An Introduction To Low-Code No-Code Test Automation Tools

Terraform: The Cloud-Agnostic Solution For Infrastructure As Code

₹120
ISSN-2456-4885

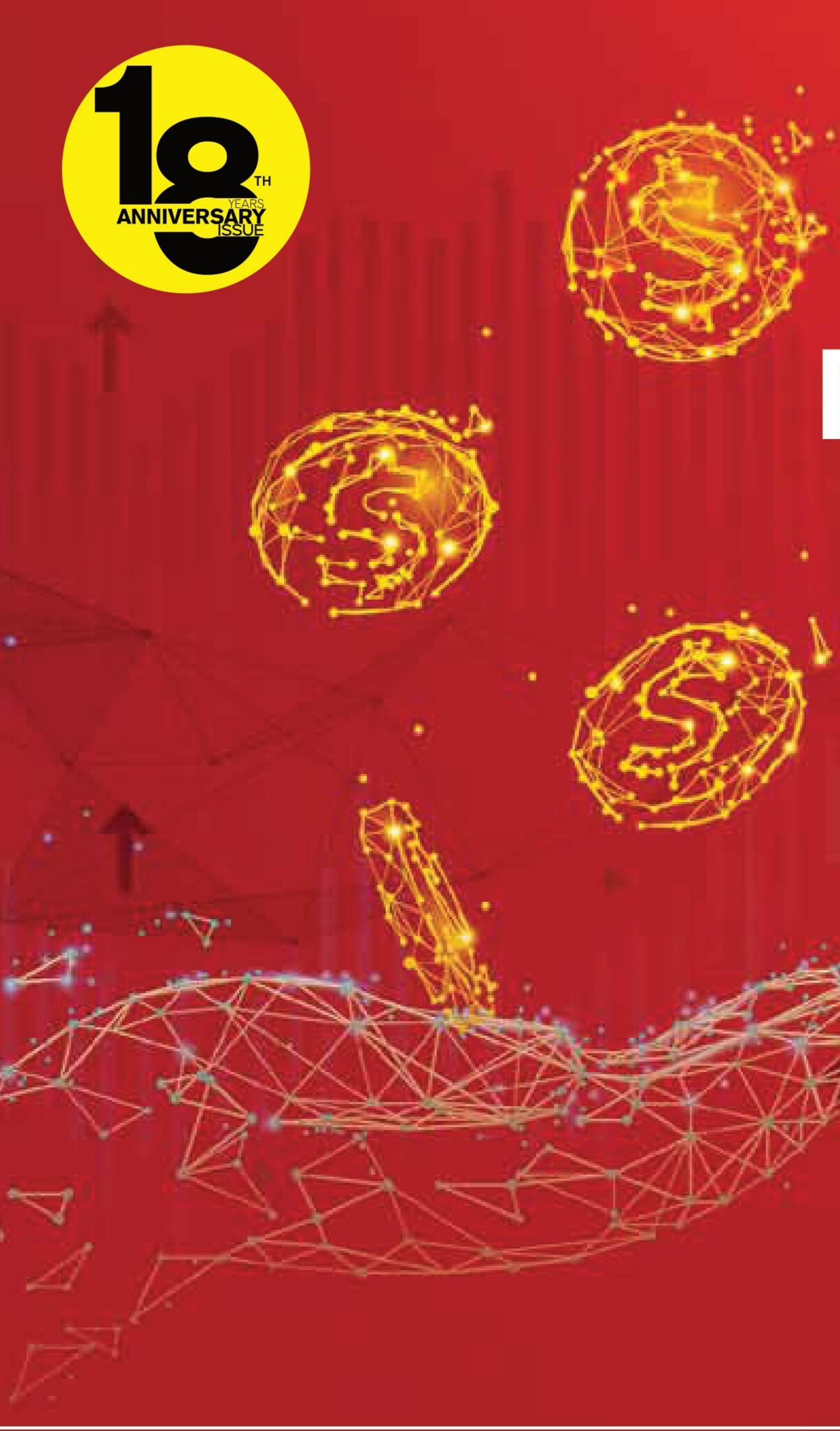
OpenSource ForYou

Volume: 10 | Issue: 04 | Pages: 108 | February 2022

THE COMPLETE MAGAZINE ON OPEN SOURCE

EFY GROUP Publication

18TH YEARS ANNIVERSARY ISSUE



How To Make Money With Open Source

Building Profitable Open Source Business Models

Why Open Source Is A Serious Business Today

Why And How To Become An Open Source Contributor

Let The Domain Drive The Design

The Future Belongs To These Open Source DevOps Tools



www.opensourceforu.com

Wanna Be Your Own Boss?

DO OPEN SOURCE.



Demand for Open Source is sky rocketing. Be it for managing IT infrastructure or development of software--Open Source solutions are what customers are seeking.

All you need to do is develop expertise in an Open Source stack, and then build a team around it!

And, Open Source For You can be your friend and a guide through this journey.

TO READ OUR PRINT EDITION Visit: https://subscribe.efyindia.com

TO READ OUR EZINE EDITION Visit: https://ezine.lfymag.com



WORLD'S LEADING PUBLICATION ON OPEN SOURCE

Looking for marketing solutions to engage with cutting edge techies?
Contact us at growmybiz@efy.in OR call us at +91-9811155335.

Does Your Antivirus Solution Provide You With Complete Protection?

Don't RELY on 20 years old technology to FIGHT current date viruses



TOP 10 things that your Antivirus solution should provide to tackle today's threat...

- Ransomware File Protection
- ATP- Advance Threat Protection
- EDR- Endpoint Detection And Response
- Application Blocking
- Deep Learning Malware Analysis
- Exploit Prevention
- URL Blocking
- Disk and Boot Record Protection
- Peripheral Control
- Respond Investigate Remove (Root Cause Analysis)

If it doesn't, contact us for a solution which can

Contact: Santosh on 9971696319 or Email at santosh.gupta@itsipl.com

ITS
A Total Solutions Company

I. T. Solutions India Private Limited
D-88/5, Okhla Industrial Area, Okhla Phase I, New Delhi -110020
Ph: 011-47695000 • Email: sales@ itsipl.com • www.itsipl.com

- Mumbai
- Jaipur
- Chandigarh

SOPHOS 2020 BEST NEXTGEN PARTNER | Delhi/ NCR
SOPHOS 2019 BEST NEXTGEN PARTNER | Delhi/ NCR
SOPHOS 2018 BEST SI PARTNER | Delhi

CONTENTS

FEBRUARY 2022 | ISSN-2456-4885

FOR U & ME

28 How to Build a Community and Why it Matters

38 Companies Don't Innovate; People Do

FOCUS

53 The Future Belongs to These Open Source DevOps Tools

DEVELOPERS

65 Use Node-RED to Build Your Brand

69 Carry Out JSONField Schema Validation Using JSON Models

81 An Introduction to Low-Code and No-Code Test Automation Tools

85 Test Your Data Science Project with Google's Caliban

COLUMNS

63 CodeSport

ADMIN

93 Installing and Using RethinkDB on Ubuntu

95 Exploring Network Statistics with 'netstat'

98 Terraform: The Cloud-Agnostic Solution for Infrastructure as Code

OPEN GURUS

102 Enhancing the Design of Autonomous Cars Using Node-RED

The Design Odyssey

14

Let the Domain Drive the Design





20

Building Profitable Open Source Business Models



23

Why Open Source is a Serious Business Today



31

Why and How to Become an Open Source Contributor



43

An Open Source Ecosystem for 5G and Telecom Networks

REGULAR FEATURES

07 FossBytes

www.opensourceforu.com

Job Guaranteed.

You don't need a training institute. Just do it yourself.

WORK WITH OPEN SOURCE.

Demand for Open Source is sky rocketing. Be it for managing IT infrastructure or development of software —Open Source experts are what organisations are seeking. Develop expertise in an Open Source stack, and watch your career graph zoom.

You don't need a training institute for this. You can do it yourself. And, *Open Source For You* can be your friend and a guide through this journey.

TO READ OUR PRINT EDITION Visit: https://subscribe.efyindia.com

TO READ OUR EZINE EDITION Visit: https://ezine.lfymag.com

WORLD'S LEADING PUBLICATION ON OPEN SOURCE

Looking for marketing solutions to engage with cutting edge techies?
Contact us at growmybiz@efy.in OR call us at +91-9811155335.



EDITOR
RAHUL CHOPRA

EDITORIAL, SUBSCRIPTIONS & ADVERTISING
Delhi (HQ)
D-87/1, Okhla Industrial Area, Phase I, New Delhi 110020
Phone: +91-9811155335
E-mail: info@efy.in

MISSING ISSUES
Phone: +91-9811155335
E-mail: support@efy.in

BACK ISSUES
Phone: +91-9811155335
E-mail: support@efy.in

NEWSSTAND DISTRIBUTION
Phone: +91-9811155335
E-mail: efycirc@efy.in

ADVERTISEMENTS
NEW DELHI (HEAD OFFICE)
Phone: +91-9811155335
E-mail: efyenq@efy.in

MUMBAI
E-mail: rmowest@efy.in

BENGALURU
E-mail: rmosouth@efy.in

CHINA
Worldwide Focus Media
E-mail: china@efy.in

GERMANY
pms Plantenberg Media Service GmbH
E-mail: germany@efy.in

JAPAN
Tandem Inc.
E-mail: japan@efy.in

TAIWAN
J.K. Media
E-mail: taiwan@efy.in

UNITED KINGDOM
ASA Media
E-mail: uk@efy.in

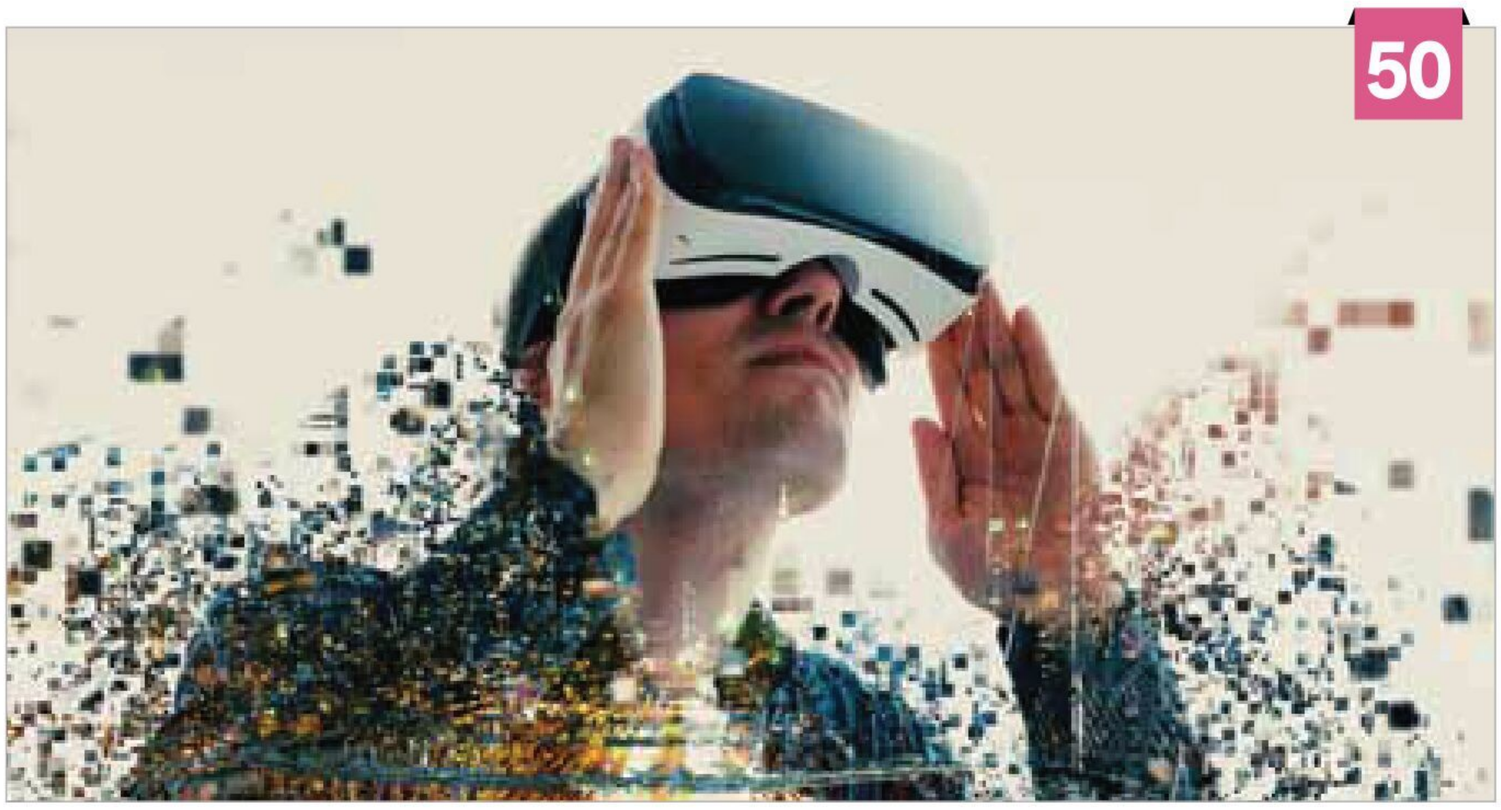
UNITED STATES
E & Tech Media
E-mail: usa@efy.in

Printed, published and owned by Ramesh Chopra. Printed at Tara Art Printers Pvt Ltd, A-46,47, Sec-5, Noida, on 28th of the previous month, and published from D-87/1, Okhla Industrial Area, Phase I, New Delhi 110020. Copyright © 2021. All articles in this issue, except for interviews, verbatim quotes, or unless otherwise explicitly mentioned, will be released under Creative Commons Attribution-NonCommercial 3.0 Unported License a month after the date of publication. Refer to http://creativecommons.org/licenses/by-nc/3.0/ for a copy of the licence. Although every effort is made to ensure accuracy, no responsibility whatsoever is taken for any loss due to publishing errors. Articles that cannot be used are returned to the authors if accompanied by a self-addressed and sufficiently stamped envelope. But no responsibility is taken for any loss or delay in returning the material. Disputes, if any, will be settled in a New Delhi court only.

SUBSCRIPTION RATES			
Year	**Newstand Price (₹)**	**You Pay (₹)**	**Overseas**
Five	7200	4320	—
Three	4320	3030	—
One	1440	1150	US$ 120

Kindly add ₹ 50/- for outside Delhi cheques.
Please send payments only in favour of EFY Enterprises Pvt Ltd.
Non-receipt of copies may be reported to **support@efy.in**—do mention your subscription number.

CONTENTS


50

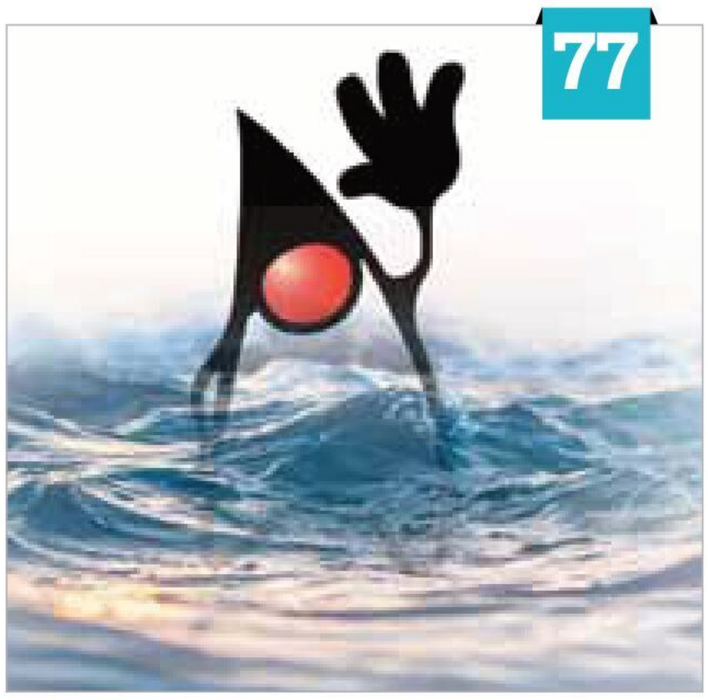
Open Source Platforms You Can Use for AR and VR


48

The Best Security Tools for Linux Administrators


59

Top 10 Open Source Deep Learning Tools


77

Is Java Dead?


87

R Testing Frameworks

Amazon Elastic Kubernetes Service to support IPv6 address spaces



Amazon's Elastic Kubernetes Service (EKS) now supports IPv6 address spaces, according to a new blog post by Sébastien Stormacq, principal developer advocate, AWS.

Amazon EKS is said to make it easy to deploy containerised workloads. It provides highly available clusters and automates tasks such as patching, node provisioning, and updates.

Kubernetes presents network management challenges. It uses a flat networking model that requires each pod to receive an IP address. This simplified approach enables low-friction porting of applications from virtual machines to containers, but requires a significant number of IP addresses that many private VPC IPv4 networks are not equipped to handle.

Some cluster administrators work around this IPv4 space limitation by installing container network plugins (CNI) that virtualise IP addresses a layer above the VPC, but this architecture limits an administrator's ability to effectively observe and troubleshoot applications and has a negative impact on network performance at scale. Further, to communicate with Internet services outside the VPC, traffic from IPv4 pods is routed through multiple network hops before reaching its destination, which adds latency and puts a strain on network engineering teams that need to maintain complex routing setups.

To avoid IP address exhaustion, minimise latency at scale, and simplify routing configuration, the solution is to use IPv6 address space.

There are a few advantages of using Amazon EKS clusters with an IPv6 network. Stormacq lists these: "First, you can run more pods on one single host or subnet without the risk of exhausting all available IPv4 addresses available in your VPC. Second, it allows for lower-latency communications with other IPv6 services, running on-premise, on AWS, or on the Internet, by avoiding an extra NAT hop. Third, it relieves network engineers of the burden of maintaining complex routing configurations."

Using Raspberry Pi for brain-computer interfaces

Brain-computer interfaces (BCIs) may sound futuristic, but are now a reality. Many approaches are being used to detect and interpret electrical activity in the

NowSecure announces new GitHub Action for mobile app security

NowSecure, the standards based mobile app security and privacy software company, has announced a new NowSecure GitHub Action to integrate mobile app security testing into mobile app developer workflows directly inside GitHub repos. With this, developers can ensure the security and privacy of mobile apps as they build them to deliver high-quality and secure releases faster.

"Developers want to deliver innovative, high-quality mobile applications fast," said NowSecure CEO Alan Snyder. "This means they need a developer-first, easy-to-use and accurate mobile app security solution that integrates directly into their dev workflows. We are excited to partner with GitHub to ensure developers have a friction-free path to coding their mobile apps with security built-in."

The NowSecure Platform is said to enable fast, automated security testing of mobile apps inside the developer workflow in both on-demand and continuous testing configuration options.



NowSecure Action for GitHub can test any Apple iOS or Google Android mobile app developed in any language or framework including Kotlin, Java, Swift, Objective-C, Apache Cordova, Flutter, Ionic, Java, jQueryMobile, Python, React Native, Xamarin, and more.

It is a part of the comprehensive suite for mobile app security, including the NowSecure Platform, for continuous security testing in the development pipeline.

propSym, a new software tool to aid material science research

A new software tool can accelerate materials science research by cutting out tedious background research on material properties. Penn State and Sandia National Laboratories researchers recently debuted propSym, an open source software on the programming platform MATLAB, to calculate the fundamental constants needed to describe the physical properties of solids, such as metals, ceramics, or composites.

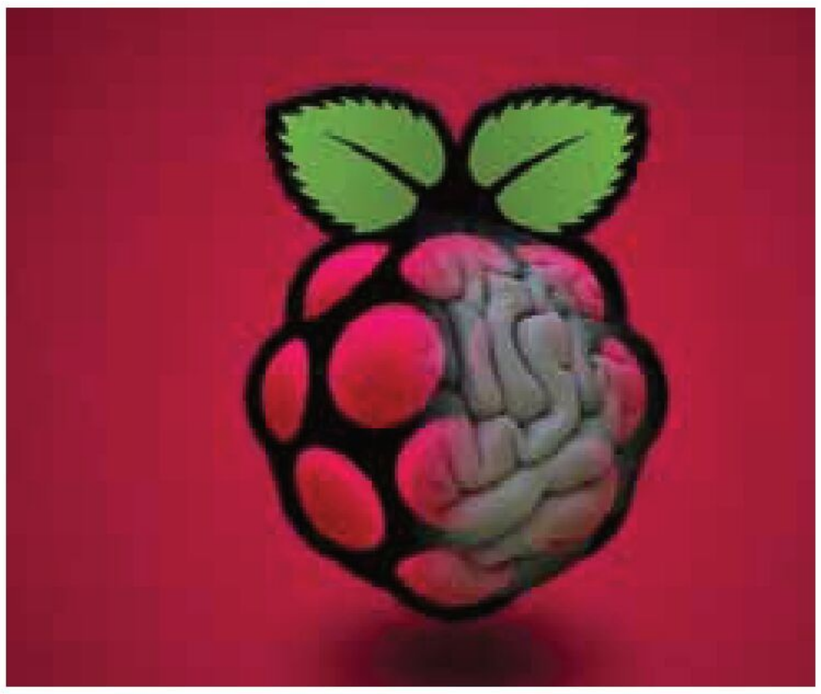
The researchers developed propSym, the details of which were published in the *Journal of Applied Crystallography*, after they could not find reliable information about langasite, a material used in sensing and energy harvesting devices in a separate joint study with Sandia National Labs.

"Traditionally, the relationships between fundamental constants and material symmetries are found only in appendices of textbooks or tables in journal articles," said Christopher Kube, assistant professor of engineering science and mechanics at Penn State, who led the project. "After a thorough search, we were not able to find reference data for several non-linear material properties for langasite. When data was available, we found instances of typos and inconsistencies across references. Incorrect input data will ruin a model."

Kube and his collaborators used propSym to determine the properties of langasite, such as elasticity and the ability to accumulate electric charge. But Kube emphasised the program is not limited to these two properties alone.

propSym is open source and available for download for anyone with MATLAB. Darren Branch and Daniel Spencer, scientists from the Sandia National Labs, contributed to this research.

human brain so that it may be utilised to control a computer. Ildar Rakhmatulin and Sebastian Volkl, authors of an article on the arXiv open access server, advocate using a Raspberry Pi instead of hyper-expensive equipment for brain-computer interfacing.



Rakhmatulin studies BCIs at Russia's South Ural State University, and Volkl works in the disciplines of AI and neural networks. He previously created a laser turret for shooting down mosquitoes using a Raspberry Pi 3 B+ and camera. They banded together to form hackerBCI, a platform that attempts to make neuroscience projects and research more accessible.

PiEEG, a Raspberry Pi project, uses C, C++, and Python to read up to eight real-time electroencephalography (EEG) signals collected from the brain by electrodes put in the subject's cap. The pair has shared use cases on GitHub, including controlling robots and drones. One of the potential applications is the control of exoskeletons.

If you're interested in using your Raspberry Pi to read your brain signals, hackerBCI will shortly begin crowdfunding the PiEEG.

LibreOffice to include two 'made up' languages to 7.3 update

Starting in February this year, LibreOffice 7.3, an open source writing platform, will offer support for two 'made up' languages: Interslavic and Klingon.

According to *neowin.com*, the decision to incorporate the Klingon and Interslavic languages was made in order to reduce user workload by allowing them to work with the languages without having to use alternative translation. Linguist Marc Orkrand created the Klingon language for the Star Trek franchise. The goal of Interslavic, on the other hand, is to "bridge the language gap between Slavic languages such as Russian and Polish."



Eike Rathke, a Red Hat employee, added the language update to the newest edition of the LibreOffice software. The program can be tweaked to incorporate feedback from a variety of users' perspectives, resulting in a product that is essentially tuned to the work at hand by those who use it.

LibreOffice 7.3 will be released in February 2022. Visit *LibreOffice.org* for further information.

CAST releases new SCA browser extension for Chrome

CAST Highlight subscribers can now install and use a new browser extension that allows them to see the legal and security exposures associated with a particular open source component as they attempt to download it from the Internet.

The ubiquitous use of open source components in custom-built applications creates intellectual property and security risks for business owners and

corporate legal teams. CAST Highlight provides an effective, fast-rollout alternative or complement to traditional software composition analysis (SCA) products for controlling the risks inherent in open source software across entire application portfolios.

The new CAST Highlight SCA browser extension for Chrome enables users to tap into CAST's knowledge base of over 100 million open source components and view insights in real-time while browsing for specific components across the Internet. The extension shows details such as licensing requirements, security vulnerabilities, and technology obsolescence directly in the browser window for the component on the current page. It also identifies if the component belongs to the 'Allow' or 'Deny' list as specified by the user's organisational policy. Users can now make smarter decisions about the use of specific open source components before downloading them from the Internet. The new SCA browser extension is available for all CAST Highlight subscribers.

According to CAST Software Inc., CAST technology is an accurate "MRI for software," which delivers actionable insights into software composition, architectures, database structures, critical flaws, quality grades, cloud readiness levels, and work effort metrics. It is used globally by forward-looking digital leaders to make objective decisions, accelerate modernisation, and raise the security and resiliency of mission-critical software.

OpenEduCat educational management system has 300 modules

OpenEduCat is an open source, cloud based educational management system for universities, colleges and institutions that is available both online and on-premise. It provides education ERP software to help institutions run their operations efficiently.



OpenEduCat, Inc. has assisted various organisations in effectively managing their day-to-day operations by offering end-to-end services for implementation, planning, strategy execution, and training. The software has set an industry standard with 300 modules, 300,000 installations, 2 million users, 45 localisations, and 65 languages.

OpenEduCat provides end-to-end services, backed by business professionals, marketing ninjas, technical gurus, project managers, and subject specialists. It has tailored solutions to handle operations easily and efficiently, with features such

Cellebrite secures a multi-year contract for its OSINT solution

Cellebrite, a provider of digital intelligence (DI) solutions for the public and private sectors, has announced that it has won a significant Open Source Intelligence (OSINT) customer contract following the company's acquisition of Digital Clues in November 2021. This multi-year contract has a value of US$ 1.5 million for the first year and is the most noteworthy SaaS deal for Cellebrite to date.

Cellebrite's OSINT solution, part of the company's investigative DI platform, is aimed at supporting the early stages of law enforcement investigations, irrespective of whether the investigative teams have a lead or not, by quickly enabling them to collect and connect large amounts of fragmented digital evidence from many dynamic sources, including the surface Web, deep Web, and dark Web.

The adoption of the OSINT solution is being driven by its use by law enforcement agencies as they increase their efforts to stop crimes against children, drug and human trafficking, organised crime, and terrorism. The new customer selected Cellebrite's OSINT solution to replace its existing system because of its ability to adapt quickly to rapidly changing open source repositories, and ease of use.

"We are pleased with the early success reflected by this significant, multi-year contract that demonstrates our ability to provide end-to-end investigative workflow solutions for law enforcement which help to make our communities safer," said Alon Klomek, Cellebrite's chief business officer.

Block Inc. to build an open source Bitcoin mining system

In a tweet, Block Inc. CEO Jack Dorsey said that the company's payments business is pushing forward with plans to establish an open source Bitcoin mining infrastructure. In October last year, Dorsey had claimed that present mining activities were overly centralised and concentrated, which he considered as incompatible with the blockchain networks' decentralised nature.



Block's general manager for hardware, Thomas Templeton, talked of the company's ambitions for the future mining system in a series of tweets. "We want to make mining more distributed and efficient in every way, from buying, to set up, to maintenance, to mining," Templeton wrote. "We're interested because mining goes far beyond creating a new Bitcoin. We see it as a long-term need for a future that is fully decentralised and permissionless."

Block is developing an engineering team led by Afshin Rezayee, its lead of Bitcoin mining, to look into the company's own specialised silicon, open source miner firmware, and other system software to accomplish this. Electrical engineers, analogue engineers, and layout engineers, according to Templeton, are still open positions.

Templeton stated that the organisation will continue to use user feedback as a main source of research in order to advance this initiative.

Block has not stated when its Bitcoin mining system will be released or provided a calendar of milestones, implying that the technology is still in an early phase of development.

as core module, essential module, advance module, LMS module, management module, communication module, ERP module, and technical module. Small businesses and educational institutions can use built-in live classrooms or third-party platforms like Google Meets, Zoom, Microsoft Teams, and Big Blue Button to organise live classrooms. The ERP software provides everything from information administration to course creation, financial management, attendance processing, LMS, and faculty management.

The founders of OpenEduCat, Parthiv Patel and Neha Patel, think that technology is a critical tool for dealing with everyday issues. After years of working with thousands of enterprises, they discovered a great need for a comprehensive solution in educational institutions. The team decided to design a product that is simple to use and comprehensive for educational institutions, based on their business domain and process understanding.

BMC introduces new capabilities in its products to accelerate software delivery

BMC, a software solutions provider for the autonomous digital enterprise, has announced new capabilities and integrations across its BMC AMI (automated mainframe intelligence) and BMC Compuware portfolios to accelerate software delivery with greater data visibility and context-based insights across IT environments.

According to the company, its Helix Discovery solution now automatically maps and visualises service dependencies across mainframe and distributed environments using application network connection details sent from the BMC AMI Ops product. This allows users to prioritise problems according to the possible business impact.



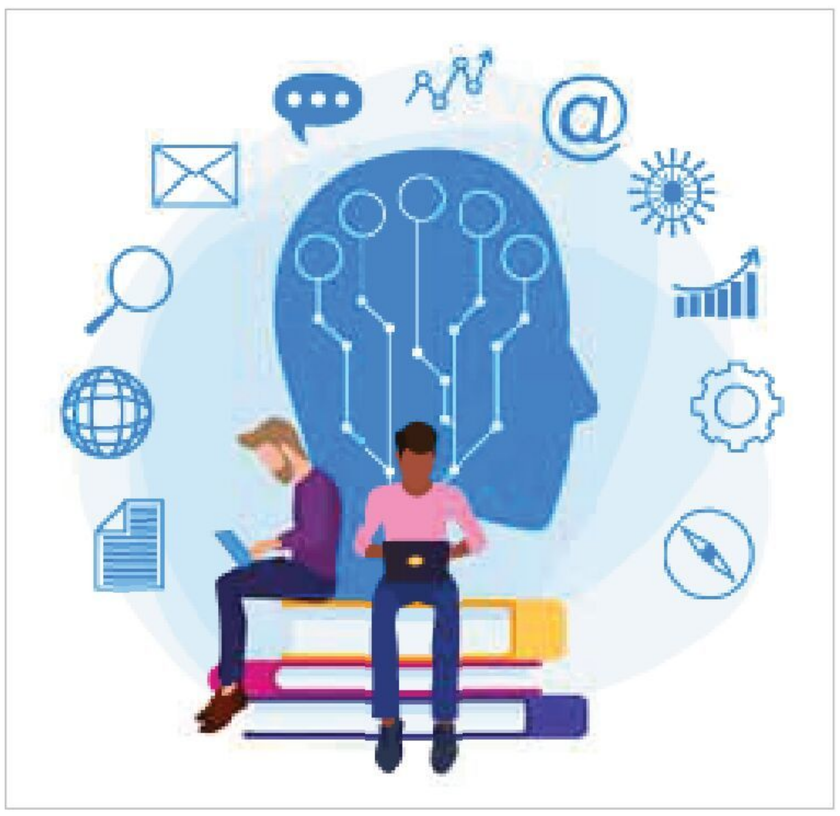
According to a Forrester study, "the ability to capture software development KPIs has become an important component in a development team's modernisation efforts." To succeed in digital transformation, it is critical to select modern mainframe development tools with the ability to capture meaningful process metrics, such as quality, velocity, and efficiency.

Diffgram training data platform for machine learning is now open source

The Diffgram training data platform is now open source – after three years of closed source work. Diffgram is similar to closed source offerings like Labelbox but is different in that it's open source.

The Open Source Diffgram community is already over 600 technical people strong. There are people submitting code for new cloud providers, creating bug

reports for issues, and more. Overall, there is a global team creating a shared repository of knowledge. So companies can rest assured they will always be able to get support and updates.



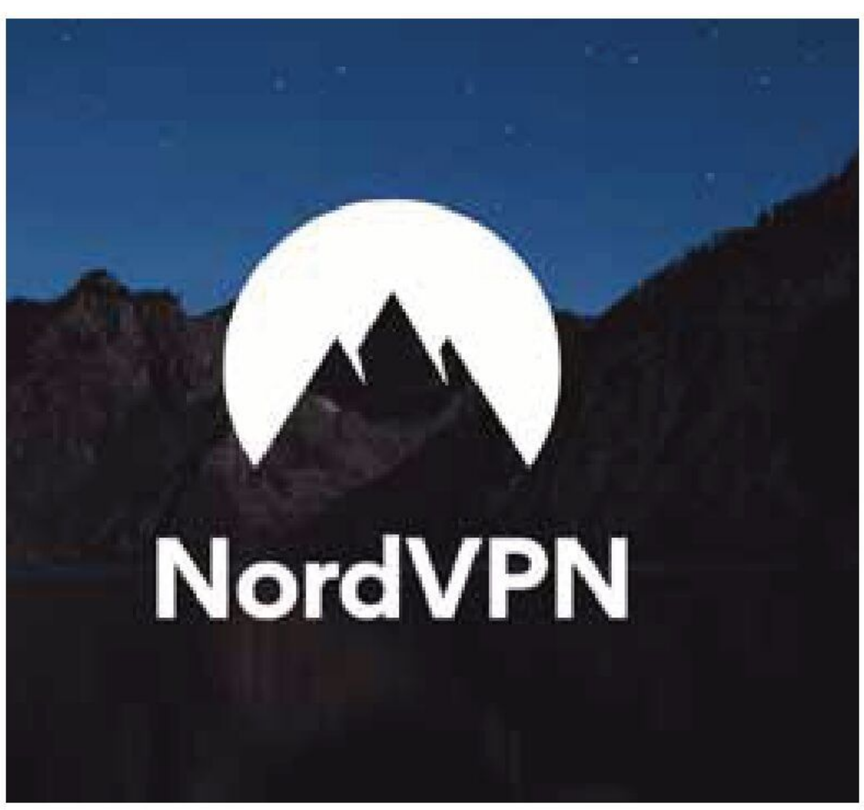
"We estimate, based on analytics from installs and public information, that Diffgram has now been installed on more systems than the nearest competitor, Labelbox. Considering that each installation of Diffgram can service hundreds, and sometimes even thousands of users, this is a substantial presence," reads a company statement.

There are many software packages in this general space that are open source. The Diffgram team said that they are the first to open source the entire platform.

NordVPN releases VPN speed testing tool

NordVPN, a virtual private network (VPN) service provider, has created an open source speed-test application that allows customers to objectively evaluate the speeds of various VPN providers.

The tool, according to the company, is designed to overcome the shortcomings of traditional VPN testing tools, such as their lack of transparency and universal application. It provides a single, consistent speed-testing approach for reliably measuring a VPN source's network throughput.



The GitHub-hosted solution provides detailed test reports that include the number of tests per hour, median download speed results, the arithmetic mean, and median winner findings. For a more realistic comparison, a VPN performance test for three days with at least one test each hour is recommended.

"Lack of a unified approach towards how performance is measured leads to the industry's current situation: different researchers via different outlets showcase different results ranked in different order," explained a NordVPN spokesperson. "This conflicting information is bad for VPN users and providers alike. The confusion can make it difficult for users to make informed decisions."

Google proposes a public-private collaboration to secure open source projects in the US

Google has urged the US government to be more proactive in identifying and protecting vital open source initiatives for Internet security. In a blog post released after the White House's Log4j vulnerability summit, Kent Walker, president of global affairs and chief legal officer at Google and Alphabet, said the US needs a public-private collaboration to properly fund and staff the most important open source projects.

Oxeye releases open source payload deobfuscation tool

Oxeye, a technology innovator in cloud-native application security testing solutions, has unveiled its first 2022 open source initiative with the introduction of Ox4Shell. This powerful deobfuscation tool is the first in a series of solutions to be developed by Oxeye to assist developers, AppSec professionals, and the open source community. Ox4Shell is designed to confront what some are calling the 'Covid of the Internet', known as the Log4Shell zero-day vulnerability. To counter a very effective obfuscation tactic used by malicious actors, Oxeye's new open source tool (available on GitHub) exposes hidden payloads that are actively being used to confuse security protection tools and security teams.

According to Jonathan Care, senior director analyst at Gartner, "The Log4j vulnerability is extremely widespread and can affect enterprise applications, embedded systems, and their sub-components. Java based applications including Cisco Webex, Minecraft, and FileZilla FTP are all examples of affected programs, but this is by no means an exhaustive list. The vulnerability even affects the Mars 2020 helicopter mission, Ingenuity, which makes use of Apache Log4j for event logging."

The new tool is the first in a series of contributions by Oxeye designed to strengthen security efforts by deobfuscating payloads often coupled with Log4J exploits. Ox4Shell exposes obscured payloads and transforms them into more meaningful forms to provide a clear understanding of what threat actors are trying to achieve. This allows concerned parties to take immediate action and resolve the vulnerability.

Progress announces support for HTTP/2 in Fiddler Everywhere 3.0

Progress, a provider of products to develop, deploy and manage high-impact applications, has announced the latest release of Progress Fiddler Everywhere, its Web debugging proxy tool.

With this release, Progress is bringing to market support for HTTP/2, a major revision of the HTTP network protocol. "Developers can now easily debug high-performing applications across Windows, MacOS and Linux securely in less time and with fewer resources than ever before," the company said in a release.

Fiddler Everywhere 3.0 enables organisations to capture all HTTP(S) traffic between a computer and the Internet from any browser and on any device. With its rich functionality, including network session analysis, debugging in production and capturing Web traffic logs, developers can trace and solve HTTP(S) issues and conduct performance evaluations. They can easily simulate external requests and responses without changing the code of the website or the mobile app.

According to Progress, Fiddler Everywhere 3.0 also provides support for WebSocket communication – developers have access to information transferred between the client and the server through a single ongoing connection. They can inspect data sent back and forth in chat applications or live update websites.

It provides improved advanced traffic filters – developers can easily apply complex filtering to their sessions list and choose from Protocol, Status, IP, URL, Cookie, Body and other values in addition to the Request and Response Headers. They can build one or multiple complex conditions that will filter the data based on their requirements.

"For too long, the software community has taken comfort in the assumption that open source software is generally secure due to its transparency and the assumption that 'many eyes' were watching to detect and resolve problems," he said. "But, in fact, while some projects do have many eyes on them, others have few or none at all."



According to Walker, the partnership would assess a project's influence and importance in order to evaluate how important it is to the ecosystem as a whole. He said the industry will require new methods for identifying software that could represent a systemic risk to Internet security in the future.

More public and private money was also needed, according to Walker, who added that Google is willing to contribute to an organisation that matches volunteers from firms like its own to essential initiatives in need of assistance. "Open source software is a connective tissue for much of the online world — it deserves the same focus and funding we give to our roads and bridges," he said.

Autoware releases Open AD Kit for autonomous EVs

The Open AD Kit is a reference framework for developing software defined vehicle (SDV) solutions for commercial autonomous driving. Autoware's description of the release stated: "The initial release of the Open AD Kit includes cloud and edge solutions from Adlink, ARM, AutoCore, AWS, SUSE, and Tier IV, as well as the implementation of Autoware supporting the containerised microservices architecture."



A wide range of organisations, from tech schools and universities to race teams and OEMs, are incorporating Autoware's AI into their ADAS development projects. Because it's open source, any gains or enhancements made by one party can benefit everyone else.

Autoware Foundation's open source AI platform, founded in 2015 by Shinpei Kato of Nagoya University, is supported by one of the largest autonomous driving open source communities, with 2,300 stars on GitHub and over 500 accounts (claimed) on Slack.

Open Source Institute to polish open source skills

The Open Source Institute (OpenSI), founded by the University of Canberra and Instaclustr, is an effort to improve open source capabilities in the IT sector. 'Developer Training and Certification for Apache Kafka' will be the first online

course it will offer. This course will teach software engineers and system administrators how to develop and use Apache Kafka.

Interactive workshops, webinars, online forums, self-paced assignments, and evaluations will all be part of the course. There will also be an optional exam to complete in order to get OpenSI certification for Apache Kafka. Other open source technologies, like Apache Cassandra, PostgreSQL, and Cadence, will be covered in future.

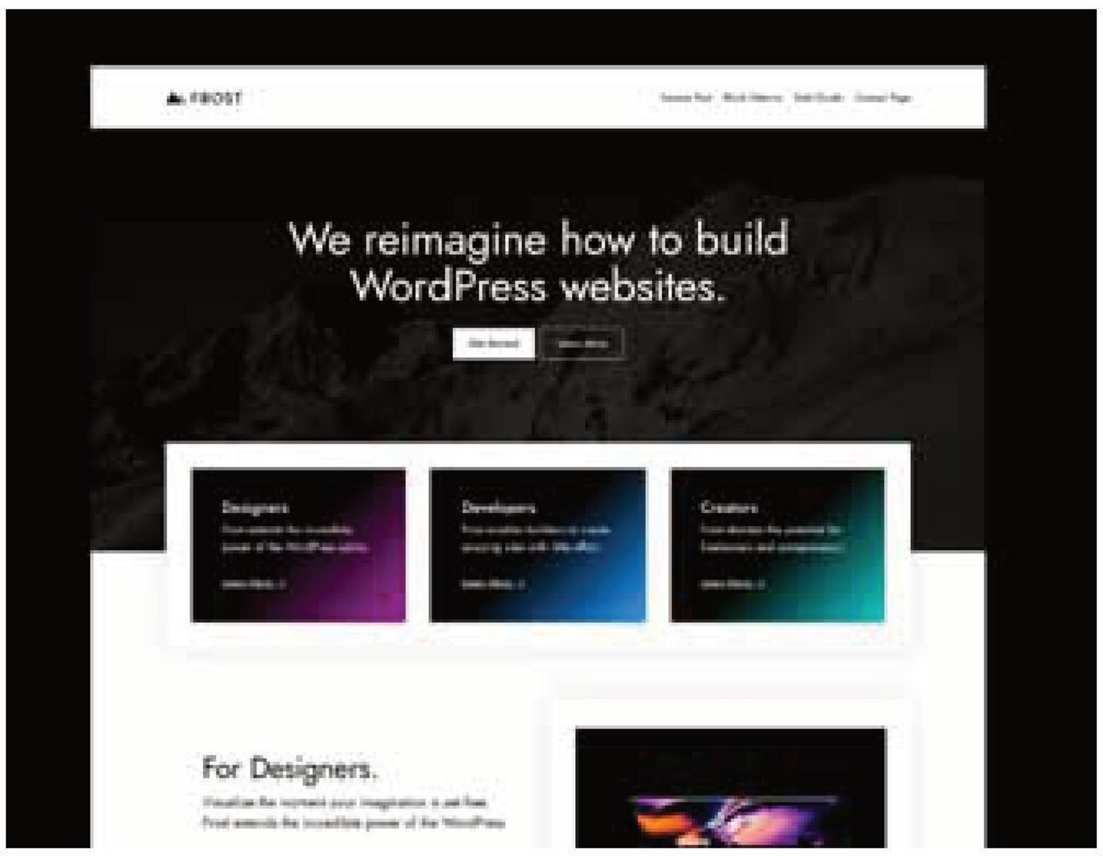


In terms of technology and digital literacy abilities, there is a large skills gap that needs to be filled. Over the next decade, technology jobs will be on the rise. Instaclustr and The University of Canberra acknowledge the importance of building these talents for a company's success, and will empower IT professionals with highly sought-after job-ready abilities.

Students will have access to experience, practical skills, and academic oversight through OpenSI, giving them a competitive advantage when they enter the workforce.

WP Engine acquires Frost, the WordPress editing theme

WP Engine has acquired Frost, the WordPress theme released by Brian Gardner early last year as the main product of a new startup business. It will now be a freely available project, with a focus on full-site editing. Frost enables builders to create sites with little effort, and elevates the opportunities for freelancers and entrepreneurs.



"When I joined WP Engine in late September, it was evident to me [WP Engine] saw the value in building relationships with designers, developers, and creators — within our Atlas product line, as well as with WordPress. While there was interest in Frost when I came on board, it wasn't until my vision for our team became clear did an acquisition enter the picture. I recommended bringing Frost into the company and hiring Nick Diego," said Gardner in a report in WP Tavern.

Diego is the creator of the 'Block Visibility' and 'Icon Block' plugins. He also began working on Frost in late September last year.

MultiPay Group assigns its database management to Percona

Percona, a provider of open source database software and services, has announced that MultiPay Group, a global payments technology company, has signed with the firm to provide managed services for its MySQL open source database deployments.

Delia Pedersoli, chief operating officer at MultiPay, said: "Assigning our database management to Percona will enable us to realise greater efficiencies as we continue to grow our technology offering and better serve consumers and merchants with fast, seamless payment experiences. We support global retail customers with their critical payment infrastructure and we plan ahead so that we can always deliver what our customers need. Percona offers that round the clock expert support to help us ensure our databases are running at peak performance, and we can rely on Percona's unbiased approach as we consider any future deployments too."



Martin James, vice president EMEA at Percona, said: "The e-commerce marketplace is competitive and customers aren't willing to put up with poor performance. MultiPay has built a unique offering that revolutionises the payment experience for consumers, and their service must keep up with what they've promised. Alongside running Percona XtraDB Cluster for availability, Percona provides 24/7/365 expertise and support to ensure MultiPay's database performance is aligned with their business goals."

For more news, visit *www.opensourceforu.com*

Let the Domain Drive the Design



Object-oriented design, when driven by the domain, decomposes huge monolith systems into a set of agile microservices. Domain-driven design has changed the narration of the object-oriented approach upside down. It has brought the domain objects onto the centrestage of design thinking, and relegated the technology to the peripheral infrastructure. In this fourth part of the 'The Design Odyssey' series, let us examine this approach to software design in some detail.

The Socratic method of learning involves continual probing and cooperative dialogue that is laced with argument. We go by that method to understand the concepts and patterns of domain-driven design. Let's go!

Domain-driven design: Concepts and patterns

What is domain-driven design?

Domain-driven design, or DDD in short, is an approach to designing software systems by applying object-oriented principles in their true spirit.

It implies that DDD is not an alternative to object-oriented design; instead, it only complements it. However, the basic principles like abstraction, encapsulation, inheritance, and polymorphism along with the celebrated SOLID principles already offer a blueprint for object-oriented design. So what else is offered by DDD?

It is true that there are several principles and patterns already in vogue. However, in practice, people often flout most of these principles and patterns, unintentionally. Let's take an example. The principle of encapsulation clearly says that the *state properties* of an object and the *behavioural methods* on those properties must be grouped together within the object itself in such a way that no outsider can directly access the *state* of the object. However, most of the objects we design either contain only state properties or only behavioural methods. The DDD attempts to correct it.

We often design entities and service objects with different responsibilities. An entity consists of just properties that are mapped to the fields of database records. And a service object holds the business logic to be applied to the

entities. In fact, JPA and Spring Framework on the Java platform offer two popular annotations @Entity and @Service to clearly mark the classes. And this practice is working. What is wrong with it?

Neither the Spring Framework nor the annotations are wrong. It is the practice of thinking from the perspective of database design that is wrong. Indirectly, we are designing the objects depending on the database technology that we choose. We freeze the database architecture, the schema, and the queries early in the design. The problem with this approach is that the rest of the system is designed at the mercy of what is offered by the database layer. Everything is retrofitted into the frozen database layer. It makes the system too rigid for a change. And, as you know, any system that is not agile enough to absorb the change obsoletes quickly.

