

Хакимова Татьяна Викторовна, магистрант, Хакасский государственный университет им. Николая Федоровича Катанова, Россия, г. Абакан

Голубничий Артем Александрович, научный руководитель, старший преподаватель кафедры ПОВТиАС, Хакасский государственный университет им. Николая Федоровича Катанова, Россия, г. Абакан

ПОПУЛЯРНОСТЬ И ПЕРСПЕКТИВЫ ЯЗЫКА ПРОГРАММИРОВАНИЯ R

Аннотация: в статье дается краткий исторический экскурс языка R, особенности расширения функционала с выходом новых версий. Приводится изменение рейтинга языка R по индексу TIOBE.

Ключевые слова: Индекс TIOBE, язык R.

Abstract: The article provides a brief historical excursus of the R language, features of the functionality expansion with the release of new versions. The change in the ranking of the language R index TIOBE.

Keywords: TIOBE programming community index, R Language.

Одним из наиболее точных методов определения популярности того или иного языка программирования может служить рейтинг, составленный по принципу подсчета количества страниц выдачи в поисковых запросах, тех или иных поисковых системах. Наибольшую популярность, в настоящее время, имеет рейтинг, ежемесячно представляемый голландской компанией TIOBE Software [1].

TIOBE programming community index (или более привычный в русскоязычной практике термин индекс TIOBE) представляет собой индекс, оценивающий популярность того или иного языка программирования на основе

количества выдачи по соответствующим поисковым запросам типа "+<language> programming". В качестве источников для формирования индекса выступают наиболее популярные поисковые сервисы и порталы, такие как Google (как по основному сайту проекта в домене верхнего уровня .com, так и по некоторым национальным доменам верхнего уровня, например .fr, .it, .cn и др.), Baidu.com, Wikipedia.org, Yahoo.com, Csdn.net, Bing.com, Ebay.com, Msn.com, Hao123.com и др. [1]. Индекс не предполагает оценку качества написанного программного кода, или даже простого подсчитывания количества строк кода в той или иной выдаче запроса. В том числе по этой причине, он дает лишь приближенное понимание о популярности языка.

ТЮВЕ представляет рейтинг языков программирования начиная с 2001 года, сам язык R появился в 1993 году, как реализация языка S, созданного Джоном Чемберсом в 1976 в Bell Labs. Язык R был создан Россом Ихакой и Робертом Джентльменом в Оклендском университете и был назван в честь первой буквы имен его разработчиков [2].

Первая стабильная бета-версия появилась в 2000 году. В рейтинг ТЮВЕ язык R вошел только в мае 2007 года с процентным соотношением запросов, в сравнении со всеми остальными языками программирования, в 0.065%. Актуальной на тот момент была версия языка 2.11. В сравнении с предыдущей версией 2.1 в ней добавилась возможность работы с кодировкой UTF-8, что послужило началом глобализации и локализации языка, в связи с возможностью работы с множеством языков, использующих не только латинский набор букв.

Некоторые особенности версий языка

Для того чтобы понять особенности популярности языка R, в тот или иной временной интервал, стоит, конечно, рассматривать особенности языка, добавленные в разные версии релиза. Для удобства восприятия особенностей версий, ключевые из них были сведены в таблицу с соответствующим описанием (таблица 1). Как видно из таблицы, последние изменения, вносимые в ядро языка, в большей степени связаны с изменениями сериализации. Такой

подход объясняется основной задачей, решаемой языком R – обработкой больших массивов данных.

Таблица 1 – Описание ключевых изменений версий языка R

Версия	Дата релиза	Описание
0.16	01.04.1997	Последняя альфа версия языка
0.49	23.04.1997	Первая версия языка по-прежнему доступна в репозитории CRAN
0.6	05.12.1997	R становится частью сообщества GNU
0.65.1	07.10.1999	Появление возможности работы напрямую с пакетами из CRAN
1.0	29.02.2000	Первая стабильная версия языка
1.4	19.12.2001	Появление S4 методов и как следствие поддержка создания сложных объектов данных с учетом парадигмы объектно-ориентированного программирования
1.8	08.10.2003	Введен гибкий механизм обработки условий
2.0	04.10.2004	Введена отложенная загрузка, которая позволяет быстро загружать данные с минимальными затратами системной памяти
2.1	18.04.2005	Поддержка кодировки UTF-8, а также начала глобализации и локализации для разных языков
2.11	22.04.2010	Поддержка 64 битной версии Windows
2.13	14.04.2011	Добавление новой функции компилятора, которая позволяет ускорить функции путем преобразования их в байт-код
2.14	31.10.2011	Добавлены обязательные пространства имен для пакетов. Добавлен новый параллельный пакет
2.15	30.03.2012	Новые функции балансировки нагрузки. Улучшена скорость сериализации для длинных векторов
3.0	03.04.2013	Поддержка числовых значений индекса 2^{31} и выше в 64-битных системах
3.4	21.04.2017	ЛГ компиляция функций и циклов в байт-код включены по умолчанию
3.5	23.04.2018	Пакеты скомпилированы при установке по умолчанию. Компактное внутреннее представление целочисленных последовательностей. Добавлен новый формат сериализации для поддержки компактных внутренних представлений.
3.6	26.04.2019	Выход последней актуальной версии интерпретатора языка

