которые создают около одной тысячи связей. Если на каждую статью назначается только один рецензент, то граф построить невозможно. Если рецензентов на статьи в журнале всегда два, то сбор данных упрощается, так как из двух человек просто создать две колонки связей в *Excel* и вставить в функцию *Force-Directed Graph* системы *Power BI*. Если рецензентов по каждой статье более двух, то их связи можно получить, используя свободную облачную версию программного продукта *Wolfram Mathematica* ([www.wolframcloud.com](http://www.wolframcloud.com)), включением функции *KSubsets*, используя следующий код (где {1,2,3} – это три рецензента на одну статью, а общее количество скобочек внутри массива – 4 статьи):

```
<<Combinatorica`
massive={{1,2,3},{4,5,6},{7,8,9,10},{11,12,13}}
Table[KSubsets[massive[[i]],2],{i,1,Length[massive]]}
```

который позволяет получить выдачу массива из сочетаний рецензентов



**Рис. 2.** Граф связей между рецензентами научно-технического журнала «Вестник Концерна ВКО «Алмаз – Антей»

**Fig. 2.** Graph of relations between reviewers of the scientific and technical Journal of “Almaz – Antey” Air and Space Defence Corporation

```
{{{1,2},{1,3},{2,3}},{4,5},{4,6},{5,6}},{7,8},{7,9},{7,10},{8,9},
{8,10},{9,10}},{11,12},{11,13},{12,13}}}
```

Для выполнения данного кода нужно вставлять его каждой строчкой отдельно и нажимать shift+enter, а также вставить двоянные квадратные скобки «и» в системе *Wolfram Mathematica* сочетанием клавиш (Esc [[ Esc) левая скобка, а правая скобка (Esc ]] Esc).

Вместо массива чисел можно использовать фамилии рецензентов. Тогда код может выглядеть примерно так.

```
<<Combinatorica`
massive={{«Иванов»,«Петров»,«Сидоров»},{«Макаров»,
«Антонов»,«Мишин»}}
Table[KSubsets[massive[[i]],2],{i,1,Length[massive]]}
```

И выдача соответственно

```
{{{Иванов,Петров},{Иванов,Сидоров},{Петров,Сидоров}},
{{Макаров,Антонов},{Макаров,Мишин},{Антонов,Мишин}}}
```

## Выводы

Применение современных вычислительных методов к анализу данных о научном журнале позволяет значительно сократить неопределенность при планировании деятельности журнала на год вперед:

1. Построение прогноза поступающих статей позволит знать заранее, какой поток придет в журнал и как его отработать с рецензентами.

2. Простой анализ конверсии, т.е. количества публикуемых статей к поступившим, дает возможность определить необходимое количество привлекаемых за год рукописей и, соответственно, воспользоваться требуемыми маркетинговыми коммуникациями для расширения потока.

3. Знание связей между рецензентами дает возможность сделать вывод о компетенциях и сильных сторонах журнала и рецензентов, а также спрогнозировать точки роста журнала, например, в новом научном направлении.

## СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Luhn H.P.A Business Intelligence System. *IBM Journal of Research and Development*. 1958;2(4):314–319. <https://doi.org/10.1147/rd.24.0314>
2. Большаков Д. Ю. Стресс-тестирование научного журнала. *Научный редактор и издатель*. 2021;6(1):18–27. <https://doi.org/10.24069/2542-0267-2021-1-18-27>
3. Большаков Д. Ю. Анализ эффективности маркетинговой политики научного журнала. *Инновации*. 2020;266(12):77–82. <https://doi.org/10.26310/2071-3010.2020.266.12.010>
4. Большаков Д. Ю. Аналитика редакционно-издательских процессов научного журнала. *Научный редактор и издатель*. 2020;5(2):102–112. <https://doi.org/10.24069/2542-0267-2020-2-102-112>
5. Большаков Д. Ю. Математическая модель прогноза поступления рукописей в научный журнал: доклад на 9-й Международной научно-практической конференции «Научное издание международного уровня – 2021: мировые тенденции и национальные приоритеты», г. Москва, 24–27 мая 2021 г. URL: <https://rassep.ru/upload/iblock/e8f/2021.05.15-Doklad-na-konferentsii-ANRI-NIMU-2021.pdf> (дата обращения: 29.04.2022).
6. Большаков Д. Ю. О связях в науке на примере редакционной коллегии научного журнала. *Наука и научная информация*. 2021;4(1–2):23–32. <https://doi.org/10.24108/2658-3143-2021-4-1-2-23-32>

## ИНФОРМАЦИЯ ОБ АВТОРЕ

**Денис Юрьевич Большаков**, кандидат технических наук, начальник отдела научно-технических изданий и специальных проектов аппарата генерального директора АО «Концерн ВКО «Алмаз – Антей», заместитель главного редактора научно-технического журнала «Вестник Концерна ВКО «Алмаз – Антей» / Journal of «Almaz – Antey» Air and Space Defence Corporation, г. Москва, Российская Федерация; ORCID: <https://orcid.org/0000-0001-7694-1454>; e-mail: [press@almaz-antey.ru](mailto:press@almaz-antey.ru)

## INFORMATION ABOUT THE AUTHOR

**Denis Yu. Bolshakov** Cand. Sci. (Eng.), Head of the Department of Scientific and Technical Issues and Special Projects of the Office of the Director General, Almaz – Antey Air and Space Defence Corporation, JSC, Deputy Editor-in-Chief of the Journal of “Almaz – Antey” Air and Space Defence Corporation, Moscow, Russian Federation; ORCID: <https://orcid.org/0000-0001-7694-1454>; e-mail: [press@almaz-antey.ru](mailto:press@almaz-antey.ru)

Поступила в редакцию 13.05.2022

Поступила после рецензирования 26.07.2022

Принята к публикации 29.07.2022

Received 13.05.2022

Revised 26.07.2022

Accepted 29.07.2022