The buck does not stop even there. The project teams often choose a platform, besides the database, before designing the objects in the business layer. For instance, the way a Python developer designs the objects is different from the way a Java developer designs them. Another disturbing observation is that even on a single platform like Java, the objects designed for the Jersey framework look completely different from the objects designed for the Spring Framework. Objects are supposed to reflect real-world entities, isn't it! However, in practice, it is the choice of database, platform, or frameworks that is driving the design. This is wrong because a rigid monolith system cannot respond to changes in business requirements. The DDD attempts to put the domain in the front seat, not the technology.

This looks reasonable. However, most of the people who are involved in software development are more technology experts, not domain experts. Isn't this an obstacle to adopting DDD?

This is one of the anomalies in the industry. The teams that gather the system requirements are often ignorant of the technology, and the teams that design and build the system solutions are ignorant of the business domain. It is a common practice for a young engineer to jump from one team to another for career growth. This often results in jumping from one domain to another domain itself. It is no surprise to see a designer moving from the financial domain to the e-commerce domain. Though this gives exposure to various domains, the person will not gain sufficient expertise in any of the domains. Ultimately, such a designer gets carried away by the technology he or she knows, rather than by the nuances of the actual domain to which the system belongs.

The DDD strongly vouches for domain expertise. The overall design of the system, granular design of the objects, and their interactions should reflect the corresponding domain. How the teams gain such expertise is left to the individuals and the organisations.

Still, it is inevitable to have separate teams for product marketing and engineering. Does DDD have any solution to iron out the communication gap between the teams?

The DDD suggests that the engineering team should refrain from technical jargon. Instead, the teams are advised to develop a consistent vocabulary with a strong affinity to the domain. The vocabulary is referred to as ubiquitous language. All the stakeholders of the team are expected to use the same language. Without such a language, a specific word means many things to many people.

For instance, an e-commerce system consists of a *subscriber* object, a database table named *customer*, and a label on the user interface that reads as the *user*. Though these look as if they are three different entities from the outset, they may all actually refer to just one entity. Instead, all the stakeholders should stick only to one word. And that word should come from the business. If that specific e-commerce company refers to them as *customers*, then the word *customer* should be chosen, nothing else. It is even better to maintain a dictionary of the vocabulary.

In summary, DDD advises modelling the objects that correctly map to the actual business domain and using the language that the business uses. In a way, it fits into the Onion Model. Isn't it?

The Onion Model consists of a domain layer, service layer, application layer, and infrastructure layer. The domain layer comprises technology-agnostic objects. The design of this layer is purely driven by the domain, nothing else. A domain object consists of both *state* and *behaviour*. However, there are places where more than one domain object needs to be involved in a business workflow. The service layer consists of objects that deal with more than one such domain object. The application layer acts as the gateway to the rest of the world. This is the layer where requests and responses are handled. The infrastructure layer is largely specific to the technology. It consists of databases, messaging systems, test environment, user interface, etc. It is natural that the DDD approach and the Onion Model go hand-in-hand.

However, DDD is not just limited to it. The DDD strongly advises decomposing a monolith domain into smaller sub-domains, and to deal with each of the sub-domains independently. In other words, the Onion Model is applied for each of the sub-domains.

If the sub-domains are handled independently, how are the domain objects that belong to more than one sub-domain handled? Are the objects duplicated or is the effort duplicated?

There are different strategies to handle such scenarios. However, it all boils down to the way the domain is decomposed into sub-domains. Often, the sub-domains are aligned to the lines of business. For instance, the catalogue service, the order service, and the fulfilment service are some of the probable sub-domains of an e-commerce domain. There are also secondary sub-domains that are not directly connected to the business, but are important, such as accounts. And we have sub-domains that are technology-specific, like messaging systems. A careful demarcation of the sub-domain boundaries is vital to get some benefit out of DDD.

Who does the domain decomposition? It looks as if it is beyond the scope of a software designer.

Yes. The domain decomposition is the responsibility of the management team. The senior solution architects are usually part of such a management team. Once the decomposition is finalised and sub-domains are identified, the engineering teams take over the responsibility to model and implement them. Speaking technically, instead of delivering one single monolith enterprise system, the engineering department delivers a set of smaller services which are now popularly known as microservices.

Does this mean that there will be no overlap between two sub-domains?

The degree of overlap between sub-domains is directly related to the nature of the business and the nature of decomposition. At one extreme, there is no overlap, which is desired. At another extreme, there is a lot of overlap that defeats the very purpose of decomposition. Both are very rare. Often, we see sub-domains with a small degree of overlap. Still, teams are advised to deal with their sub-domains independent of others, even if it results in duplicate effort. In the long run, a set of loosely coupled services are more beneficial.

Other than high-level concepts and best practices, does DDD offer any handy patterns that are easy to follow and replicate?

Of course, yes! Entities, value objects (VOs), data transfer objects (DTOs), aggregates, and repositories are very popular though a good number of people may not really be aware of the fact that they are DDD patterns. An entity is a domain object that can be uniquely identified, whereas a VO is a domain object that does not have a unique identity. A DTO is an object that is transferred across the domain layer. An aggregate is a cluster of domain objects with a root, and a repository is an abstraction of the aggregate store.

Is the entity of DDD the same as the entity of ORM, E-R models, etc?

They are not exactly the same. An entity in DDD is not necessarily associated with any persistence. They may also exist only during runtime. The point is that any domain object that is uniquely identified is called an entity. In other words, the state of an entity can change over a period of time, but it can still be identified uniquely. For example, an *order* in an e-commerce system is identified by something like *order-id*, though the value of the order and status of the order keeps changing.

An employee with an employee-id, a customer with a customer-id, an account with account-number, and a book with ISBN are all entities, then! They may or may not be persisted. When persisted, they may be mapped to a record in the RDBMS, or a document in NoSQL, etc. Entities that have relations with other entities form aggregates. Can the entities also have behaviour?

Of course, they can! Entities are first-class domain objects. They can have *behaviour* besides the *state*. For example, an *order* object may have operations like *applyDiscount(), cancel(), deliver()*, etc. For many, this looks a bit odd since somehow they are under the impression that an entity is just a database record without any operations. And that is a misplaced understanding.

Well, a domain object with a unique identity is an entity, and a domain object without a unique identity is a value object. Is that right?

That's not the way to look at it. A value object is identified by looking at the whole of the object, not just one property. For example, two event objects in a UI system are indistinguishable if the type, time, and coordinates of the event are exactly the same. The value objects must be immutable, since any change in the value of the object makes it unrecognisable. Usually, it is encouraged to model properties of primitive types like phone numbers, email, addresses, etc, as value objects. Like the entities, the VOs also may or may not be persisted.

How are the entities and value objects persisted? Do they use data access objects?

One of the hallmarks of the DDD is that the domain model is technology-agnostic. In other words, the objects in the domain layer do not really know about the outside world. Even the storing and retrieving of the domain objects is done through an abstraction called repository. A repository gives an impression as if it is in-memory storage of the domain objects. Some of the usual methods of a repository include *save(), find(), update(), delete()*, etc.

Unlike DTOs, which are more inclined towards databases, repositories may be backed by anything such as databases, file systems, REST endpoints, etc. A notable

EFYGROUP
YOURS SINCE 1969

Wanna Support Open Source For You?

Subscribe to the Magazine— so that we can keep promoting Open Source.



AMAZING OFFER

Pay for 6 Issues Get 12!

(Buy One Get One Free!)

WORLD'S LEADING PUBLICATION ON OPEN SOURCE

TEAR OUT — **ORDER FORM** — TEAR OUT

Please Tick (✓) Your Choice

☐ **Pay for 12 Issues Get 24 Issues**
and save 50%
(Buy 12 Issues Get 12 Issues Free!)
₹1440

☐ **Pay for 24 Issues Get 48 Issues**
and save 50%
(Buy 24 Issues Get 24 Issue Free!)
₹2880

☐ **Pay for 36 Issues Get 72 Issues**
and save 50%
(Buy 36 Issues Get 36 Issue Free!)
₹4320

To subscribe online, visit
https://tinyurl.com/y5kuv4la

OR SCAN THIS CODE

Name_____Organisation_____Mailing Address_____

_____City_____

Pin Code_____State_____Phone No._____ Email_____

Subscription No. (for existing subscribers only)_____. **I would like to subscribe to the above (✓)marked Open Source For You magazine starting with the next issue.** Please find enclosed a sum of Rs_____by DD/MO/crossed cheque bearing the No._____ dt._____in favour of **EFY Enterprises Pvt Ltd,** payable at Delhi. *(Please add Rs 50 on non-metro cheque)*

Please mark one (nearest) relating to your subscription: ☐ **Indian Company** ☐ **MNC** ☐ **R&D organisation** ☐ **Engineering institute** ☐ **College/School** ☐ **Any other (specify):**_____

Send this filled-in form or its photocopy to : **EFY Enterprises Pvt Ltd,** D-87/1 Okhla Industrial Area, Phase 1, New Delhi 110 020 | Ph: **011-40596600** | e-mail: **support@efy.in**

Terms:- # These rates are applicable for new subscribers as well as renewal by existing subscribers. # Can access ezine till your subscription is active # The rates are valid for subscribers within India only. # Please allow 4-6 weeks for processing of your subscription. # The subscription copies will be dispatched through ordinary post only # Subscription Agents will not get agency commission against this scheme # Disputes, if any, are subject to exclusive jurisdiction of competent courts and forums in Delhi/New Delhi only. * Replacement will be made if intimation of damaged / non-receipt of copies is received within 30 days of its publication ** After three months, if you are not satisfied with the magazine, your balance amount will be returned (Not applicable for gift offer)

point is that the repository interface belongs to the domain layer whereas its implementation belongs to infrastructure. Anything that is related to the infrastructure layer is expected to be pluggable.

And finally, the data transfer objects, as the name suggests, are special kinds of objects that are meant for request and response payloads. They do not belong to the domain layer; instead, they are useful in the application layer. But, what is the need for DTOs? Why can't entities and VOs be used for transferring the data?

The reason is self-explanatory. The entities and value objects are domain objects. A domain object is useful and relevant within its own domain or sub-domain. Exposing the domain model outside of the domain creates a needless coupling between the domain and the outside world. For instance, the name of a customer might be modelled as a value object within the domain model with *FirstName* and *LastName* as the fields. However, for the outsiders it is immaterial. They may see the name as just a character sequence. A DTO does the job of converting a name object into a string and vice versa.

Illustration

Let us look at some examples to understand the domain-driven design patterns. These examples have nothing to do with any specific programming language. Consider them as a pseudo-language to convey the idea.

Following are the classes *Order* and *OrderManagement*:

```
class Order {
  private double amount;
  private Date date;
  private String status;
  ....
  getAmount();
  getDate();
  getStatus();

  setAmount(double);
  setDate(Date date);
  setStatus(String status);
}

class OrderManagement {
  public void cancel(Order order) {
    apply business rules;
    order.setStatus("cancelled");
  }
  ....
}
```

The problem with the above design is that the *status* property of *Order* is being manipulated by an outsider *OrderManagement*. Where is the encapsulation? This is more like global variables and global functions with a bit of sugar-coating in the form of setters and getters. The status of the *Order* must be handled by the *Order* itself since it is its state property. The following is a better version of the design:

```
class Order {
  private double amount;
  private Date date;
  private String status;
  ....
  getAmount();
  getDate();
  getStatus();

  setAmount(double);
  setDate(Date date);

  public void cancel() {
    apply business rules
    status = "cancelled";
  }
}
```

The above *Order* class does not expose any direct setters. However, it should not be understood as if setters are bad. The object is still mutable, but through more meaningful methods like *cancel()*.

The next enhancement to the above class is to add an identity. Remember that it is not a technically-driven design decision. In the e-commerce domain, every order is given a unique identity. We are merely mapping to the real world.

```
class Order {
  private int id;
  private double amount;
  private Date date;
  private String status;
  ....
  getId();
  getAmount();
  getDate();
  getStatus();

  setId(int id);
  setAmount(double);
  setDate(Date date);

  public void cancel() {
```

```
    apply business rules
    status = "cancelled";
  }
}
```

A blunder in the above design is that the property ID is made mutable. Identities are never mutable. The *setId()* method is meaningless. The *Order*, being an entity, must be associated with any identity at the time of creation and must never be changed after that.

```
class Order {
  private int id;
  private double amount;
  private Date date;
  private String status;
  ....
  getId();
  getAmount();
  getDate();
  getStatus();

  setAmount(double);
  setDate(Date date);

  constructor() {
    generate and assign id;
  }

  constructor(int id) {
    assign id;
  }
  public void cancel() {
    apply business rules
    status = "cancelled";
  }
}
```

So far, so good. The next betterment is to replace the primitive types with value objects wherever they are appropriate. For instance, the above *Order* object accepts any double as the order amount. How is that acceptable? Isn't it reasonable to say that the order amount must be above zero? Let's refactor.

```
class Money {
  private double value;
  getValue();
  setValue(double);
}
```

```
class Order {
   private int id;
   private Money amount;
   ...
}
```

This looks a bit more reasonable. However, we again committed a blunder. Isn't *Money* a value object? Can a value object be modelled as mutable? No, right? Here is the correct version:

```
class Money {
  private double value;
  getValue();
  constructor(value) {
   validate the value;
   assign value;
  }
}
```

And, finally, a repository for *Order*:

```
interface OrderRepository {
  save(order)
  load(id);
  delete(id);
  loadAll();
  ...
}
```

There can be several implementations like *JpaOrderRepostory*, *MongoOrderRepository*, *RedisOrderRepository*, and so on.

UMS as microservices

The user management system that we have been exploring in the last three parts of this series of articles is a single monolith system that offers multiple features. By applying the principles of domain-driven design, it can be decomposed into a set of microservices. And that is what we are going to see in the next part. END

By: Krishna Mohan Koyya

The author is the founder and the principal technology consultant at Glarimy Technology Services, Bengaluru. He has delivered more than 500 programs to corporate clients on design and architecture in the last 12 years. Prior to that, he held various positions in the industry and academia for a decade.

Building Profitable Open Source Business Models



The world is moving towards deploying more open source solutions, but are there any proven business models that can help open source companies sustain themselves? Should new companies venture into the world of open source? This article tries to answer these questions.

Open source software, about a decade back, was looked upon as something solely meant for geeks, programmers and educational institutions. But a few years ago, it started gaining recognition as software that every organisation could leverage. Today, even large proprietary organisations like Microsoft are openly talking about it, and investing in the same as well!

"A lot has happened and evolved in open source. The top view shows it is free. However, even if it is being made freely available, there are billions of dollars riding on the ecosystem. Not only are there open source companies valued at more than a billion dollars, industry giants like Google and Microsoft are investing millions of dollars in it," says Mukul Mahajan, MD, Tetra Information Services Pvt Ltd.

Most of these investments focus on hiring manpower for the open source ecosystem. Business models like professional services and cloud hosting are evolving, but the big question is: what are the opportunities available to organisations that aim to, or are already working on, being an open source player? And what are the business models they can pursue for profits?

"I think the key aspect is to have a business model. But here's the thing — the vast majority of people or organisations who are investing in open source are not investing for the sake of it. I don't think there is something directly attracting them to invest in open source," says Hadi Hariri, VP, developer advocacy, JetBrains.

He adds, "They are not investing to make a better world using open source. There are key indicators that force them to use open source, which is good. But they need a sustainable business model, which unfortunately many do not have."

Cue: Microsoft'shere!

We have all used software developed by Microsoft, and the biggest cue that open source is indeed the future, can be derived from what this organisation is doing on the open source front. But why is Microsoft in open source?

"Satya (Nadella) was very clear when he took over Microsoft that it has to be more open than it ever was. This actually goes back to our core thinking of wanting to do more for every person and organisation, which cannot be achieved by being closed source," says Sanjeev Sharma, director, new ISV partnerships, Microsoft.

He adds, "Today, 40 per cent of CIOs have Linux as one of their prime strategies, 60 per cent of developers want to see cloud providers working with open source, and 90 per cent of enterprises deploy Linux in some shape and form. This indicates an absolutely clear market decision."

Sharma feels there are tremendous opportunities around open source, and affirms that Microsoft is committed to staying open in the future. Fifty per cent

of Azure's VM cores and 60 per cent of the images on it are Linux based. In fact, more than half of the ecosystem at Microsoft is based on Linux. SaaS and various other applications are leveraging open source too.

He says, "Any new offering coming to the market can continue to leverage its core DNA while also getting exposed to the larger market worldwide. This gives an opportunity to experience the full might of larger cloud providers."

Solving a problem using open source can help solve many other problems. For example, a normal customer service app has to touch different platforms, and be compatible with different types of devices. It may also have to browse through multiple databases. The idea, says Sharma, is to slowly and steadily multiply what is being offered to clients.

"You begin by offering one solution to one type of customer base and then you can pivot to a larger market. Today's cloud provides you that opportunity to experiment," he adds.

Microsoft had acquired GitHub for US$ 7.5 billion back in 2018. Then, in March 2021, it announced a grant of Rs 10 million to fund India's developers who are building software using open source tools.

Subsidise and maximise

It is pretty clear that the open source business model of tech giants like Microsoft, Amazon and Google is subsidised by the cloud services they sell. Facebook, as per Hariri, is subsidised by data analytics and advertising. Organisations like JetBrains subsidise the business of open source by selling tools.

"I don't think there are any key parameters. You should look at how you are going to sustain your open source contributions. The best example that can be thought of here is Red Hat," explains Hariri.

Picking the right open source business model

"You can always look at it as a half-full or half-empty glass. We can leverage a larger ecosystem. We can reach out to companies all over the world by leveraging the public clouds."

—Sanjeev Sharma, director, new ISV partnerships, Microsoft

"Focus more on the pain of the consumer, and then make a business model that can help you survive. Times are changing rapidly. It's good to be on the dynamic side, as these changes will help you survive in the long run."

—Samir Doshi, co-founder and director at iNextrix Technologies

"It's important to keep figuring out things on the way. It's critical to always remain agile and flexible. Everyone should nurture their curiosity and keep testing waters to see what works best for them. The market is optimistic; so why shouldn't you be."

—Kabir Chandhoke, COO, SourceFuse Technologies

"If you want to leverage open source, you should first mull over your decision. Think about whether open source will provide any benefits or not. Once you are certain, make sure you develop a business model around it. If you want to make it successful, make sure you have a licence that is not overly permissive."

—Hadi Hariri, VP, developer advocacy, JetBrains

The Red Hat business model, as most industry *gurus* will agree, is complicated yet successful enough to require consulting, which is provided by the company team as part of the model. This business model may become less sustainable if the consulting is taken care of by a third party.

Open source, as Mahajan points out, might be more of an enabler that helps get traction and build newer business models. What helps new business models based on open source to sustain is the demand they get from different business verticals.

"When running a business, sometimes you need to subsidise what you are doing for the greater good," says Kabir Chandhoke, COO, SourceFuse Technologies.

Marketplace, enterprise and community software; and the role of the cloud

Most open source business models in use today are offered as enterprise and community editions. The community edition provides free usage of open source with limited capabilities, whereas the enterprise edition offers a lot more features at a cost. "These two models are becoming very popular. On top of these, you can also offer complimentary services such as certification, development services and support contracts," says Samir Doshi, co-founder and director at iNextrix Technologies.

But it's open source marketplace software that joins all the dots together.

"Marketplace is definitely a brilliant business model, as it adds a lot of value to all the open source products. You have brilliant brains coming up with add-ons and applications," says Doshi.

It is critical to understand the relationship between open source and the cloud in order to understand business models better. The debate around whether the cloud is making open source popular, or vice-versa, is something that experts still lock horns on.

"Open source is what actually drove the cloud in the first place. Can you imagine the public clouds we have today and companies paying licence fees for their operating systems?" asks Chandhoke. "We would never have

achieved the scale of, say, AWS, if licence fees for millions and millions of servers had to be rolled out."

He adds, "The reverse is also true as the public cloud is responsible for making open source projects popular. Another way public cloud has accelerated the monetisation of open source is through the SaaS route."

Cloud and open source can enjoy a mutually beneficial relationship as long as the former does not start eating into the latter's business model. This happens when an open source product becomes tremendously popular. "I definitely see the partnership between the cloud and open source evolving. It will be interesting to see how this happens," says Chandhoke.

Elasticsearch needs no introduction. This search engine, based on the Lucene library, has been the go-to favourite of many open source developers. But, in early 2021, the company came out with the groundbreaking Elastic License 2.0.

The licence allows users the right to use, modify, create derivative works, and redistribute software. However, there are 'three limitations':

1. Users may not provide products to others as a managed service.
2. Users may not circumvent the licence key functionality or remove/obscure features protected by licence keys.
3. Users may not remove or obscure any licensing, copyright or other notices (as published by Elastic.co).

This means that both Elastic License 2.0 and SSPL have not been approved by the OSI. Hence, neither Elasticsearch nor Kibana are now considered open source. This is an example of how fast the world of open source is evolving, and how quickly existing business models are shrinking while new ones are making their way!

"This has impacted the community in two ways. The first is that contributors were suddenly told that what they have worked for is no longer truly open source. The second is that the solutions big and small players are developing cannot be made public now, keeping in mind the announcements made by Elasticsearch," says Chandhoke.

Partnerships with big players

The last decade has been all about collaborations and mergers in the world of technology, and the open source vertical is not an exception. To partner with a big player like Microsoft can be a big opportunity for startups and organisations already operating in the domain of open source.

Sharma says, "If you are a services provider, there are opportunities available with Microsoft as the company has a specialisation in Linux and Kubernetes."

There are many opportunities for ISVs too. Sharma gives an example of an assessment engine based on open source notes. "It scaled up very easily. However, the problem was finding customers beyond India. So they leveraged the 60+ data centres that Microsoft has."

Whether the solutions are for education, health, finance or any other industry, big players can not only help scale the offerings across geographies but can also help make these solutions more secure.

"The core of the solutions that Microsoft offers is actually coming via partners. We are only providing them the platform. The solutions are coming through organisations that are based around the world," explains Sharma. More than 60 per cent of the solutions being offered are based on open source!

Why do open source companies have high valuations?

Partnering with big names also brings forth a question: Why is the value of open source companies so high? The answer lies in the consumer base that many of these organisations have, and how brilliantly they are handling the complex layers of architecture.

"The conflict between Amazon and other companies is also one of the reasons for the high valuation. There still exist companies that believe in hosting everything on their own server rather than on the cloud," says Doshi.

Then, why do developers work for free? Doshi says, "Because you are using that tech for yourself. For example, if I am working on a project, I will want to use that work for myself as well. I may want to build my UI using React and many such things. Once you start doing that and notice a missing feature, or find a bug, you end up fixing it. And that creates a solution for the greater good."

"The same also allows young and experienced developers to polish their skills. It's the best CV you can have, and probably serves better than most internships. Where else will you have the opportunity to have your code checked by some of the best and brightest in the world," he adds.

Big players (the one's providing cloud services) have now also started facilitating open source stacks. Going by the trend, it looks as if every big company will soon join this race, which could lead to big changes in the open source ecosystem. But according to Doshi, "I don't think there is space for any major changes. The community will always be there, whether it is from the cloud provider or open source companies. I do not think there will be any threat to communities or business models as such." END

Note: *This article is based on a panel discussion held at Open Source India 2021.*

By: Mukul Yudhveer Singh

The author is a business editor at Open Source For You.

Why Open Source is a Serious Business Today



Open source software has been around for a few decades now, but has recently become a popular choice for many organisations, small and big. What has changed? Why is open source, which was looked upon as software to be used by people who did not want to or couldn't spend money, in demand today? Let's find out.

Open source has been defined by some as "technology that is run by great people, who do not care much about money and make cheap business solutions." That may no longer be true today. Open source is turning into a serious business, but does face a few challenges.

Testing times

The open source software industry is going through testing times due to technical changes, recruitment challenges, etc.

Ten years ago, new code was being written for retail, healthcare, education, banking, etc. Now that all the software has been written, it just needs a few enhancements. We don't actually develop new software anymore; rather, we enhance the software to get better output from it. While working on the product, testing teams that used to stay outside are now doing automation within the software code itself. Infrastructure, design, testing, documentation, and code reviews are all codes today.

A solution that fits all was the concept used for building big monolithic products earlier, but nowadays, the focus is on developing one solution to solve one problem perfectly, which fits with all the other solutions, so that together they offer value to customers. This is what API

based developments, services, and products are all about.

Whatever is being designed today is getting outdated in no time. For instance, if you had developed excellent software for a fax machine, you may need to follow up with software for a cordless machine, and then perhaps for a mobile phone. So whether you like it or not, your software is going to be destroyed and decommissioned sooner or later. The bottom line is that change is constant, and we have to be ready for it.

How can India be a powerhouse of open source?

- India has a large pool of developers. However, its contribution to open source is abysmally low for various reasons. Open source is not charity, and companies can ask their people to focus on it wherever it will be helpful. This is definitely not a waste of time, as there will be an outcome.
- Meetups should be considered for recruitment.
- Open source should be part of the college syllabus. All projects in engineering colleges can be developed on open source.
- Government welfare schemes can be made available through open source communities. Quality solutions can be provided without a lot of investment.

The key ingredients

If viewed as a recipe, there are three key ingredients needed to build good software.

Skill: Recruiters in India are having a tough time hiring people with the right skills. On the flip side, over 80 per cent of engineers are unemployed mainly because of the lack of new age technology skills. At the same time, the attrition rate for experienced people is now extremely high. The pandemic has created a huge demand for experienced talent, and employers are finding it hard to retain such people.

Long notice periods create additional problems as they give a candidate time to look around, get a job that pays better, and then come back and negotiate with his or her current employer. With a shortage of talent, the employer very often gives in, and offers a higher salary to retain the employee.

Then there is a proliferation of 3-month courses that offer placements, where there is no skill assessment involved. After the course, you suddenly become a geek because of the job you got, irrespective of what you have been taught. However, how far you succeed in this job is anyone's guess.

A few companies have set up their own colleges and hire these students after they graduate, so that their remuneration can be less. Even though this may be a good strategy from a business standpoint, the graduates may not be able to fit in the outside world.

Mindset: Let's take an example to understand this better. Vijay, a school kid extremely interested in computers and an experienced Linux admin, will not be able to find a job in the industry because he is still at school. Divya, an experienced quality analyst, even if extremely talented, will find it difficult to find a job once she takes a career break for personal reasons. Sekar, who is an excellent designer but cannot communicate well, may not find a perfect job as well. An organisation with the right mindset will gladly hire all three individuals and give them the right training.

The technical decisions made by a team will solely depend on the mindset of its members, and the future of any project depends on that.

Purpose: Each software design has a purpose, but the process of digital transformation can have a few challenges.

- Developers should not be disturbed while they are working. However, there is a delay in a lot of releases and, very often, teams are internally divided among themselves.
- If the progress of the project is slow and deadlines are not met, the team still works extra hours, based on the need. But these extra hours come with a cost.
- There is some internal resistance from the design team itself when the same people end up working together always.
- If technical decisions are made by non-technical people, it can lead to confusion.

Is open source perfect?

When it comes to choosing between writing your own application programming interface (API) and taking one from open source, as much as it entirely depends on the use case, open API is generally used for its security, good quality code, feature availability, extensibility, etc. So, is open source perfect in all areas? Does it have no disadvantages? Not at all. There should be a perfect balance between interfaces and open source, in order to make good software.

But first, the positives…

- ***Simplicity:*** The skillsets involved are simple and don't demand too many machines, infrastructure, etc. Whatever is available is used.
- ***Attracts talent:*** Open source attracts talent. Only the passionate opt for it.
- ***Developers are important people:*** Decisions in the open source domain are mostly made by developers.
- ***Remote work:*** Work in the open source domain is done remotely, which is a great boon during this

pandemic. Developers work on multiple open source projects without meeting anyone in person, and continue to innovate that way. Open source is naturally built for remote work.

- ***Availability of infrastructure:*** If an open source project is being developed, there are a lot of free resources offered.

And, here are the negatives…

- ***Security:*** With a lot of things moving forward, security can be an issue. Initiatives like Whitesource, etc, help in collective knowledge of open source security, which may be helpful later.
- ***Licensing issues:*** After what happened to Elasticsearch and MongoDB, licensing issues can create challenges.
- ***Documentation:*** Very often, documentation is poor and not made available to the community that wants to contribute to projects.
- ***Single owner components:*** People who work with Node.js might have faced this problem of depending on a single owner component. If the component provider does not support your project after some time, it could lead to a lot of problems.

Open source strategy is definitely not a cakewalk, and it's not safe to get into it blindly. It is better to consider the good and bad facets of open source, and make a navigation strategy before jumping into it.

'Sharing knowledge doubles it' and 'Great things are invented, but for them to run, it needs the SUM of great people'. With these two philosophies backing it, open source can be beneficial both for the organisation as well as the individual. It would be nice if more contributions were made from India. But for that to happen, skills need to be developed.

The earlier definition of open source does not hold true now. The right definition today is: 'Open source is run by great people who do care about money and make affordable business solutions'. END

This article is based on a talk by Sathish Kumar Thiyagarajan, Open Source Advocate and Co Organizer of Bangalore Open Source Java User Group, at OSI 2021.

By: Sharon Abhignya Katta

The author is a technology content and promotions executive at OSFY.



An EFY GROUP EVENT

5243
ATTENDEES

70+
SESSIONS HOSTED

100+
SPEAKERS

18
COUNTRIES REPRESENTED



18th Edition

OPEN SOURCE INDIA

VIRTUAL EVENT

Thank You

VISITORS, for making the event a huge success

SPEAKERS, for contributing your valuable thoughts

PARTNERS, for your support

Microsoft Azure | soda foundation | Google | aws

MySQL | HUAWEI | iTEXT | GitHub

SUSE | ENTREPRENEUR FIRST JOINEF.COM | TRUEFOUNDRY BY ENSEMBLE LABS | Application Library Engineering Group

openLooKeng | EDGE GALLERY | soda foundation meetupindia | BANGALORE CNCF MEETUP

www.OpenSourceIndia.in

For more details, call on +91-9667743666 or email us at info@opensourceindia.in



How to Build a Community and Why it Matters

Building and sustaining an open source community is certainly not a bed of roses. This article outlines a few important steps you can follow to do so.



The demand for open source communities is growing. Community development is like an adventure in the deep woods. You need to start with a lot of preparation, especially to face the unknown. Technology and tools continue to evolve, so you cannot rely on them completely. On your journey, you will find fellow travellers who support you, but you should be ready to deal with surprises.

Why build a community?

It's often thought that a community is needed to address problems or offer support to people. These are valid reasons, but building a thriving community will entirely depend on you as a leader. There are four common challenges that have to be overcome for people to start or join a community.

Why?: This is an important question that people have, as they feel community contribution yields no returns.

Awareness: There are a lot of communities that need a supporting hand, but there is not much awareness about them.

Where to start?: There are many who do not know where to start, despite wanting to do so.

How to start?: People are unsure of what role they can play and where their contributions can be helpful.

These challenges can be addressed if people realise the wide range of valuable skills they will gain by building and participating in an open source community. These range from exposure to new technology and improved communication skills, to development of strategic thinking and influencing large numbers. To build a community, it will be helpful to start discussions with volunteers, note down your purpose, explore opportunities, and connect with fellow teammates.

Five steps to help you build a community

1. **Research for a common interest:** The very first thing to do is carry out research, which can be done in a day, a week, or a month at the maximum. Do not spend too much time on it. Think of it as a need for a group to be formed, be it technical or non-technical. You can visit two places to figure out whether it would be worth your while to start or not. The first is *Meetup.com*, where there are many communities worldwide, not limited by location or technology. The second place is *Discord*. Even though it was started for gamers, it is

PIONEERS IN TECHNOLOGY.
PARTNERS IN SUCCESS.

EFY GROUP
YOURS SINCE 1969





Mr Ajwani is not the only one who owes his success to *Electronics For You*. The technology magazine that started off in 1969 is today, South Asia's most popular one. Covering the latest on emerging technologies, changing industry trends, evolving trade practices...

Not surprisingly, top decision makers and those high up on the corporate ladder in the electronics fraternity treat EFY as their Bible.

- **Over half a million readers**
- **India's largest selling technology magazine**
- **Caters to the entire electronics fraternity**

D-87/1, Okhla Industrial Area, Phase 1, New Delhi-20 • Ph.: 011-26810601/02/03 • Email: info@efy.in • www.electronicsforu.com

a vibrant site for non-gamers also.

2. **Start small:** Do not stop after the analysis. You can start a meet-up with a small group. Be aware, for instance, that the registrations for your first event may fall between 50-100, but the actual turnout may be just 10 people. Don't be disappointed by this. It is a pattern seen very often. We have seen a ratio of one person attending for every three people registering for an event. Hence, do not give up early. Conduct the sessions at least a couple of times, before you decide on whether the community build-up is working or not. The net takeaway is to put your thoughts in action and not just have an intention to build.
3. **Communicate for impact, participate for feedback:** It is important to develop the skills of communication and participation. A good communicator will create an impact that leads to action. For instance, let's say you are writing an email for people to register. A lengthy email on the benefits of the topic and a small link for registration may not help. You should begin with a big popup that contains the link for registration, along with an agenda repeating the process of registration. Also, work on the time of the day and the day of the week that may help you fetch more registrations. Though this will not guarantee people coming in, you will learn to communicate precisely over a period of time.
Participation is important too. For example, I once asked people about the topics they would like to be spoken about in the next three sessions. A lot of ideas poured in, but the same set of people did not attend the next session. People give feedback because you ask for it, and it does not mean they are committed to being a part of the community. As time passes, you will slowly see a group of people participating repeatedly, indicating they are committed to the community.
4. **Adapt to evolve:** You have to be nimble and adaptive. When the pandemic hit, we were forced to shift from offline sessions to remote interactions. We looked at options like Zoom, Slack, etc, to sail through it. So it is important for you to keep an eye on where your set of community members walk around or spend most of their time. With trials on apps like WhatsApp, Facebook and Telegram for our community, we figured out *Meetup.com* gained us 9,500 registrations spanning across 15 cities and eight countries approximately. We also found that the WhatsApp group is the most active after *Discord.* We have around 1,800 members in this group, where the maximum interaction happens.
5. **Grow and expand:** This is where you start thinking like an entrepreneur. You get an entry as a member of a credible community and run multiple meetups consistently every single month, say for 3-5 years. That's when people start to take the community seriously. You may even have a good number of fans and a great set of volunteers that may get you sponsorships. Credibility comes with numbers and fetches you a good amount of sponsors too. Hence, build and boost your numbers for others to talk about the community.
If you want to expand your community even further, you can think about spot leaders. As mentioned earlier, a few members will attend your meetups repeatedly and be active in various forums. They can be trained to be leaders, which will help when you cannot devote time as a community leader. These spot leaders can help with events and keep the community thriving. Just remember, the willingness and commitment of the persons is more important than their skill or expertise.

Once you get sponsors, you can think about the purpose of the community to decide on tie-ups. For example, at JavaScript, we are tying up with React India. Such tie-up strategies help you grow at a different scale, as long as your intentions match those of your partners.
So, if you are building a community and have plans to expand, you should also imbibe the skills of entrepreneurship. Understand your market, identify more places where you can attract people, boast your numbers, sell what you are doing, and create win-win scenarios.

What can you expect as a community leader?

Community leaders not only influence the behaviour of the community, but also that of the corporates. Microsoft, Google and Amazon are all seeking leaders for new positions and roles to influence the community. You could either be directly associated with them as an employer or contractor, or simply have a tie-up to get sponsorships for your events. The demand for community leaders will only increase in the next 5 to 10 years, and it can turn into a full-time vocation.

Community leadership has improved my strategic thinking and my ability to adapt to different situations. In spite of the highs and lows, a leader has the backing of a community of around 10,000 people. There is no single route to be followed in the journey to community leadership, and you may make a few wrong turnings on the way. Just remember that when you act, you are taking a step forward. END

The article is based on a talk given by Janardan Revuru, community leader, JavaScript Meetup, at Open Source India 2021.

By: Abbinaya Kuzhanthaivel
The author is assistant editor, OSFY.

Why and How to Become an Open Source Contributor



Open source is often associated with coding, giving the impression it's meant only for computer geeks. But anybody can become an open source contributor — even if they haven't written a line of code their entire life. Sounds interesting? Then read on...

The beginning of the New Year calls for bringing about a change in your life. It could be learning a new skill, becoming healthier, performing your best at work or studies, or anything fruitful that will help you develop as a person. And it would be great if that change could benefit others too. So here's an interesting proposition — become an open source contributor.

But why?

To which, I say why not? After all, it is exciting and will help you achieve your objectives. And most importantly, you don't need to be an expert in coding. You can contribute without possessing any prior coding experience.

Interested? Let's do it. But first, a bit of the basics.

The formal definition of open source software goes like this: Open source software is software whose *source code* is released under a *licence* that grants anyone the right to inspect, modify and distribute it for any purpose. The two important things to note in this definition are source code and licence.

Releasing a source code for free doesn't hinder you from selling the result of your creation. It's possible to earn money based on open source software and make a living out of it. Companies, including GitHub and MongoDB, to name a few, are doing just that. Nevertheless, there's a lot more to simply giving away all your software for free.

There's also a difference between free software (not as in free beer) and open source software. While free

software is part of the larger open source community, the opposite is untrue. Open source software allows everyone the right to inspect, modify and distribute its code. However, free software also forces you to *redistribute any changes you made to it* under the original licence.

Popular open source licences

There are essentially two different categories of licences. On the one hand, there are permissive licences, allowing you to make some modifications to the original software code or sell the modified version as you wish. You can also combine other open source codes and release the final result with little to no constraints, as long as you do not alter the original licences and copyrights. Some common examples of this license category are the MIT license (MIT), Apache License 2.0 (Apache 2.0) and 3-clause BSD license (BSD-3-Clause).

On the other hand, you'll find copyleft licences. The most commonly used one is the GNU General Public License (GPL). If you are including a code in your project that is under the GPL licence, then you must release the final modified version under GPL too. You cannot mix GPL code with any other proprietary source code. Some common examples of this licence category are GNU Lesser General Public License (LGPL) and Mozilla Public License 2.0 (MPL 2.0).

So it is the licence that separates the general open source community from the free software community.

Why should you contribute to open source?

You may not be aware, but open source software plays an essential part in the functioning of devices that you use daily such as PCs, smartphones, and even cars. Thanks to open source software, global tech has evolved at a rapid pace since 2000. Without this transformation, the startup ecosystem wouldn't have thrived.

A key benefit of taking the help of open source while starting a new project is that it saves time and energy spent reinventing everything from scratch. You can simply reuse some of the previous work released under an open source licence. By building upon it, you can create something new and innovative. With open source, you can focus on what's important for your work instead of going through the basics every time.

Another benefit is that it can open many new opportunities in your life and career by helping you become a part of several open knowledge communities out there. Open source is one of the most efficient ways to learn and hone your technical skills and knowledge, because you get to work on highly-technical projects.

You may have heard of open data for sharing data sets for machine learning model training and to access any previous research works. This 'open community' approach helps support various sectors to achieve their objectives. Imagine reviewing and sharing codes for such purposes while also learning to collaborate with very smart people around the globe. Sounds fun.

Speaking of community interaction, by collaborating with other project developers, you come across different people and ideas. And that helps you build your network.

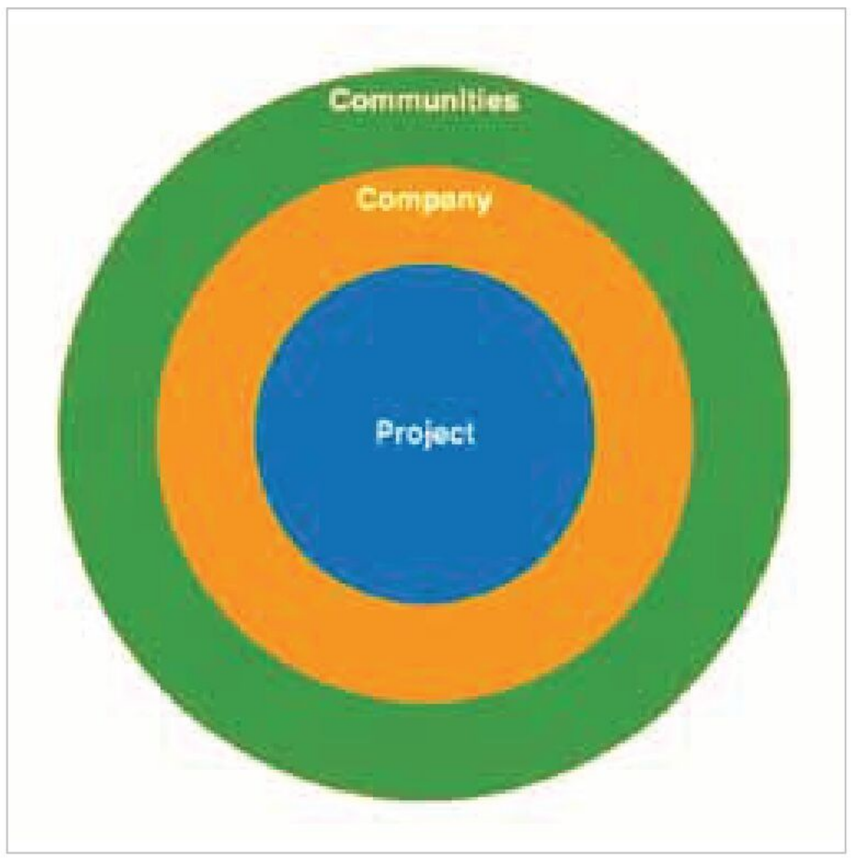


Figure 1: Circles of communication (Credit: OSI 2021)

Websites such as Bountysource and IssueHunt offer 'bounties' or compensations for fixing certain problems of various open source projects. This means you can make money by assigning yourself to specific issues and contributing to them. There are other websites as well where you can search for issues that have a bounty.

Far from the stereotype

Whenever someone is told to describe a developer, most people think of a person working alone with machines. But the truth is the opposite when it comes to real-life work. A big part of software development is communication. Yes, you have to write code for machines. But more importantly, you need to work with your peers and other developers.

Communication also entails getting information and feedback from end users, UX teams, product designers and everyone else directly involved in your project. Basically, a developer acts as a central point of communication of a project. All this can be regarded as the first circle of communication.

Outside the immediate circle, developers are expected to interact with different people outside their project or company, especially with bigger development or open source communities. Doing so exposes them to different opinions and constructive feedback, which is very important for improving and building key technical skills.

More contributions = More visibility

As a member of the open source community and contributor to open source projects, you will get more visibility. That means your professional profile will get more desirable views and requests from people looking to hire techies/developers.

EFY GROUP
YOURS SINCE 1969

Stay **Connected.**
Stay **Informed.**
Stay **Ahead.**



SUBSCRIBE AND **SAVE**

ORDER FORM

PRINT MAGAZINE	1 YEAR (12 copies each)	3 YEARS (36 copies each)	5 YEARS (60 copies each)
Electronics For You (Rs 100/copy)		**WITHIN INDIA (IN RUPEES)**	
	840 ☐	2150 ☐	3000 ☐
		SAARC COUNTRIES (IN US$ BY AIR)	
	50 ☐	135 ☐	NA
		OTHER COUNTRIES (IN US$ BY AIR)	
	100 ☐	270 ☐	NA

e-magazine subscriptions within India are available at half the rates mentioned here.
Overseas rates for each e-magazine in US$: 1 year: $12; 3 years: $33; 5 years: $50 only

PRINT MAGAZINE SUBSCRIBERS GET:
- Free e-magazine every month
- Free delivery of print magazine by post
- And much more (check: subscribe@efy.in)
- For delivery by courier, please add Rs 50 for each copy

To subscribe online, visit https://payment.efyindia.com | OR | **SCAN THIS CODE**

Name_____ Organisation_____ Mailing Address_____

_____ City_____

Pin Code_____ State_____ Phone No._____ Email_____

Subscription No. (for existing subscribers only)_____. **I would like to subscribe to the above (✓)marked magazine(s) starting with the next issue.** Please find enclosed a sum of Rs_____ by DD/MO/crossed cheque bearing the No._____ dt._____ in favour of EFY Enterprises Pvt Ltd, payable at Delhi.

Please mark one (nearest) relating to your subscription: ☐ **Indian Company** ☐ **MNC** ☐ **R&D organisation** ☐ **Engineering institute** ☐ **College/School** ☐ **Any other (specify):**_____

Send this filled-in form or copy to : EFY Enterprises Pvt Ltd, D-87/1 Okhla Industrial Area, Phase 1, New Delhi 110 020 | Ph: 011-40596600 | e-mail: support@efy.in

Terms:- # These rates are applicable for new subscribers as well as renewal by existing subscribers # Please allow 4-6 weeks for processing of your subscription.
Please include your pincode for prompt delivery of your copy.

This gives you an edge as it allows you to easily showcase your work results. And if lucky, your open source work can help you skip past some of the tricky parts of interviews upon selection for a tech job.

A few ground rules before you make your first contribution

1. Don't fear making your first contribution. You don't have to be an expert to begin.
2. Even if you haven't written a line of code in your entire life, there are still different ways to contribute to open source. You don't necessarily have to learn to code.
3. You don't need to ask anyone's permission before contributing to an open source project. But it's nice to chat with the project maintenance team to have a better chance of your contribution being accepted.
4. Always start small rather than making a big contribution. If you don't know the people maintaining the project, their purpose for the project and their expectations from your contribution, chances are your contribution will get rejected. You'll get less frustrated if you decide to work small instead of working big for several weeks, only to get refused later.
5. Never give up if your first try doesn't work out. Learn from the feedback on your contribution and make relevant improvements to your style of work. Your contribution will surely get accepted next time.

Steps to make your first open source contribution

Choose a project: If you don't know where to start, then there are a few different open source tools you can use.

One of them is a website called *First Contributions*, where you can find a lot of open source projects with specific issues that are friendly for first-time contributions. Simply go through them and select.

On selecting a project, you will be taken to the GitHub page of the project and into the *Issues* section, which is filtered with good first issues for first-time contributors (Figure 2).

Go through the different issues and select the one that interests you.

After selecting the issue, read the general description to know what you are expected to do. If someone has already assigned himself/herself to the project, then look for any other issue to work upon.

If you don't wish to go through GitHub, then there is another method, for which Node.js should be installed in your machine. Type *npx good-first-issue* in the terminal and it will give you a list of available projects (Figure 4).