Последней актуальной версией, на момент написания данного текста, является версия 3.6, однако на центральном сайте проекта [3] уже содержится анонс версии 3.6.1, получившей название Action of the Toes, планируемой к выходу 5 июля 2019 года.

Популярность языка согласно индексу ТЮВЕ

Сводные данные о популярности языка представлены на рисунке 1. На вершине своей популярности (январь 2018 года) язык имел значение индекса немногим превосходящее 2,5% и занимал восьмую строчку в рейтинге языков

программирования, уступая Java, C++, Python и ряду других языков, стандартно занимающих лидирующие строчки практически всех рейтингов.

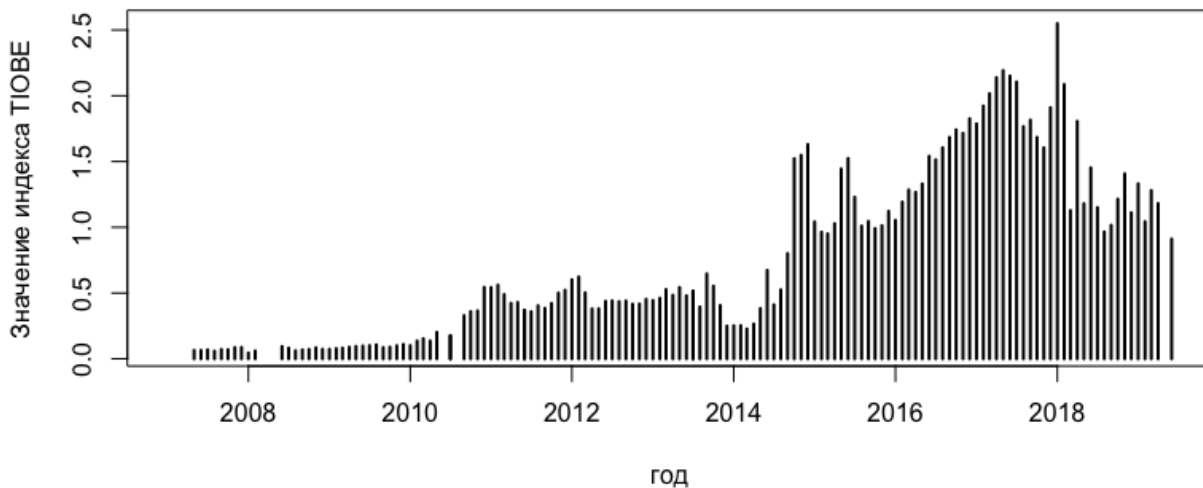


Рисунок 1 – Рейтинг языка R на основе индекса ТЮВЕ.

При этом из рисунка 1 видно, что в настоящее время, несмотря на некоторое уменьшение интереса к языку в Интернет-сообществе, индекс по-прежнему близок к значению 1%, что позволяет занимать 21-22 место в общем рейтинге языков программирования.

Соотнося изменения, появившиеся в данном языке в разных версиях, и популярность данного языка в тот или иной интервал времени, стоит сделать вывод о независимости этих процессов.

Основным конкурентом языка R в области анализа данных, в настоящее время, является язык Python. Большое количество библиотек, используемых для анализа данных, часто пишутся именно на данном языке, таким образом, одной из причин снижения интереса к R является переход части разработчиков на более простой в синтаксисе язык. При этом особенности векторизации языка R позволяют реализовать многие решения с большей скоростью, что актуально для особо сложных и высоконагруженных проектов. Таким образом, на наш взгляд, данный язык прочно занял свою нишу и в ближайшее время также

активно будет использоваться определенной частью программистов, веб-аналитиков и вообще специалистов в области анализа данных.

Библиографический список:

1. Programming Languages Definition | TIOBE – The Software Quality Company [Электронный ресурс] URL: <https://www.tiobe.com/tiobe-index/programming-languages-definition/> (дата обращения 17.06.2019).

2. Hornik, K R FAQ Frequently Asked Questions on R Version 2018-10-18 [Электронный ресурс] URL: <https://cran.r-project.org/doc/FAQ/R-FAQ.html> (Дата обращения 17.06.2019).

3. R: The R Project for Statistical Computing [Электронный ресурс] URL: <https://www.r-project.org> (дата обращения 17.06.2019).