Figure 2: Screenshot of the GitHub page showing different good first issues (Credit: OSI 2021)

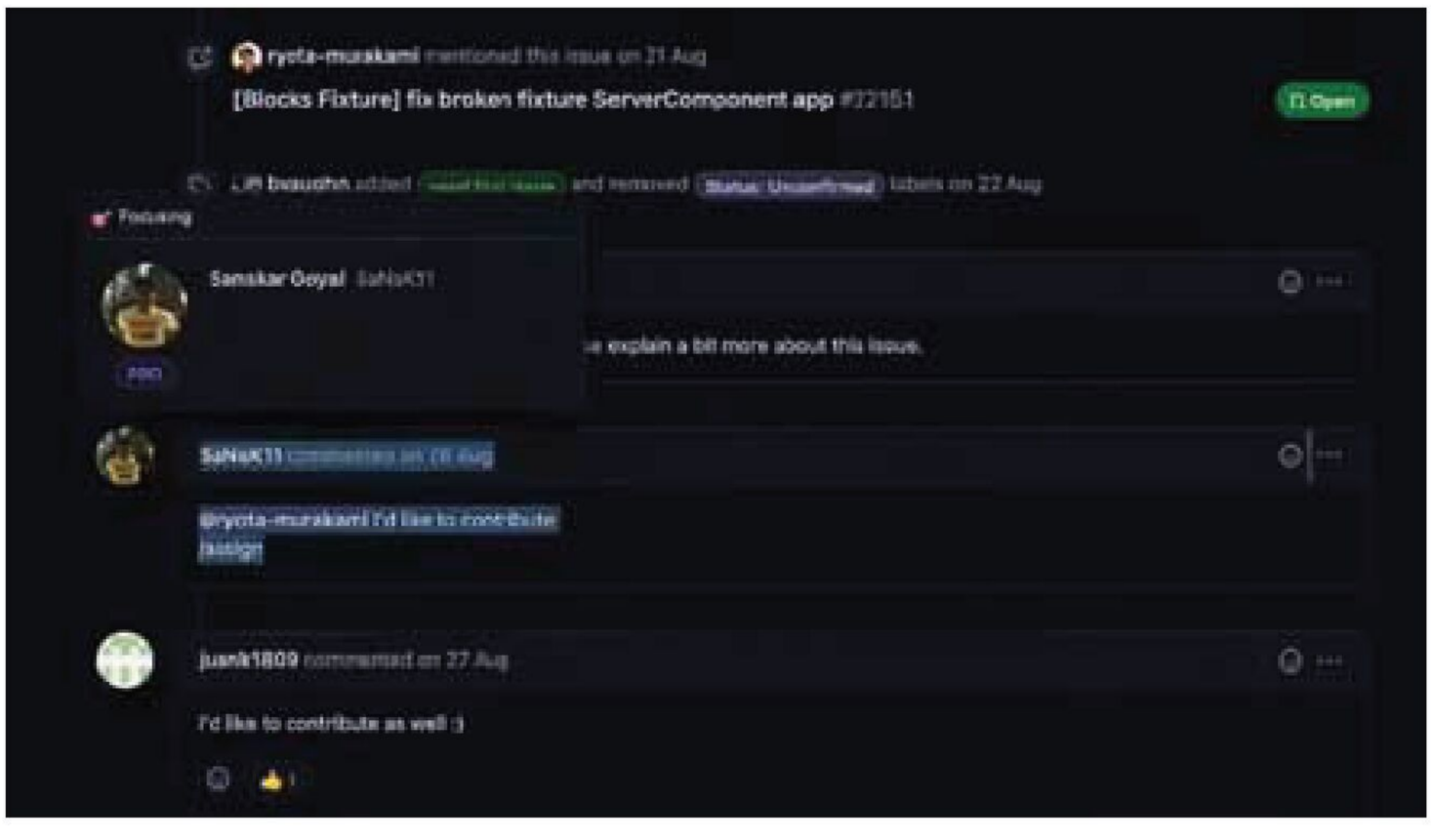


Figure 3: Screenshot showing a user who is already assigned to an issue (Credit: OSI 2021)

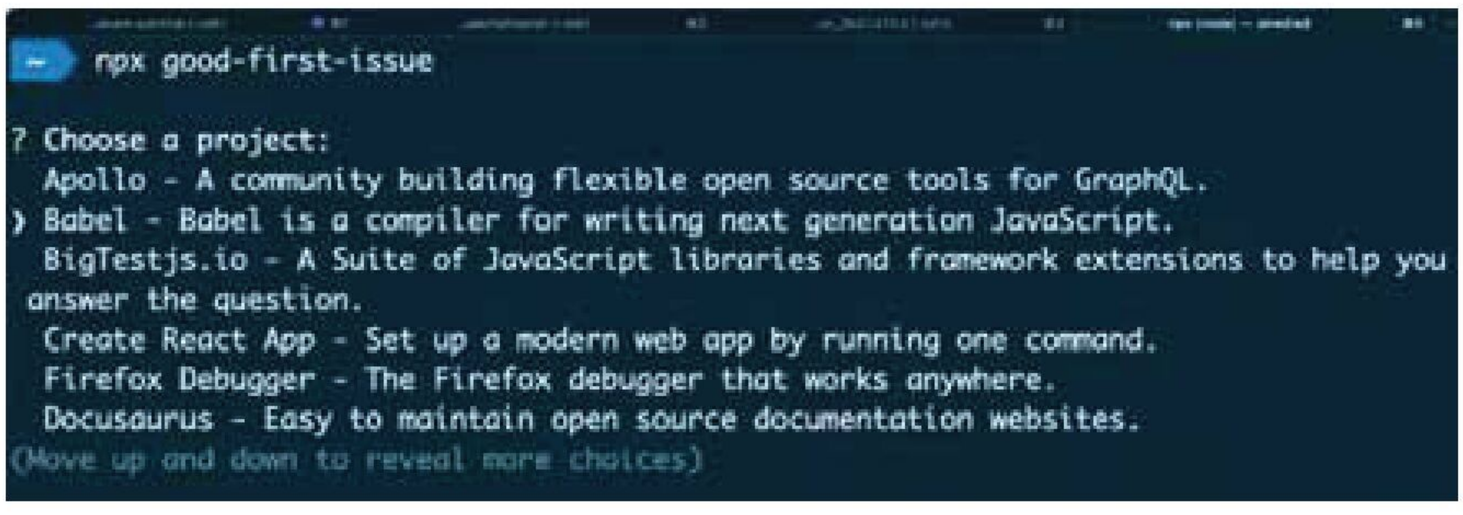


Figure 4: Command type and project selection options



Figure 5: Example of a good first issue in Babel (Credit: OSI 2021)

For instance, if you choose to contribute to Babel, a random issue will be automatically picked. To start working on the issue, go to the given GitHub link.

Tip: Before making any contribution, check whether the project has been actively maintained over time so that your contribution request gets accepted soon.

Work on a project: To make a pull request for starting your first contribution, go to the link *aka.ms/first-pr* that will take you to the GitHub repository called 'First Contributions'. There you can access a tutorial on how to create your first pull request, which includes forking the repository, cloning the repository on your local machine, and making a change to a file — for example, adding your user name by creating a new branch, committing the change to a local copy of the project, pushing those changes to GitHub and then creating the actual pull request where the project maintenance will be able to review it, give a feedback and accept/reject your changes.

As stated earlier, coding isn't the only way to contribute to an open source project. There are many other ways. The first one is to promote the project and share it on different social media platforms. If you have experience with a particular project, then you can answer some of the questions related to it on Stack Overflow or GitHub. It will be very helpful for project maintenance.

Another good way to make the first contribution is to go through the project documentation and fix errors. You can also write tutorials and translate the documentation into different languages, making the project accessible to a lot of different people.

If you come across any bug, then do report it. If you don't do so, nobody's going to have a look and fix it in time. Writing unit tests for projects, especially after witnessing a bug that escaped someone else's notice, is also a good way to contribute to any open source project. And if you are a good designer, then you can design a nice logo for a particular open source project.

Besides the above, there are multiple ways to contribute. You can try to come up with some other helpful ones.

Tip: If you plan to work on an issue for a specific project, do assign yourself to it so that other people can know that someone else is working.

Submit your pull request: Certain things can increase the chances of your pull request getting accepted. The first thing is to read the contribution guidelines before submission. It's good to describe the changes made and why for the project.

Just be mindful that some of the people maintaining an open source project won't be able to regularly check the repository. So be patient and allow them to go through your pull request. If not now, they will do so in a few days.

Giving and receiving feedback is an essential part of the open source cycle. It helps to understand what needs to be changed. If you receive criticism, don't take it as a personal attack but as a means to further improve your skills. Since hundreds of big open source projects get reviewed every day, feedback on your contribution may take days — even months. Nevertheless, don't worry. You will eventually receive it.

Get ready for a new adventure

After all your hard work, unfortunately, everything may not work out the first time. Maybe your contribution didn't fit the project's objective or someone else contributed in a better way. Don't worry, it happens. Remember to never give up. Once your contribution gets accepted, you'll be motivated to contribute even more.

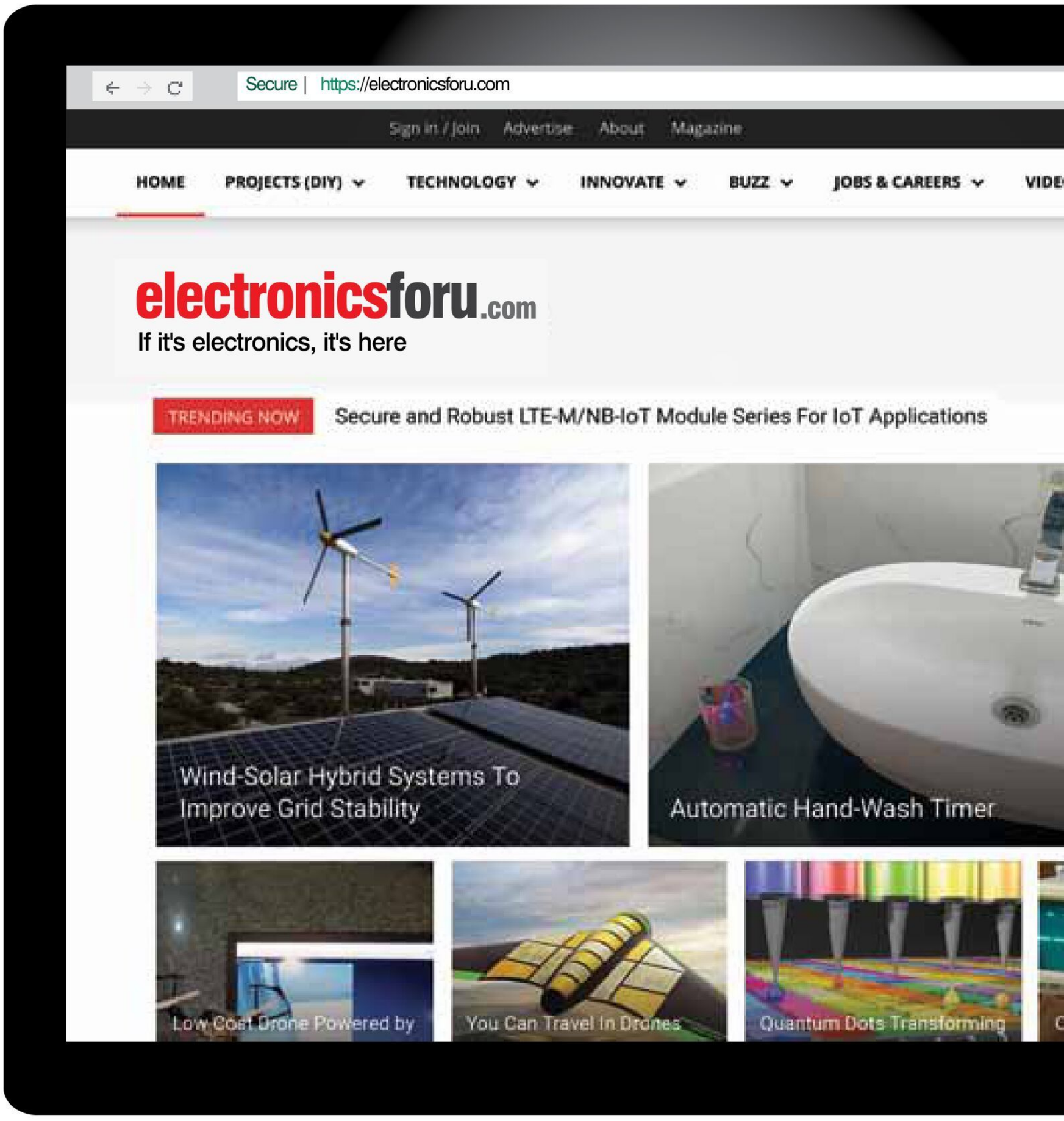
So, if you love the philosophy of sharing content, receiving feedback and being a part of a larger tech ecosystem, then the life of an open source contributor is meant for you. END

The article is based on a talk given by Yohan Lasorsa, senior developer advocate, Microsoft, at Open Source India 2021.

By: Vinay Prabhakar Minj

The author is a technology journalist at *OSFY*.

Your favourite website has



electronicsforu.com

THANKS TO YOU—OUR ONLINE NETWORK IS

FACTS & FIGURES

- 4 websites (two more coming soon)
- Five major Facebook communities
- Seven major LinkedIN groups & pages
- Million-plus active users (monthly)
- Million-plus reach through Facebook
- Fifty-thousand-plus industry connections through LinkedIN

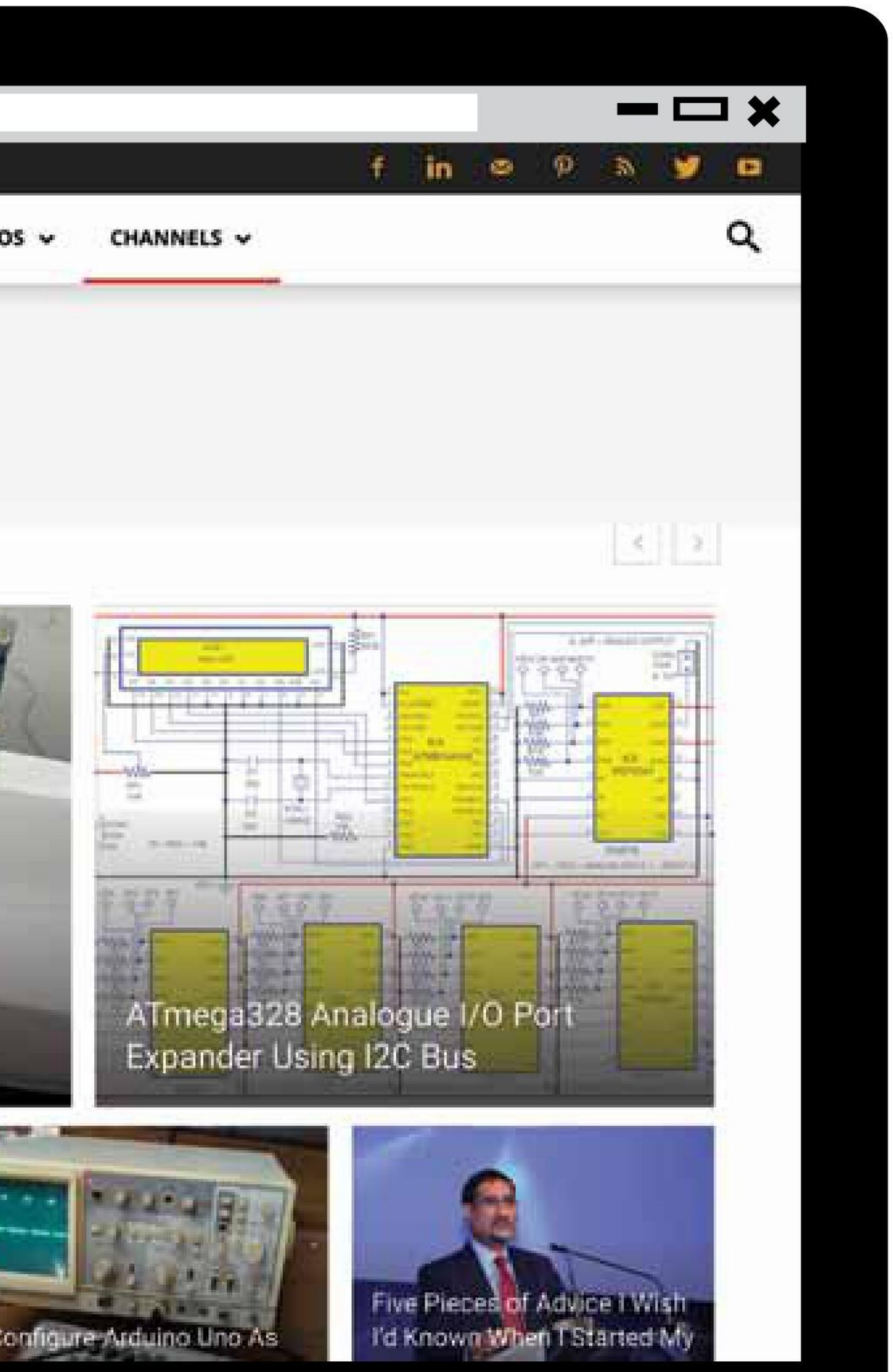
READERS

- You can access all content for FREE
- You can subscribe to newsletters for FREE --on most websites
- Register on our websites to get free invites to technical webinars and seminars

EXPERTS

- Experts who want to share their knowledge through articles, DIY Projects, etc are welcome
- We also welcome experts who want to share their knowledge through webinars or seminars
- You can contact us at editop@efy.in

fast growing peers now...





The Latest in IOT.
A platform for enablers, creators and providers of IOT solutions.



India. Electronics. Directory.
Enabling commerce between buyers & sellers of electronics in India.



Business. Electronics. India.
Everything you want to know about India's electronics industry.

Amazing DIY Projects. Latest Tech trends.

The hang-out for electronics enthusiasts.

AMONGST THE WORLD'S TOP 5 AND GROWING!

INDUSTRY

- You can advertise for as little as US$ 100 per month
- Special combo offers for advertisers in our print publications
- We've now enabled flexible CPM-based advertising
- You can advertise on the platform of your choice (based on your target audience)
- We invite press releases at efy-edit-team@efy.in
- Press releases are published free of cost, subject to discretion of the editorial team

RESPONSE GUARANTEED SOLUTIONS

We now also act as marketing partners for our clients and drive entire marketing for them, where we charge them on basis of results and not efforts!

CONTACT US: Shrikant Rao • growmybiz@efy.in • +91-98111 55335

Companies Don't Innovate; People Do

When you are hiring people, some characteristics are useful for particular roles and not for others. Knowing what to look for besides just the hard and soft skills is very important. The article describes the characteristics that people working on innovation should have. So, when you are hiring for innovation, these are the ones you should be looking for.



Being innovative means doing things differently or doing things that have never been done before. And doing so needs a different mindset and skillset. But the fact remains, companies do not innovate; people in those companies do.

The sourcing and development of innovative ideas are essential to a business' long-term success. Therefore, hiring managers need to keep innovation in mind when reviewing candidates. Candidates possessing innovative qualities and experience contribute to a successful company. They directly impact the work culture and many business practices.

However, hiring for innovation is different from all other hiring criteria, processes, and decisions. This article talks about at least two dozen characteristics of employees who are best suited for innovation.

1. They are not yes-masters

Employees who are not ready to say 'no' when needed do not fit the innovation team quite well. Innovation requires an ability and willingness to challenge the status quo. People who are 'yes masters' lack this ability.

2. They are not nay-sayers

Employees who are constant nay-sayers impede innovation at best. Their negative outlook and attitude can stifle innovative ideas and innovation overall. These candidates may be suited in more cautionary positions where a negative outlook may work as an asset. Unfortunately, innovation is not one of them.

3. They are marathon runners, not sprinters

Innovation is not a short-term game. You are not developing an off-the-shelf product or an already known process. Working on something that has not been tried before needs several iterations, trials, errors, failures, and whatnot. People who prefer shorter projects and sprints are not well suited for innovation because they might get bored sooner and lose interest.

4. They are self-starters and independent thinkers

Innovation often requires independent thinking and taking the initiative. Someone who prefers waiting on managers or leaders for taking action may slow down the innovation faster. Moreover, if they cannot think independently, it is unlikely that they will work on original and creative ideas.

5. They are first-principle thinkers

If you were to check whether someone is the first-principle thinker or not, try this. Ask them a question and demand reasoning for their answer. See how they reason their thinking. The first-principle thinkers almost always will start from the ground up or fundamental principles.

On the contrary, others will do so with associative reasoning. Associative reasoning and heuristics are good approaches for faster decision-making. However, when it comes to innovation, they often result in a lack of originality.

6. They are creative thinkers or better implementers

While creative thinking gets the limelight in innovation, rigorous implementation is what makes it successful. So, you may find it difficult to find a candidate who has both capabilities. It is not impossible for some candidates, but it is surely rare. Having a balanced team is essential, and that means you need both types of candidates. Decide what you want in your innovation team and hire accordingly. Expecting both characteristics from the same candidate might not be an effective strategy.

7. They have excellent communication skills

It is often said that better innovators are not better communicators, which in my opinion is just a red-herring. For any innovation to succeed, it must get communicated well at all stages and to everyone. So, checking for excellent communication skills is going to be highly important. See how the candidate responds when you challenge their ideas and how they convince or try to convince you.

8. Be okay with outside-sector experience

This point is perhaps one of the most counterintuitive ones on the list. Most people would think that having industry experience is a plus as that can help in innovating further. However, in my experience, it often backfires.

Institutional blindness is a real thing. It means that once someone has stayed in a particular industry for a long time, they often develop an ability to ignore the most obvious things, which is often detrimental to the innovation. Institutional blindness may be useful when it comes to speed, efficiency, and productivity. However, specific to innovation, it works against that.

9. They are good at adapting

Adaptability is an important trait in general, and it is even more important in the innovation game. The ability to respond to continuously changing situations and unexpected outcomes is the key to running successful innovation projects. Why not change your hiring methodology and see how people adapt to your process. That itself will filter a few candidates.

For example, instead of following a traditional hiring process of tests and interviews, do something different like running a contest or asking them to do something. Put the candidates in a novel scenario and ask them to make decisions. When you change the hiring methodology, only those who are good at adapting will excel and thrive. Rest will fall off eventually and make it easier for you to select.

10. They are more curious and less judgmental

Curiosity is often the beginning of the innovation journey. If someone is judgmental and has already made up their mind, they are unlikely to investigate the matter further. And if that is the case, you are not going to see anything innovative.

On the contrary, curiosity often leads to the quest to know more and find out 'Why?' This quest then can result in deeper dive into the root causes of the matter at hand, or it may pose a 'Why not?' question. Both ways, it leads to innovation. Hire curious people for innovation and avoid judgemental ones.

11. Their resume most likely has a lot but can be confusing

Over the last two decades, I have noticed that most (or almost all) innovators have quite a non-linear career path. That often means that

the traditional resume format does not do any justice when documenting that path. Suppose you come across a slightly off resume, which does not look like a traditional one; probe further. You may be dealing with a cross-functionally experienced person. Although their career path looks non-linear and does not make sense at first glance, talking to them can often give you a better perspective. In most cases, they can draw a connecting line between all those experiences. And that can help you make the decision.

12. They are good at problem-solving but better at problem finding

Yes, problem-solving is a commonly revered skill for all cadre of employees. And when you deal with well-defined problems, that's all you need from someone. However, finding better problems to solve is becoming a forgotten skill. When you have someone who can find better problems to solve, it can open up a new potential opportunity for your business. It may help you tap into a completely new market, which can be a serious advantage for your business.

Testing a candidate for problem-finding ability is relatively easier. Of course, you want a meaningful problem finder who is willing to solve the problem too, not a complainer who finds faults but does nothing to fix them.

13. Avoid specialists; prefer more synthesists

If you ask someone, "What is the opposite of a specialist?" You will most likely hear, "A generalist." And that is incorrect. The real opposite of a specialist is synthesist. Specialists are good at breaking down things and working on them at a deeper level. On the other hand, synthesists are good at combining things and looking at the big picture.

The success of meaningful innovation is in defining its higher-level context in which your innovative product or solution will eventually operate. Only synthesists can see the big picture and explain how a particular innovation fits a bigger system or situation. It also means that synthesists can often connect the dots and derive innovation from different fields. More often than not, innovation is less about original ideas and more about original combinations.

14. They know what they don't know

'Know it all' people do not make a good fit for innovation teams. This point is a corollary to being more curious and less judgmental. When people are self-aware, they know their limits, and they also know what to seek. Knowing what you do not know pushes you to learn that skill or find that information. It develops an attitude of continuous learning.

15. They challenge the status quo with logic

As I mentioned before, innovation requires the ability and willingness to challenge the status quo. However, there is a major difference between being a rebel for its sake and being pragmatic with logic. It can be slightly tricky to test it out when you are in the hiring process, mainly because everyone wants to show their best-case persona in the hiring process. These nuances are easy to miss.

16. They are modest about achievements

This one is a bit odd. In any hiring process, both parties want to highlight their achievements and boost their advantage as much as possible. However, this is when you are most likely to miss better candidates for your innovation journey because these candidates are modest about their achievements.

I have seen in many cases people jump the gun and celebrate their innovation win too early. In other cases, the proportion of their celebration does not even match the outcome. It can be detrimental to your innovation programme in the long run.

Candidate's modest nature often puts them behind the queue. So, next time you find a candidate being modest about their achievements, chances are they are a good fit for the innovation role.

17. They are experimenters and action takers

As a corollary to point 4 and 6 above, innovative people are avid experimenters and have a strong bias towards taking action. You will find them constantly experimenting and testing out various assumptions. And when they see a positive outcome or find a way through that experimentation, they take action.

It often needs tenacity, perseverance, and patience. Many see patience as a roadblock in a world where everything is moving fast or appears to be moving fast. But not the innovators.

So, in your hiring process, see how you can check these qualities. Particularly, see how you can make them show the experimenter in them. It will be interesting as well as supportive of the process of selection.

18. Look for diversity of thinking, experience, and skills

When we talk about diversity, we usually think of gender balance or demographics. However, for innovation, we also need diversity of thinking, experience, and skills.

Diversity of thinking comes from the fact that you have learned to think differently. You do not just follow what others are saying. Your opinions are based on what you see around you, and you can

think independently on your feet. And therefore, in most cases, this thinking is different from others.

The other part is about perspective. Because of the diversity of experience, you are able to see the same issue from different angles.

Talking about the experience, if you have a candidate who has experience in the same industry for a very long time, the chances are that they haven't seen enough. Different perspectives from other industries, whether they are better or maybe worse, are important. Regardless, these candidates won't be aware of it. A candidate from different industry experiences or different types of roles has gathered so much variety that it helps them think innovatively. So, look for a variety of experiences.

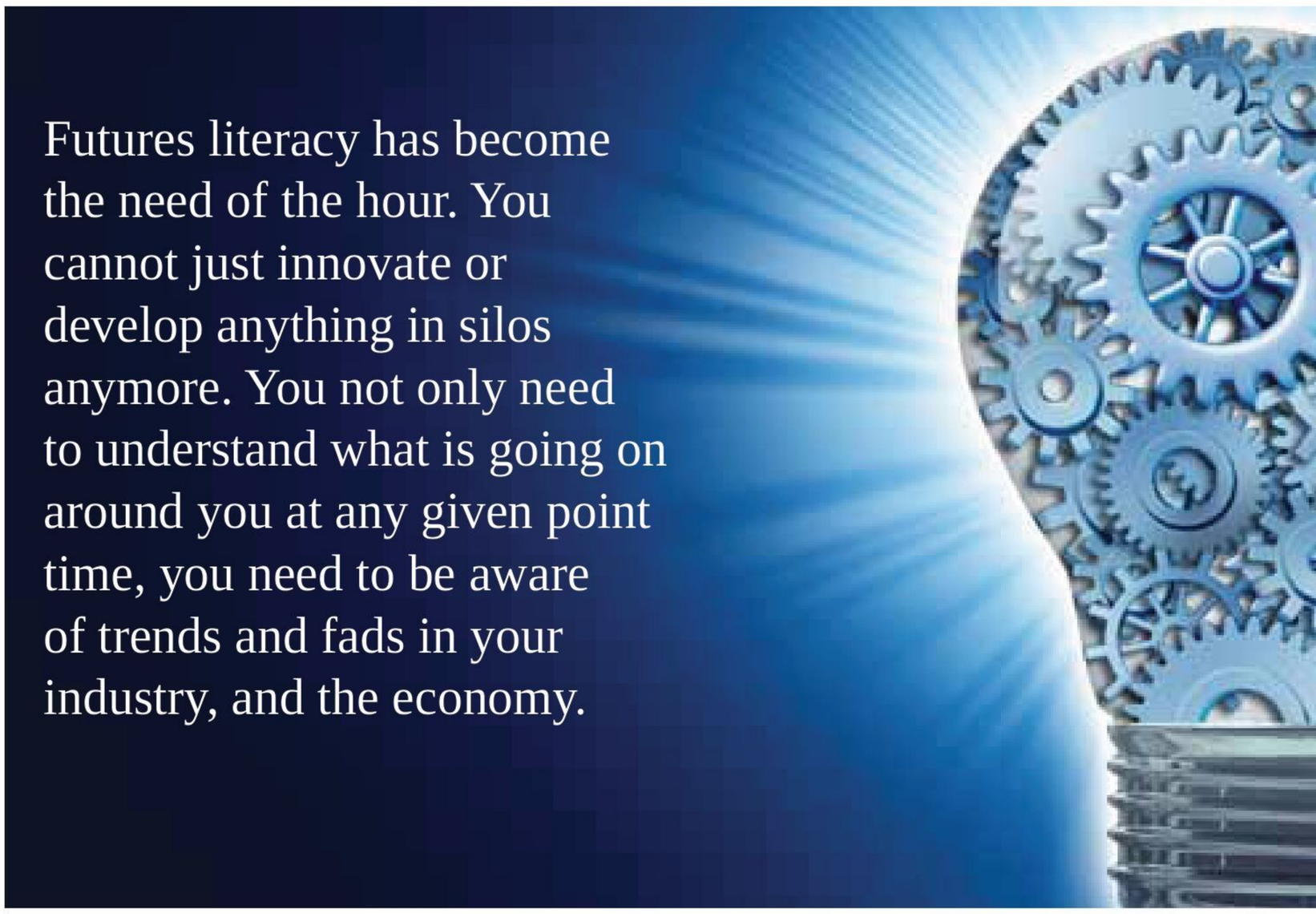
And of course, for the skills, you need someone who is not just a specialist in one thing but in many. It is important because innovation is often about finding new ways to do something, new combinations, and original combinations. Apparently, that is only possible when someone has a diverse set of skills; so go for it.

19. Entrepreneurs should be welcomed

Now, this one is interesting. I want to emphasize this point because I have often seen companies being skeptical of hiring entrepreneurs. They think that entrepreneurs may leave early or will be more distracted thinking about their next venture. Well, this is a very negative outlook on entrepreneurs.

First of all, entrepreneurs are and can be more loyal to you than an employee can. They understand it because they have had their employees, and they exactly know what you expect of them.

But more importantly, entrepreneurship teaches you so many different skills and techniques. They have gathered so much expertise through their entrepreneurial venture. So, once you have an entrepreneur in your innovation journey, you have multiple employees packed together in that one person. Therefore, if you get a chance to hire an entrepreneur for your innovation plans, jump on it; do not lose that opportunity. Without question, it is a very good bargain.



20. They are future-literate

Futures literacy has become the need of the hour. You cannot just innovate or develop anything in silos anymore. You not only need to understand what is going on around you at any given point in time, you need to be aware of trends and fads in your industry, neighbouring industries, and the economy. But you also have to be aware of what is coming and what might be coming in time to come.

That means people need to be aware of future scenarios. So, futures literacy is becoming more and more important. The complexity and volatility of the markets and economies force people to stay two steps ahead at any time. Being future-literate will help you do that, which is why you must look for this trait.

21. They love problems instead of solutions

People love what they create as a solution, product, service, or software program. And this is a trap. If you love your solutions, you are less likely to change when the problem changes and when the context changes. So, my strong advice to anyone is this: Not just for innovation, but a successful and sustainable business, always love the problem you are working on. Love the fact that you have an opportunity to work on that problem.

Your solutions may change over time as technology evolves and context changes. But if you are stuck with just the solution and are in love with your solution, it may soon become redundant.

Check out these attributes when hiring for innovation. Give them scenarios when they have to pivot and have to adapt to the problem. You need people who are okay with adapting to the problem and changing solutions. They are better suited than someone who refuses to change the solution despite seeing that everything else has changed.

22. They are aware of the ethical challenges of technology

Ethical challenges are the new minefield in technology. Everything in technology development may look cool, but it is very likely to have some ethical angle. That means they might face significant challenges down the line.

So, is the candidate aware of those challenges? Do they understand the importance of ethics in dealing with powerful technologies? Do they have a plan or at least an idea to deal with it if that arises? How would they handle it? Ethical maturity is a new soft skill now, and almost everyone needs to have it.

23. They have a track record

A good candidate for innovation usually does not have only one celebrated win. They have a track record. They have a track record of several failures, along with a few successes. Not every win needs to be a bumper success. Some can be small, while others might be much celebrated ones.

But, in general, a good track record does have lots of failures and lots of learnings from those failures. So, check for those. Ask them if they have multiple failures? What did they learn from them? Do they have multiple smaller wins? There may be a few bumper wins, but they are not necessary to have. What you are looking for is not a one-time stroke of luck. You are looking for a track record.

24. They are resourceful

This one goes without saying. Resourcefulness is the ability to get what you want with what you have. Being resourceful is quite an important skill in innovation, business in particular. However, it is a highly revered one for every employee, in general. But specifically in innovation, you are often stranded by resources, budget, or even by lack of information.

Being a resourceful person comes in handy when resources are scarce, and the goal is important and timely. This one is easier to check with a few exercises, scenarios, and carefully designed tests. But being a critical one, you must prioritize it when you can.

What about you?

The hiring process is not a one-way street. As much as you evaluate candidates and their fitment for your organisation, they are doing the same with you.

So, what are a few things you must do to show you are a good fit for innovative candidates?

Demonstrate. Make sure you can demonstrate company capabilities and willingness to innovate. Most of the time, companies show that they care about employee's opinions and are innovative. However, once the candidate becomes an employee, things change. They realize that it is not as rosy as the picture they were shown in the hiring process. So, you need to find ways to demonstrate your ability, willingness, culture, and approach towards innovation.

Pay: This one is straightforward and applies across the board. Top talent is seldom cheap. You often get what you pay for. So, be ready to pay for the talent. Some companies tend to think from an ROI (return on investment) perspective and calculate it that way.

I think it is a bit limiting and short-term view. It is not wrong as such, but there is more to it. When you hire candidates who are suitable for innovation, they often impact your company culture and several business practices. So, the impact they can have is beyond just an ROI.

Respect: You may hear or see it sometimes. Many strong candidates for innovation are comparatively different, and people call them eccentric, which is an essential trait to have. Eccentricity is what makes them good at what they do. So, being respectful of their eccentricities is a must.

Listen: Even if it is too early in the process, and the candidate is not your employee yet, being receptive to their inputs and ideas can go a long way. People who are good at the innovation game appreciate someone listening to their point of view. Demonstrating that you care will instill confidence and will attract them.

Do not be offended: A lot can be told about company culture based on how people respond when they are challenged. When someone challenges your legacy, and so-called holy-cow system or process, do not be offended by it. Stay curious and dig deeper. They may have a point. At least it is worth listening to their point of view. If you feel offended, it sure sends a wrong signal and shows that you are not as flexible as you need to be for an innovative organisation. It is often a tell-tale sign for candidates not to go ahead.

The point is

You may have observed quite a bit of overlap amongst all these characteristics. But the thing is, when you are hiring people, some characteristics are useful for particular roles and not for others. Knowing what to look for besides just the hard and soft skills is very important. These are the characteristics people working in innovation should have. So, when you are hiring for innovation, these are the ones you must be looking for.

End of the day, we all must always remember that companies do not innovate; people do. END

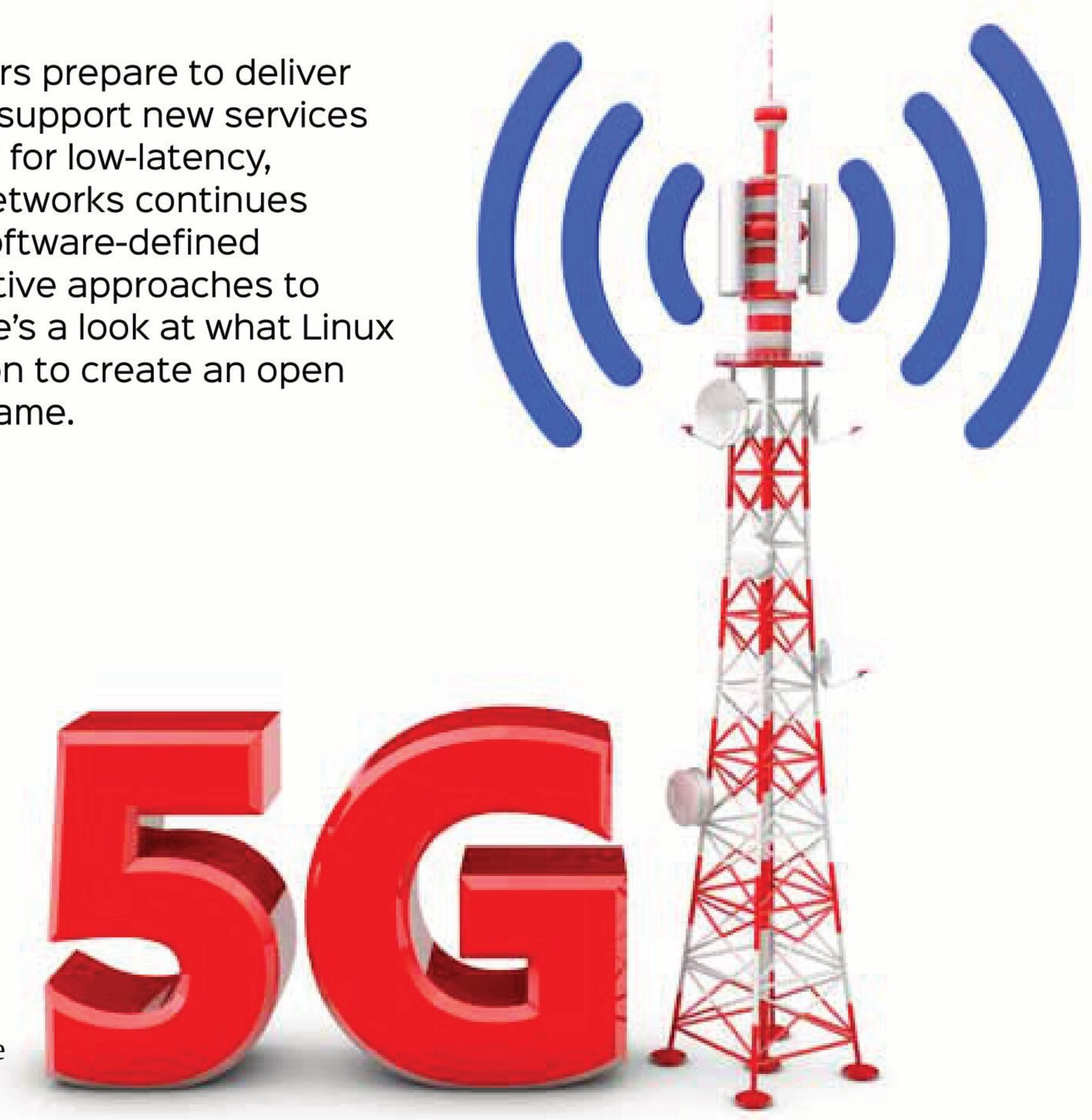
By: Anand Tamboli

The author is a serial entrepreneur, speaker, award-winning author, and an emerging-technology thought leader.

The article was originally published in the November 2021 issue of Electronics For You.

An **Open Source** Ecosystem for **5G** and **Telecom Networks**

As communications providers prepare to deliver high-speed connectivity to support new services and use cases, the demand for low-latency, high-bandwidth, scalable networks continues to grow. This may lead to software-defined infrastructure and cloud native approaches to meet 5G requirements. Here's a look at what Linux Foundation (LF) is working on to create an open source ecosystem for the same.



Successful open source developments depend on complete life-cycles of the projects that the market will adopt and deploy. There is an exciting convergence in the networking industry around open source, and the energy is palpable. In early 2021, the Linux Foundation (LF) announced a collaboration with Facebook to be managed under a neutral governance framework.

Magma, an open source software platform, enables operators to build and augment modern and efficient mobile networks at scale. It features an access-agnostic mobile packet core, advanced network automation and management tools, and the ability to integrate with existing LTE networks. It has use cases across both virtual and container network functions (xNFs) including fixed wireless access, carrier Wi-Fi, private LTE and 5G, network expansion, and mobile broadband.

In all these major areas, the LF hosts over 700 projects that are either in a technology area or a market area where innovation can speed up ongoing work in India. The numbers are just staggering these days. Most of the Fortune 500 companies use our quarter of a million developers, contributing code to almost 500 plus projects. And then if you look at just the stats through LFX insight, 31 million lines of code were added between 1.1 million poll requests and 10,000 repositories.

Technology disruption

The open source journey has been quite significant in the past decade, starting with virtualisation, desegregation, and software defined networking to control plane databases. We have come a long way in the Kubernetes cloud native and are now moving to intent based.

While earlier focus was to develop projects in technology, building blocks for open source, in the next decade we are going to see all these open source projects working together for a common solution. The problem we are trying to solve this year is how to move code into production as quickly as possible.

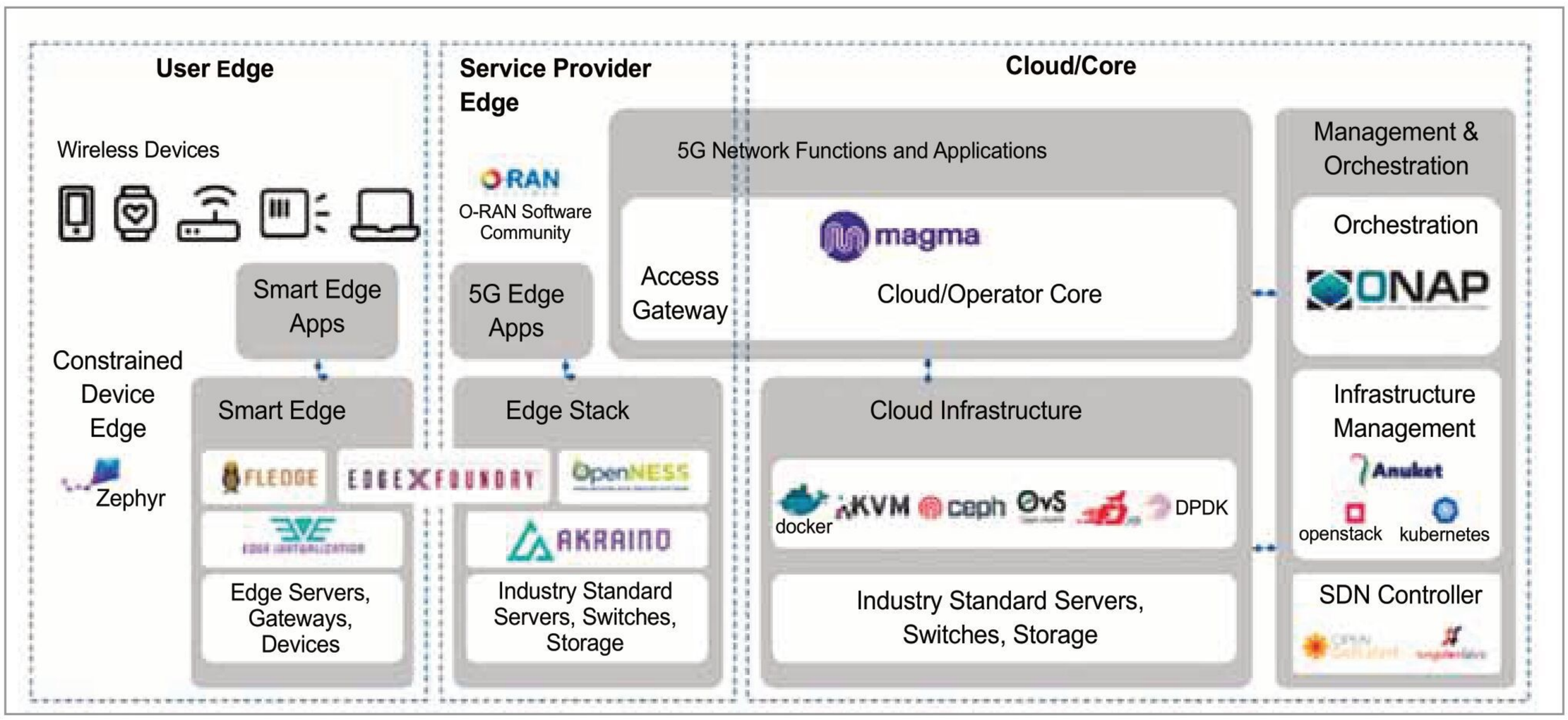


Figure 1: LF open source component projects for 5G

While there are projects like Anuket, which is in LF networking, projects which are LFH can solve exactly this problem. They look at blueprints and how to comply with a spec, do open verification, submit the test results and do open interrupt, testing, training, and certification. So that's kind of how we are seeing most of the evolution in open source this year.

Collaboration

LF networking has a unique perspective with the broadest set of projects that make up the diverse and evolving open source networking stack. It provides platforms and building blocks across the networking industry that enable rapid interoperability, deployment, and adoption, and is the nexus for 5G innovation and integration. It has now tapped into industry efforts to structure a new initiative to develop 5G Super Blueprints for the ecosystem. Major integrations between the building blocks are now underway between ONAP and ORAN, Akraino and Magma, Anuket and Kubernetes, and more.

The 5G super blueprint is what is going to integrate a set of projects all the way from access to core. 'Super' means that we are integrating multiple projects, umbrellas (such as LF Edge, Magma, CNCF, O-RAN Alliance, and LF Energy) with an end-to-end structure for the underlying facilities and application layers throughout edge, gain access to, and core. This end-to-end combination allows leading market usage cases such as repaired wireless, mobile broadband, personal 5G, multi-access, IoT, voice services, and network slicing. In short, 5G Super Blueprints are a vehicle to team up and produce end-to-end 5G options.

With availability of customised solutions, there is a tremendous opportunity for the market. This is how the verticals are taking advantage of edge computing. Now healthcare and smart cities are towards the lower end of the spectrum. But it will need a little bit of regulations and process controls to update before adoption happens.

LF edge projects

There are two types of edges—user edge and service provider edge. A user edge is dedicated and operated by the user. A service provider edge is shared as a service, but then not all usages are the same. You can either have a very constrained device edge, which could be microcontroller based embedded compute, or you could have a gateway type edge. You know in the user control or an on-prem data centre edge, which is secure in a factory building.

And what separates the two is the last mile networks. So, under the base station or in a smart central office, which is the access edge is not hard to know. Edge compute is for applications that need 20 milliseconds or less latency.

The edge project stack is a very important attribute. For example, if a sensor wakes up and dumps data to an Amazon cloud every week, it's an IoT application. It's not an edge application, and vice versa. Hence, the LF edge as an umbrella is unifying the frameworks.

Akraino R4 blueprints

Akraino is a blueprint concept. The open community comes together to define a use case. And then, through open governance, tests out the use case and interoperability all the way from hardware to the application,

and then submits the results so that developers can replicate it.

That's a blueprint, a declarative configuration, with CICD documentation. The one that has become very popular with telcos is PCEI (public cloud edge interface).

So, the question arises, how do the public cloud and the service provider at the edge have a common set of interfaces? There are no standards that exist.

Use cases

ORAN use cases include the slicing and the disaggregation that comes with the open RAM. And then on the Kubernetes, it's the cloud, multi-cloud hybrid layer below with the standard compliance and then ONAP, which is the largest and the most important. It goes through 5G network slicing and all the use cases for telecom. This is really the fundamental platform, 10 plus million lines of code that is at the heart of network automation and network automation is a must for 5G.

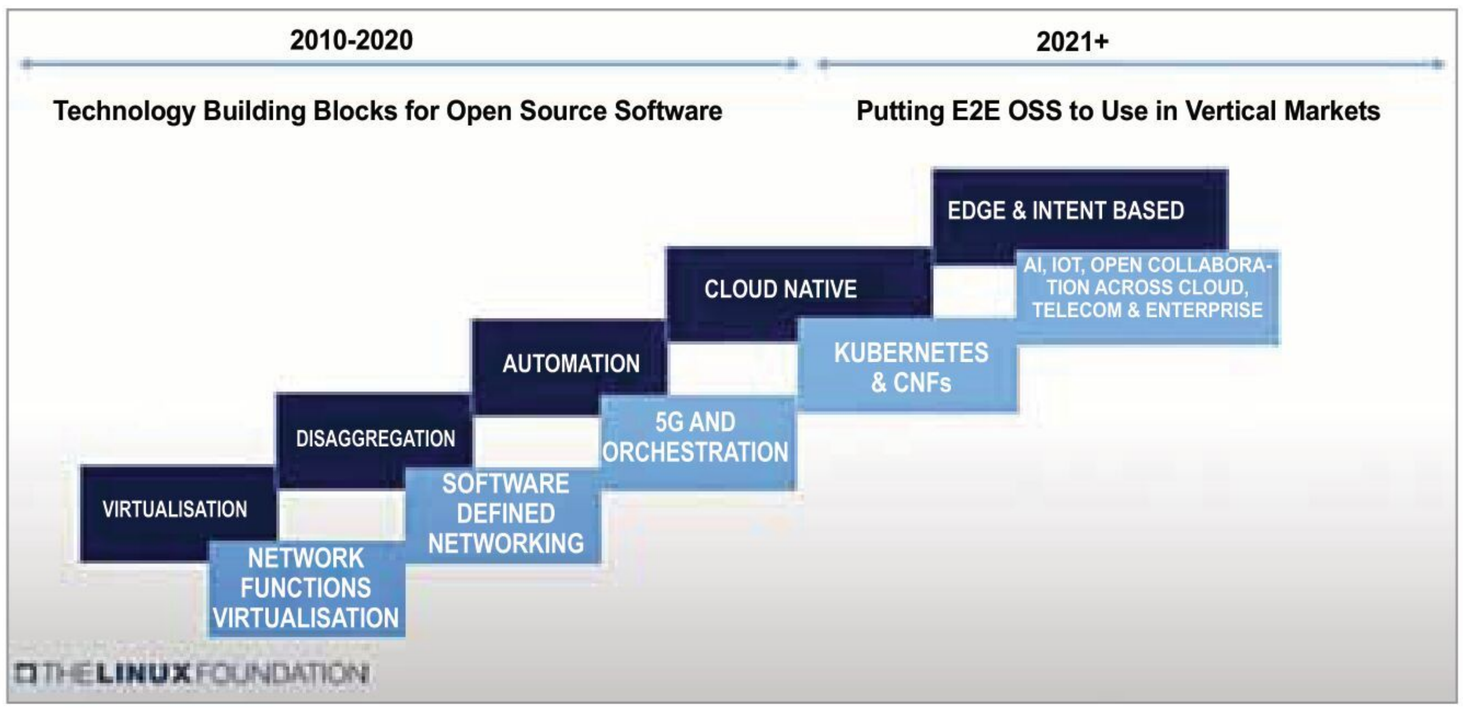


Figure 2: Focus shifting from building blocks to putting E2E OSS to use

So as a global community, ONAP has matured and is one of the most important projects to enable end to end systems.

In case of intermittent connectivity, the user needs an on-prem framework, which is Eve or edge virtualisation. EdgeXFoundry, the open source vendor neutral, is by far the largest framework for IoT. It abstracts southbound and northbound—all the messaging protocols and everything northbound—and this is where building automation, smart process control, water cities happens.

LF networking may now be working on about ten projects—Magma cloud native computing foundation, Kubernetes, or a software community. We are taking these communities and doing an open 5G super blueprint integration.

How it comes together

The 5G super blueprint comes pre-baked and only a small set of integration and support is needed by

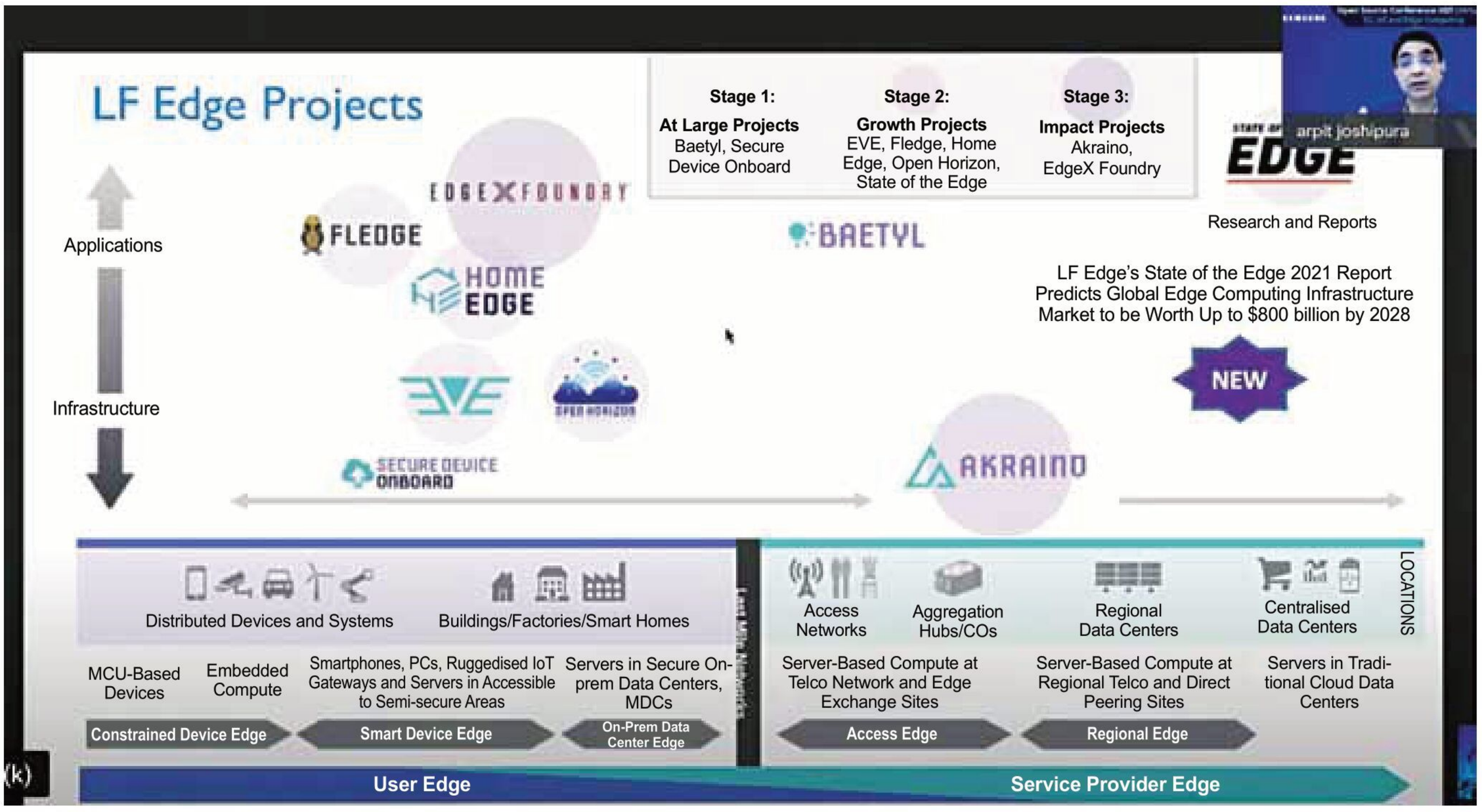


Figure 3: LF edge projects

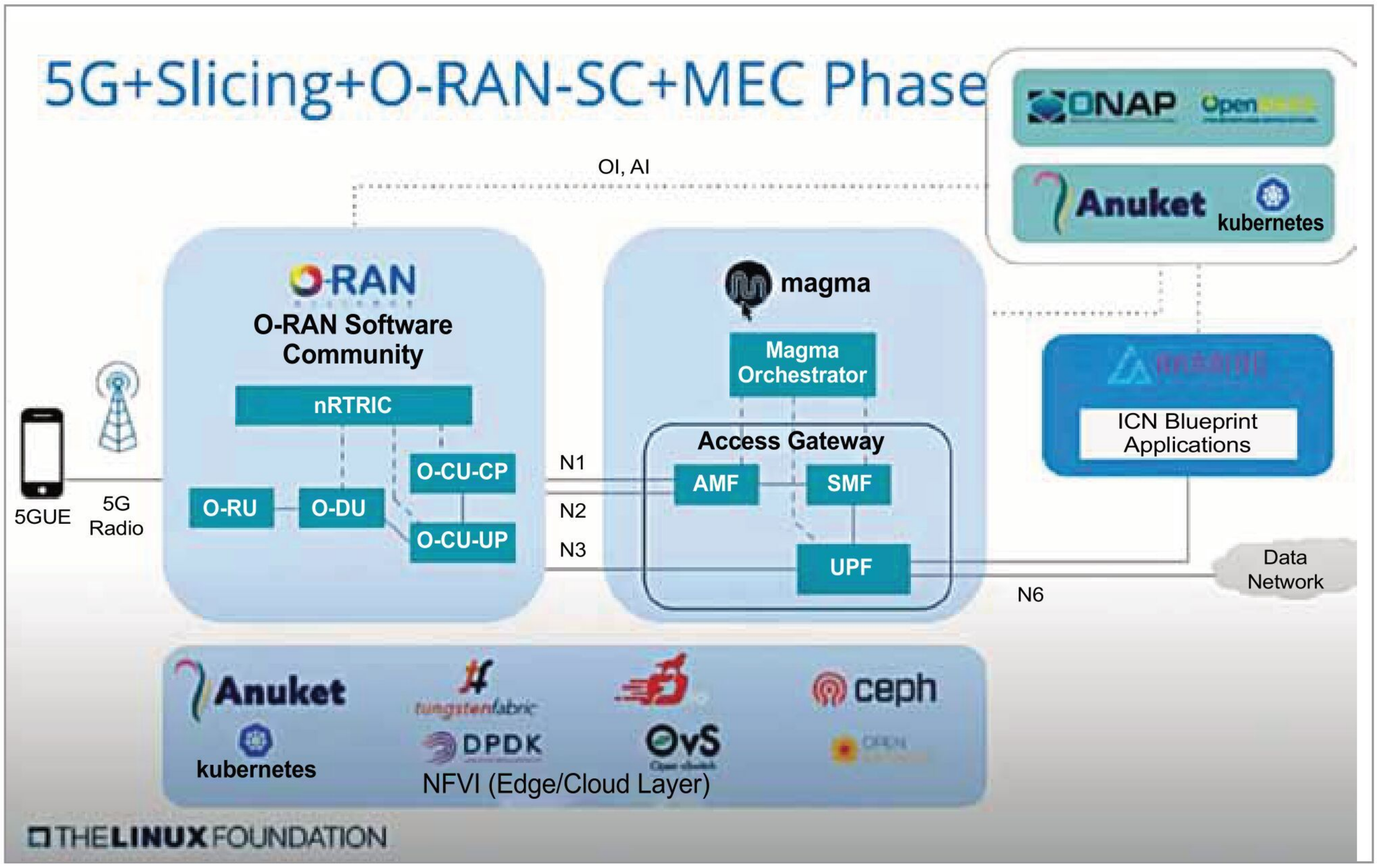


Figure 4: The third phase of 5G modernisation

the community or their vendors. So, instead of multiple vendors doing things in different ways, now you have it done in the community and then the vendor can provide innovation and differentiation on top.

5G will modernise today's energy grid with awareness monitoring across distribution systems.

It will roll out in three phases: 5G core+MEC phase, 5G E2E+MEC phase, and 5G+Slicing+O-RAN-SC +MEC phase.

In the first phase the communities are using the emulators on the RAN side, but Magma is being inserted with ONAP and Anuket on Kubernetes. There are some components from Intel called openness also coming in but, in general, those are the big new components.

The second phase is bringing in a commercial genome piece. The RAN side of things is integrated, and then you run a slice through it. So, the end-to-end network slicing comes in.

The last phase, which is towards the end of the year, is when ORAN comes in. What's happened is that ONAP and ORAN are already deeply integrated with interfaces. There is a lot of component reuse between projects because these are all microservices, containerised modules, and you can just use them as needed and it will just get you a head start. And that's already happening.

Summing up

Completing the solution stacks to integrate a fully open networking and edge stack, speeding up 'backend interior and deployment process' with open compliance, integrated deployment in vertical markets to start embracing, utilising and building on networking, cloud, edge stacks are seen as the top priorities.

To conclude, blueprints are fairly matured and deployed. Most of the projects are live in production. The first step is that it needs to happen. In several use cases, the 5G node is now merged from the open-ended interface side of things. And so, Magma itself, as an open mobile packet core in the Q3 timeframe, will be fairly mature.

The testing of these blueprints should be ready by the end of the year. Till then it is a proprietary mobile packet core. Pretty much everything that is at the higher layers of the stack is mature. At the lower layers, radio will stay proprietary, but the disaggregation concept is mature. END

This article is based on the tech talk by Arpit Joshipura, GM, networking edge and IoT, Linux Foundation, at Open Source Conference 2021 (Virtual).

The article was originally published in the December 2021 issue of Electronics For You.

FOCUS

Trending Tools

Open Source Platforms You Can Use for AR and VR



Augmented reality and virtual reality are being used in various domains and have many useful applications today, apart from gaming. A few useful open source platforms that help to implement these technologies are listed in this article. *Read more on page.....50*

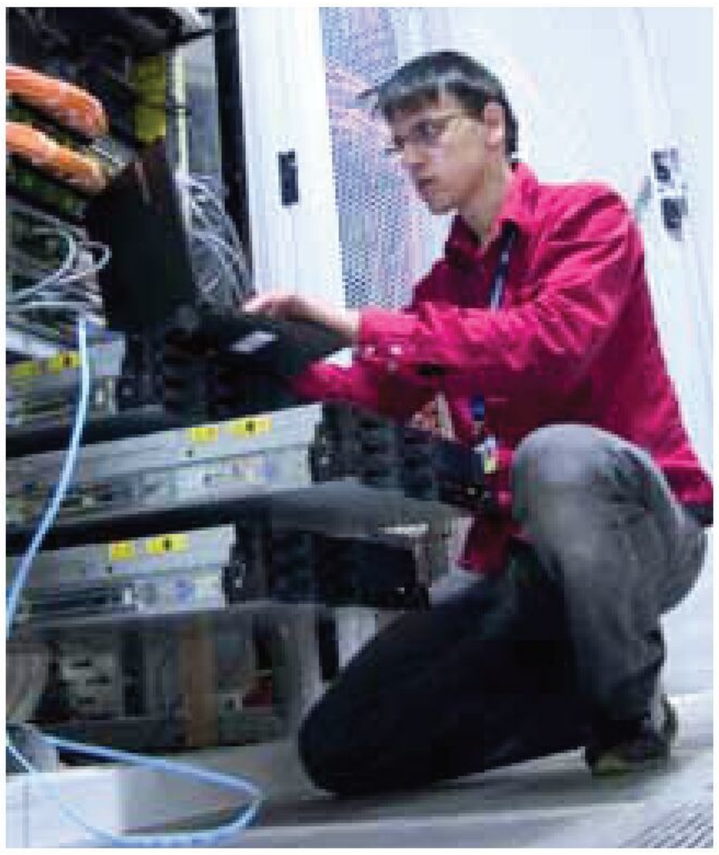
Top 10 Open Source Deep Learning Tools



Deep learning is deeply intertwined with our lives today, and promises to enter many more domains in the near future. This article briefly describes the best open source deep learning tools.

Read more on page.....59

The Best Security Tools for Linux Administrators



Security is of utmost importance in any computer system. In this article, we will be looking at some amazing open source security tools that can beused on a Linux desktop by anyone.

Read more on page.....48

The Future Belongs to These Open Source DevOps Tools



Organisations are adopting DevOps today for their own good. This article briefly discusses a few open source DevOps tools that will be used a lot in the near future.

Read more on page.....53

The Best Security Tools for Linux Administrators

Security is of utmost importance in any computer system. In this article, we will be looking at some amazing open source security tools that can be used on a Linux desktop by anyone.



The security tools listed in this article span a wide range — from network security tools to intrusion detection and prevention systems.

Nikto

The first one that we will be looking at is Nikto. This tool is mainly used for Web security. More than 6000 tests can be automatically performed by Nikto to detect Web vulnerabilities in a system.

We can install this tool on our system using the following command:

```
sudo apt-get install nikto
```

Nmap

Nmap is a tool used to scan networks. Developed by Gordon Lyon, it helps us get the list of open ports and analyse them by sending and receiving packets and responses. Various services and hosts can be detected or discovered using this tool. Many more services like OS detection, etc, are provided by Nmap.

For installation, just enter the following command:

```
sudo apt-get install nmap
```

Now check that Nmap has been installed properly by using the following command. Once you give this command, you should get the version of Nmap, which will indicate that the installation has been done correctly.

```
nmap --version
```

ClamAV

One of the most important tools to keep your system secure is ClamAV. This open source and free toolkit helps our machines fight against malware. Many different types of malware can be successfully detected by ClamAV. It was first developed by Cisco for only UNIX systems but is now available for Linux, MacOS, etc, with the help of some third parties.

The command for installing ClamAV is given below:

```
apt-get install clamav clamav-daemon -y
```

Snort

Snort is an open source intrusion detection and prevention system that was developed by Martin Roesch, but is now owned by Cisco. It has many inbuilt tests to detect different kinds of attacks including Web based attacks.

Installation can be done by executing the following commands in the terminal:

```
wget https://www.snort.org/downloads/
snort/snort-3.tar.gz (check for the
latest version)
tar xvzf snort-3.tar.gz
cd snort-3
./configure –enable-sourcefire && make &&
sudo make install
```

rkhunter

This tool can scan exploits, backdoor attacks, rootkits, etc, and is based on UNIX. It can identify which files or software are dangerous and which are not. It also has special tests for Linux.

The command for installation is:

```
apt install rkhunter -y
```

Lynis

Lynis is a free, open source and extensible security auditing tool used by many security auditors across the world. Security specialists, network administrators and penetration testers are all the intended users of this tool. With Lynis, you can add your own custom test.

To install, execute the command given below:

```
sudo apt-get install lynis
```

We can also perform a basic test using the following command:

```
lynis audit system --quick
```

OSSEC

The seventh tool in this list of top security tools is OSSEC, which is basically an open source and free host based intrusion detection system. Its many special features include integrity checking, rootkit detection, time based alerting, log analysis, etc.

Run the following commands to install OSSEC:

```
apt install gcc make libevent-dev
zlib1g-dev  libssl-dev libpcre2-dev
wget tar -y
wget https://github.com/ossec/ossec-
hids/archive/3.6.0.tar.gz -P /tmp
#cd /tmp
tar xzf 3.6.0.tar.gz
cd ossec-hids-3.6.0/
./install.sh
```

Wazuh

The eighth tool is Wazuh. This is also an open source, free and enterprise wide security monitoring solution that provides a host based security system similar to OSSEC, but is much lighter in weight.

Installation commands are:

```
curl -s https://packages.wazuh.com/
key/GPG-KEY-WAZUH | sudo apt-key
add -
echo “deb https://packages.wazuh.
com/4.x/apt/ stable main” | sudo tee
/etc/apt/sources.list.d/wazuh.list
sudo apt update
sudo apt install wazuh-manager
sudo systemctl daemon-reload
sudo systemctl enable --now wazuh-
manager
```

Radare2

Radare2, which is also known as r2, is a reverse engineering and command-line based tool. It works on different operating systems.

You can install this tool by executing the command given below on a Debian based system:

```
sudo apt-get install -y radare2
```

OpenVAS

This is a part of the GVM (Greenbone vulnerability manager); it provides vulnerability scanning and vulnerability management services. The installation command is:

```
sudo apt-get install openvas9
```

Maltrail

The last top tool in this list is Maltrail. This open source high traffic detection system helps keep malware away from our system. It helps us avoid malicious threats. All the traffic sources and data from the Internet are used to compare and understand where the malware traffic is.

The installation commands are:

```
git clone https://github.com/
stamparm/maltrail.git
cd maltrail/
python sensor.py &
```

I have listed what are, according to me, the best and most famous tools that can be used on Linux systems to improve, test and audit security. Visit their websites for complete documentation and try them out. Exploring them could be a lot of fun. END

By: Jishnu Saurav Mittapalli

The author is a full stack Web developer, interested in modern technologies like artificial intelligence and machine learning. He is currently working in the field of natural language processing.

Open Source Platforms You Can Use for AR and VR



Augmented reality and virtual reality are being used in various domains and have many useful applications today, apart from gaming. A few useful open source platforms that help to implement these technologies are listed in this article.

The global market for augmented reality (AR) and virtual reality (VR) is increasing at an exponential rate. Consumers and enterprises are benefiting, as they can access and experience a remote environment in real-time.

Figure 1 depicts Statista.com's research analytics with respect to the AR and VR market, both from historical as well as future perspectives.

Virtual and augmented reality complement each other, as they enhance people's perception of their surroundings by incorporating digital elements that look like natural features. Virtual Fixtures, a system created by the US Air Force's Armstrong Laboratory in 1992, was the first fully functional AR system that enabled an immersive mixed reality experience for users.

The entertainment and gaming industries were the first to offer commercial augmented reality experiences. Many commercial businesses have adopted AR applications for education, healthcare and entertainment since then. With AR techniques, content can even be accessed by scanning or viewing a picture with a mobile device.

AR provides users with new levels of sensory stimulation. Its usage (e.g., integrating computer vision, putting AR cameras into smartphone applications, and object identification) can help you digitally modify and interact with the surroundings. An example is superimposing additional measured information like electromagnetic radio waves, in precise alignment with where they are in space, to the existing reality. There is a huge potential for AR to be used in the collection and dissemination of tacit knowledge. Real-time and semantic settings that contain environmental factors are frequent when using augmentation techniques. As an example, a live video broadcast of a sporting game can provide both immersive and perceptual information along with additional data like scores. Technologies that use both AR and heads-up display (HUD) technology can be combined in this manner.

AR vs VR

Virtual information permeates the user's experience of reality in VR. In AR, the user's view of reality is enriched by including computer-generated information inside the data gathered from the actual world.

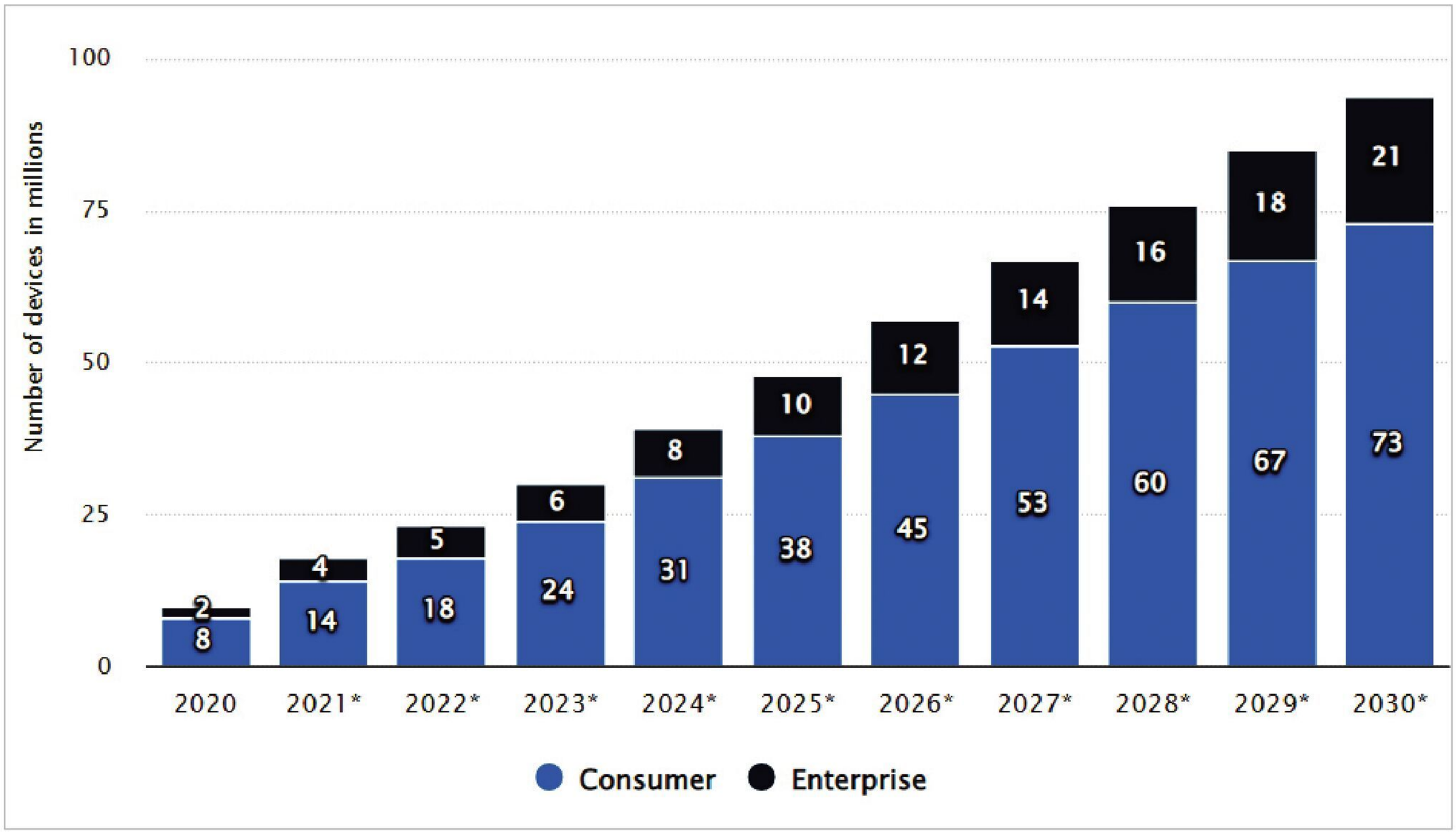


Figure 1: Market analytics for AR and VR devices (2020-2030) (*https://www.statista.com/statistics/1259882/human-machine-interface-augmented-virtual-reality-devices-worldwide*)

Architects, builders and real estate companies can utilise VR to generate a virtual tour of a new building or layout, while AR can help to illustrate the structure and systems of a building superimposed over a real-world scenario.

Key implementation domains in AR and VR

AR and VR have a wide and huge scope in education, e-governance, the corporate sector and many other domains, as listed below:

- Architecture or building plans
- Qualitative analysis in chemistry
- Assembly/disassembly of an electrical engine
- Exploration of human anatomy
- Simulation of circulation of blood in heart
- Simulation of flight/vehicle/space stations
- Building electronic circuits in zero physical defects
- Virtual classrooms for teaching and learning

Myths regarding AR and VR implementations

AR and VR can be implemented using minimum resources as well as very limited infrastructure and hardware. There are a lot of myths associated with AR and VR, some of which are listed below:

- Need for huge resources and costly hardware
- Specialised graphic cards are needed
- Specialised processors are needed
- Integration of GPU/TPU is required
- Need for high performance programming frameworks

None of the above is true.

Key features required for effective use of AR and VR

To reap the maximum benefits of AR and VR, a few key features must be integrated into the devices, including software and hardware infrastructure:

- Drag-and-drop
- Dynamic editing
- Simulation
- Real-world backdrop
- Superimposed objects
- Hardware integration
- SLAM (simultaneous localisation and mapping) support
- Cross-device publishing
- Real-time analytics
- Uploading dynamic content
- 3D object tracking
- Support for smart glasses
- Content creation
- Geolocation tracking
- Cloud storage

A few open source high performance platforms used for implementing AR and VR are listed below.

A-Frame *(aframe.io):*

- Open source platform for VR with entity component system
- Java based programming functionality for high performance implementations
- No need for any third party application/hardware
- All we need is a Web browser and JavaScript library

OpenIllusionist *(https://www.openillusionist.org.uk/home/):*

- Dynamic integration of objects
- Real-time user inputs
- Dynamic modelling
- Virtual objects analytics

ApertusVR *(http://apertusvr.org/):*

- High performance platform for AR, VR and mixed reality (MR)
- HTTP rest API integrated
- Embeddable framework-independent
- Network topology-independent

ARToolkit+ *(http://www.arreverie.com/artoolkit.html):*

- Modular open source code
- Smart glasses
- Content library
- Mobile focus on secure development
- Android and iOS support
- Content creation
- Simulation
- Superimposed objects
- Drag-and-drop
- Virtual reality integration

Holokit *(https://holokit.io/):*

- Virtual reality integration
- Real-world backdrop
- Superimposed objects
- Inside-out tracking function
- 3D objects
- Multi-degree diagonal field of view
- Content creation
- Merges the real and virtual

BRIO *(https://experience.brioxr.com/):*

- Drag-and-drop interface
- Multiple file format support
- Virtual reality integration
- 3D objects
- Content creation
- Cross-device publishing
- Content library
- Real-world backdrop
- Simulation
- Tracking and analytics
- Superimposed objects



Figure 2: Augmented and virtual reality

Adobe Aero *(https://www.adobe.com/in/products/aero.html):*

- Dynamic interactive elements
- Content creation
- Set pathways for animations
- Drag-and-drop
- Content library
- Tracking and analytics
- Simulation
- Virtual reality integration
- Cross-device publishing

Vuforia Engine *(https://developer.vuforia.com/):*

- Unparalleled accuracy
- Dynamic simulation
- Virtual reality integration
- 3D objects
- Creative empowerment
- Superimposed objects
- Content creation
- Maximum reach
- Content library
- Tracking and analytics
- Real-world backdrop
- Advanced vision
- Drag-and-drop

Scope for research and development

AR and VR help in providing a multisensory experience in which real-world items are improved with computer-generated perceptual information, including visual, aural, haptic, somatosensory, and olfactory information. Combining the real and virtual worlds, these technologies allow in-the-moment interaction, and accurately register 3D models of virtual and actual items.

Researchers and practitioners can use open source platforms of AR and VR to provide an interface between real-time experiences and remote infrastructure. There is a huge scope for research as there is no need to move physically to any location and remote points are always accessible.

Scientists, medical practitioners, surgeons, and experts in space exploration missions can use AR, VR and MR (mixed reality) to have a simulated view or remote experience of the actual infrastructure. This can be beneficial for multiple applications. END

By: Dr Gaurav Kumar

The author is a director at Magma Research and Consultancy Services, Haryana. He is associated with various academic and research institutes for delivering expert lectures and conducting technical workshops on the latest technologies and tools.

The Future Belongs to These Open Source DevOps Tools

Organisations are adopting DevOps today for their own good. This article briefly discusses a few open source DevOps tools that will be used a lot in the near future.

DevOps is a culture, and implementing its practices helps manage the application life cycle in an effective manner using automation. It has three pillars — people, processes, and tools. Tools play a critical role in DevOps or culture transformation initiatives. People and processes help in maintaining uniformity and sustainability across business units or projects in an organisation that aims for faster time to market with high quality for end users.

Category-wise list of open source DevOps tools

Category	Tools
Source code management	GitLab, Git, GitHub, Atlassian BitBucket,
Code analysis tools	SonarQube
CI/CD	Jenkins, Microsoft Azure DevOps, TeamCity, GitLab, GitHub, Atlassian Bamboo
Test automation	Selenium, Appium, JMeter
Configuration management	Ansible, Puppet, Chef
Containerisation and orchestration	Docker and Kubernetes
Cloud service providers	Amazon Web Services, Microsoft Azure, Google Cloud Platform
Monitoring	Prometheus, Grafana
Logging	Elastic Search, FluentD, FluentBit, Kibana
Communication and collaboration	Microsoft Teams, Slack, Skype and Zoom

I have used a few open source tools and I believe that these are very important for implementing DevOps practices in 2022 and beyond. Let's take a brief look at them.



Figure 1: GitLab CI/CD pipeline



GitLab

This DevOps platform provides the ability to plan, develop, secure, and operate software in a single application. Its essential features include cache and artifact management.

In my experience, GitLab has one of the most futuristic features for implementing DevOps practices.

Initial release	2014
Stable release	14.6.0
Written in	Ruby, Go and JavaScript
Licence	Community Edition - MIT License
Website	*about.gitlab.com*
GitHub repository	*https://gitlab.com/gitlab-org/gitlab-foss/*
Features	• Highly scalable • Supports collaboration • Supports on-premise and cloud installation
How is it useful in implementing DevOps practices?	DevOps platform • Version control • CI/CD • Package management • Resource management • Security • Distributed architecture • Cloud support
Can we integrate it with Pipeline as Code?	Yes. The command is: `Gitlab-ci.yaml`
Is a commercial flavour available?	Yes
Pricing	*https://about.gitlab.com/pricing/Free, Premium, Ultimate*
Community	2500+ code contributors
Use cases	• Value stream management • Agile development • Source code management • Continuous integration (CI/CD) • Out-of-the-box pipelines • Security (DevSecOps) • GitOps • The DevOps platform

Important Git commands

Description	Command
Set a user name	`git config --global user.the name "Mitesh Soni"`
Set email address	`git config --global user.email "xxxxxxxxx@xxxxx.com"`
Initialise a Git repository	`git init`
Clone a repository	`git clone https://github.com/xxxxxxxx/Angular9786-DevOps`
Verify the status of files in the repository	`git status`
Add files in the current directory for Commit operation	`git add .`
Commit files	`git commit -m "First commit in the repository"`
List your branches	`git branch`
Switch branch	`git checkout`
Create a new branch	`git branch new-branchname`
Merge current branch with the target branch	`git merge target-branch`
Show all commits	`git log`
Save modified and staged changes	`git stash`
Fetch and merge any commits from the remote branch	`git pull`

Jenkins

Jenkins is an open source tool that provides integration with existing tools used in application life cycle management to automate all activities, starting from build to test and deploy. Many beginners start their journey of automation using Jenkins. It is easy to learn and has a huge community for support.

Jenkins offers a simple way to create a pipeline that includes continuous integration and continuous delivery. Plugin based architecture helps to use existing tools in the pipeline; hence, too many changes in culture can be avoided in one go.



Figure 2: Jenkins Blue Ocean

Initial release	2011
Stable release	2.324
Written in	Java
Licence	MIT
Website	*www.jenkins.io*
GitHub repository	*https://github.com/jenkinsci/jenkins*
Features	• Jenkins installation is supported for Docker, FreeBSD, Gentoo, MacOS X, OpenBSD, OpenSUSE, Red Hat/Fedora/CentOS, Ubuntu/Debian, Windows, and Generic Java package (*.war*) • Multiple views are supported along with folders to manage jobs in an easier manner • New UI for Jenkins dashboard that is easy to use • New UI to create Pipeline as Code in an easier way • Pipeline as Code/multibranch pipeline to manage pipelines easily in repositories • Distributed architecture to manage Jenkins effectively • Role based access to users based on build jobs/folders/ pipelines/configurations/actions • Authentication - Integration with active directory, LDAP and with internal database
How is it useful in implementing DevOps practices?	Automation server helps to automate activities involved in application life cycle management
Can we integrate it with Pipeline as Code?	Yes; Pipeline as Code using Jenkinsfile
Is a commercial flavour available?	CloudBees Jenkins
Awards	• DevOps Dozen 2017: Best DevOps Open Source Project • DevOps Dozen 2016: Best DevOps Open Source Project • Bossie Awards 2014: The best open source application development tool • Geek Choice Award 2014 • SDTimes 100, 2014: DevOps & SCM • SDTimes 100, 2013: Development Tools & CI • SDTimes 100, 2012: Development Tools & CI • InfoWorld Bossies Award, 2011 • O'Reilly Open Source Award, 2011 • SDTimes 100, 2011: ALM & SCM • SDTimes 100, 2010: ALM Platforms • GlassFish Community Innovation Award 2008 • Duke's Choice Award
Third party integrations	• More than 1500 plugins to integrate Jenkins with many open source and commercial tools (*https://plugins.jenkins.io/*) ○ Platforms (iOS development, .NET, Android development, Ruby development) ○ User interface (User Interface, List view column plugins) ○ Administration (Agent controllers, Page decorators, Users and security, Cluster management, CLI extensions) ○ Source code management (SCM connections, SCM related) ○ Build management (Build triggers, Build wrappers, Build notifiers, Deployment plugins, Build parameters, Clean-up actions, Build tools, Build reports, Artifact uploaders)

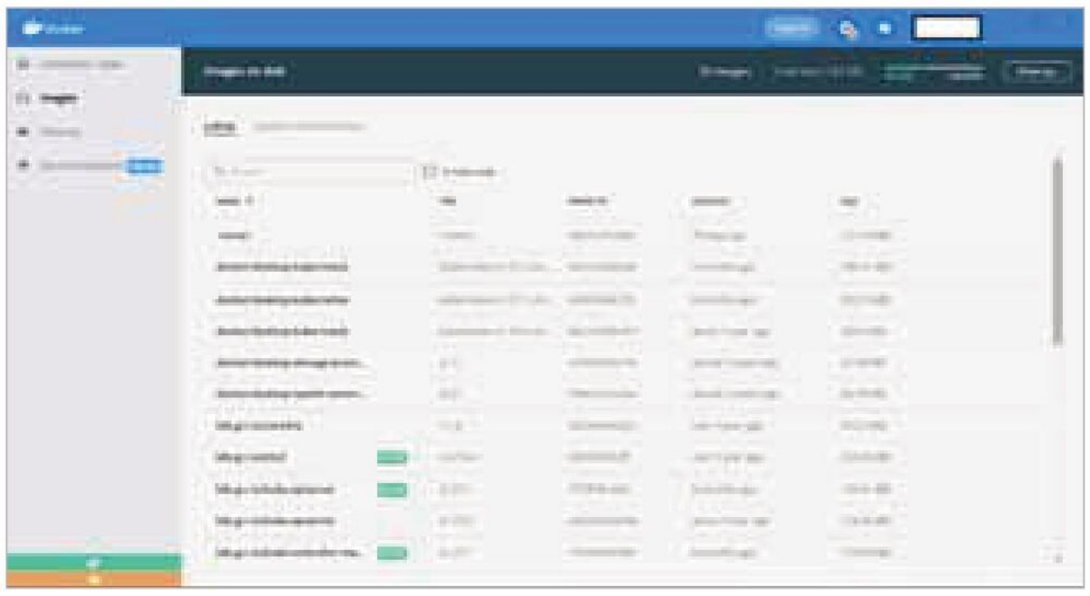


Figure 3: Docker desktop images

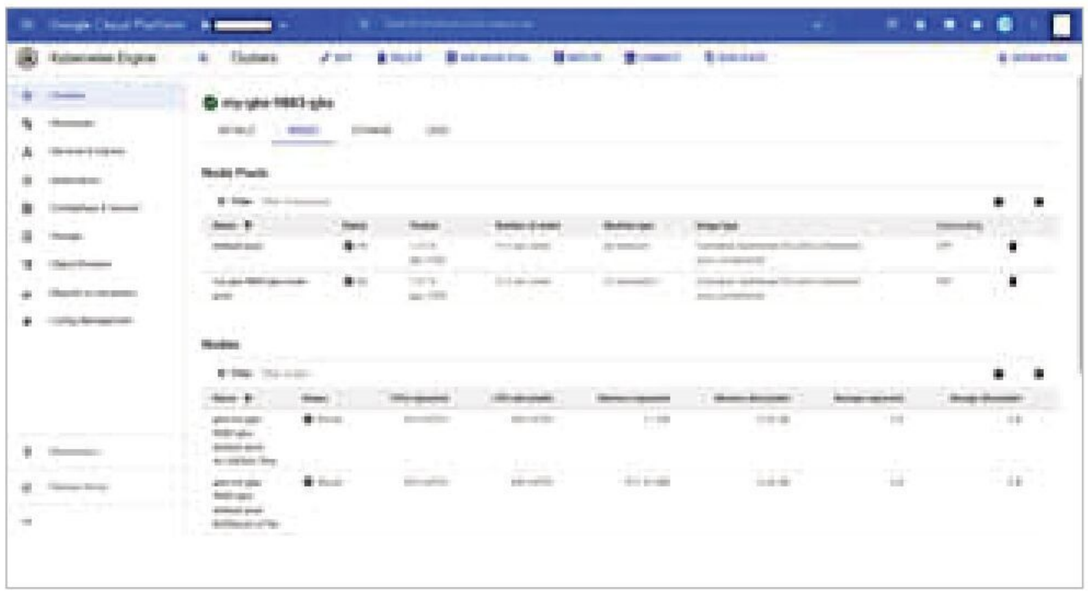


Figure 4: Google Kubernetes engine

Docker

Docker is a lightweight DevOps tool that helps to speed up and ease different workflows in your software development life cycle with a combined methodology. A Docker container image is a standalone package that contains all the dependencies required for running an application. It is an open source lightweight tool to create, deploy, and manage containers on a different host operating system using resource isolation features, such as cgroups and Linux kernels.

A container package can have application code, libraries, and configurations. The container engine is installed on the host OS.

Initial release	2013
Stable release	20.10.11
Written in	Go
Licence	Apache License 2.0 Docker Community Edition
Website	*https://hub.docker.com/* *https://www.docker.com/*
GitHub repository	*https://github.com/docker*
Features	• Easy management of applications • Docker hub – public registry • Uniform packaging • Integrates well with Kubernetes
How is it useful in implementing DevOps practices?	OS-level virtualisation, Containers
Can we integrate it with Pipeline as Code?	Yes
Is a commercial flavour available?	Docker Enterprise Edition
Docker commands	**Description** / **Command** Log in to the registry: `docker login` Build an image from the Docker file: `docker build -t sampleapp :1.0` List all the images: `docker images` Delete an image: `docker rmi imagename` Create a container: `docker run --name containerName dockerImage` Pull an image from the registry: `docker pull imageName` Push an image: `docker push imagename: tag` Clean an unused image: `docker image prune`

Kubernetes

Kubernetes or k8s is an open source, extensible and portable container orchestration platform developed by Google. It offers declarative configuration and automation for deployment. It has become one of the most popular tools for DevOps practices because of its capabilities and features for deployment, management, and scaling applications in Kubernetes clusters. It helps to manage resources effectively.

We can create a Kubernetes cluster in an on-premise environment as well as on private and public clouds. Public cloud provides managed services for Kubernetes. Google Cloud Platform has Google Kubernetes Engine, AWS has Amazon Elastic Kubernetes Services (EKS) and Azure has Azure Kubernetes Service (AKS) for managing Kubernetes clusters in a cloud environment.

Initial release	2014
Stable release	1.22.2
Written in	Go
Licence	Apache License 2.0
Website	*https://kubernetes.io/*
GitHub repository	*https://github.com/kubernetes/Kubernetes*
Features	• Distributed architecture • Easy integration with Docker • Supports hybrid, on-premise, or cloud environments
How is it useful in implementing DevOps practices?	Container orchestration and management
Can we integrate it with Pipeline as Code?	Yes
Is a commercial flavour available?	Cloud flavours are available
Important Kubernetes commands	**Description** / **Command** Get Kubectl version: `kubectl version` Get details on Kubernetes cluster: `kubectl cluster-info` Get details about a node: `kubectl describe nodeName` Get details about a pod: `kubectl describe podName` List of all pods: `kubectl get pods` List of all services: `kubectl get services` List of all deployments: `kubectl get deployments`

Prometheus

This is an open source tool for monitoring and alerting. The Prometheus ecosystem has components such as a Prometheus server, client libraries, a push gateway, exporters for services, an alert manager, and support tools.

Prometheus has features such as a multi-dimensional data model, a flexible query language, and multiple modes of graphing and dashboarding support.

Initial release	2012
Stable release	v2.32.0
Written in	Go
Licence	Apache License 2.0
Website	*https://prometheus.io/*
GitHub repository	*https://github.com/prometheus/prometheus*
Features	• A multi-dimensional data model • A flexible query language (PromQL) • Support for service discovery or static configuration
How is it useful in implementing DevOps practices?	Monitoring
Is a commercial flavour available?	Prometheus is 100 per cent open source and community-driven

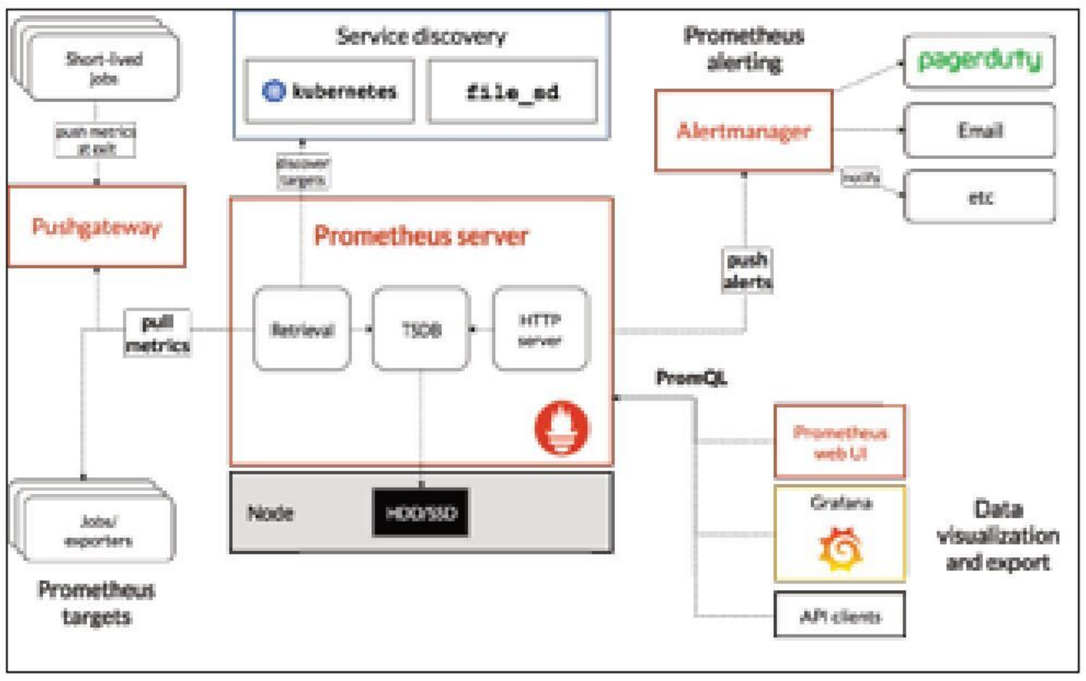


Figure 5: Prometheus architecture
(Source: *https://prometheus.io/docs/introduction/overview/)*

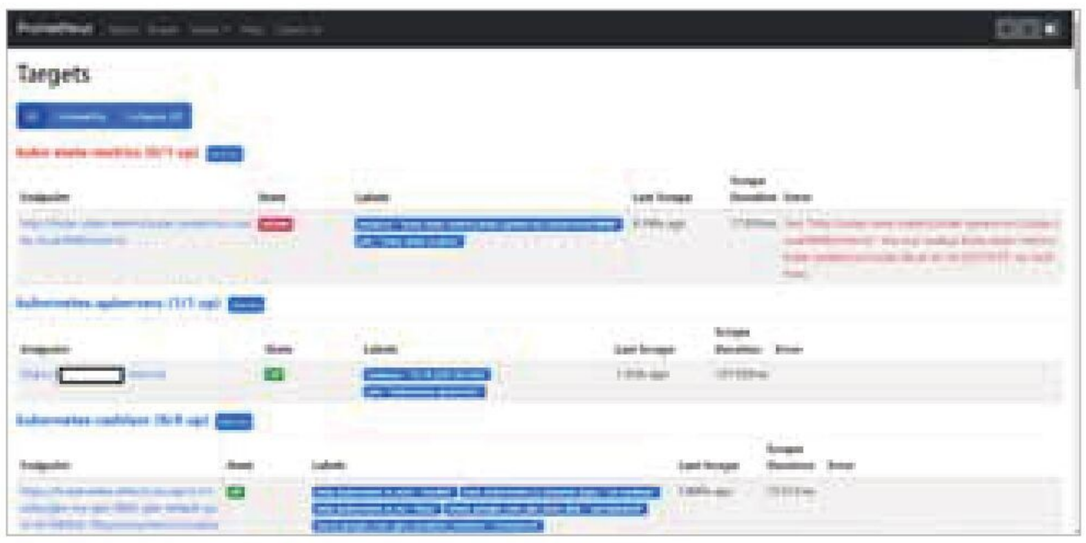


Figure 6: Prometheus dashboard

Terraform

This open source 'Infrastructure as Code' tool, developed by HashiCorp, helps in creating immutable infrastructure to avoid the problem of configuration drift. Terraform is a platform-agnostic declarative coding tool that helps developers to use a high-level configuration language to manage infrastructure. It supports different cloud providers like AWS, Azure, Google Cloud Platform, and so on.

Initial release	2014
Stable release	1.0.10
Written in	Go
Licence	Mozilla Public License v2.0
Website	*https://www.terraform.io/*
GitHub repository	*https://github.com/hashicorp/terraform*
Features	• Manages Infrastructure as Code using declarative config files • Dependency graphs • State management
How is it useful in implementing DevOps practices?	Infrastructure as Code
Can we integrate it with Pipeline as Code?	Yes. It can be part of Pipeline as Code, and different strategies can be used to create resources and deploy applications in different cloud platforms.
Important Terraform commands	See below
Tutorials	*https://learn.hashicorp.com/terraform*

Description	Command
Verify Terraform version	`terraform --version`
Install Terraform modules	`terraform init` `terraform init -input=false`
Check configuration and create a plan	`terraform plan`
Execute the plan	`terraform apply` `terraform apply -auto-approve`
Delete all resources	`erraform destroy` `terraform plan -destroy`
Validate/check the syntax of the Terraform files	`terraform validate`

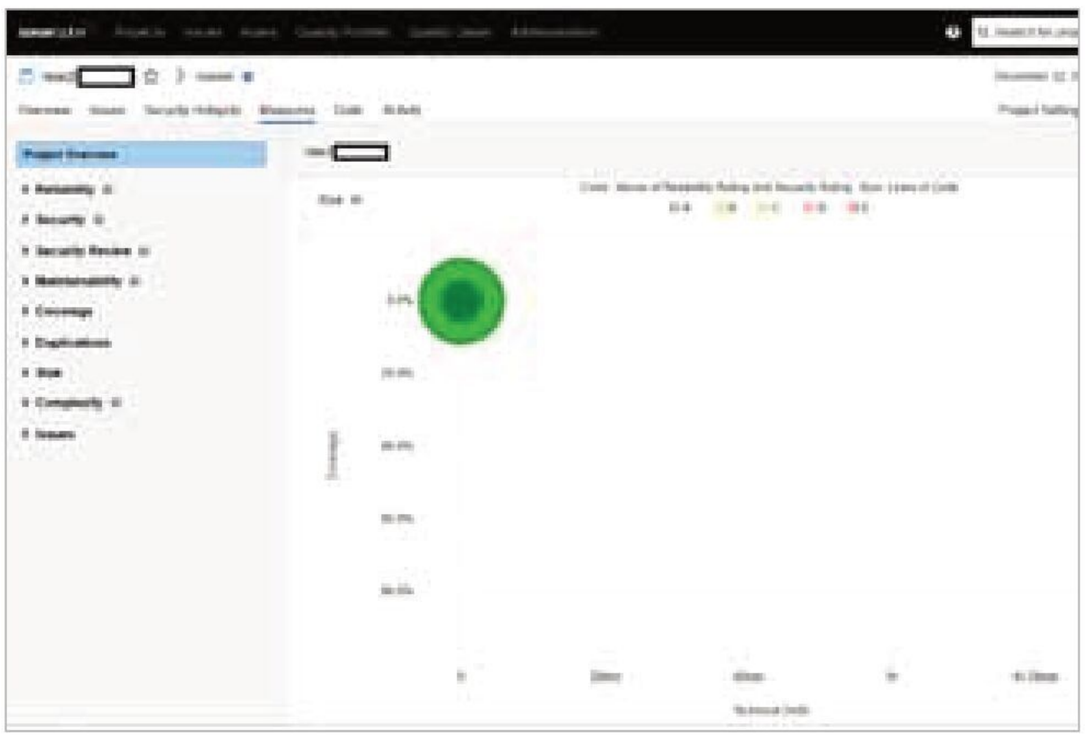


Figure 7: SonarQube dashboard

SonarQube

SonarQube works like a sanitiser and helps clear bugs, vulnerabilities and code smell in the code. It supports more than 20 programming languages for static code analysis (SCA).

It is much easier to integrate static code analysis using SonarQube in Pipeline as Code. This tool helps to transform the culture of an organisation from manual to automated code review using quality gates and profiles.

Initial release	2006-07
Stable release	8.9
Written in	Java
Licence	Lesser GNU General Public License
Website	*https://www.sonarqube.org/*
GitHub repository	*https://github.com/SonarSource/sonarqube*
Features	• Support for the programming languages Java (including Android), C#, C, C++, JavaScript, TypeScript, Python, Go, Swift, COBOL, Apex, PHP, Kotlin, Ruby, Scala, HTML, CSS, ABAP, Flex, Objective-C, PL/I, PL/SQL, RPG, T-SQL, VB.NET, VB6, and XML • Quality gates • Quality profiles
How is it useful in implementing DevOps practices?	Static code analysis
Can we integrate it with Pipeline as Code?	Yes
Is a commercial flavour available?	Yes

Grafana

This open source tool is fully managed by Grafana cloud, but you can run the enterprise version on your infrastructure with self-managed options.

Grafana is one of the most popular tools when we are using Kubernetes. It is easy to install, configure and manage Grafana and monitoring metrics keeping Prometheus as a data source.

Initial release	2014
Stable release	8.3.3
Written in	Go, TypeScript
Licence	GNU Affero General Public License, version 3.0
Website	*https://grafana.com/*
GitHub repository	*https://github.com/grafana/grafana*
Features	• Panels - heatmaps to histograms • Plugins - https://grafana.com/grafana/plugins/ • Alerts • Annotations
How is it useful in implementing DevOps practices?	Visualisation of monitoring metrics
Is a commercial flavour available?	*https://grafana.com/pricing/*



Figure 8: Grafana dashboard

Ansible

Ansible is an open source IT automation/configuration management tool that provides support for cross-platforms. It is easy to use, and provides security and reliability for safe adoption in an organisation.

It is also popular for deployments and rolling updates in release management activities. Ansible is different from other popular configuration management tools as it is agentless. It can be integrated with Kerberos, LDAP, and other authentication management systems.

Initial release	2012
Stable release	5.0.1
Written in	Python, PowerShell, Shell, Ruby
Licence	GNU General Public License
Website	*https://www.ansible.com/*
GitHub repository	*https://github.com/ansible/ansible*
Features	• Distributed architecture • Easy integration with Docker • Supports hybrids, on-premise, or cloud environments
How is it useful in implementing DevOps practices?	• Configuration management tool • Infrastructure as Code
Can we integrate it with Pipeline as Code?	Yes
Is a commercial flavour available?	Yes *https://www.ansible.com/products/pricing*

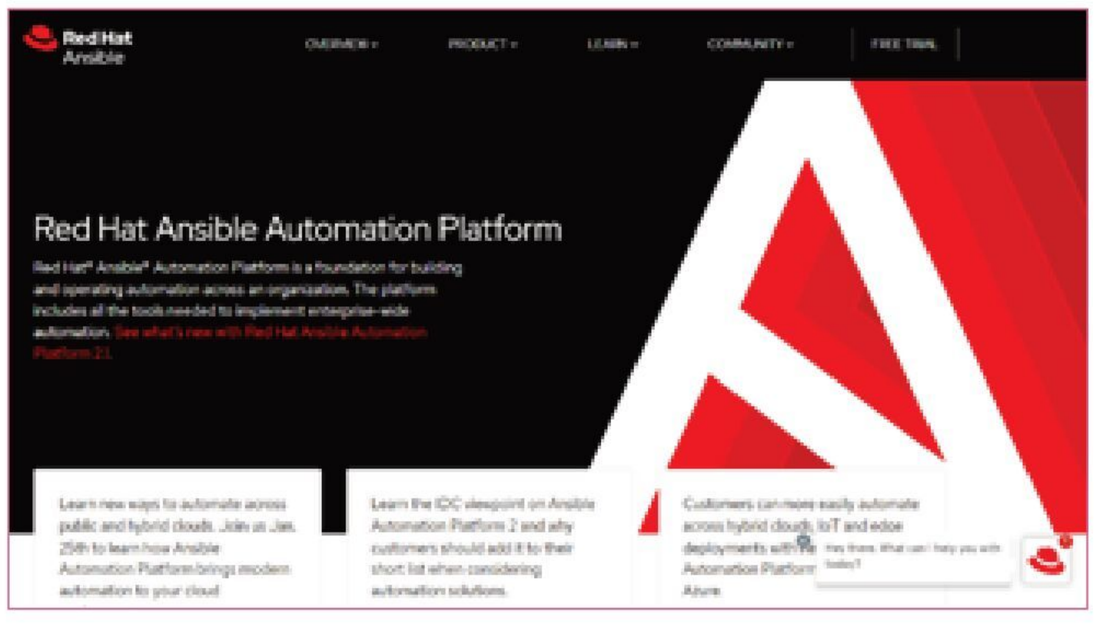


Figure 9: Ansible

Trivy

Trivy is an open source yet simple and comprehensive vulnerability/misconfiguration scanner for containers and other artifacts.

It detects vulnerabilities of the following:

- OS packages: Alpine, RHEL, CentOS, etc
- Language-specific packages: Bundler, Composer, npm, yarn, etc.
- Infrastructure as Code (IaC): Terraform and Kubernetes

Stable release	V0.22.0
Written in	Go
Licence	Apache-2.0 License
Website	*https://aquasecurity.github.io/trivy/v0.22.0/*
GitHub repository	*https://github.com/aquasecurity/trivy*
Features	• Vulnerability detection for Docker Images, Red Hat Enterprise Linux, CentOS, Oracle Linux, Debian, Ubuntu, Amazon Linux and other OS packages • Simple, fast and easy to install
How is it useful in implementing DevOps practices?	• To find vulnerabilities of Docker images • DevSecOps
Can we integrate it with Pipeline as Code?	Yes. The pipeline stage can be used to analyse Docker images and publish the report.
Example	`trivy image python:3.4-alpine`
Docker image	`docker pull aquasec/trivy:0.22.0`

END

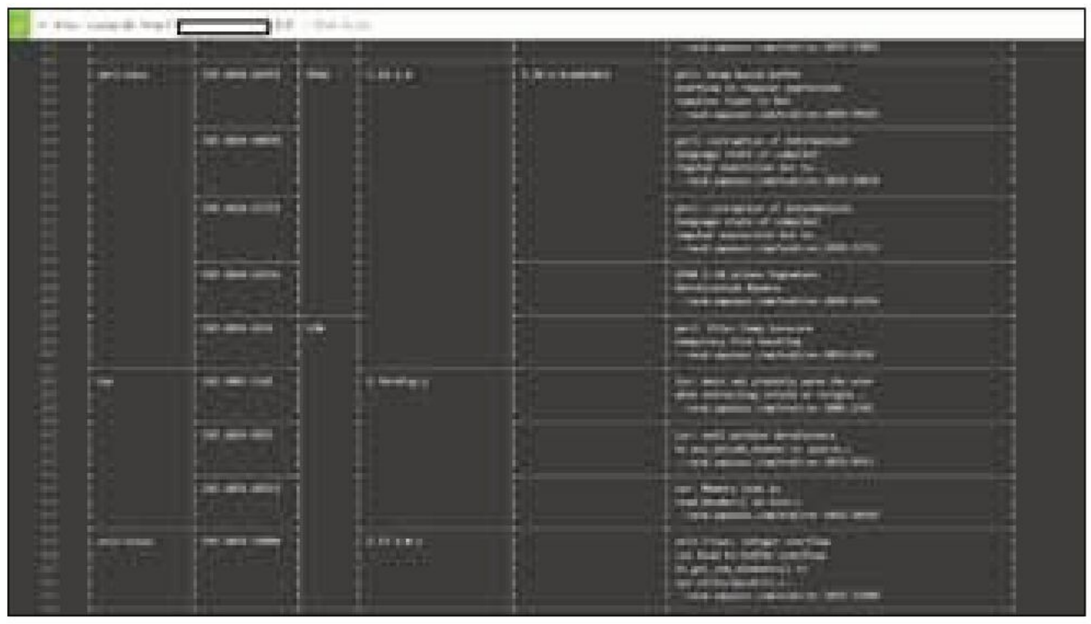


Figure 10: Trivy Docker image analysis

References

[1] Hands-on Pipeline as Code with Jenkins, *https://www.amazon.com/gp/product/B08WJ99GJF*

[2] Hands-on Pipeline as YAML with Jenkins, *https://www.amazon.com/gp/product/B0976SB2YG*

[3] Agile, DevOps and Cloud Computing with Microsoft Azure, *https://www.amazon.com/gp/product/B07WXJ4TD9*

[4] Hands-on Azure DevOps, *https://www.amazon.com/gp/product/B08BLG5QM4*

[5] GitLab, *https://en.wikipedia.org/wiki/GitLab*

[6] Jenkins, *https://en.wikipedia.org/wiki/Jenkins (software)*

[7] Overview of Prometheus, *https://prometheus.io/docs/introduction/overview/*

By: Mitesh Soni

The author has written the books 'Hands-on Azure DevOps, Agile, DevOps and Cloud Computing with Microsoft Azure', 'Hands-on Pipeline as Code with Jenkins', and 'Hands-on Pipeline as YAML with Jenkins'.

OSFY Magazine Attractions During 2021-22

Month	Theme
March 2021	AI, Deep learning and Machine Learning
April 2021	DevOps Special
June 2021	Cyberscurity, Open Source Firewall and Network security and Monitoring
July 2021	Mobile/Web App Development and Optimisation
August 2021	Blockchain and Open Source
September 2021	Data Management, Storage and Backup
October 2021	Cloud Special: BigData, Hadoop, PaaS, SaaS, Iaas and Cloud, Automation in Cloud
November 2021	Open Source on Windows
December 2021	Database management and Optimisation
January 2022	Open Source and IoT and Edge
February 2022	Alternatives: Best in the world of Open Source (Tools and Services)

Top 10 Open Source Deep Learning Tools



Deep learning is deeply intertwined with our lives today, and promises to enter many more domains in the near future. This article briefly describes the best open source deep learning tools.

Alexa on your side table playing your favourite song, autonomous cars on the road, or Netflix suggesting shows based on what you have watched...the power of deep learning is evident all around us! A number of open source deep learning tools are used in this field. Let us take a brief look at some of the best ones.

Table 1: Popular deep learning tools and their developers

Deep learning tool	Developed as a part of
TensorFlow	Google Research
Keras	Open-ended Neuro-Electronic Intelligent Robot Operating System (ONEIROS)
PyTorch	Facebook's AI Research (FAIR) lab
OpenNN	ArteInics
CNTK	Microsoft Research
MXNet	Apache Software Foundation (ASF)
DeeplearningKit	Memkite
Deeplearning4J	Eclipse Machine Learning Group
Darknet	Joseph Chet Redmon
PlaidML	Vertex AI, acquired by Intel

Top 10 open source deep learning tools

1. TensorFlow

TensorFlow *(www.tensorflow.org)* is an open source deep learning framework for the development of machine learning and deep learning models, founded on November 9, 2015.

Features

- Provides an end-to-end model, from building to deployment
- Supports model deployment on mobile devices and embedded devices also
- Good documentation and help from the community
- Extensive support for multiple GPUs
- Graph visualisation and queues
- Supports image processing, computer vision, natural language processing, text classification, speech recognition, etc

2. Keras

Keras *(keras.io)* is an open source deep learning library in Python, released on March 27, 2015.

Features

- Extensive documentation and developer guides
- Easy to learn and use
- Offers clear and comprehensive error messages
- Supports models for mobile devices, the Web as well as for Java Virtual Machine
- Enables distributed deep learning models on graphics processing units (GPUs) and tensor processing units (TPUs)
- Broadly adopted in the research community as well as in the industries

3. PyTorch

PyTorch *(pytorch.org)* is an open source machine learning library



Figure 1: Diverse deep learning applications

released on September 2016 and authored by Adam Paszke, Sam Gross, Soumith Chintala, and Gregory Chanan. The essential elements (cheat sheet) for new users are available at *pytorch.org/tutorials/beginner/ptcheat.html.*

Features

- Broadly accepted across major cloud platforms to provide easy development and straightforward scalability
- Facilitates end-to-end pipeline flow, from Python development to mobile device deployment (iOS and Android)
- Active research community and industry developers provide robust ecosystem libraries that enable development in various domains — from computer vision to reinforcement learning
- Provides direct interface with ONNX (open neural network exchange) compatible systems, environments, visualisers, and more, by exporting models in the standard ONNX format
- Offers multiple GPU support for the implementation of deep learning models
- Supports C++ interface for high performance and low latency applications
- Supports data sharing with external libraries

4. OpenNN

OpenNN *(www.opennn.net)* is an open source neural networks library for machine learning and deep learning. Its first version (0.1) was released on November 22, 2018.

Latest applications that use deep learning

- Autonomous vehicles: Tesla and Waymo
- Polybee: Pollinating bee drone in greenhouses
- RNNoise: RNN based noise suppression library
- Grammarly: NLP based text summarisation
- Perlmutter: An Nvidia project that targets 3D mapping of the universe
- MERIDIAN: Helping compile a catalogue of all known fish sounds and creating a benchmark data set for marine bioacoustics
- Neurala Brain: Neural network and software that emulate the functions of the human brain for navigation and perception
- Zeroeyes: Real-time fire arms detection, prevention (police) and protection
- WaveNet: Generates raw audio from speech patterns using CNNs
- 6sense: A B2B sales and marketing platform that predicts buying behaviour
- Robbie AI: Patient monitoring platform to prevent falls and delirium in ICU patients

Features

- It allows users to create a neural network model without programming
- Offers a good user interface in terms of data management and result interpretation using the neural designer tool
- Provides faster training compared to other deep learning frameworks
- Outstanding performance in terms of execution speed and memory allocation

5. Microsoft Cognitive Toolkit (CNTK)

Microsoft Cognitive Toolkit (CNTK) is a commercial grade open source deep learning framework. Its primary release was on January 25, 2016. For more details, visit *docs.microsoft.com/en-us/cognitive-toolkit.*

Features

- CNTK APIs are available in Python, C# as well as C++. Also provides its own model programming language such as BrainScript
- Evaluation of the CNTK model feature is supported by Java Apps
- Provides low-level and high-level APIs
- Provides an easy and simple way to combine various deep learning models like a convolutional neural network, recurrent neural network, and deep feed-forward neural network
- Autonomous differentiation and parallelisation over several GPUs and servers
- Supports implementation of supervised learning, unsupervised learning, reinforcement learning, as well as generative adversarial networks

6. Apache MXNet

Apache MXNet *(mxnet.apache.org)* is an open source deep learning framework released (version 0.7.0) on May 26, 2016.

Features

- Provides hybrid front-end with seamless transition between Gluon eager imperative mode and symbolic mode; high speed and flexibility
- Supports dual parameters server and Horovod (distributed deep learning framework) to offer scalable distributed training and performance optimisation
- Provides comprehensive and flexible Python APIs and supports other languages such as Scala, Julia, Clojure, Java, C++, R and Perl
- Supports computer vision, natural language processing, text classification, and timeseries

7. DeepLearningKit

DeepLearningKit is an open source deep learning framework developed for the Apple iOS, OS X, and tvOS. For more details, visit *github.com/DeepLearningKit/DeepLearningKit.*

Features

- Developed using Metal and Swift for GPU acceleration and app integration, respectively
- Uses convolutional neural networks for image recognition on Apple devices

8. Deeplearning4J

Deeplearning4J is an open source deep learning framework that was primarily released in 2014, with a stable release on May 13, 2020. The main authors are Alex D. Black, Adam Gibson, Vyacheslav Kokorin, and Josh Patterson. For more details, visit *deeplearning4j.konduit.ai.*

Features

- Supports all the Java virtual machine based languages such as Java, Scala, Clojure, and Kotlin
- Provides a distributed deep learning framework by felicitating training in a cluster. The framework is supported by distributed systems Apache Spark and Hadoop
- Has a designed vector space model and topic model to manage large size text sets and perform natural

Table 2: A comparison of open source deep learning tools

Tool	Licensed by	Written in	CUDA support
TensorFlow	Apache 2.0	C++, Python, CUDA	Yes
Keras	MIT	Python	Yes
PyTorch	BSD	Python, C, C++, CUDA	Yes
OpenNN	GNU LGPL	C++	Yes
CNTK	MIT	C++	Yes
MXNet	Apache 2.0	Small C++ core library	Yes
Deeplearning-Kit	Apache 2.0	Metal, Swift	No
Deeplearning4J	Apache 2.0	C++, Java	Yes
Darknet	YOLO	C, CUDA	Yes
PlaidML	Apache 2.0	Python, C++, OpenCL	No

language processing tasks smoothly
- Includes implementation of the restricted Boltzmann machine, deep belief networks, deep autoencoder, recursive neural tensor network, stacked denoising autoencoder, word2vec, doc2vec, and GloVe.

9. Darknet

Darknet is an open source neural network framework primarily released in 2013. For more details, visit *pjreddie.com/darknet/*.

Features

- The framework is written in C and CUDA to provide good performance
- Fast and easy installation
- Supports CPU and GPU computations
- Facilitates image classification, time series prediction, natural language processing, etc

10. PlaidML

PlaidML is an open source tensor compiler for the machine learning framework, released on October 20, 2017. For more details, visit *github.com/plaidml* or *www.intel.com/content/www/us/en/artificial-intelligence/plaidml.html*.

Features

- Added nGraph compatibility support for new GPUs and platforms
- Offers modular hardware support from embedded to new processors
- Works on major operating systems such as Linux, macOS and Windows END

By: Geetali Saha and Dr. Safvan Vahora

Geetali Saha is a member of the faculty of the Department of Electronics and Communication, G.H. Patel College of Engineering and Technology (GCET), Gujarat.

Dr. Safvan Vahora is Assistant Professor, Government Engineering College, Modasa, Gujarat.

FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK

We Value Your Feedback

We love to hear from you as **OSFY** consistently strives to make its content **informative, interesting** and **engaging.**



Please share your feedback/ thoughts/ views via email at osfyedit@efyindia.com.

We welcome your comments/ suggestions and encourage you to send them to:
The Editor, D-87/1, Okhla Industrial Area, Phase 1, New Delhi-110020.

OpenSource For You
THE COMPLETE MAGAZINE ON OPEN SOURCE

FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK FEEDBACK

CODE SPORT



Sandya Mannarswamy

In this month's column, we discuss a medley of computer science interview questions.

Last month, we had discussed named entity recognition and relation extraction problems in natural language processing. In this month's column, we take a break from this topic and feature a bunch of interview questions on algorithms, data structures and machine learning.

1. You are given a data set for a binary classification problem. The input data is a set of features associated with that problem and the ground truth is a binary label 0/1. Your initial solution gave you a test set F1 score of 0.8. In order to improve this F1 score, you decided to try to improve the performance by creating additional features – using polynomial functions of the input variables. You expected your new solution to give a better F1 test set score. But you got only 0.69. Can you explain what may be causing this?
2. How do you decide what may be a good activation function for the deep learning model you are building? There are a number of activation functions – RelU, Tanh, Leaky RelU, Swish etc. How do you choose the right activation function?
3. What is a bagging classifier and what is a stacking classifier? If you have a handful of weak base classifiers for a binary classification problem, will you combine them using a bagging classifier or a stacking classifier? Explain the reason behind your choice.
4. Let us assume that you decided to use a bagging classifier for a binary classification problem. You decided to combine the classifiers using majority voting. Will you combine them using majority voting on the binary labels or by using majority voting on the class probabilities? Or will you use majority voting on the weighted average of class probabilities? Explain the reason behind your choice.
5. We all know about classifier ensembles. You are given only one neural network model based on, say, BERT. Now you are asked to train this model and build an ensemble classifier. Explain how you can do that. This type of ensemble classifier built using a single neural network topology is known as snapshot ensemble classifier.
6. Explain the K-fold cross validation and stratified cross validation approaches. When will you choose one over the other? And how does one choose the value of K in K-fold cross validation?
7. When designing a convolutional neural network (very often used for image processing problems) layer, in addition to filter sizes, there are two additional hyper parameters – namely, stride and padding. Can you explain what these parameters are and how to select their value?
8. We are all familiar with the stochastic gradient descent method for training neural networks. Let us assume that you have two training schedules, one where the batch size is 32 samples and another where it is 256 samples. Which of these training schedules will converge faster (i.e., will require lesser number of epochs to reach the specified loss threshold)? Can you explain the reason behind your answer?
9. Can you explain the advantage of a masked language model as opposed to a standard left to right language model?
10. You have been given all the stock market transactions made by a user in his stock exchange account, while buying and selling shares. Each transaction consists of the details on: (a) name of the company, (b) type of transaction — buy/sell, (c) number of shares bought/sold, (d) price of each share bought/sold, and (e) date of the transaction. You also have access to stock market information in terms of the price of each company's stock at the end of every day. The user would like to know which of his shares he can sell at a profit of 10 per cent or above on each day. You are asked to write a program, which examines the closing price

of stocks the previous day, and recommends which of his equity holdings the user can sell that day for a profit of 10 per cent or more. Your recommendation should contain the company name, number of shares to sell, the sale price and the percentage of profit that the user will make if he sells at that price. What is the time complexity of your algorithm, assuming that you are given N transactions in total, corresponding to M different companies?

11. Consider the above question. Stock markets are quite volatile and the price of company stocks can fluctuate considerably during the day. You are now told that instead of using the previous day's closing price, your application should read in real-time the current price of the stocks that the user holds, and then recommend which stocks the user can sell at that point of time for a 10 per cent profit. Assume that you have stock market prices available as a real-time feed at every 60 second interval. How will you modify your solution for question (10) to satisfy this criterion?
12. You are given a search query and you can use any of your favourite Web search engines to find the query results. Web search engines use their own ranking algorithm to rank the query search results. However, you are asked to apply a filter on top of the search engine query results such that the results are ordered temporally. For example, consider an example query of the form: 'Donald Trump US Presidential election'. Given a set of search engine query results for this query which will contain news articles, events, etc, you need to sort them temporally. For instance, a query result link on the news story from July 2015 describing Trump's intention to run should be ranked before the news story featuring his 2016 New Year wishes to the American electorate. Essentially, your task is to organise the search engine query results on a timeline. What are the challenges here?
13. You are given four strings S1, S2, S3 and S4. S1 is a large string and S2, S3, and S4 are comparatively small strings. You are asked to find the smallest length sub-string 't' of S1 such that it contains S2, S3 and S4 as its sub-strings. For example, given string S1 'BytheriversofBabylon', and S2=er, S3=of and S4=so, your function should return 'ersof'. How will you modify your solution if you are now told that the ordering of S2, S3, and S4 should be maintained in the sub-string 't' (i.e., S2 should appear before S3, and S3 should appear before S4 in sub-string 't')?
14. When we input a query into a search engine such as Google, we get an 'auto complete' suggestion for it based on past queries that have been tried on the search engine. Can you explain how you will design an auto-complete mechanism that will scale well across the millions of queries that can happen concurrently at a time? What is the data structure that you will use which can store past queries and automatically provide suggestions? Will you store the complete query as is or will you break it into its constituent words? Given that scalability and correctness is key for this problem, can you explain how the choice of your data structure will address both?
15. You are given a set of documents and asked to perform their clustering based on document level similarity. You are familiar with different clustering algorithms such as 'kmeans', so you are confident that you can do the clustering. However, the interviewer now adds a second part to the problem where he wants you to label the clusters. How will you generate labelled clusters? What are the means of generating meaningful labels for the clusters? Can the clusters be labelled post the clustering process, prior to it, or during the cluster process itself? Describe your choice of the solution and why you made it.
16. You are given an array A of N integers. You are also given a number X. You are asked to find the element in A that is closest to X. You are given the constraint that you should not sort the array A. (a) Can you write a C program to solve this problem? (b) Now you are told that the set S consists of floating point values instead of integers. How will you modify your solution for (a) to solve (b)?
17. Given a classifier, can you define the terms 'sensitivity' and 'specificity' of your system? Let us assume that you have built a classifier that can predict whether a tumor is benign or malignant based on certain tumor characteristics. Your intention is to build a classifier that can detect all true malignant tumors, even if some of the benign tumors get falsely reported as malignant. Which of the two metrics — sensitivity or specificity — should you try to improve in this case? Now assume that you have been asked to use a classifier to predict whether a student will pass the examination or not. Your system should be designed such that it does not report a failing student as passing even if it mispredicts certain passing students as potential failure students. In that case, which of these two metrics should you try to improve?
18. What is a maximum margin classifier? Can you give an example? What is the advantage of using this classifier?
19. Graph neural networks are becoming popular in machine learning and natural language processing. Can you explain the details of a graph convolutional network (GCN)? How is it different from a standard convolutional network (CNN)?
20. While transfer learning has become very popular and common in both computer vision and NLP, can you list some scenarios where it may not be useful and you may need to train a model from scratch?

Feel free to reach out to me over LinkedIn/email if you need any help in your coding interview preparation. If you have any favourite programming questions/software topics that you would like to discuss on this forum, please send them to me, along with your solutions and feedback, at *sandyasm_AT_yahoo_DOT_com*. Wishing all our readers a very happy new year 2022! Stay healthy and stay safe. END

By: Sandya Mannarswamy

The author is an expert in natural language processing and is currently working as an independent researcher. Her interests include natural language processing, machine learning and AI.

Use Node-RED to Build Your Brand



Online media monitoring helps businesses to build their brands and manage the way their products and services are being perceived in the market. This article gives a short tutorial on how to build an automated online media monitoring system using Node-RED.

Online media monitoring refers to the process of scanning open or public data on digital platforms by looking for keywords on a particular topic. It helps you to manage the way your product or service is perceived in the market. Even though many premium media monitoring tools are available and come with a free trial period, the 'feedparser node' provided in Node-RED makes it an advisable alternative to the commercial ones. This node monitors an RSS (really simple syndication) or Atom feed for new entries. RSS refers to simple text files that contain updated information like news pieces, articles and dialects. An RSS feed first takes the headline and summarises it. Then it updates notices, and links them back to articles on the required website. This Web feed allows users and applications to access website updates in a standardised, computer-readable format. Figure 1 shows the relationship between RSS and Node-RED.



Figure 1: RSS and Node-RED

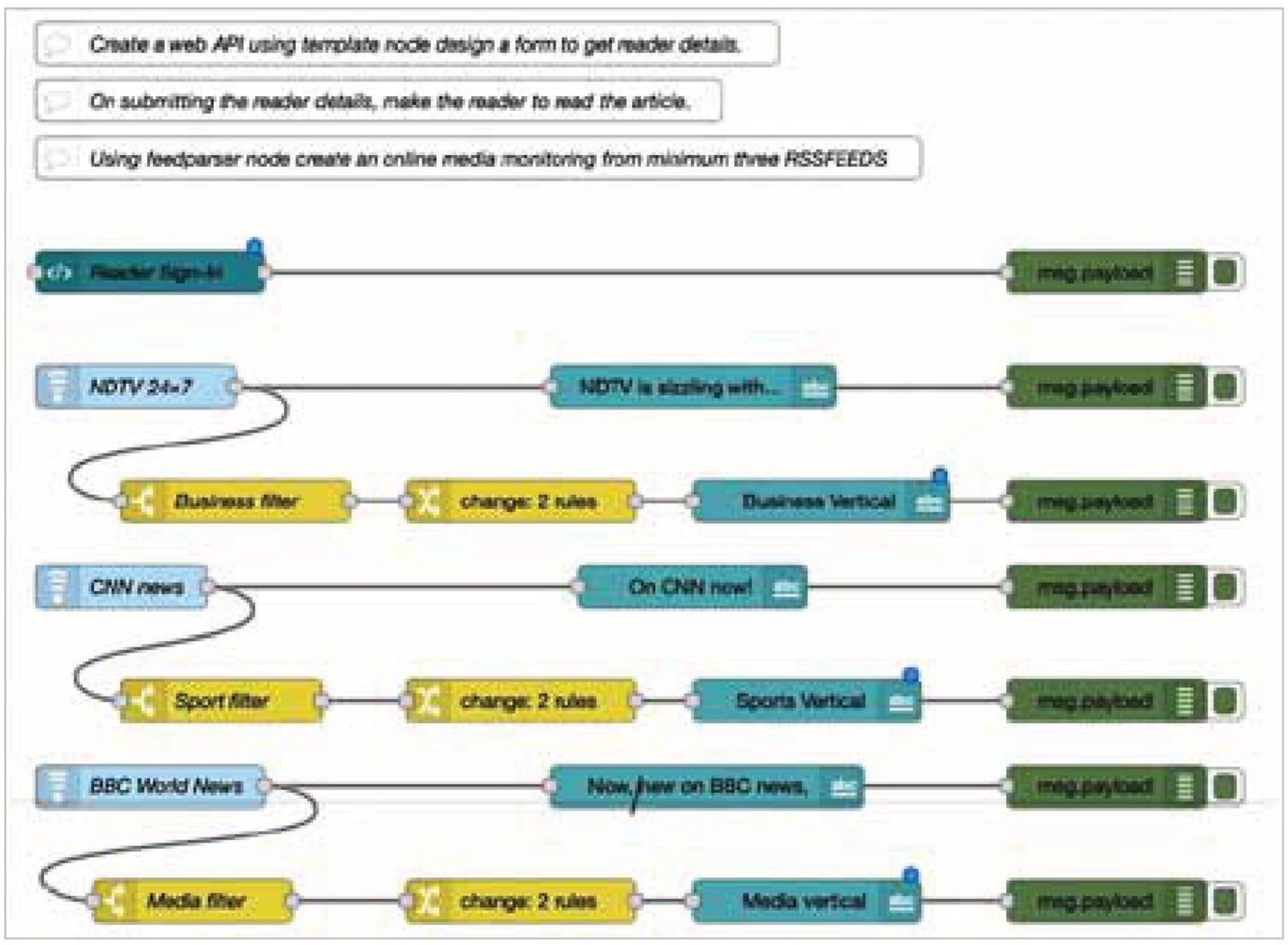


Figure 2: Complete Node-RED flow of the online media monitoring system



Figure 3: *Feedparser* node property configurations

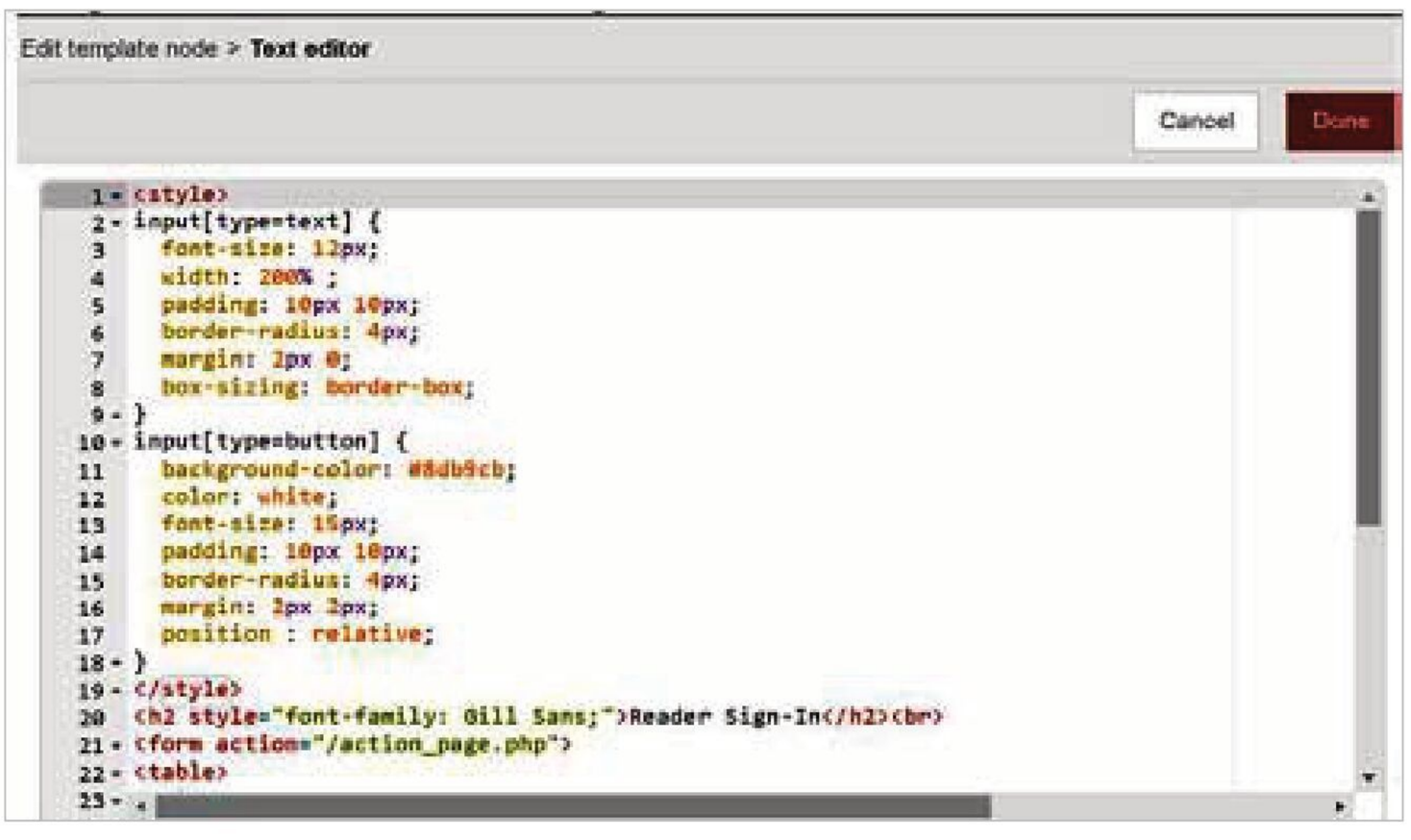


Figure 4: Template node property configurations

This article focuses on creating an automated online media monitoring system using Web APIs (application programming interfaces). This is done by incorporating a *template* node to design a user form that gets readers' details. After the details are submitted, the system enables the reader to read the article. Using the *feedparser* node, this automated system is created for three RSS feeds for the popular news channels BBC, NDTV and CNN. From every feed, news on three verticals (such as sports, economics, and education) is filtered. The outputs are displayed in the Node-RED debug monitor and UI dashboard

Procedure to construct and build the flow using Node-RED

- Install the latest version of Node.js from *https://nodejs.org/en/download/*
- Install Node-RED through the command window by giving the following command:

```
npm install -g node-red
```

- Through the browser, get access to the Node-RED flow using the IP address *http://127.0.0.1:1880*
- From the 'Node palette' of Node-RED, insert the appropriate nodes, configure their properties, and wire them.
- Confirm deployment by clicking on the 'Deploy' option.
- View the output through the 'Debug' monitor window and dashboard UI.

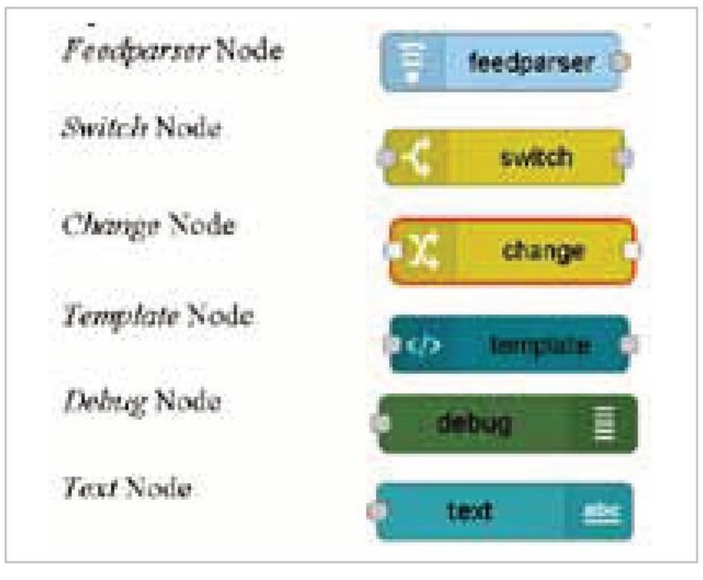


Figure 5: Nodes for online media monitoring

Online media monitoring – node flow design

1. Install and add the *feedparser* node from *Manage palette* using the following command:

```
node-red-node-feedparser
```

2. Drag and drop *feedparser* node from the *Node palette*.
3. Browse the various RSS feeds of media companies like BBC, NDTV and CNN.
   - Paste the RSS feed URL in the *Feed URL* box.
   - Customise the refresh time if needed.
4. Connect the feed parser to a *Text* dashboard widget so that the XML is converted to JS object, and is displayed properly as text.
5. Add a *switch* node and configure the message.
6. Connect the *feedparser* node to *switch* node to change payload value to the news feed article value.
7. Type in the word you want to find in the switch box and set configuration to find regex (regular expression).
8. Click on the *Ignore uppercase dialogue* box to include matches regardless of the uppercase and add the change node.
9. The *change* node is to be configured in such a way that the *msg.payload* is set to *msg.article. title.* Set a no reply email in the *msg.from object.* This *feedparser* node is connected to a *switch* node and *change* node in order to be used as a news filter, i.e., to read a particular section of the news.
10. Apply the same rules for other news channels to add them with different titles or headlines like sports, education, and finance, and follow what was done above.
11. The *template* node is configured with HTML code to create a form that requires readers to sign in before they can continue reading. Connect the *function* node and the *template* node to create a media monitoring UI using HTML/CSS code and display the news articles. This *template* node helps in creating a form that asks for user details, which has to be filled in before viewing the news feed.
12. Insert a *text* node and set the data to be printed as the title of the article.
13. Insert *debug* nodes appropriately to view the output in the debug monitor. After creating the complete flow, click on 'deploy' from the top right corner of the window and deploy the system.
14. View the results in the debug console. To check each node's output you can use node filters.
15. To view the user interface, go to the link *127.0.0.1:1880/ui/.* Enter the required details and click on 'Submit'. It will take you to the next page, where all the news articles can be read.

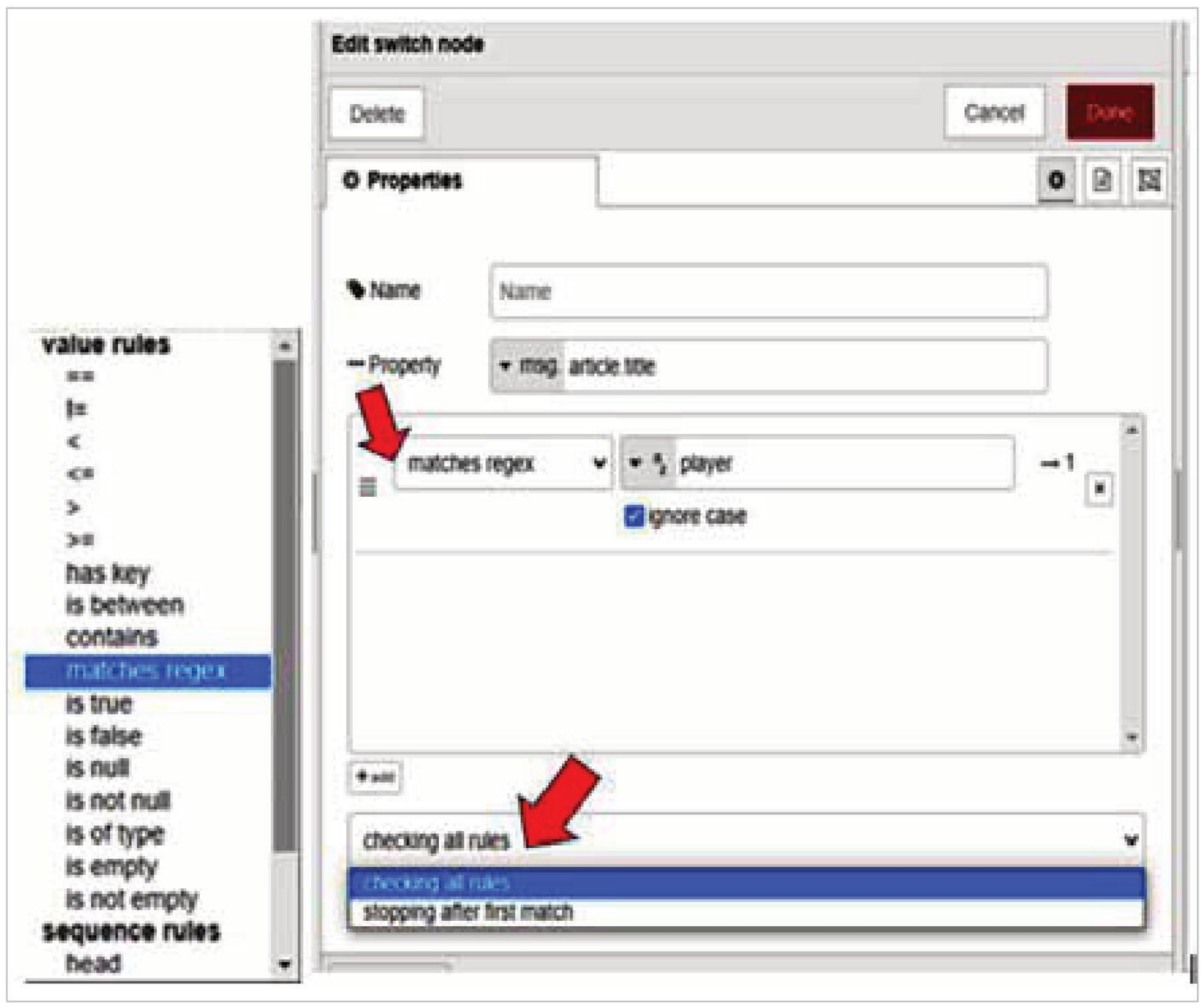


Figure 6: Switch node property configurations

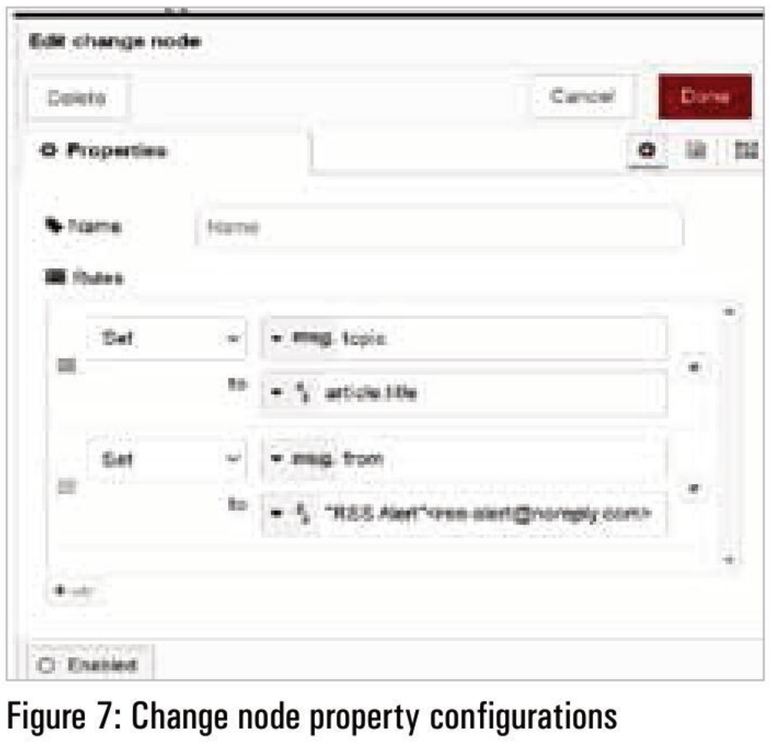


Figure 7: Change node property configurations



Figure 8: Text node property configurations

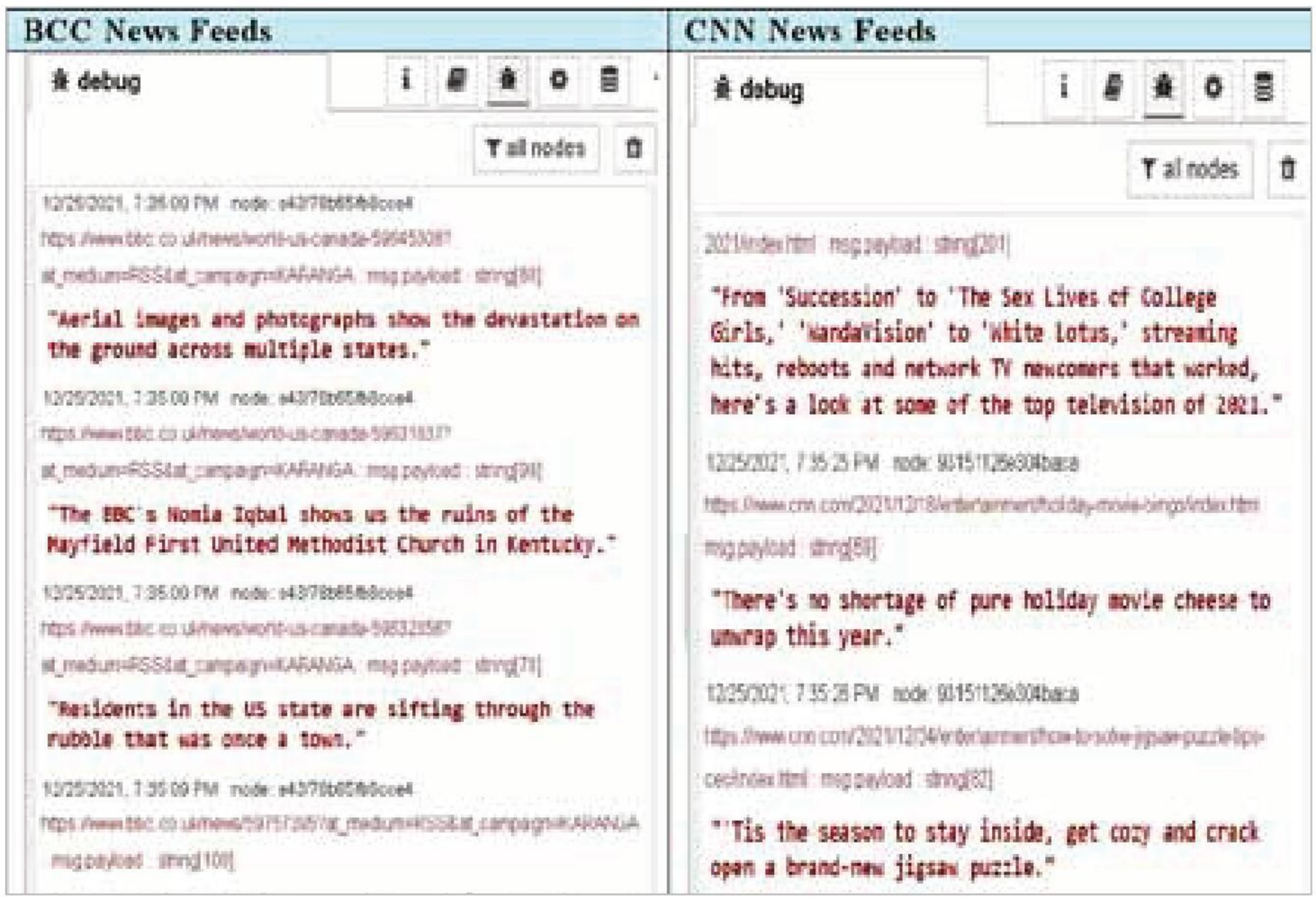


Figure 9: Debug console output screenshot



Figure 11: Web API online media monitoring system

Inference

The *feedparser* node monitors RSS, Atom, and RDF feeds in Node.js. The polling time can be customised, though the default value is 15 minutes. Feed-parser is a transform stream operating in 'object mode': *XML in -> JavaScript objects out.* Every readable chunk is an object representing an article in the feed.

The flow in Node-RED for online media monitoring

Property configuration

(i) *Feedparser* node: Three such nodes, each with different feed URLs are used in this flow.

- BBC: *http://feeds.bbci.co.uk/news/world/us_and_canada/rss.xml;*
- NDTV: *http://feeds.feedburner.com/ndtvprofit-latest?format=xml;*
- CNN: *http://rss.cnn.com/rss/edition_entertainment.rss;*

As the remaining configurations are the same, only BBC World News is shown in Figure 3.

(ii) *Template* node - dashboard widget: This is shown in Figure 4.

(iii) *Switch* node: The filter used is *Matches regex*, which is not case sensitive, as shown in Figure 6.

(iv) *Change* node: Three such nodes, each with different switch inputs, are used in the flow for BBC, CNN and NDTV, respectively, as shown in Figure 7. The same configuration is applicable for all.



Figure 10: Reader login UI

(v) *Text Dashboard* node: Three such nodes, each with different feedparser inputs, are used in the flow.

- **BBC:** Belongs to webAPI tab BBC group; label: *Now, new on BBC news*
- **NDTV:** Belongs to webAPI tab NDTV group; label: *NDTV is sizzling with..*
- **CNN:** Belongs to webAPI tab CNN group, label: *On CNN now!*

Since the other configurations are the same, only NDTV News is shown in Figure 8.

Screenshots (outputs)

(i) The debug console output is shown in Figure 9.

(ii) The dashboard output has a Web user interface. This involves the creation of a Web interface, which consists of a reader's login form to enter user details such as name, mobile number, and email ID (Figure 10) .

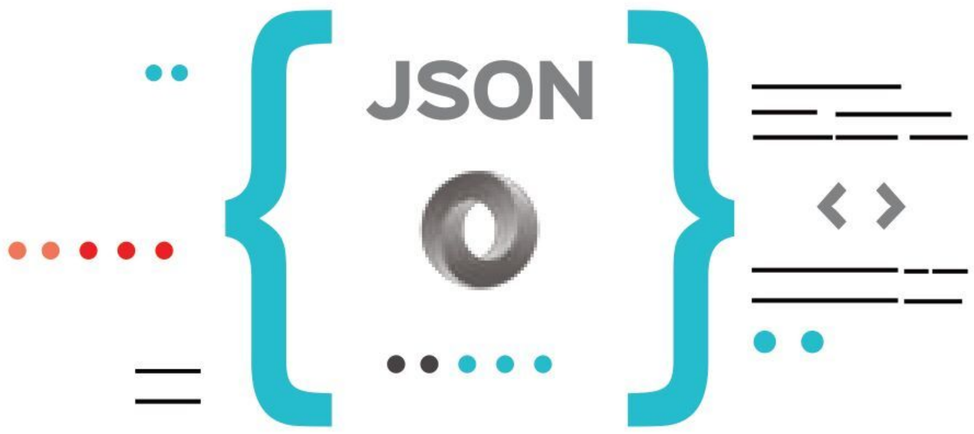
An RSS newsfeed dashboard with three different news sources and news categories is presented in Figure 11. END

By: Dr Maheswari R.

The author is an associate professor at the Vellore Institute of Technology in Chennai, Tamil Nadu.

Carry Out JSONField Schema Validation Using JSON Models

This article shows how JSONField schema can be both manually written and automatically generated, and entails modifying the same in the latter case.



Among the numerous fields that exist for the models in the Django REST framework, JSONField is one where data is often required to abide by a strict format (or) schema, if included in the data used for testing generated APIs. Surely one can manually write the entire schema, but what if there was a way to skip writing it yourself and have it generated? This is where JSON models come to the rescue. In this article, we will be using example data, represented in snippets, on the software 'Insomnia', a popular platform that helps design and test APIs. We will cover two scenarios — writing the schema ourselves, followed by using JSON models to have the schema handed to us without much manual work.

Initial setup

Before getting familiar with the measures that can be taken for the validation of JSONField entry in Django REST framework (not bringing a defined schema into the equation just yet), it can be pointed out that in addition to other nonJSON fields, a POST operation can successfully be carried out if the following are included:

- An empty JSONField
- A JSONField filled with various types of values

The following examples will help understand this better.

models.py

```
from django.db import models
class Student(models.Model):
     student_name=models.CharField(max_length=100)
roll_no=models.IntegerField()
    detailed_data=models.JSONField()
```

serializers.py

```
from rest_framework import serializers
from studentapp.models import Student

class StudentSerializer(serializers.ModelSerializer):
     class Meta:
     model=Student
                                   fields='__all__'
```

views.py

```
from studentapp.serializers import StudentSerializer
from rest_framework.views import APIView
from rest_framework.response import Response
from django.http import JsonResponse,HttpResponse
from rest_framework import status

     class StudentView(APIView):
     def post(self,request):
               serializer=StudentSerializer(data=request.
data)
               if serializer.is_valid():
                    serializer.save()
                    return   Response(serializer.
data,status=status.HTTP_201_CREATED)#inside if serializer.
is_valid()
                  return Response(serializer.
errors,status=status.HTTP_400_BAD_REQUEST) #when serializer
is not valid
```

urls.py

```
from django.contrib import admin
from django.urls import path,include
from studentapp.views import StudentView
urlpatterns = [
    path('student/',StudentView.as_view()),
]
```

As can be seen in the Student model above, we have three fields: a character field, an integer field and a JSONField. Let us enter three types of data in our JSONField: an empty value, a dictionary and a list of dictionaries. With the help of the code snippets given in Figures 1, 2 and 3, let's see how the two points mentioned in the beginning come into play.

Unlike our 'StudentSerializer', which strictly checks the presence of and types of values entered for the 'student_name' and 'roll_no' fields within our Student model (while merely checking if the JSONField 'detailed_data' exists regardless of the type of data entered for it, as observed in Figures 1, 2 and 3), we do not currently possess a defined format that our JSONField must follow. This is where the keyword 'schema' kicks in. Schema is the outline or the structure that the data within JSONField needs to abide by. If the data/values we enter into 'detailed_data' in Insomnia abide by the schema, then we can say that the validation of 'detailed_data' is successful. However, if the value entered for a field within the 'detailed_data' is of a type different from the one defined in the schema, or if any field(s) within 'detailed_data' is/are missing, it implies that we have encountered validation error(s). Let's try to understand this by deep-diving into the 'detailed_data' field within our Student model.

Let's follow the layout given below for the fields within our 'detailed_data' JSONField:

1. Date of joining — Date format
2. Mother's name — Character type
3. Father's name — Character type
4. Teacher information — A dictionary comprising: Name, Date of birth, Department, Date of joining
5. Semester 1 subjects — A list of dictionaries, with each dictionary comprising subject name, score
6. Semester 2 subjects — Another list of dictionaries with the same structure and subfields as mentioned in point 3.



Figure 1: When entering an empty value into our JSONField 'detailed_data'

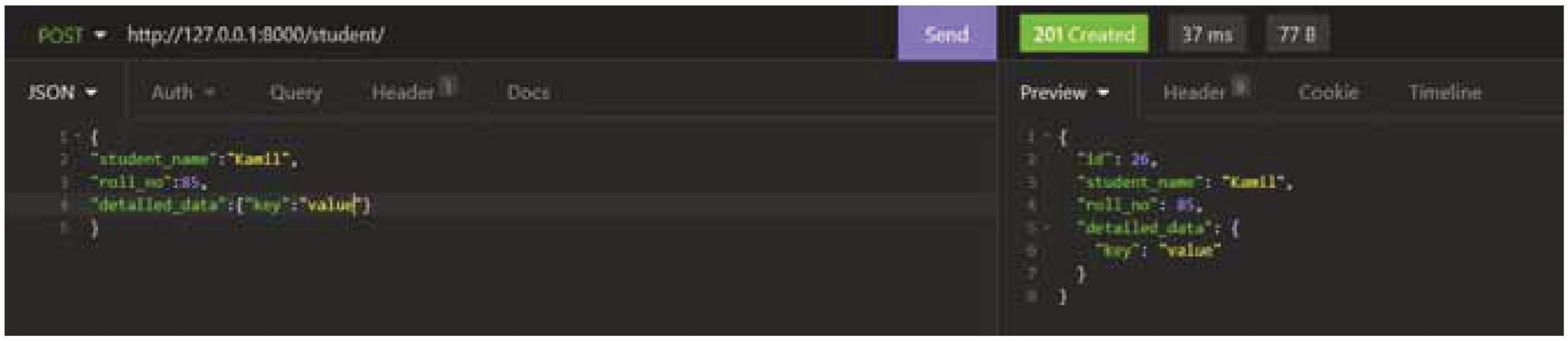


Figure 2: When entering a key-value pair into our JSONField

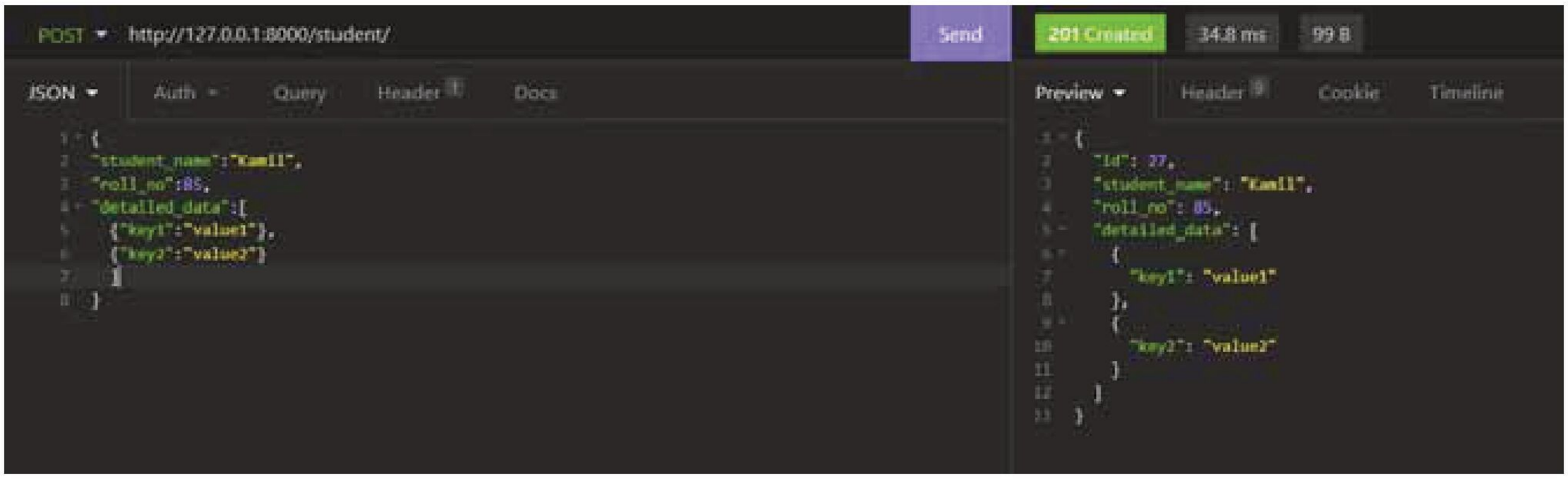


Figure 3: When entering a list of dictionaries

Given the above, our objective is to first check if the serialisation of ‘student_name’ and ‘roll_no’ is valid, followed by checking for the validity of the contents of ‘detailed_data’ by comparing with the schema. We can do either of the following: manually write the schema or use JSON models. The latter is recommended since it auto-generates the schema, but to get an idea of what the schema looks like, we will go through both methods.

Methods of schema validation

Writing the schema manually

To map our schema onto the target data that we want to enter in the JSONField, consider Figure 4.

POST http://127.0.0.1:8000/student/

JSON Auth Query Header 1 Docs

```
{
"student_name":"Kamil",
"roll_no":85,
"detailed_data":{
  "date_of_joining":"01-05-2017",
  "mother_name":"Alice",
  "father_name":"Matthew",
  "teacher_information":{
    "name":"Ms. Anne",
    "date_of_birth":"09-10-1978",
    "department":"102",
    "joining_date":"05-07-2007"
    },
  "semester_1_subjects":[
    {
      "subject_name":"Maths 1",
      "professor":"Mr. Adam",
      "score":75
    },
    {
      "subject_name":"English 1",
      "professor":"Ms. Evelyn",
      "score":80
    }
  ],
  "semester_2_subjects":[
    {
      "subject_name":"Maths 2",
      "professor":"Mr. Ram",
      "score":75
    },
    {
      "subject_name":"English 2",
      "professor":"Mr. Peter",
      "score":80
    }
  ]
}
}
```

Figure 4: Example data for JSONField

The schema suitable to fit the type of data is as follows:

```
{
“type”:”object”,
“properties”: {
    “date_of_joining”: {“type”:
“string”,”pattern”:’^[0-9]{2}-[0-9]{2}-[0-9]{4}$’},
    “mother_name”:{“type”:”string”},
    “father_name”:{“type”:”string”},
    “teacher_information”: {
            “type”:”object”,
            “properties”:{
              “name”: {“type”: “string”},
                “date_of_birth”: {“type”:”string”,”patt
ern”:’^[0-9]{2}-[0-9]{2}-[0-9]{4}$’},
                “department”:{“type”:”string”},
                “joining_date”:{“type”:”string”,”patte
rn”:’^[0-9]{2}-[0-9]{2}-[0-9]{4}$’}
            },
            “required”:[“name”,”date_of_
birth”,”department”,”joining_date”],  --(1)
            “additionalProperties”: False
    },
    “semester_1_subjects”:{
            “type”:”array”,
            “items”:{
              “type”:”object”,
                “properties”:{
                    “subject_name”:{“type”:”string”},
                    “professor”:{“type”:”string”},
                    “score”:{“type”:”integer”}
            },
            “required”:[“subject_name”,”professor”,”sco
re”],          --(2)
            “additionalProperties”:False
                }
    },
    “semester_2_subjects”:{
            “type”:”array”,
            “items”:{
                “type”:”object”,
                “properties”:{
                    “subject_name”:{“type”:”string”},
                    “professor”:{“type”:”string”},
                    “score”:{“type”:”integer”}
                },
            “required”:[“subject_name”,”professor”,”sco
re”],          --(3)
            “additionalProperties”:False
            }
    },
            “required”:[“date_of_joining”,”mother_
```

```
name","father_name","teacher_information","semester_1_
subjects","semester_2_subjects"],        --(4)
"additionalProperties":False
}
```

The list of fields which serve as 'value' to the 'required' key in line (1) imply that those fields (within *teacher_information*) MUST be included. Even if one of them is missed, a validation error occurs. The same applies for lines (2) and (3). However, line (4) deals with the outer layer of fields. Lines (1), (2) and (3) deal with the compulsory inclusion of fields within 'teacher_information', 'semester_1_subjects' and 'semester_2_subjects'. Line (4) also deals with the same, but for the inclusion of the aforementioned fields on the whole.

Next, we include our schema in *views.py* and use *Draft7Validator* to compare what is being entered in the defined schema.

Updated views.py

```
import jsonschema
from jsonschema import Draft7Validator

    class StudentView(APIView):
      def post(self,request):
            data=request.data                    ------(1)
            serializer=StudentSerializer(data=request.
data)
            if serializer.is_valid():
                  myschema= #insert above defined schema
here
                  v=Draft7Validator(myschema)       -----
--(5)
                  if len(list(v.iter_
errors(data["detailed_data"])))!=0:
                        return
Response({"error":str(list(v.iter_errors(data["detailed_
data"])))})
                  else:
                     serializer.save()
                     return Response(serializer.
data,status=status.HTTP_201_CREATED)  -----(10)
            else:
                  return Response(serializer.
errors,status=status.HTTP_400_BAD_REQUEST)
```

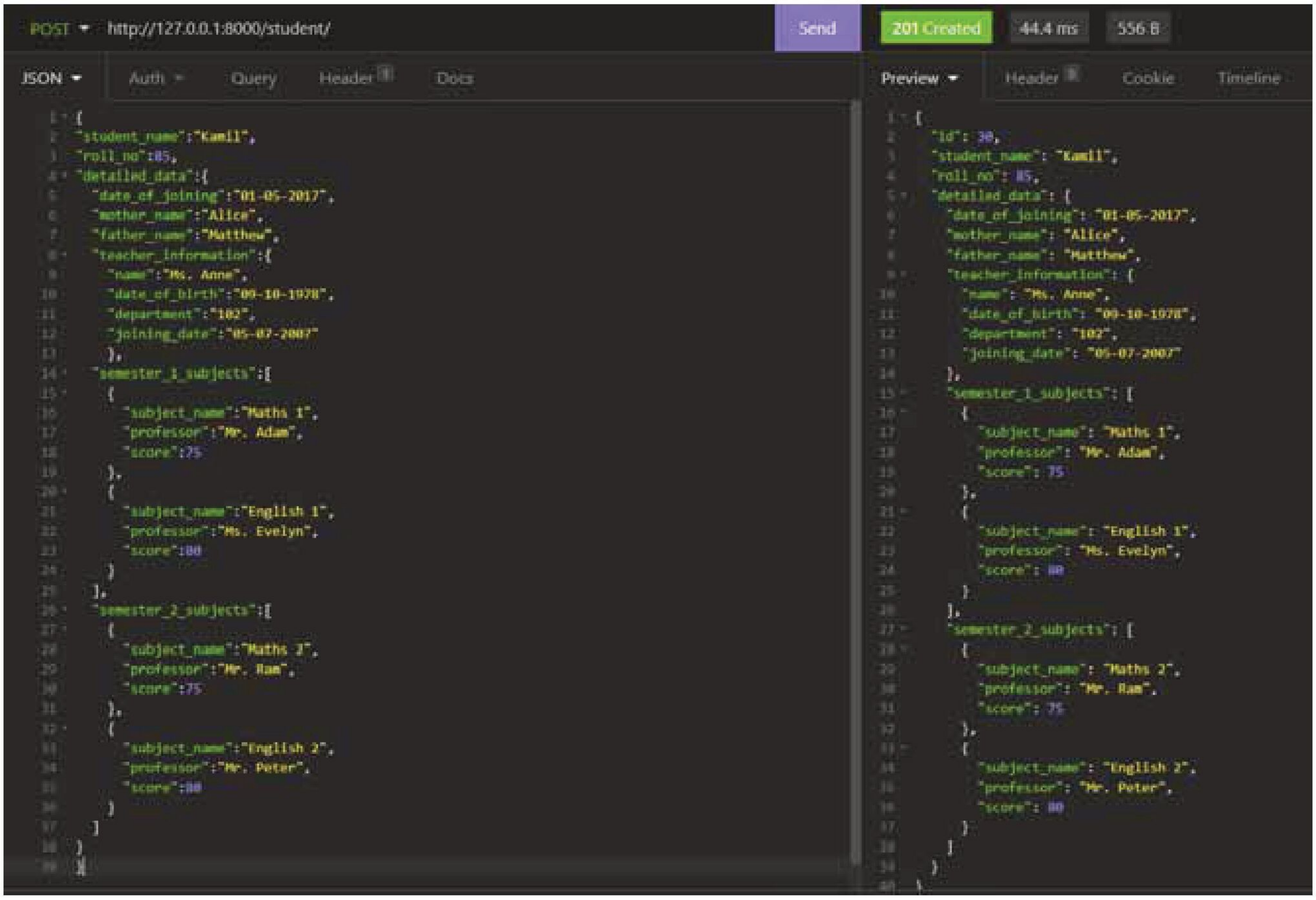


Figure 5: An example of a flawless POST operation

The aim of line (1) is to obtain the data entered by the user and store it in a variable, namely, 'data'. In line (2), we aim to map this data against the Student model via our StudentSerializer and check for the presence of 'student_name', 'roll_n' and 'detailed_data' fields, while also validating the type of values entered for the first two fields. If the 'serializer' is valid, i.e., if none of the three fields are missing and there is no objection to the type of values entered for the former two fields, we check the data within the 'detailed_data' JSONField. After assigning the manually-written schema to 'myschema' in line (4), line (5) passes the schema to Draft7Validator and the same is assigned to 'v'. The 'v.iter_errors(data["detailed_data"])' in line (6) intends to point out any validation errors, should they exist, such as missing fields, wrong value entered for fields, etc. We simply append these errors in a list as 'elements' and check the length of the list. If the length is 0, it means the data entered in the JSONField abides by the schema and validation is successful. If not, the validation errors are displayed as a list, serving as 'value' to the 'error' key mentioned in line (7).

In Figure 5, we can see that values for all the fields are exactly of the types written in the schema. Additionally, there are no missing fields, which results in the particular entry being saved without any validation errors. Let's try making a few changes to see how the validation errors are shown. Let's not include the first 'professor' key-value pair in 'semester_2_subjects'. Do the same for the second 'subject_name' within 'semester_1_subjects'. Let's write an integer type value for 'department' within 'teacher_information' and a wrong pattern for 'date_of_joining'.

We have covered the various types of validation errors. This method requires us to write out the schema manually, but what if you didn't want to do so and what if it could be done for you instead? This is where JSON models come into play.

Using JSON models

To generate the previous schema, the following additions have to be made in *models.py*.

models.py

```
from jsonmodels import models as m1,fields

class teacherField(m1.Base):
    name=fields.StringField(required=True)
    date_of_birth=fields.DateField(required=True)
    department=fields.StringField(required=True)
    joining_date=fields.DateField(required=True)
class sem1Field(m1.Base):
    subject_name=fields.StringField(required=True)
    professor=fields.StringField(required=True)
    score=fields.IntField(required=True)
class sem2Field(m1.Base):
    subject_name=fields.StringField(required=True)
    professor=fields.StringField(required=True)
    score=fields.IntField(required=True)
class detailed_data_field(m1.Base):
    date_of_joining=fields.DateField(required=True)
    mother_name=fields.StringField(required=True)
    father_name=fields.StringField(required=True)
```

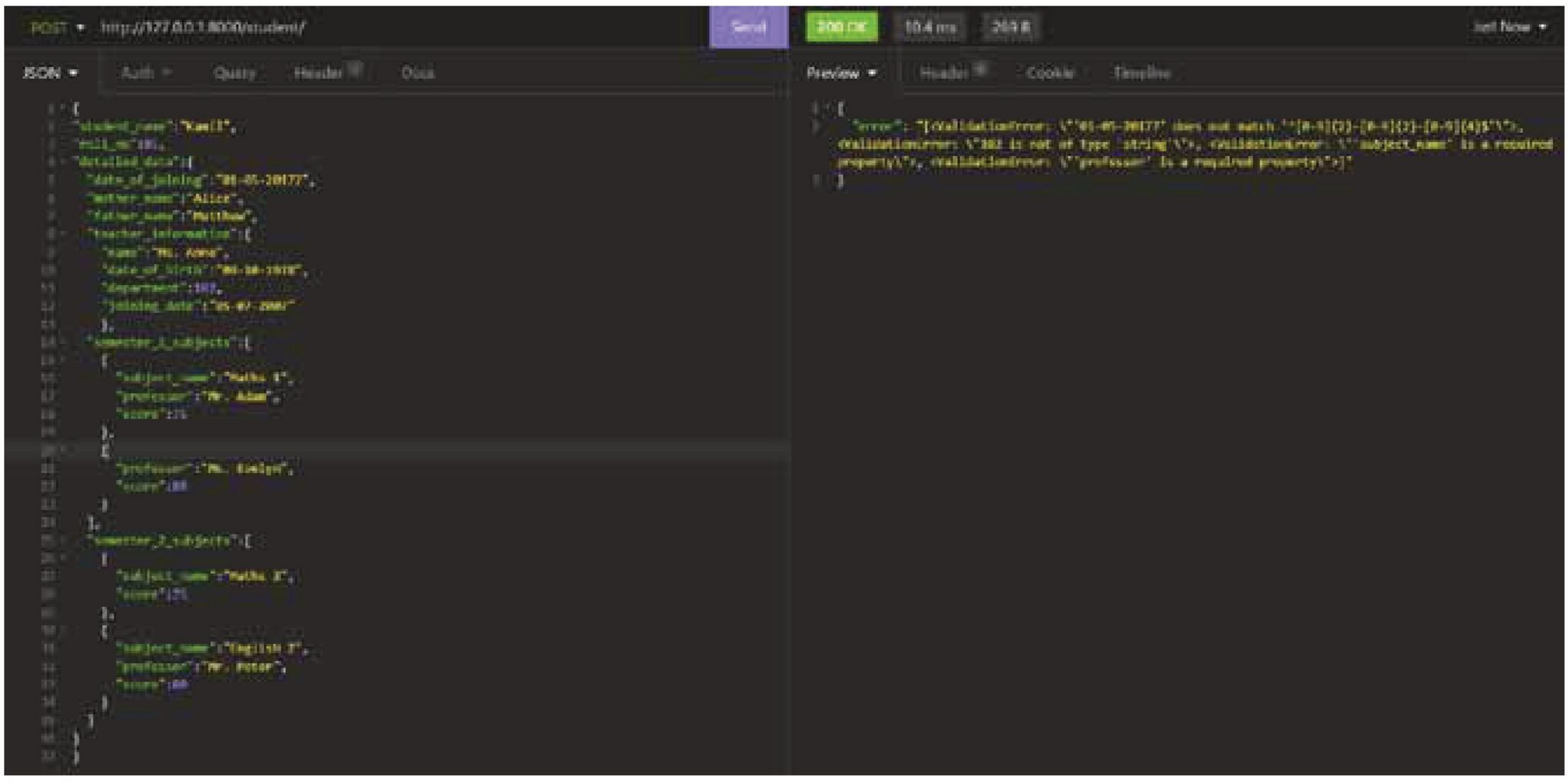


Figure 6: Display of list of validation errors

```
teacher_information=fields.EmbeddedField(teacherField)
semester_1_subjects=fields.EmbeddedField(sem1Field)
semester_2_subjects=fields.EmbeddedField(sem2Field)
```

We are treating the distinct fields within 'detailed_data', notably 'teacher_information' (a dictionary), 'semester_1_subjects' and 'semester_2_subjects' (both are lists of dictionaries), as separate JSON models and then giving a call to them using the embedded field within the outlying 'detailed_data_field' JSON model. The 'EmbeddedField' mentioned in the last three lines above performs the same function as that of foreign keys in databases.

Updated views.py

```
from marksapp1.models import detailed_data_field      ----(1)
class StudentView(APIView):
    def post(self,request):
        data=request.data
        serializer=StudentSerializer(data=request.data)
----(5)
        if serializer.is_valid():
            data1=detailed_data_field()
            myschema=data1.to_json_schema()
            print(myschema)
            v=Draft7Validator(myschema)                    ----
(10)
            if len(list(v.iter_errors(data["detailed_
data"])))!=0:
                return Response({"error":str(list(v.iter_
errors(data["detailed_data"])))})
            else:
                serializer.save()
                return Response(serializer.
data,status=status.HTTP_201_CREATED)  ----(15)
        else:
            return Response(serializer.errors,status=status.
HTTP_400_BAD_REQUEST)
```

In line (7), our overall JSON model 'detailed_data_field' is assigned to variable 'data1', followed by schema generation of the same in line (8). A glimpse of the schema generated is seen below.

```
{                                                  ----(1)
'type': 'object',
'additionalProperties': False,
'properties': {
     'date_of_joining': {'type': 'string'},
----(5)
     'father_name': {'type': 'string'},
     'mother_name': {'type': 'string'},
     'semester_1_subjects': {'type': 'object',
              'additionalProperties': False,
              'properties': {'professor': {'type':
'string'},      ----(10)
                     'score': {'type': 'number'},
                     'subject_name': {'type': 'string'}
                     },
              'required': ['professor', 'score', 'subject_
name']
     },                                                ----(15)
     'semester_2_subjects': {'type': 'object',
              'additionalProperties': False,
              'properties': {'professor': {'type':
'string'},
                     'score': {'type': 'number'},
                     'subject_name': {'type': 'string'}
----(20)
                     },
              'required': ['professor', 'score', 'subject_
name']
     },
     'teacher_information': {'type': 'object',
              'additionalProperties': False,
----(25)
              'properties': {'date_of_birth': {'type':
'string'},
                     'department': {'type': 'string'},
                     'joining_date': {'type': 'string'},
                     'name': {'type': 'string'}
                     },                  ----(30)
              'required': ['date_of_birth', 'department',
'joining_date', 'name']
     }
},
'required': ['date_of_joining', 'father_name', 'mother_
name']}
```

We can observe in the auto-generated schema that for 'semester_1_subjects' and 'semester_2_subjects', the 'type' is 'object', whereas it should actually be 'array'. Among the fields that exist for JSON models, there are none that help develop a *type:'array'* in the schema. To visualise this, let us compare the 'semester_1_subjects' structure from the auto_generated schema with the manually written one.

Auto-generated schema

```
'semester_1_subjects': {'type': 'object',

              'additionalProperties': False,
              'properties': {'professor': {'type':
```

```
'string'},
                    'score': {'type': 'number'},
                    'subject_name': {'type': 'string'}
                    },
                'required': ['professor', 'score', 'subject_
name']
            },
```

Manually written schema

```
"semester_1_subjects":{"type":"array",
                    "items":{"type":"object",
                        "properties":{
                                "subject_
name":{"type":"string"},
                                "professor":{"type":"string"},
                                "score":{"type":"integer"}
                    },
                    "required":["subject_name","professor","sc
ore"],
                    "additionalProperties":False
                    }
        },
```

We need to make the auto-generated structure of 'semester_1_subjects' look like that of a manually written schema. All we have to do is make the entire dictionary of the auto-generated schema serve as a 'value' to a key titled 'items' within the dictionary, which acts like a 'value' to 'semester_1_subjects', followed by adding a key-value pair 'type':'array' to the dictionary, i.e., 'value to 'semester_1_subjects'. Hence, the structure of 'semester_1_subjects' will look like that defined in the manually written schema. We do the same for 'semester_2_subjects'. We also need to append 'semester_1_subjects', 'semester_2_subjects' and 'teacher_information' to the 'required' list mentioned towards the end of the auto-generated schema. The 'pattern' for the date-type fields needs to be included as well. Let us take a look at the new view, which will factor in all the aforementioned changes.

```
class StudentView(APIView):
----(1)
     def post(self,request):
               data=request.data
               serializer=StudentSerializer(data=request.
data)
               if serializer.is_valid():
     ----(5)
                         data1=detailed_data_field()
                         myschema=data1.to_json_schema()
                         myfields=['semester_1_
subjects','semester_2_subjects']
                         for i in myfields:
                              d={}
----(10)
                              d["items"]=myschema["properties"]
[i]
                              myschema["properties"][i]=d
                              myschema["properties"][i]
["type"]="array"
```

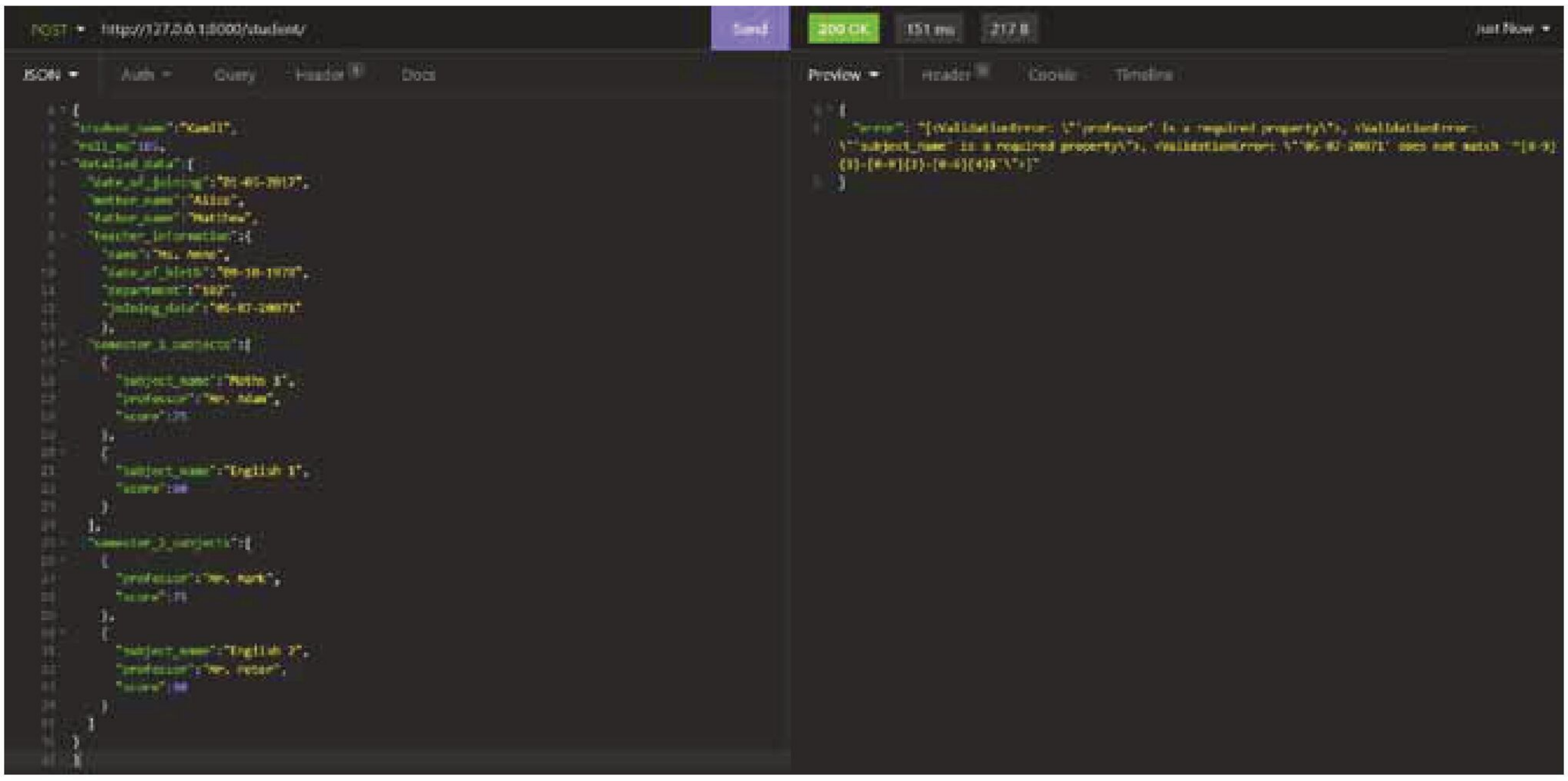


Figure 7: Listing of validation errors

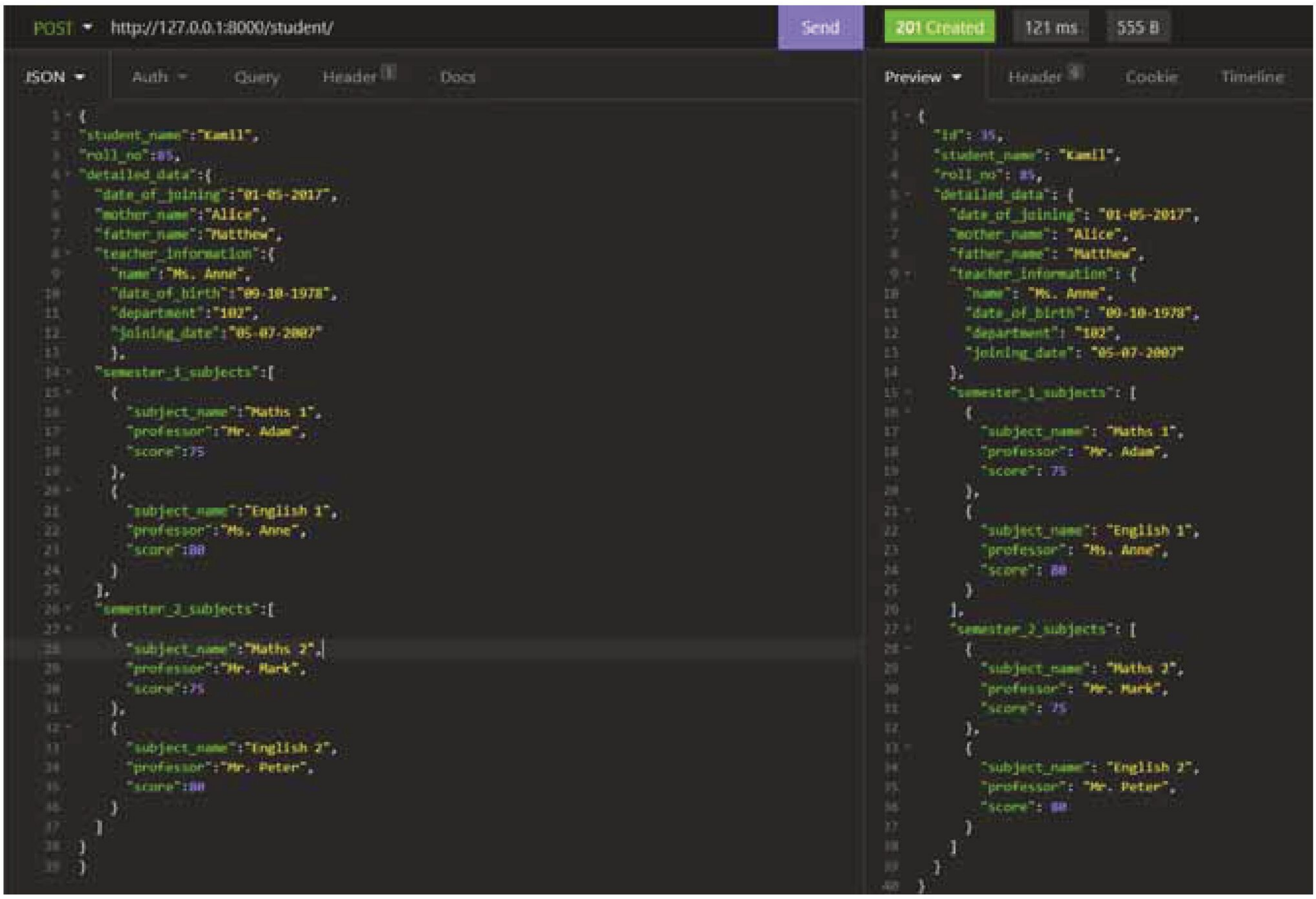


Figure 8: Successful validation of JSONField schema

```
                    myschema["required"].extend(["teacher_
information","semester_1_subjects","semester_2_subjects"])
                    myschema["properties"]["date_of_
joining"]["pattern"]='^[0-9]{2}-[0-9]{2}-[0-9]{4}$'  ----(15)
                    myschema["properties"]
["teacher_information"]["properties"]["joining_date"]
["pattern"]='^[0-9]{2}-[0-9]{2}-[0-9]{4}$'
                    myschema["properties"]
["teacher_information"]["properties"]["date_of_birth"]
["pattern"]='^[0-9]{2}-[0-9]{2}-[0-9]{4}$'
                    v=Draft7Validator(myschema)
                    if len(list(v.iter_
errors(data["detailed_data"])))!=0:
                        return
Response({"error":str(list(v.iter_errors(data["detailed_
data"])))})    ----(20)
                    else:
                        serializer.save()
                        return Response(serializer.
data,status=status.HTTP_201_CREATED)
                else:
                    return Response(serializer.
errors,status=status.HTTP_400_BAD_REQUEST)      ----(25)
```

The section of views from lines (10) to (13) aims to convert the structure of 'semester_1_subjects' and 'semester_2_subjects' from 'type':'object' to 'type':'array'. Line (14) intends to append 'teacher_information', 'semester_1_subjects' and 'semester_2_subjects' to the outermost 'required' list. Lines (15) to (17) help include a 'pattern' key-value pair in the schema structure for date-related fields.

So you now have a basic understanding of how the structure of data of a JSONField within a Django REST framework model can be cross-checked against a well-defined schema. Covering the lengthier method, i.e., writing the schema yourself, was essential to understand how a schema looks. You also have an idea of how JSON models help generate the schema for you, and even though the auto-generated schema is not an exact reflection of the schema you may write yourself, you can manipulate it to suit your needs utilising a fundamental data structure in Python: dictionaries, and its basic operations. END

By: Mohammed Kamil Khan

The author is currently working at Caravel.AI. His aim is to make available the knowledge of topics he comes across in a beginner-friendly manner.

Is Java Dead?

Java has been one of the most in-demand languages for over 20 years now. It is widely used across the industry and is often a compulsory part of the syllabus in most computer science courses across the world. But many developers are beginning to see it as an aged and primitive programming language, which, they predict, will soon become obsolete. This article tries to get to the heart of this issue by analysing popular developer surveys.



In the last week of December 2021, I received a call from a friend of mine. She is a programmer who just completed her master's degree in computer science. She was setting her learning goals for the upcoming year and asked me to recommend five programming languages that she should focus on. One of the recommendations I gave was to learn Java. "But Mir," she said, "Isn't Java dead?" I was taken aback by this statement. How can someone claim that Java is dead? It is one of the most widely adopted programming languages in the industry. It is also very much in use and in demand in the current market.

Later that day, I Googled 'Is Java dead?' just to see if there was any meat in this argument. This took me down a rabbit hole filled with a great many developers who do, in fact, believe that Java is mostly a redundant language with not much use in the future. Amongst the critics of Java were some great names in the industry such as Linus Torvalds. "What a horrible language," said Linus in an interview when asked about Java. That got my attention and I started to dig deeper. I noticed a pattern in the arguments that the developers shared against Java. The vast majority of Java critics seemed to have issues with the way this language is structured.

The best way to understand how the developer industry perceives a technology is to follow the data published on popular developer survey reports. In this article, I have researched data from GitHub Octoverse Report 2021, JetBrains State of Developer Ecosystem 2021 and StackOverflow Developer Survey 2021. The insights that I drew from these reports hold the key to understanding what has gone right or wrong with the Java ecosystem.

A brief overview of Java

My assumption in this article is that the reader already has some notion of Java programming and how it functions. As such, I shall not go into too many details about the history of Java. Doing so will take us off topic. But it is always a good idea to share the basics so that we can brush up on what we already know about Java. Feel free to skip this section if you have a good understanding of the language.

Java is a high level, class based, object-oriented programming language. Built on the write once, run anywhere (WORA) model, it was developed by Sun Microsystems in 1995. Oracle acquired Sun Microsystems in 2010 and took charge of the development process of Java.

From 2010 onwards, the development of Java sped up under Oracle with new versions being released almost every six months. The latest version at the time of writing this article is Java 17, released in October 2021.

Design principles of Java

Before heading into the analysis of the developer surveys, I find it necessary from the context of this article to share with you Oracle's vision for Java. Listed below are the key points from the design principles of Java, as mentioned on the official website of Oracle.

Java is designed to be:

1. Simple, object-oriented and familiar
2. Robust and secure
3. Architecture-neutral and portable
4. High performance
5. Interpreted, threaded and dynamic

These points are about to become very relevant to this article. I shall be connecting the dots in the last section. For now, let us begin the analysis of the developer surveys.

GitHub Octoverse Report 2021

Every year, GitHub publishes a survey of its own platform to gain insights on developer trends. It is called the GitHub Octoverse Report. The following are the key insights from the 2021 report with respect to Java, as shown in Figure 1.

Java ranked second in the top ten programming languages used in the GitHub repositories from 2014 to 2018. JavaScript remained the most popular programming language.

Java ranked third in the top ten programming languages used in the GitHub repositories from 2018 to 2021. It was preceded by JavaScript in the first place and Python in the second.

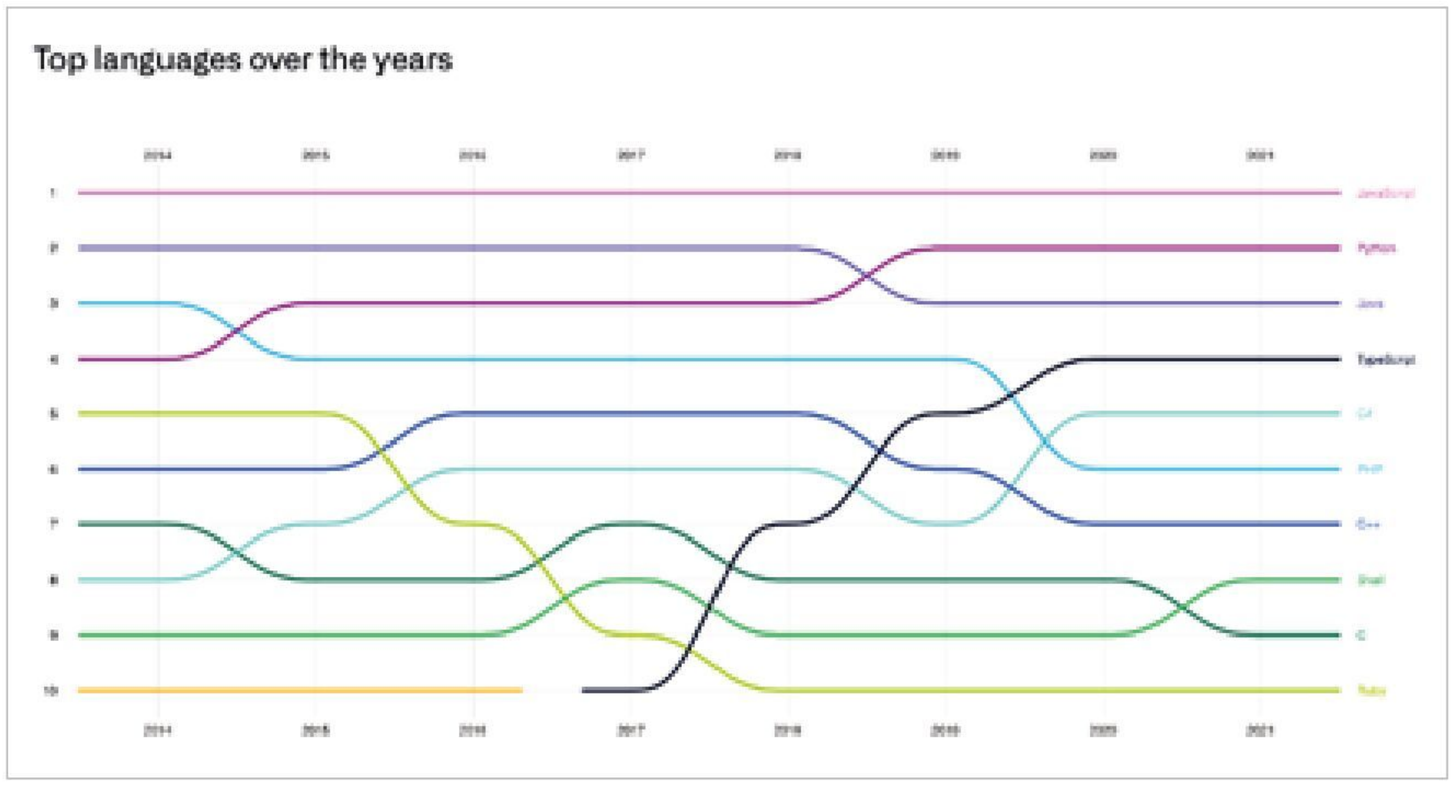


Figure 1: Key insights on Java in the GitHub Octoverse Report of 2021

JetBrains State of Developer Ecosystem 2021

JetBrains conducts its own developer survey each year to understand the upcoming trends in the developer community. For 2021, 31,743 developers from 183 countries participated in the survey. The following are the key insights on Java in this report.

1. Java is the fifth most popularly used programming language, as shown in Figure 2.
2. Java ranks second in terms of the primary programming language used in a project.
3. Fifty-six per cent of Java developers build Web applications.
4. Of the Java developers that build Web applications, 77 per cent are backend Web developers.
5. Java is more popularly used in South Korea, China and India.

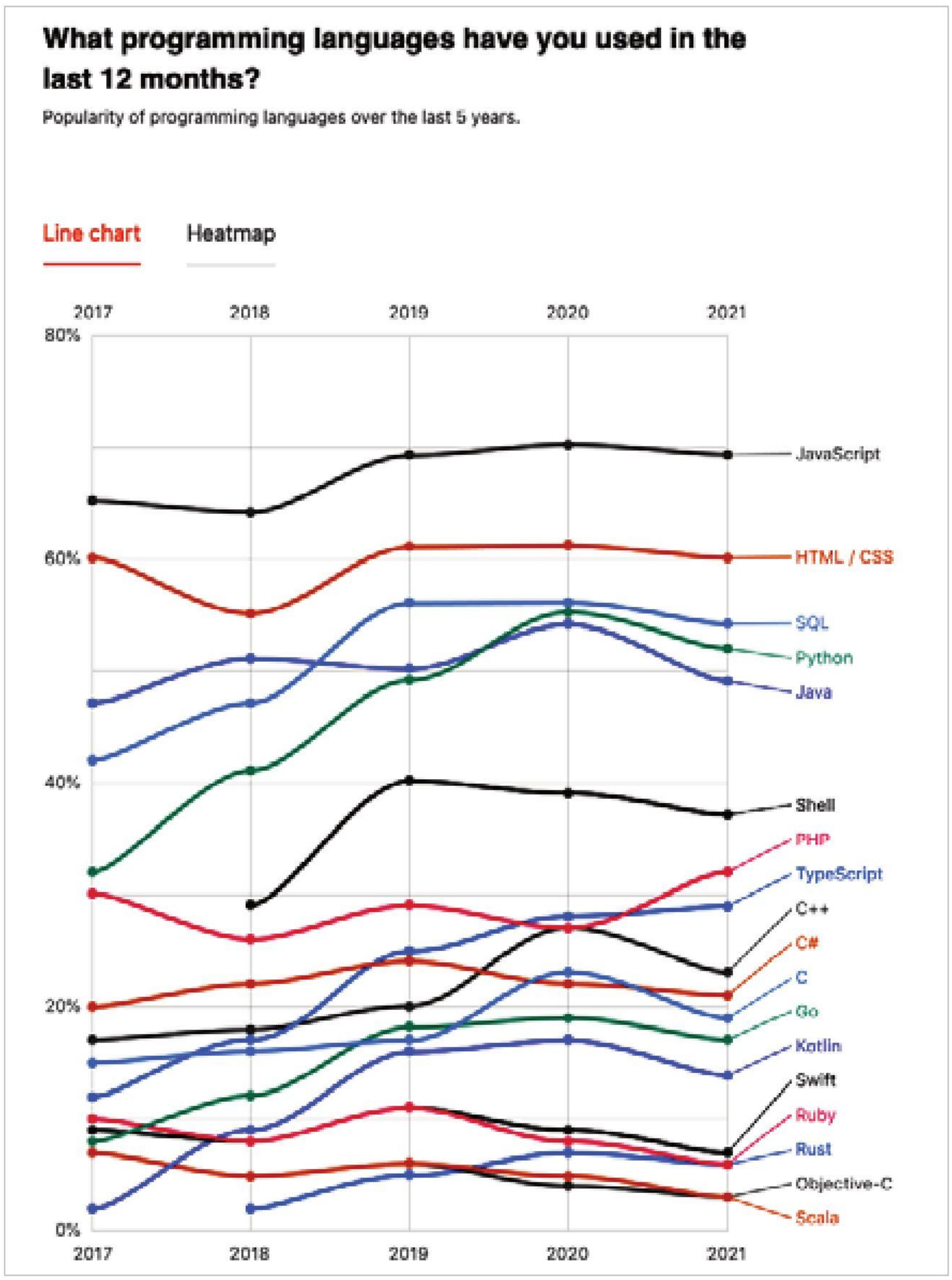


Figure 2: Key insights on Java in the JetBrains State of Developer Ecosystem 2021

Stack Overflow Developer Survey 2021

Stack Overflow conducts the most comprehensive developer survey in the tech space. In 2021, over 80,000 developers were surveyed to gather key insights and trends. A great convenience provided by this report is that it lets you filter data by the parameters of ‘all respondents’ or ‘professional developers’. This means that you have a choice to see the data given by all respondents or you can filter out the data of the respondents who are not yet professional developers, i.e., haven’t worked in the industry yet. Given the topic of this article, I have only taken the responses of professional developers into account.

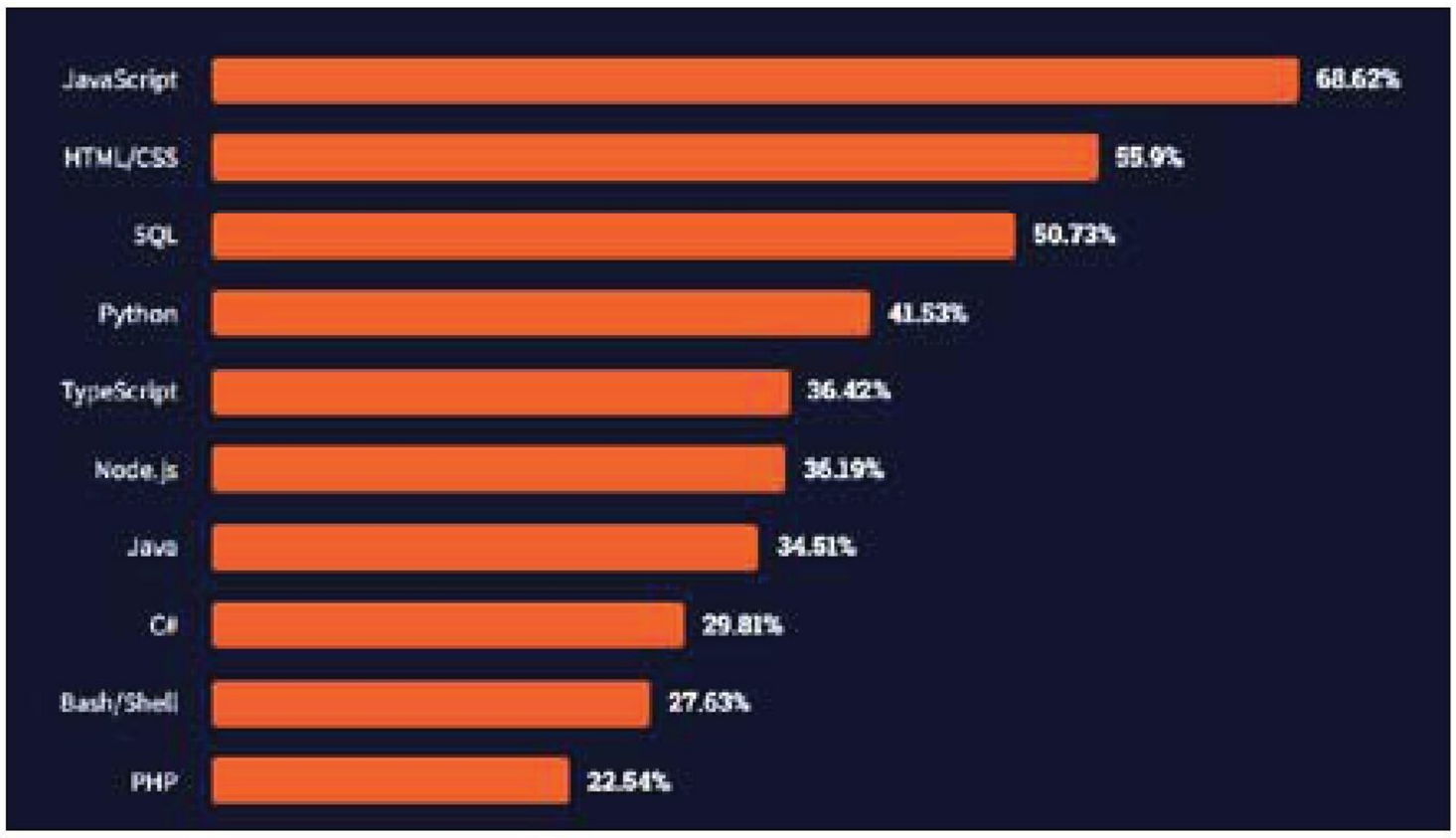


Figure 3: Key insights on Java in the Stack Overflow Developer Survey 2021

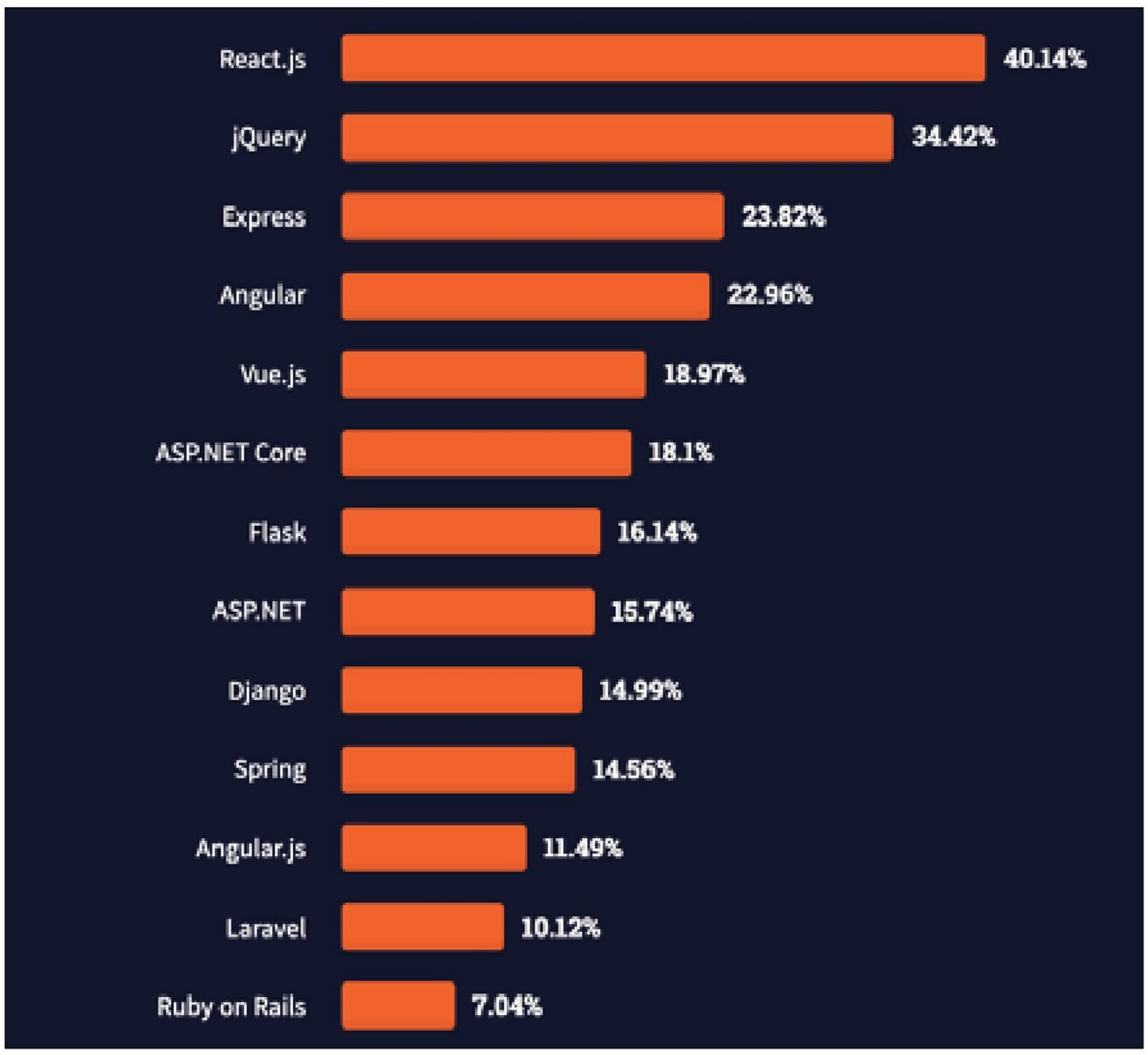


Figure 4: Key insights on Spring in the Stack Overflow Developer Survey 2021

Given below are the key insights on Java from this report.

1. Java is the seventh most popular programming language according to 58,031 respondents, as shown in Figure 3.
2. On the question of 'loving/dreading' a programming language, 29,162 developers responded the following about Java:
   a. 52.8 per cent (15,413) responded by saying they dread using Java.
   b. 47.1 per cent (13,749) responded by saying they loved using Java.
3. On the question 'What programming language do you want to learn', Java was ranked ninth according to 82,914 respondents.

There are a few more insights within this survey that I felt were necessary to take note of in relation to Java.

First, a few insights on the Java based Spring Web development framework.

1. Spring ranks as the eighth most popular Web development framework according to 49,941 responses, as shown in Figure 4.
2. On the question of loving/dreading a framework, 9177 developers responded the following about Spring:
   a. 59.8 per cent (5,490) responded by saying they love Spring.
   b. 40.1 per cent (3,687) responded by saying they dread Spring.
3. Spring ranks eleventh in the list of frameworks that the developers "want to learn." Of 66,202 respondents, only 3.8 per cent showed interest in learning Spring.

Second, some insights on the Java based Hadoop Big Data management library.

1. Hadoop ranks last in the 'most popular libraries' to use according to 41,760 respondents, as shown in Figure 5.
2. On the question of loving/dreading a library, 2,155 developers responded the following about Hadoop:
   a. 54.7 per cent (1,179) responded by saying they dread using Hadoop.
   b. 45.2 per cent (976) responded by saying they love using Hadoop.
3. Hadoop ranks tenth in the list of 13 libraries the developers 'want to learn' according to 58,282 respondents.

Two additional key insights to take note of from this survey are:

1. Of the 61,216 professional developers that participated in this survey, the majority have been working professionally for 10 years or less.
2. Out of the 81,641 total respondents in this survey, over 50 per cent have been programming for 10 years or less.

What do these reports say about Java?

The insights drawn from the reports discussed above can be summarised in the following points:

1. Java remains a popular programming language in the developer community, although its popularity has seen a slight decline due to the rise of programming languages like Python and Rust.
2. Although Java is used in many industries, its ecosystem primarily consists of backend Web developers. This is in huge part due to the popularity and demand of the Spring Web development framework, which is the most popular Java based Web development framework.
3. The developer community is losing interest in learning Java or investing time into learning frameworks that require Java. The lack of interest in learning Hadoop and Spring is a case in point.
4. Most developers coding professionally find Java an intimidating language. This resistance spills over in the statistics for Java based frameworks as well. A significant majority of Hadoop developers dread using Hadoop, whereas a significant minority of Spring developers dread using Spring.

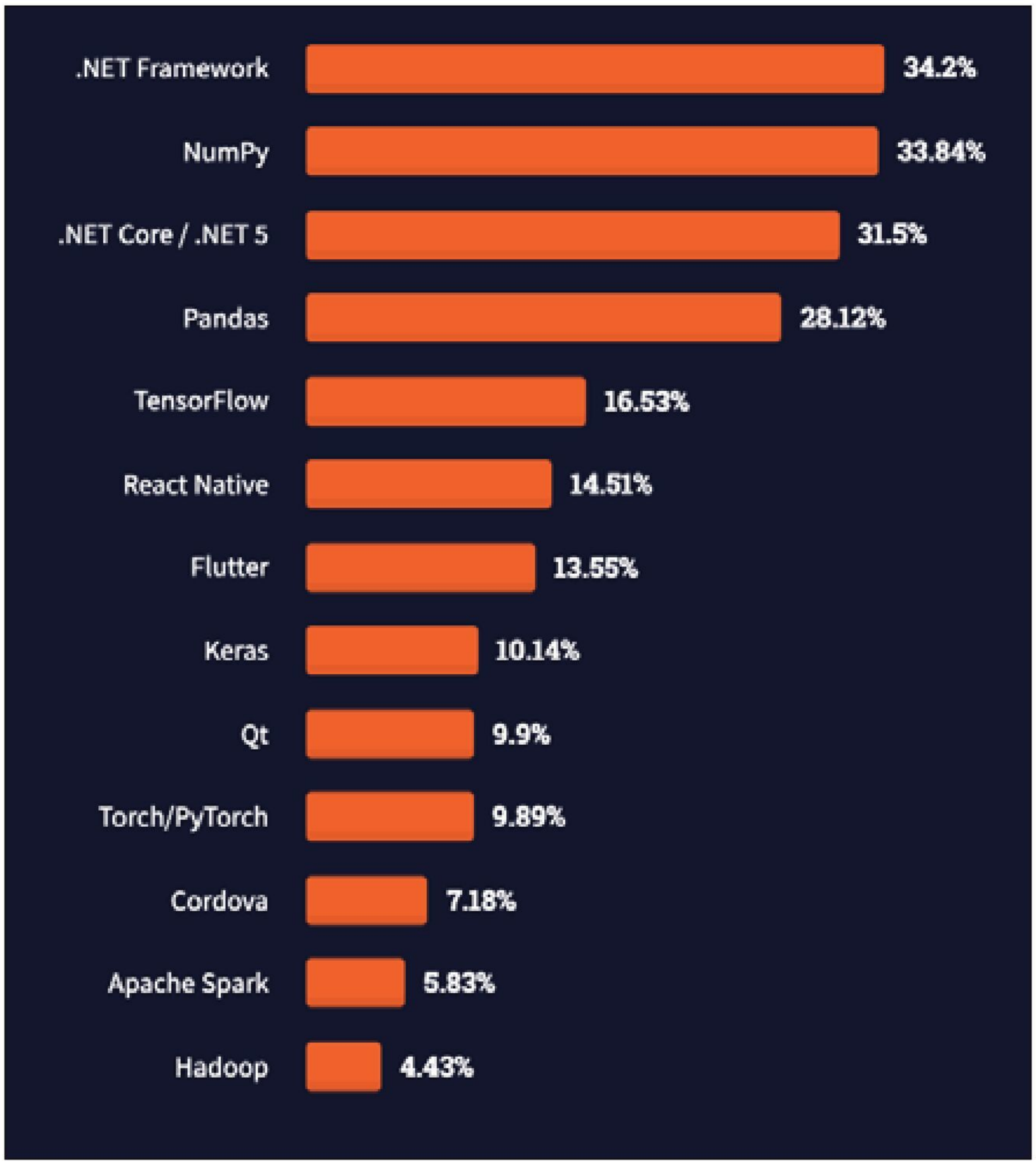


Figure 5: Key insights on Hadoop in the Stack Overflow Developer Survey 2021

Summing it up

By no means is Java a dead programming language. But it is most certainly a programming language that is facing trouble finding its footing in the ever evolving developer community. In the beginning of the article, I had shared the design principles of Java as described by Oracle. As far as the utility of Java is concerned, Oracle has done a great job. Java can be used for a wide variety of applications efficiently. However, the price of this utility has been the sacrifice of simplicity. The first design principle of Java is that it should be *simple*. Unfortunately, developers do not seem to find it simple to use.

Most of the developers who are participating in these surveys have been working for 10 years or less. This new generation of developers is having trouble with the strict implementation of the object-oriented design pattern of Java. With the rising popularity of Python (thanks to its many libraries) and the immense rise of computational power that allows many companies to not sweat over the time complexity of programs, Java is beginning to be seen more as an obstacle rather than a solution.

The beauty of Python and other such programming languages that do not strictly implement object-oriented programming is that due to their syntactic simplicity, it is easier to write semantically complex code with them.

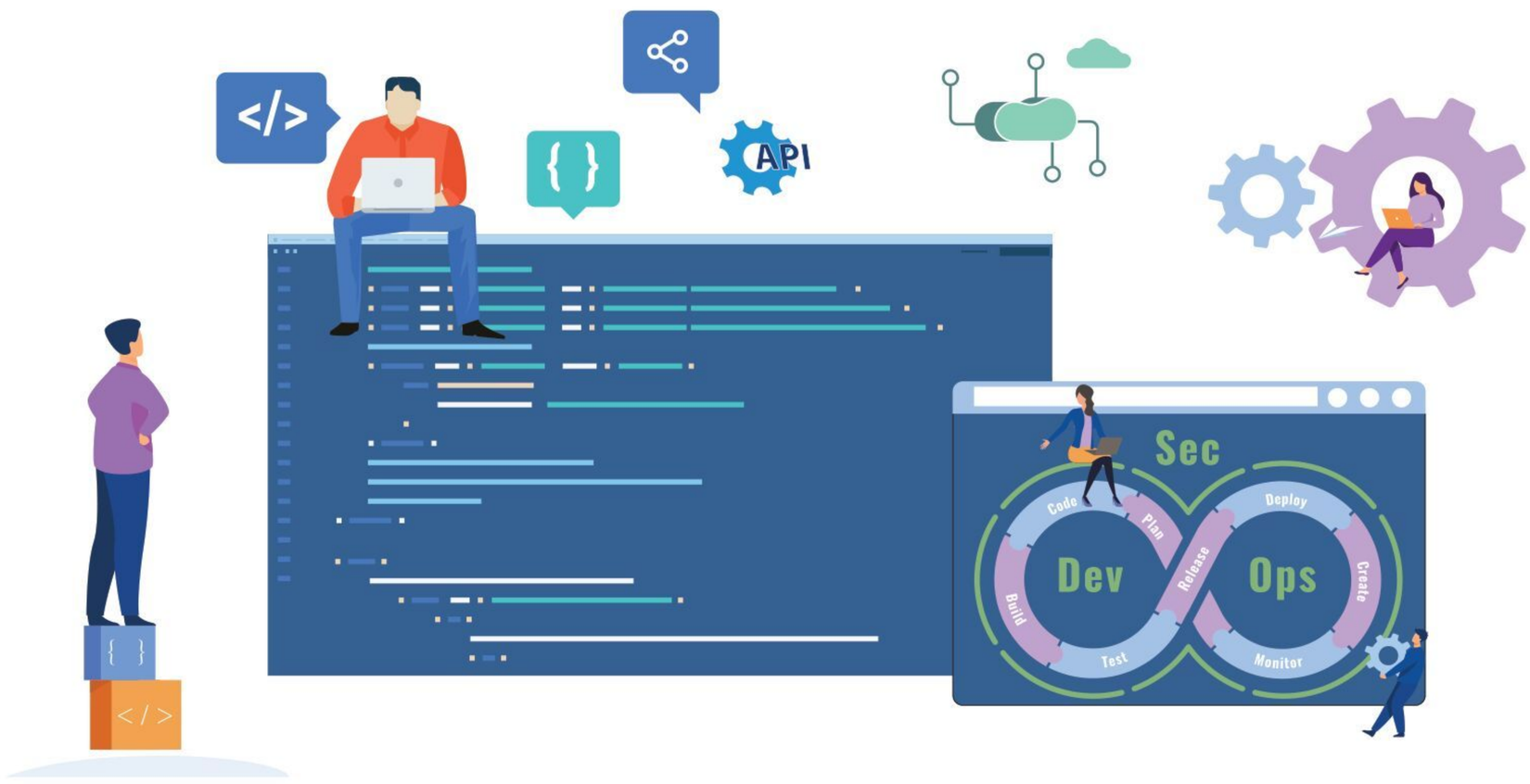
So to conclude, I would say that Java is most certainly not dead, YET. END

By: Mir H. S. Quadri

The author is a research analyst with a specialisation in artificial intelligence and machine learning. He is the founder of Arkinfo.xyz. He loves analysing technological trends and understanding their implications. Being a FOSS enthusiast, he has contributed to several open source projects.

An Introduction to Low-Code and No-Code Test Automation Tools

In today's application development world, low-code and no-code software are omnipresent. As developers have to build applications to work across a wide range of devices, including smartphones, understanding the market requirements and the implications of no-code and low-code software is very important today. Low-code and no-code test automation tools help to overcome software development issues across different platforms.



In this article, we will be talking about the technologies and techniques that can help to accelerate test automation. If you are a developer, you will need to shift focus to adapt to some significant changes and if you are running a business, then these new approaches may open up opportunities that will allow you to move faster and be more innovative than ever before.

Low-code (LC)

Low-code is a visual development approach to application development. It enables developers with different experiences to create applications for the backend, Web and mobile, using drag-and-drop components and model driven logic through a graphical user interface. Low-code platforms enable people to write code and build applications with minimum technical knowhow.

No-code (NC)

No-code software, on the other hand, takes the low-code concept to a higher level, allowing literally anyone to create and modify an application as per their needs without any programming knowledge.

Table 1 lists the different components of the LCNC platform along with a brief description of each component.

Now that we know about the different components of LCNC platforms, let us see how they can help in enhancing the quality of the software and in its testing.

Generic problems with continuous testing

Before going ahead, let us list a few of the common problems with continuous testing processes in the software development life cycle.

- **Skilling of testers:** There are many firms that still rely on manual software testing professionals. They face challenges in meeting the testing needs of agile development cycles as their people are not skilled to support handwritten codes. Learning new technologies involves a lot of training, which adds to the cost and time needed for firms.

Table 1: Components of the LCNC platform

Components	Description
Continuous integration (CI)	CI is a DevOps practice of merging code changes into a version control repository upon task completion. Committing changes triggers the automated build system to use the latest version from the repository, in order to build and run all the automation tests. It finally validates and checks if the code is suitable for release before final deployment to the master branch.
Reusability	As LCNC platforms are majorly generalised, reusability makes it easy for developers to use the code from one project to another or create templates. At times, a developer will only need configuration changes to build several systems from the original one. The underlying system does all the work, which allows the developer using LCNC to easily build applications.
Omni-channel	Rather than developing and supporting independent code and toolsets for each incompatible target, developers are looking for ways to unify development activities and serve many targets from a single, modular code base. And only make cosmetic changes at a high level.
Rules	RPA (robotic process automation) is perhaps the fastest growing LCNC platform. It helps perform tasks that mimic what the user does using rules, with each developer making his or her own decisions. Developers can automate their tasks without writing source code and allow them to run whenever needed.
Performance	The LCNC platform has built-in automated testing, proactive quality monitoring, and real-time performance management. But, at times, this can be a bottleneck for developers if they fail to understand the underlying architecture.
Cloud based	LCNC platforms provide hosting services and take advantage of a public cloud that automatically manages application reliability and scalability with a simple configuration. This reduces the cost and the effort needed to maintain infrastructure. The solution can also be provided in the cloud (Software as a Service).

- **Constant testing:** Due to agile development, the progress of the automation tests needs to be revisited — it's continuous action that requires continuous efforts. This requires significant time, which leads to a delay in the project as developers rely upon testers to begin the next phase of development.
- **Repetitive regression cycles:** Most of the time, developers add new features to the software that was last tested by the tester. This can cause relapses in past highlights. Testers use regression tests to distinguish this issue. However, manual regression testing is illogical in a fast-paced agile methodology.
- **Frequently changing requirements:** Many times, senior management in an organisation changes the prerequisites or drops stories during a sprint cycle, despite the fact that this isn't empowered in an agile/scrum system and no scrum master can facilitate such a request. This means that work half-done pretty well has to be disposed of or altered, which changes the extent of testing out of the blue.

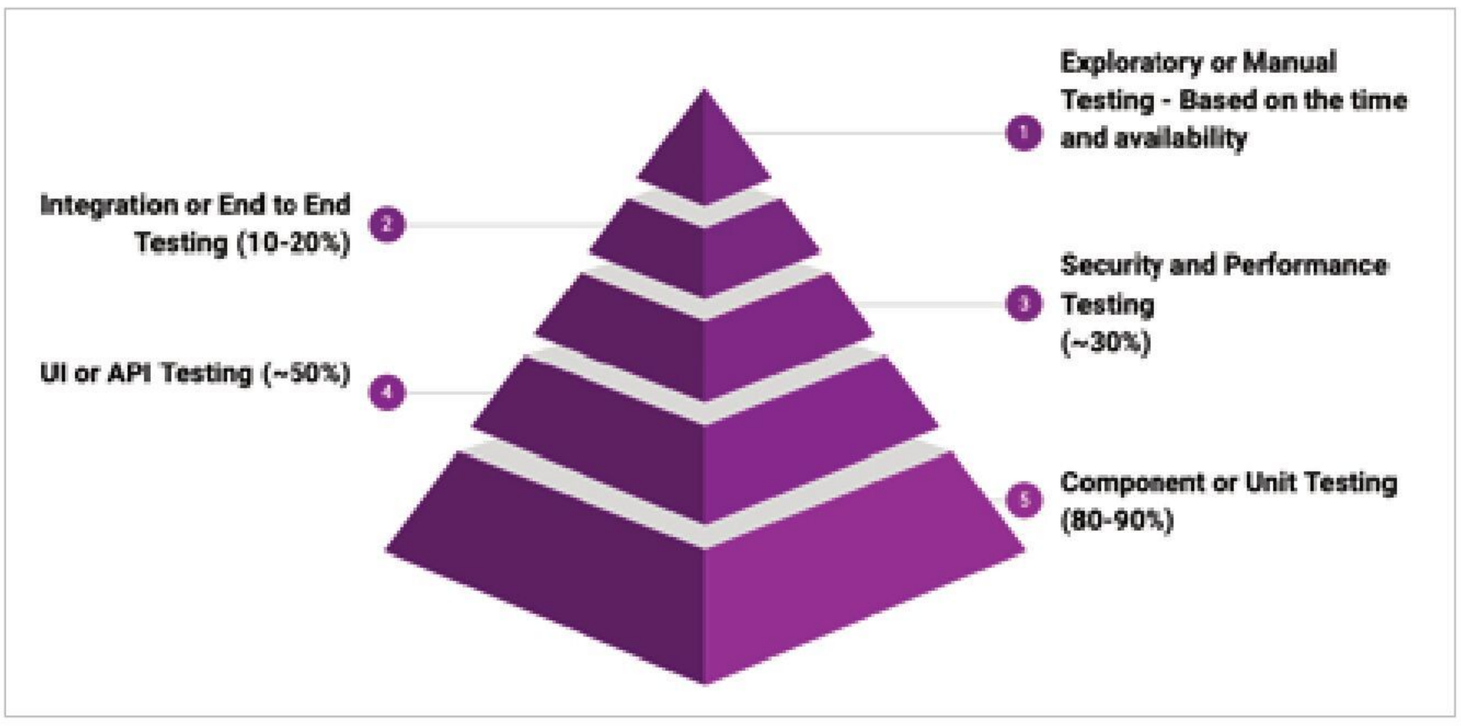


Figure 1: Types of software testing

Role of low-code or no-code platforms in agile testing

Agile testing starts at the beginning of a software development phase and is incorporated in the process of test and development. Generally, the testing is done after the coding stage, but in agile methodology, testing is persistent. So testers mediate between project owners and developers. This makes testing a progressive process, and improves the quality of code while making it resilient.

Improved communication and collaboration, and careful delivery and automation of processes results in:

- Delivering high quality code in faster time
- Reduced chances of failure in new releases
- Improved time for fixes and better recovery

The use of efficient tools saves time and improves the quality of the software. This is where automated testing comes in.

Automated testing

Automated testing is a software testing technique to test and compare the actual outcome with the expected outcome. This can be achieved by creating test scripts or using any automation testing tool to create tests. Test automation can be used to automate repetitive and other testing tasks which are difficult to perform manually.

The LCNC approach can speed up the process of automating, executing and getting a rapid feedback of the application tests.

Low-code or no-code automation tools

Low-code or no-code tools help to overcome software development issues across different platforms. The tables that follow list these tools and give important details about them. Further details can be taken from the documentation links.

Component testing tools

Product	Features	Supported platforms	Ratings	Price	Documentation
Behave	BDD testing	Python	2.5K stars on GitHub	Open source	*https://behave.readthedocs.io/en/stable/index.html*
Hypothesis	Property based testing	Python	5.6K stars on GitHub	Open source	*https://hypothesis.readthedocs.io/en/latest/*

API testing tools

Product	Features	Supported platforms	Ratings	Price	Documentation
SoapUI	API test automation framework for soap and REST	Windows, Linux, Mac OS	1.1K stars on GitHub	Open source	*https://www.soapui.org/getting-started/*
Rest Assured	Testing REST API	Windows, Linux, Mac OS – Needs to be Java enabled	4.8K stars on GitHub	Open source	*https://rest-assured.io*

UI testing tools

Product	Features	Supported platforms	Ratings	Price	Documentation
Selenium	Web automation tool	Windows, Linux, Mac OS	18.1K stars on GitHub	Open source	*https://selenium-python.readthedocs.io*
Katalon	Web, API, mobile and desktop app test automation	Windows, Linux, Mac OS	152 stars on GitHub	Free licence, with paid support services	*https://docs.katalon.com/katalon-studio/docs/index.html*
Appium	Mobile automation testing tool	Windows, Linux, Mac OS	11.7K stars on GitHub	Open source	*https://appium.io/docs/en/about-appium/intro/*

Performance testing tools

Product	Features	Supported platforms	Ratings	Price	Documentation
JMeter	Load and performance testing	Windows, Linux, Mac OS	4.3K stars on GitHub	Open source	*https://jmeter.apache.org/usermanual/index.html*
Locust	Load testing tool	Windows, Linux, Mac OS	2.3K stars on GitHub	Open source	*http://docs.locust.io/en/stable/*

Security testing tools

Security testing ensures that Web applications and information systems remain secure.

Product	Features	Supported platforms	Ratings	Price	Documentation
ZAP	Security testing tool	Windows, Linux, Mac OS	7.4K stars on GitHub	Open source	*https://www.zaproxy.org/docs/*
w3af	Allows testers to find over 200 types of security issues in Web applications	Linux, Mac OS	3.2K stars on GitHub	Open source	*http://docs.w3af.org/en/latest/*

AI-driven testing (AI-DT) tools

Tools	Features	Documentation
Applitools Eyes	Offers visual testing with automation. When applications are developed for different browsers, mobiles and tabs, it becomes difficult to ensure they provide the same features in all. Applitools Eyes is a powerful platform for automated visual testing to help find the differences in the features available in two different kinds of devices.	*https://applitools.com/tutorials/*
Sauce Labs	Provides a testing platform for Web and mobile applications. It runs Selenium tests from a cloud based server. Even if they have low or no knowledge of programming languages, testers can write fully automated test suites, as Sauce Labs is AI powered.	*https://docs.saucelabs.com*
Testim	In Testim, test cases can be created with or without code, or both. It can record tests and export the code, and is very easy to use even by non-developers.	*https://help.testim.io/docs/testim-overview*
SeaLights	SeaLights' technology automatically identifies, analyses, and communicates every perceivable quality risk across the entire delivery pipeline by continuously collecting telemetry data from all stages of the SDLC. Its machine learning-like technology examines both your code and the tests that run on it, letting you know exactly what your tests cover and what they don't.	*https://www.sealights.io*
Test.ai	Test.ai creates test cases for UI based applications. Using AI and ML for identification of elements, it also provides a simple drag-and-drop codeless interface.	https://www.test.ai
mabl	mabl is a low-code and no-code cross-browser testing tool. Test cases are added based on AI and ML using the UI.	*https://help.mabl.com*
retest	retest is used to detect functional differences, and also visual differences.	*https://docs.retest.de*
ReportPortal	The ReportPortal AI-powered test automation dashboard acquires, aggregates and analyses test reports across all environments. It provides visibility of the test execution results throughout the pipeline and gives quick feedback.	*https://reportportal.io*
Appvance IQ (AIQ)	Appvance IQ is an AI driven autonomous test creation and continuous testing system. It creates several test scripts and data driven tests speedily using AI.	*https://www.appvance.com/docs/UserGuide/*
Functionize	Functionize is a cloud based automated testing tool for functional, performance, and load testing – it's a one-stop-shop for all this testing. It is perhaps the most powerful enterprise AI-powered testing tool/platform in the industry. It assists teams in breaking down testing barriers, allowing enterprises to release the product more quickly.	*https://www.functionize.com/*

All these different types of tools are of great use for testing software in really minimal time, as they help testers to match the pace of CI and CD.

Future of software quality assurance

The process of continuous change will continue to be an important part of software development in the future. However, artificial intelligence will change the software development industry completely, and will, in time, be a part of about every piece of software.

The table above lists ten AI-driven testing tools that make testing more simple, without requiring developers or testers to write a lot of code.

Challenges with low-code or no-code

Every solution comes with some hidden problems, and LCNC is no exception. A few of these are:

- Hard to debug and get to lower level of code
- Less control on creating scenarios
- Integration with other tools (not all tools are supported)
- Customisation is difficult
- Vendor lock-in if proprietary tools are being used
- Security risk

Investing in test automation is crucial in order to ensure quality software is delivered in quick time. We should make the best use of low-code and no-code tools for faster automation and release, and also explore the AI-DT tools. END

References

[1] *https://www.techaheadcorp.com/blog/how-ci-cd-save-app-development-time/*
[2] *https://www.mendix.com/low-code-guide/*
[3] *https://dzone.com/articles/top-5-challenges-in-agile-testing*
[4] *https://www.softwaretestinghelp.com/open-source-testing-tools/*
[5] *https://qainfotech.com/top-ai-powered-test-automation-tools/*

By: Thangaselvi Arichandrapandian

The author works in a leading bank as AVP. She is a perennial learner and loves the quote: "The more I learn, the more I realise how much I don't know."- Albert Einstein.

Test Your Data Science Project with Google's Caliban

Data science projects need a real-time test environment to see if they are useful in real-life situations. Caliban solves this hurdle for a data scientist. It acts as a testing tool that tracks the environmental properties of your project during execution, and reproduces a specific running environment.



(Source: https://github.com/google/caliban)

Data science is an emerging field of advanced data analysis comprising collection, preparation, exploration, analysis, testing, visualisation, management and preservation of large data sets. In some cases, even a well-executed data science project may fail due to limitations in the test environment. Testing the project in all scenarios is mandatory to make it ready-to-use in reality. Caliban is a great open source tool developed by Google. It helps develop research workflows in isolated and reproducible learning environments. Caliban makes it easy to develop any machine learning prototype locally and run the code on your device. A similar code can be tried in a cloud environment for execution on giant machines. It is best suited with JAX, TensorFlow and PyTorch for machine learning workflows.

Caliban requires Python 3.6 or above as a prerequisite. Docker is another important component. Install nvidia-docker2 version, if you have a CPU and want to use Caliban.

Installing Caliban

Once the prerequisites are in place, Caliban can be installed using the following commands:

```
pipx install caliban
pip install caliban
```

The same commands can also be used to upgrade Caliban. After it's installed, you can check Caliban on the device by giving the following command:

```
caliban - help
```

This open source tool contains several useful sub-commands, which are essential for the execution of any project.

- **shell:** This generates a Docker image with the required dependencies. This environment is similar to what is available with the code.
- **notebook:** This starts a Jupyter notebook within a Docker image containing dependencies.
- **run:** This command packages directory code into a Docker image and executes it locally.
- **cloud:** This command allows submitting jobs to the cloud and runs inside the same Docker image that is used with the *run* command.
- **build:** This builds a Docker image used in the cloud and runs without submitting any code.
- **cluster:** This command creates Google Kubernetes Engine (GKE) clusters and submits jobs to them.
- **status:** This displays information about all the jobs submitted by Caliban. It is used to cancel pending jobs, re-build containers, or resubmit a batch after fixing errors.
- **stop:** This command stops running jobs.

- **resubmit:** This resubmits the batch of experiments.

Getting started

First, create an empty directory and write a script in Python to train a basic model. Next, create a file called *requirements.txt* and declare dependency. Caliban automatically makes an entry in the 'requirements.txt' file when the code is running. Train your model with the command:

```
caliban run –nogpu  script.py
```

Check the accuracy of the model and if it is not up to the mark, improve it by setting the default learning rate with the following command:

```
caliban run --nogpu script.py --
--learning_rate 0.01
```

You can experiment with broadcasting, which allows many jobs with different sets of parameters to run at the same time:

```
caliban run --experiment_config
experiment.json --nogpu script.py
```

The job can be submitted to the cloud by giving the following command:

```
caliban cloud --nogpu script.py --
--learning_rate 0.01
```

With the cloud you can use on-demand GPUs and TPUs to train models on huge data sets at very high speeds.

As can be seen, Caliban is very helpful for data scientists as it provides a robust testing environment. END



Reference

[1] *https://www.kdnuggets.com/2021/06/5-data-science-open-source-projects-contribute.html*
[2] *https://caliban.readthedocs.io/en/latest/getting_started/*
[3] *https://github.com/google/caliban#caliban*

By: Dr Dharmendra Patel and Dr Atul Patel

Dr Dharmendra Patel is a professor at Smt. Chandaben Mohanbhai Patel Institute of Computer Applications, Gujarat. His areas of interest are data mining, Web mining, image processing, Internet of Things and software engineering.

Dr Atul Patel is a principal and professor at Smt. Chandaben Mohanbhai Patel Institute of Computer Applications, Gujarat. His areas of interest include ad hoc networks, artificial intelligence, data mining, image processing and Internet of Things.



THE COMPLETE MAGAZINE ON OPEN SOURCE



» **To find dealers** to buy copy from news-stand, **visit: www.ezine.lfymag.com**

» **To subscribe**, **visit: pay.efyindia.com**

» **To buy an ezine edition**, **visit: www.magzter.com** choose **Open Source For You**

R Testing Frameworks

In this seventh article in the 'R, Statistics and Machine Learning' series, we shall explore the various testing frameworks available in R, and how to use them.



We will be using R version 4.1.0 installed on Parabola GNU/Linux-libre (x86-64) for the code snippets.

RUnit

RUnit is a unit testing framework for R written by Thomas König, Klaus Jünemann, and Matthias Burger and is similar to the xUnit testing tools. The package is released under the GPLv2 license, and the last major version is 0.4.32. It is not actively developed, but a number of R packages use the RUnit framework and hence the package is maintained. You can install the package in an R console as shown below:

```
> install.packages("RUnit")
  ...
  *** installing help indices
  ** building package indices
  ** installing vignettes
  ** testing if installed package can be loaded from
temporary location
  ** testing if installed package can be loaded from final
location
  ** testing if installed package keeps a record of temporary
installation path
  * DONE (RUnit)
```

You can load the library to the R session using the following command:

```
> library(RUnit)
```

Consider a square function:

```
> square <- function(x) return (x * x)
```

A *test.square* function can be written to use the *checkEquals* function to assert expected values, as shown below:

```
> test.square <- function() {
    checkEquals(square(2), 4)
  }
```

The test function can now be executed in the R console, as illustrated below:

```
> test.square()
[1] TRUE
```

You can now increase the scope of the tests for zero, negative numbers, fractions and complex numbers as well.

```
> checkEquals(square(0), 0)
[1] TRUE
> checkEquals(square(-1), 1)
[1] TRUE
> checkEquals(square(1/2), 1/4)
[1] TRUE
> checkEquals(square(1i), -1+0i)
[1] TRUE
```

The special constants in R can also be tested with the *checkEquals()* function as follows:

```
> checkEquals(NA, NA)
[1] TRUE
> checkEquals(Inf, Inf)
[1] TRUE
> checkEquals(NaN, Nan)
Error in mode(current) : object 'Nan' not found
> checkEquals(NaN, NaN)
[1] TRUE
> checkEquals(NULL, NULL)
[1] TRUE
```

The 'i' for complex numbers in R must be prefixed with a number; otherwise, it will throw an error. This error can be caught using the *checkException* function, as shown below:

```
> checkException(square(i), -1+0i)
```

```
Error in square(i) : object 'i' not found
[1] TRUE
```

A number of other assert functions are provided by the RUnit package. The *checkEqualsNumeric* function checks if both the compared elements are equal and numeric. Otherwise, it returns an error, as illustrated below:

```
> checkEqualsNumeric(square(2), 4)
[1] TRUE

> checkEqualsNumeric(square(2), 3)
Error in checkEqualsNumeric(square(2), 3) :
  Mean relative difference: 0.25

> checkEqualsNumeric(square(2), 'a')
Error in checkEqualsNumeric(square(2), "a") : Modes: numeric,
character
target is numeric, current is character
```

The *checkIdentical* function is used to verify if the observed and expected values are equal, and it also has an additional logging mechanism. In the following example, two vectors, 'u' and 'v', are compared with each other as well as a list 'l'. A log message is printed if the compared elements are not identical.

```
> u <- c(1, 2, 3, 4)
> v <- c(1, 2, 3, 4)
> l <- list(1, 2, 3, 4)

> checkIdentical(v, l, "The types are different")
Error in checkIdentical(v, l, "The types are different") :
FALSE
 The types are different

> checkIdentical(u, v, "The types are different")
[1] TRUE
```

The *checkTrue* function takes a Boolean expression and a message, and asserts to see if the expression is true. Otherwise, it prints the message as shown below:

```
> checkTrue(TRUE, "True")
[1] TRUE

> checkTrue(FALSE, "False")
Error in checkTrue(FALSE, "False") : Test not TRUE
False

> checkTrue(5 > 2, "True")
[1] TRUE

> checkTrue(5 < 2, "False")
Error in checkTrue(5 < 2, "False") : Test not TRUE
False
```

The RUnit package also provides a tool for code inspection to obtain a test coverage report. This is useful to know which lines of code have been executed during test runs, and what additional tests need to be written. A tracker needs to be initialised, and the *inspect()* method needs to be invoked on the code with arguments. For example, the following commands demonstrate the sequence to run the inspection tool for the square function.

```
> track <- tracker()
> track$init()
> x <- 1:3
> resultSquare <- inspect(square(x), track=track)
> resTrack <- track$getTrackInfo()
> printHTML.trackInfo(resTrack)
```

A *results/* folder is then generated with the following file contents:

```
$ ls reports/
con1.html  con1.png  index.html  result1.html
```

The *result1.html* file that provides the code coverage report is shown in Figure 1.

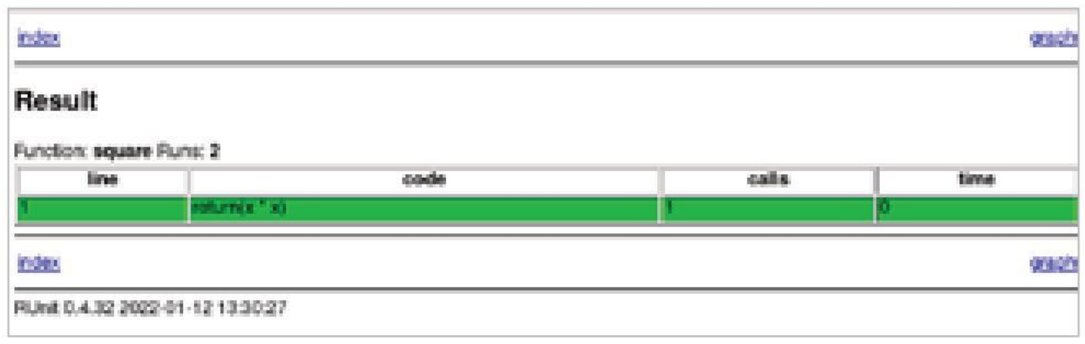


Figure 1: RUnit inspector report

For larger file modules, the connection graph between the functions and modules is also presented in the HTML output.

The *R CMD check* command can be used to build and also run the tests in an R source package. The *aucm* R package implements functions to identify linear and non-linear marker combinations. This maximises the area under the curve (AUC). It is released under the GPLv2 license and the tests are implemented using RUnit. A sample run of the tests is shown below:

```
$ R CMD check aucm
 ...
 * using R version 4.1.2 (2021-11-01)
 * using platform: x86_64-pc-linux-gnu (64-bit)
 * using session charset: UTF-8
```

```
 * checking dependencies in R code ... OK
 * checking compilation flags in Makevars ... OK
 * checking for GNU extensions in Makefiles ... OK
 * checking for portable use of $(BLAS_LIBS) and $(LAPACK_
LIBS) ... OK
 * checking use of PKG_*FLAGS in Makefiles ... OK
 * checking include directives in Makefiles ... OK
 * checking compiled code ... OK
 * checking examples ... OK
 * checking for unstated dependencies in 'tests' ... OK
 * checking tests ...
   Running 'doRUnit.R'
  OK
 * DONE

Status: OK
```

testthat

testthat is another unit testing framework similar to the xUnit family of test packages. You can install the same in an R session using the following command:

```
> install.packages("testthat")
```

The library needs to be loaded into the R session, as shown below:

```
> library (testthat)
```

Consider the same square function example:

```
> square <- function(x) return (x * x)
```

You can use the testthat library's *expect_equal* assertion to validate the function execution:

```
> test_that ("square of two", {
   expect_equal(square(2), 4)
  })
Test passed
```

An error is reported if there is a mismatch in the *expect_equal* check, as shown below:

```
> expect_equal(square(2), 3)
Error: square(2) not equal to 3.
1/1 mismatches
[1] 4 - 3 == 1
```

The *expect_equal()* function is more suitable for numeric values only. The *expect_identical* function is useful to compare objects, and checks if they have the same values as well. A couple of examples are shown below:

```
> u <- c('one', 'two', 'three')
> v <- c('one', 'two', 'three')
> a <- list(1, 2, 3, 4)

> test_that ("expect identical", {
    expect_identical(u, a)
    })
── Failure (Line 2): expect identical ──
`u` not identical to `a`.
Types not compatible: character is not list
Error: Test failed

> test_that ("expect identical", {
   expect_identical(u, v)
   })
Test passed
```

The *expect_is* function checks if the argument belongs to an object or inherits it from a specific class. For example:

```
> test_that ("expect is", {
    expect_is(u, "character")
  })
Test passed

> test_that ("expect is", {
    expect_is(a, "character")
  })
── Failure (Line 2): expect is ─────────────────────────────

`a` inherits from `'list'` not `'character'`.
Error: Test failed
```

A character vector can be matched to a regular expression using the *expect_match* function. A few examples are given below:

```
> quote <- "Science is curiosity, testing and experimenting"
> test_that ("expect match", {
    expect_match(quote, "Science")
  })
Test passed

> test_that ("expect match", {
    expect_match(quote, "science")
  })
── Failure (Line 2): expect match ──────────────────────────

`quote` does not match "science".
Actual value: "Science is curiosity, testing and
```

```
experimenting"
Backtrace:
 1. testthat::expect_match(quote, "science")
 2. testthat:::expect_match_(...)
Error: Test failed

> test_that ("expect match", {
    expect_match(quote, "science", ignore.case = TRUE)
  })
Test passed
```

You can pass an *ignore.case=TRUE* third argument option to the *expect_match()* function for a case insensitive match.

The *expect_message*, *expect_warning*, *expect_error*, and *expect_output* functions are a variant of *expect_match* that examine the output. Examples of their usage are given below:

```
> test_that ("expect message", {
    expect_message(message("2021-02-11 Summary"), "2021-02-11
Summary")
  })
Test passed

> test_that ("expect message", {
    expect_message(message("2021-02-11 Summary"), "Summary
Statistics")
  })
── Failure (Line 2): expect message ────────────────────────

`message("2021-02-11 Summary")` produced unexpected messages.
Expected match: Summary Statistics
Actual values:
 * 2021-02-11 Summary
Error: Test failed
> test_that ("expect warning", {
    expect_warning(warning("Danger!"), "Danger")
  })
Test passed

> test_that ("expect warning", {
    expect_warning(warning("Danger!"), "Danger Alert!")
  })
── Failure (Line 2): expect warning ────────────────────────

`warning("Danger!")` produced unexpected warnings.
Expected match: Danger Alert!
Actual values:
 * Danger!
Error: Test failed

> test_that ("expect error", {
    expect_error(10 / quote, "Numeric argument required")
  })
── Failure (Line 2): expect error ──────────────────────────

`10/quote` threw an error with unexpected message.
Expected match: "Numeric argument required"
Actual message: "non-numeric argument to binary operator"
Backtrace:
 1. testthat::expect_error(10/quote, "Numeric argument
required")
 2. testthat:::quasi_capture(...)
 5. rlang::eval_bare(quo_get_expr(.quo), quo_get_env(.quo))
Error: Test failed
> test_that ("expect output", {
    expect_output(str(quote), "Science is curiosity, testing
and experimenting")
  })
Test passed

> test_that ("expect output", {
    expect_output(str(quote), "A diagnostic message")
 })
── Failure (Line 2): expect output ─────────────────────────

`str\(quote\)` does not match "A diagnostic message".
Actual value: " chr "Science is curiosity, testing and
experimenting""
Backtrace:
 1. testthat::expect_output(str(quote), "A diagnostic
message")
 2. testthat::expect_match(...)
 3. testthat:::expect_match_(...)
Error: Test failed
```

The *expect_lt*, *expect_lte*, *expect_gt*, and *expect_gte* functions compare two values to be less than or greater than and equal to each other. Examples for the same for the success and failure cases are illustrated in the following examples:

```
> test_that ("expect lt", {
    expect_lt(1, 2)
  })
Test passed

> test_that ("expect lt", {
    expect_lt(2, 1)
 })
── Failure (Line 2): expect lt ─────────────────────────────

2 is not strictly less than 1. Difference: 1
Error: Test failed

> test_that ("expect lte", {
```

```
    expect_lte(1, 1)
  })
Test passed

> test_that ("expect lte", {
    expect_lte(1, 2)
  })
Test passed

> test_that ("expect lte", {
    expect_lte(2, 1)
  })
── Failure (Line 2): expect lte ──────────────────────────

2 is not less than 1. Difference: 1
Error: Test failed

> test_that ("expect gt", {
    expect_gt(2, 1)
  })
Test passed

> test_that ("expect gt", {
    expect_gt(1, 2)
  })
── Failure (Line 2): expect gt ──────────────────────────

1 is not strictly more than 2. Difference: -1
Error: Test failed
> test_that ("expect gte", {
    expect_gte(2, 2)
  })
Test passed

> test_that ("expect gte", {
    expect_gte(2, 1)
  })
Test passed

> test_that ("expect gte", {
    expect_gte(1, 2)
  })
── Failure (Line 2): expect gte ──────────────────────────
1 is not more than 2. Difference: -1
Error: Test failed
```

The *expect_silent* function can be used to assert if a code produces any output or not. For example:

```
> test_that ("expect silent", {
    expect_silent(quote)
  })
Test passed

> test_that ("expect silent", {
    expect_silent(str(quote))
  })
── Failure (Line 2): expect silent ──────────────────────

`str(quote)` produced output.
Error: Test failed
```

The *expect_true* and *expect_false* checks can be used if none of the above validations are suitable for your requirements.

```
> test_that ("expect true", {
    expect_true(TRUE)
  })
Test passed

> test_that ("expect true", {
    expect_true(FALSE)
  })
── Failure (Line 2): expect true ──────────────────────────
FALSE is not TRUE
`actual`:   FALSE
`expected`: TRUE
Error: Test failed

> test_that ("expect false", {
    expect_false(TRUE)
  })
── Failure (Line 2): expect false ──────────────────────────
TRUE is not FALSE
`actual`:   TRUE
`expected`: FALSE
Error: Test failed

> test_that ("expect false", {
    expect_false(FALSE)
  })
Test passed
```

At times, when you want to omit a test from execution, you can use the *skip* function, which takes an argument message. A meaningful text can be used for the message to be printed during execution of the tests to indicate the reason for skipping the test.

The *R CMD check* command can again be used to run the testthat tests for a package. Consider the high dimensional numerical and symbolic *calculus* package that provides high-order derivatives, ordinary differential equations, tensor calculus, Einstein summing convention, Taylor series expansion and numerical integration. The output of the check command that runs the tests is shown below:

```
$ R CMD check calculus
 * using R version 4.1.2 (2021-11-01)
 * using platform: x86_64-pc-linux-gnu (64-bit)
 * using session charset: UTF-8
 * checking compiled code ... OK
 * checking installed files from 'inst/doc' ... OK
 * checking files in 'vignettes' ... OK
 * checking examples ... OK
 * checking for unstated dependencies in 'tests' ... OK
 * checking tests ...
   Running 'testthat.R'
  OK
* checking for unstated dependencies in vignettes ... OK
* checking package vignettes in 'inst/doc' ... OK
* checking running R code from vignettes ...
   'derivatives.Rmd' using 'UTF-8'... OK
   'differential-operators.Rmd' using 'UTF-8'... OK
   'einstein.Rmd' using 'UTF-8'... OK
   'hermite.Rmd' using 'UTF-8'... OK
   'integrals.Rmd' using 'UTF-8'... OK
   'ode.Rmd' using 'UTF-8'... OK
   'taylor.Rmd' using 'UTF-8'... OK
  NONE
* checking re-building of vignette outputs ... OK
* DONE

Status: OK
```

tinytest

The tinytest R package is meant to be a lightweight, no-dependency, complete package for unit testing. It is released under the GPLv3 license, and is meant to be easy to use. A list of available test functions provided by the package is given in the following table:

Function	Purpose
expect_match	A string regular expression match.
expect_true	Check that the argument is TRUE.
expect_false	Check that the argument is FALSE.
expect_null	Check that the expression is NULL.
expect_equal	The data and argument attributes must be equal.
expect_inherits	The object must inherit from the specific class.
expect_error	The expression must be an error.
expect_warning	The expression must be a warning.
expect_message	The expression must be a message.
expect_silent	There should not be an output.

The syntax for the test functions is similar to that of testthat. Consider the same square function once again:

```
> square <- function(x) return (x * x)
```

You can use the tinytest library's **expect_equal** assertion as follows:

```
> tinytest::expect_equal(square(2), 4)
----- PASSED      : <-->
 call| tinytest::expect_equal(square(2), 4)

> tinytest::expect_equal(square(2), 3)
----- FAILED[data]: <-->
 call| tinytest::expect_equal(square(2), 3)
 diff| Expected '3', got '4'
```

The objective here is to easily set up the test cases with simple R scripts. You can run the tests for the tinytest package from the R console itself, and produce a summary output as shown below:

```
> library(tinytest)

> out <- run_test_dir(system.file("tinytest",
package="tinytest"), verbose=0)

> summary(out)

File                    Results fails passes
  test_call_wd.R              1     0      1
  test_env_A.R                2     0      2
  test_env_B.R                6     0      6
  test_extensibility.R        1     0      1
  test_file.R                23     0     23
  test_gh_issue_17.R          2     0      2
  test_gh_issue_32.R          3     0      3
  test_gh_issue_51.R          1     0      1
  test_gh_issue_58.R          2     0      2
  test_gh_issue_86.R          4     0      4
  test_init.R                 1     0      1
  test_RUnit_style.R          5     0      5
  test_tiny.R                74     0     74
  test_utils.R                4     0      4
Total                       129     0    129
```

You are encouraged to read the reference manuals for the above unit testing frameworks, and use them in your R packages. END

References

[1] RUnit - A Unit Test Framework in R, *https://cran.r-project.org/web/packages/RUnit/vignettes/RUnit.pdf*
[2] Testthat, *https://testthat.r-lib.org/reference/index.html*
[3] Tinytest, *https://github.com/markvanderloo/tinytest*
[4] Using Tinytest, *https://cran.r-project.org/web/packages/tinytest/vignettes/using_tinytest*.pdf

By: Shakthi Kannan

The author is a free software enthusiast.

Installing and Using RethinkDB on Ubuntu



Here's a short tutorial on how to install and use RethinkDB, the popular NoSQL database, on Ubuntu machines.

RethinkDB is an open source, free, distributed, and document based database server. This NoSQL database has been developed by Rethink, and can store JSON files with schemas that are dynamic. More importantly, real-time updates for query results can be pushed for applications to use. Since it is a distributed database, it is highly scalable. It has many automatic functions, making it a highly available database. As it is a popular database today, it is important that we learn how to use it.

Let's start with how we can install and use it on our Ubuntu machines. First, to add the packages required to install RethinkDB to our updates, we need to run the following command, as shown in Figure 1.

```
source /etc/lsb-release && echo"
deb https://download.rethinkdb.com/
repository/ubuntu-$DISTRIB_CODENAME
$DISTRIB_CODENAME main" | sudo tee /
etc/apt/sources.list.d/rethinkdb.list
```

The following command adds it to our repositories (Figure 2):

```
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$ source /etc/lsb-release &
& echo "deb https://download.rethinkdb.com/repository/ubuntu-$DISTRIB_CODENAME $
DISTRIB_CODENAME main" | sudo tee /etc/apt/sources.list.d/rethinkdb.list
deb https://download.rethinkdb.com/repository/ubuntu-bionic bionic main
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$
```

Figure 1: Adding the packages to updates

```
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$ wget -qO- https://downloa
d.rethinkdb.com/repository/raw/pubkey.gpg | sudo apt-key add -
[sudo] password for jishnusaurav:
OK
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$
```

Figure 2: Adding RethinkDB to our repositories

```
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$ sudo apt update
[sudo] password for jishnusaurav:
Hit:1 http://in.archive.ubuntu.com/ubuntu bionic InRelease
Get:2 http://in.archive.ubuntu.com/ubuntu bionic-updates InRelease [88.7 kB]
Hit:3 https://download.docker.com/linux/ubuntu bionic InRelease
Get:4 http://in.archive.ubuntu.com/ubuntu bionic-backports InRelease [74.6 kB]
Hit:5 http://downloads.metasploit.com/data/releases/metasploit-framework/apt luc
id InRelease
Get:6 http://security.ubuntu.com/ubuntu bionic-security InRelease [88.7 kB]
Get:7 http://in.archive.ubuntu.com/ubuntu bionic-updates/main i386 Packages [1,3
70 kB]
Hit:8 http://packages.microsoft.com/repos/code stable InRelease
Hit:9 https://linux.teamviewer.com/deb stable InRelease
Hit:10 https://packages.microsoft.com/repos/ms-teams stable InRelease
Err:11 https://apache.bintray.com/couchdb-deb bionic InRelease
  403  Forbidden [IP: 52.37.4.158 443]
Ign:12 http://ppa.launchpad.net/couchdb/stable/ubuntu bionic InRelease
Hit:13 https://deb.nodesource.com/node_14.x bionic InRelease
Get:14 http://in.archive.ubuntu.com/ubuntu bionic-updates/main amd64 Packages [2
,277 kB]
Get:15 http://in.archive.ubuntu.com/ubuntu bionic-updates/main amd64 DEP-11 Meta
data [303 kB]
```

Figure 3: Updating package manager

```
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$ sudo apt-get install reth
inkdb
Reading package lists... Done
Building dependency tree
Reading state information... Done
rethinkdb is already the newest version (2.4.1~0bionic).
The following packages were automatically installed and are no longer required:
  libevent-core-2.1-6 linux-hwe-5.4-headers-5.4.0-73
  linux-hwe-5.4-headers-5.4.0-74 linux-hwe-5.4-headers-5.4.0-77
  linux-hwe-5.4-headers-5.4.0-80 linux-hwe-5.4-headers-5.4.0-81
  linux-hwe-5.4-headers-5.4.0-84 linux-hwe-5.4-headers-5.4.0-86
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 76 not upgraded.
N: Ignoring file 'nvi' in directory '/etc/apt/sources.list.d/' as it has no file
name extension
jishnusaurav@jishnusaurav-IdeaPad-L340-15IRH-Gaming:~$
```

Figure 4: RethinkDB installation

```
jlshnusaurav@jlshnusaurav-IdeaPad-L340-15IRH-Gaming:~$ sudo systemctl start rethinkdb
jlshnusaurav@jlshnusaurav-IdeaPad-L340-15IRH-Gaming:~$ sudo systemctl status rethinkdb
● rethinkdb.service - LSB: This starts a set of rethinkdb server instances.
   Loaded: loaded (/etc/init.d/rethinkdb; generated)
   Active: active (exited) since Thu 2021-10-28 00:57:14 IST; 1min 2s ago
     Docs: man:systemd-sysv-generator(8)
    Tasks: 0 (limit: 4915)
   CGroup: /system.slice/rethinkdb.service

Oct 28 00:57:14 jlshnusaurav-IdeaPad-L340-15IRH-Gaming systemd[1]: Starting LSB:
Oct 28 00:57:14 jlshnusaurav-IdeaPad-L340-15IRH-Gaming rethinkdb[26969]: rethink
Oct 28 00:57:14 jlshnusaurav-IdeaPad-L340-15IRH-Gaming rethinkdb[26969]: rethink
Oct 28 00:57:14 jlshnusaurav-IdeaPad-L340-15IRH-Gaming systemd[1]: Started LSB:
lines 1-11/11 (END)
```

Figure 5: Starting and checking the status of RethinkDB

```
jlshnusaurav@jlshnusaurav-IdeaPad-L340-15IRH-Gaming:~$ rethinkdb
Running rethinkdb 2.4.1~0bionic (CLANG 6.0.0 (tags/RELEASE_600/final))...
Running on Linux 5.4.0-89-generic x86_64
Loading data from directory /home/jlshnusaurav/rethinkdb_data
Listening for intracluster connections on port 29015
Listening for client driver connections on port 28015
Listening for administrative HTTP connections on port 8080
Listening on cluster addresses: 127.0.0.1, ::1
Listening on driver addresses: 127.0.0.1, ::1
Listening on http addresses: 127.0.0.1, ::1
To fully expose RethinkDB on the network, bind to all addresses by running rethi
nkdb with the `--bind all` command line option.
Server ready, "jlshnusaurav_IdeaPad_L340_15IRH_Gaming_2wg" 4ecf96c2-727b-4a8a-a1
3d-0e60b5481c80
```

Figure 6: Setting the RethinkDB server

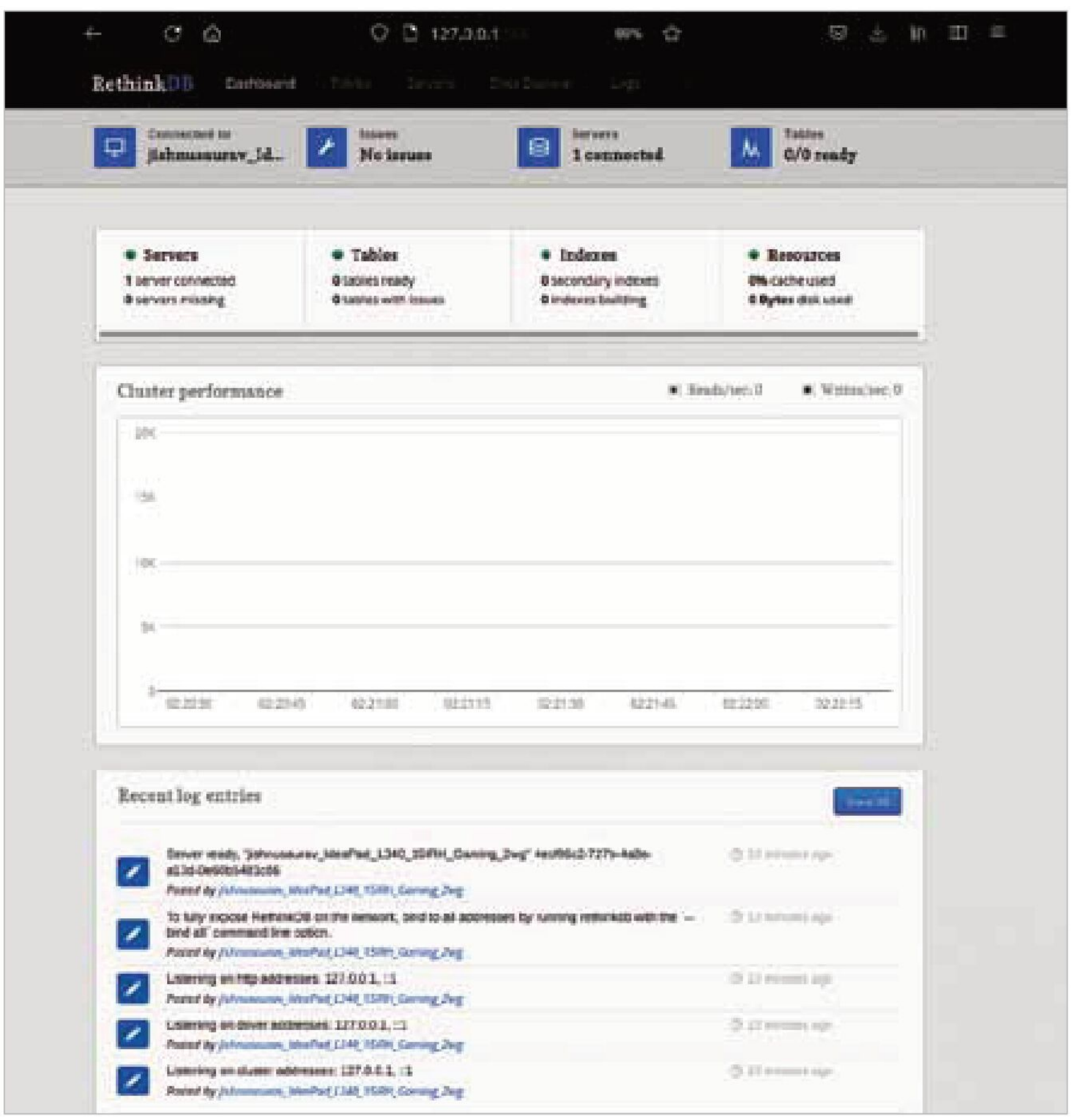


Figure 7: Opening the RethinkDB UI on the browser

```
wget -qO- https://download.
rethinkdb.com/repository/raw/
pubkey.gpg | sudo apt-key add -
```

After adding new repositories, our package manager has to be updated using the following command, as shown in Figure 3.

```
sudo apt update
```

Once this is done, we can install RethinkDB using the following command, as shown in Figure 4.

```
sudo apt-get install rethinkdb
```

After installation, we first need to start it by running the following command:

```
sudo systemctl start rethinkdb
```

To check if it has started properly, enter the following command, as shown in Figure 5:

```
sudo systemctl status rethinkdb
```

We can now start using it. To set it up, enter ‘rethinkdb’ on the terminal, as shown in Figure 6.

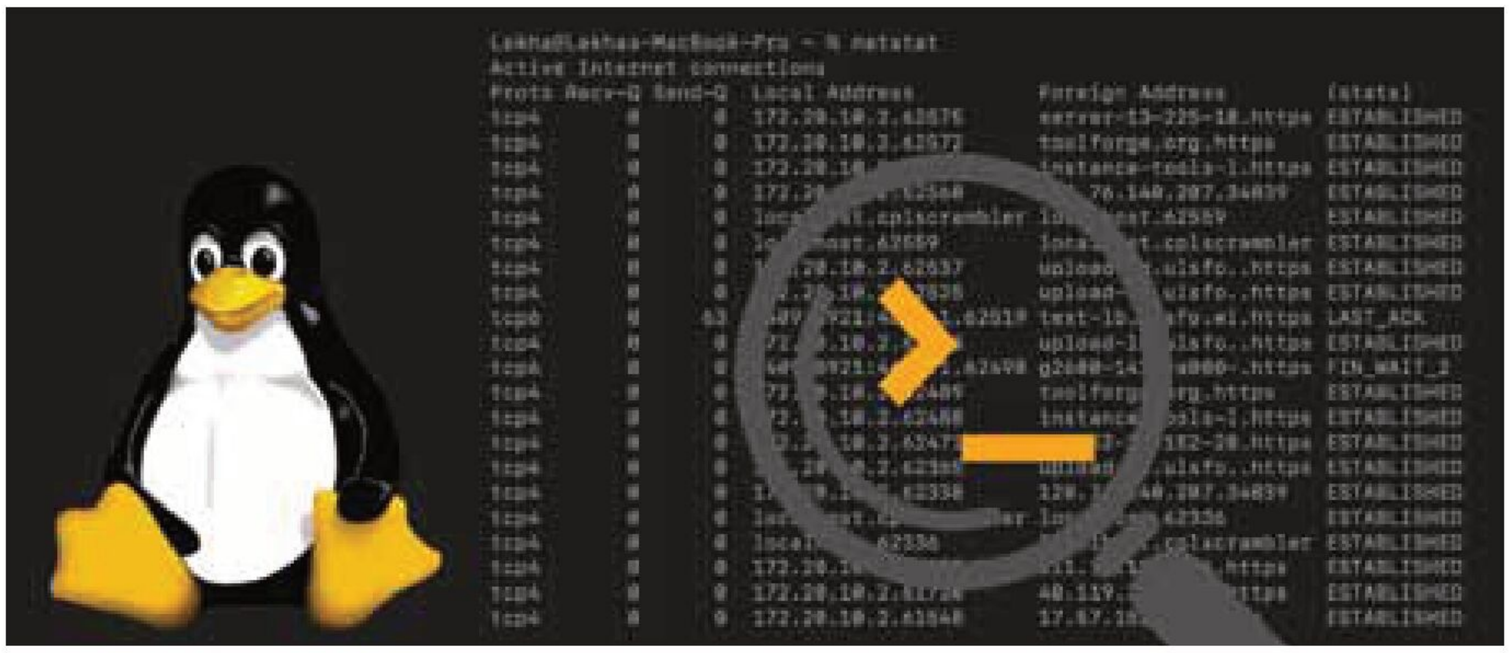
To open up the database server, enter the IP address with port number 8080 on the browser, as shown in Figure 7.

RethinkDB can now be used as any NoSQL database. It’s great fun to explore it and use it to its fullest potential!! END

By: Jishnu Saurav Mittapalli

The author is a full stack Web developer, interested in modern technologies like artificial intelligence and machine learning. He is currently working in the field of natural language processing.

Exploring Network Statistics with 'netstat'



netstat is generally used to find the various network connections a machine has, but has other functionalities too. It may be considered obsolete by some, but can be a handy tool when network issues have to be debugged.

'netstat', also known as network statistics, is most commonly used for displaying current network connections in devices/machines. However, it can also be used for printing the routing table, network card statistics, etc, of the machines.

In Linux, netstat is obsolete and hence it is recommended to use a combination of *ss* and *ip cli* utilities. That being said, it is still shipped by default with major Linux distributions, and is also a cross-platform tool available in BSD systems, MacOS and even Windows. So it's a good tool to keep in your toolset.

Understanding netstat output

While using a browser, you will definitely have a bunch of network connections. netstat is generally used to find all the network connections a machine has.

```
> netstat | head
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address   Foreign Address     State
tcp  0      0 arya:54960   a104-113-212-245.:https ESTABLISHED
tcp  0      0 arya:54958   a104-113-212-245.:https ESTABLISHED
tcp  0      0 arya:40910   52.98.58.34:https       ESTABLISHED
tcp  0      0 arya:38398   52.114.159.60:https     ESTABLISHED
tcp  0      0 arya:55748   13.69.109.131:https     ESTABLISHED
tcp  0      0 arya:48510   52.98.59.18:https       ESTABLISHED
tcp  0      0 arya:44802   lb-140-82-113-26-:https ESTABLISHED
tcp  0      0 arya:46554   ec2-65-1-252-10.a:https ESTABLISHED
```

The meaning of the individual columns in the output above is as follows.

Column name	Description
Proto	The protocol used by the socket; e.g., TCP, UDP
Recv-Q	If the connection has already been established, as is the case with the above output, then it refers to the count of bytes that are not copied by the user program that is connected to this socket
Send-Q	Refers to the count of bytes not acknowledged by the remote host in this established connection state
Local address	Address and port number of the local end of the socket
Foreign address	Address and port number of the remote end of the socket
State	State of the socket: Listen, Established, etc.

In the above output, you can see the various values in the 'foreign address' column. These are the different IP addresses the local machine is speaking to. We can also find out the process that initiated the above socket connections via the *--program* or *-p* option:

```
> netstat -p | head
(Not all processes can be identified; non-owned process info
will not be shown, you should have the root access to view
all of them.)
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name

tcp    0   0 arya:40620              51.105.71.137:https
ESTABLISHED 2049/chrome --type=

tcp    0   0 arya:34078              a104-96-204-135.d:https
ESTABLISHED 2049/chrome --type=

tcp    0   0 arya:36442              52.113.206.194:https
ESTABLISHED 2049/chrome --type=

tcp    0   0 arya:36440              52.113.206.194:https
ESTABLISHED 2049/chrome --type=

tcp    0    0 arya:35088             52.114.133.162:https
ESTABLISHED 2049/chrome --type=

tcp    0    0 arya:48510             52.98.59.18:https
ESTABLISHED 2049/chrome --type=

tcp    0    0 arya:44802             lb-140-82-113-26-:https
ESTABLISHED 2049/chrome --type=

tcp    0    0 arya:46554             ec2-65-1-252-10.a:https
ESTABLISHED 2049/chrome --type=
```

Though a warning is given that not all processes could be identified, the output does seem to have program names in it. The name 'arya' under 'local address' is the host name in this instance. If you wanted to display IPs, you could use *-n* or *--numeric* instead:

```
> netstat -pn | head
(Not all processes could be identified, non-owned process info
will not be shown, you would have to be root to see it all.)
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name
tcp        0      0 192.168.3.107:40826
13.232.113.235:443      ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:51168     140.82.112.25:443
ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:48510     52.98.59.18:443
ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:46554     65.1.252.10:443
ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:35838     13.107.6.171:443
ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:60932     52.114.40.56:443
ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:35130
13.126.138.201:443      ESTABLISHED 2049/chrome --type=
tcp        0      0 192.168.3.107:45854
52.114.159.213:443      ESTABLISHED 2049/chrome --type=
```

In this case, it can be confirmed that the local IP address is 192.168.3.107.

```
> ip a show eno1
2: eno1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc
fq_codel state UP group default qlen 1000
    link/ether 94:c6:91:a6:33:df brd ff:ff:ff:ff:ff:ff
    altname enp0s31f6
    inet 192.168.3.107/24 brd 192.168.3.255 scope global
dynamic noprefixroute eno1
       valid_lft 69032sec preferred_lft 58232sec
    inet6 fe80::96c6:91ff:fea6:33df/64 scope link
noprefixroute
       valid_lft forever preferred_lft forever
    inet6 fe80::4ed1:9a96:4747:eb59/64 scope link
noprefixroute
       valid_lft forever preferred_lft forever
```

One of my favourite use cases for using netstat is to see which of the ports from my machine are being exposed locally to my LAN or the Internet. You can use *-l* to just list the listening sockets and couple it with *-t* to filter the TCP protocol.

```
> netstat -nlt --program
(Not all processes could be identified, non-owned process info
 will not be shown, you would have to be root to see it all.)
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name
tcp        0      0 0.0.0.0:22              0.0.0.0:*
LISTEN      -
tcp6       0      0 :::22                   :::*
LISTEN      -
```

As in this case we are unable to see the program name, we

can elevate our permission using *sudo* to get the result.

```
> sudo netstat -nlt --program
[sudo] password for sibi:
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name
tcp        0      0 0.0.0.0:22              0.0.0.0:*
LISTEN      1110/sshd: /nix/sto
tcp6       0      0 :::22                   :::*
LISTEN      1110/sshd: /nix/sto
```

You can see that it's the SSH daemon that has opened the port on 22. You can even remove the *-n* option so that it prints the name of the port.

```
> sudo netstat -lt --program
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name
tcp        0      0 0.0.0.0:ssh             0.0.0.0:*
LISTEN      1110/sshd: /nix/sto
tcp6       0      0 [::]:ssh                [::]:*
LISTEN      1110/sshd: /nix/sto
```

The 0.0.0.0 at the local address indicates that it is listening on all the network interfaces (your network card, 127.0.0.1, etc). If this was 127.0.0.1, it would mean that connections are from the machine itself, and not from the network or Internet, are being listened to. If the local network IP (192.168.3.107 in this case) is displayed, it means that connections will be allowed from the local network as well.

In fact, you can make a program listen to your local network and use netstat to confirm it:

```
> python3 -m http.server --bind 192.168.3.107
Serving HTTP on 192.168.3.107 port 8000
(http://192.168.3.107:8000/) …
Verification can be done on another terminal session:

> sudo netstat -lnt --program
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State       PID/Program name
tcp        0      0 0.0.0.0:22              0.0.0.0:*
LISTEN      1110/sshd: /nix/sto
tcp        0      0 192.168.3.107:8000      0.0.0.0:*
LISTEN      21410/python3
tcp6       0      0 :::22                   :::*
LISTEN      1110/sshd: /nix/sto
```

Looking further

While netstat is obsolete, learning it is still valuable because of the ubiquitous nature of this tool. Having a good grasp of how to read its output will be quite handy when networking issues have to be debugged. With various functionalities, netstat has a wide scope and can be used to view even the routing table (which is usually done by IP r):

```
> netstat -rn
Kernel IP routing table
Destination   Gateway   Genmask  Flags  MSS Window irtt Iface
0.0.0.0   192.168.3.1    0.0.0.0    UG    0 0   0 eno1
0.0.0.0   192.168.0.1    0.0.0.0    UG    0 0   0 wlp0s20f3
0.0.0.0   192.168.0.1    0.0.0.0    UG    0 0   0 wlp0s20f3
172.17.0.0  0.0.0.0     255.255.0.0   U   0 0   0 docker0
192.168.0.0  0.0.0.0    255.255.255.0  U  0 0   0 wlp0s20f3
192.168.0.0  0.0.0.0    255.255.255.0  U  0 0   0 wlp0s20f3
192.168.3.0  0.0.0.0    255.255.255.0  U  0 0   0 eno1
```

The manual page is given under 'References' to understand the other functionalities that netstat provides. If you are comfortable reading its output, you will be better equipped to debug your networking problems. END

Reference

Manual page: *https://linux.die.net/man/8/netstat*

By: Sibi Prabakaran

The author is a senior engineer working at FPComplete.

THE COMPLETE MAGAZINE ON OPEN SOURCE
OpenSource ForYou
The latest from the Open Source world is here.
OpenSourceForU.com
Join the community at facebook.com/opensourceforu
Follow us on Twitter @OpenSourceForU

Terraform:
The Cloud-Agnostic Solution for Infrastructure as Code



Image Source: freepik.com

Infrastructure as Code (IaC) makes the life of a cloud engineer easier in many ways. But it is still complex to manage the cloud and its operations in large enterprises. When we think of a cloud-agnostic solution for IaC in a multi-cloud environment, we can think of HashiCorp's Terraform.

Infrastructure as Code (IaC) automates, templatizes and reuses infrastructure configuration and setup through code in a cloud platform. It is useful for setting up different environments (dev, test, UAT, PROD) quickly in a cloud environment and when you want to replicate a setup for disaster recovery or multi-region setup of application instances. With its template based scripting for infrastructure provisioning, service configuration and logic development, IaC makes the life of a cloud engineer easier.

It is easy to develop and maintain IaC on the cloud in the case of greenfield software development. But when you already have cloud architecture ready and deployed, IaC can be developed manually. However, this could be tedious if you are using more cloud resources or many architectures for the various applications developed.

Creating IaC from infrastructure services and cloud resources can be done using a reverse engineering tool like TerraCognita, which was developed as an open source facility by Cycloid. TerraCognita can be used to develop IaC code from live cloud accounts in AWS, Azure or Google and from private cloud setups in OpenStack and VMWare. Cycloid has multiple open source projects including TerraCognita, InfraMap (for cloud diagrams) and TerraCost (for predictive cloud cost estimation). TerraCognita gets IaC scripts as Terraform templates from infrastructure services and uses InfraImport to get this infrastructure service converted as code.

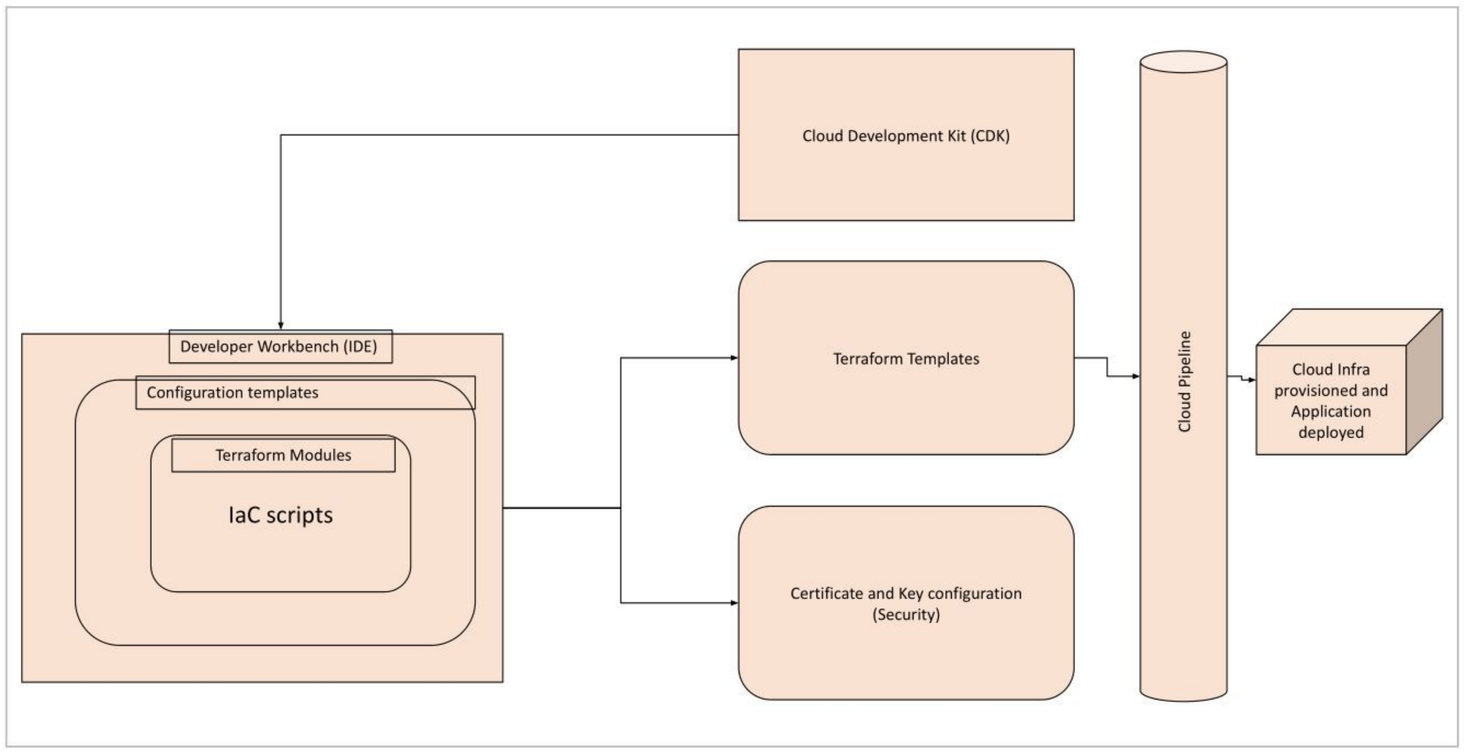


Figure 1: Components for developing IaC using Terraform templates

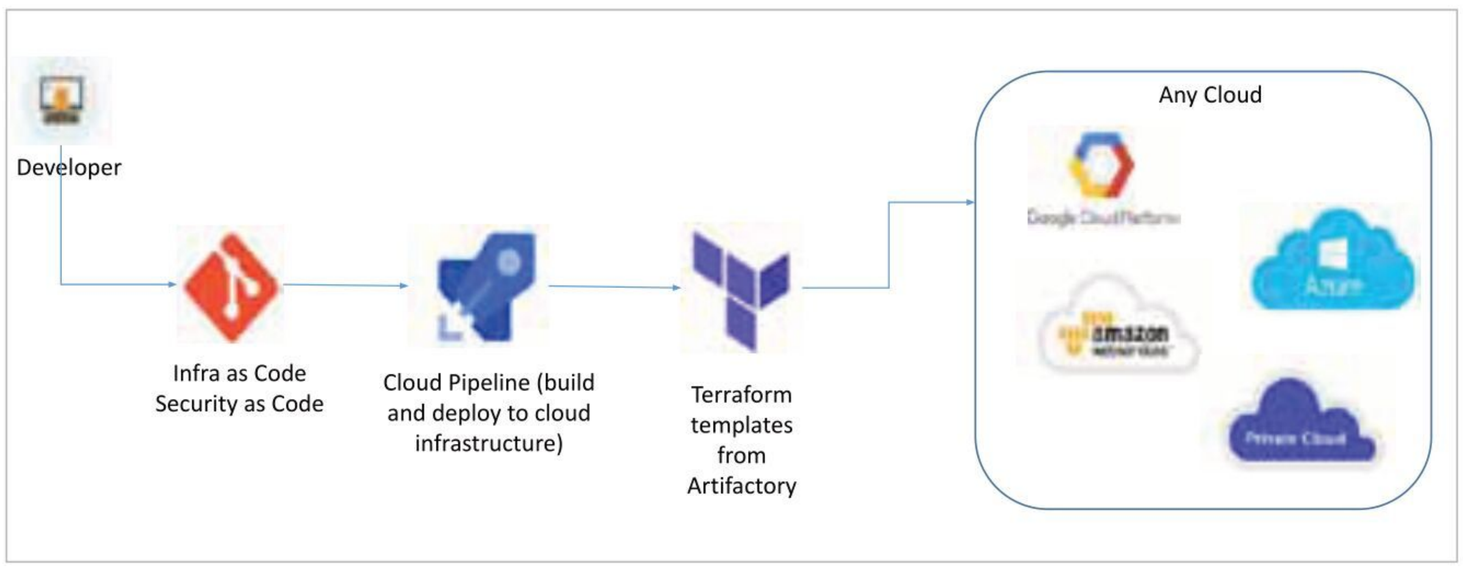


Figure 2: Implementation flow using Terraform templates

Why we need Infrastructure as Code

During cloud adoption, we have repeated activity for infrastructure provisioning, which could be simple or complex when it comes to multiple environment provisioning. Hence, automating and reducing the manual effort in such activities (including compute, storage, network and security configurations) helps to reduce overall DevOps operational activities and improve efficiency in engineering activities (Figure 1).

This kind of IaC can be available as a native facility in Azure as Azure Resource Manager (ARM) templates, in AWS as CloudFormation templates and in GCP as Google Deployment Manager (GDM). All these are in JSON format to prepare scripts for infrastructure provisioning and configuration.

For instance, an AWS feature like CloudFormation accomplishes this by defining provisioning code in a script using Typescript or Python for resource development and deployment. With a cloud development kit (CDK), this CloudFormation is taken to the next level of IaC. By using CDK, we can develop AWS constructs embedding CloudFormation templates, and configure services and business logic in native code using Java, Python, JavaScript or Typescript.

When we have a large number of microservices in terms of containers and serverless applications, it would be a Herculean task to handle code build, deployment, infrastructure provisioning, and platform management of cloud services across different stages, monitoring and management activities, and platform updates.

Platform teams automate many of these activities through Infrastructure as Code (IaC), but it is still a complex activity to manage the cloud and its operations in large enterprises. In the case of Microsoft Azure, using a templatized solution with Azure Resource Manager (ARM) helps to replicate configuration and application infrastructure provisioning. Here, pre-defined code scripts are used for building a descriptive model that defines VMs, networks, security and connectors easily.

Azure DevOps can be integrated with the ARM group for secure deployment of applications to different stages by selective users, and for tag deployment and code promotion with different parallel releases. This is so that we can promote (when regression checks are successful) or demote (when there is a reason to rollback to the previous release) the code build, as needed, quickly. It can also be used to store reusable artifacts like library or container images, so that it can be reused for deployment hooks without being built again.

Foundational services of Terraform

When we think of an agnostic solution for IaC in a multi-cloud environment, we can think of HashiCorp's Terraform

template. This can be a JSON (to export/reuse from native scripts) or YAML script based simplified solution that can be integrated with any cloud service provider's native deployment pipeline or third-party pipeline like Ansible to enable infrastructure provisioning activities on-the-go with flexibility.

It is easy to do large scale infrastructure creation, spin-up development, test instances during the implementation cycle or create pre-production or production instances during horizontal scaling, with Terraform scripts. These can be run with resource schedulers, and can create network configuration or secured environment setups for software defined networks (SDN) using Terraform templates.

Terraform scripts can be used to spin-up instances or to destroy environments for disposable temporary test environments in a cloud platform. Terraform templates can be validated before actual provisioning. Terraform workflow has three steps — write (create Terraform templates for IaC), plan (preview the changes before applying them in the cloud environment to validate) and apply (to provision infrastructure quickly).

For a cloud-agnostic solution like multi-cloud architecture, it is easy to adapt Terraform templates for IaC to quickly create a uniform setup and configuration for multiple environments, reduce complexity in creating multiple parallel scripts, and reuse these scripts as much as possible for each cloud service provider in the multi-cloud setup.

Some claims on the Internet, though, state that Terraform is not cloud-agnostic but a wrapper running on top of the native cloud services of any cloud service provider.

Cloud-agnostic provisioning using Terraform

Infrastructure as Code (IaC) helps to look at infrastructure as a software module and programmatically handle it through scripts. This can be quite useful for cloud platforms to quickly provision virtual infrastructure components like compute, storage, network, security and databases.

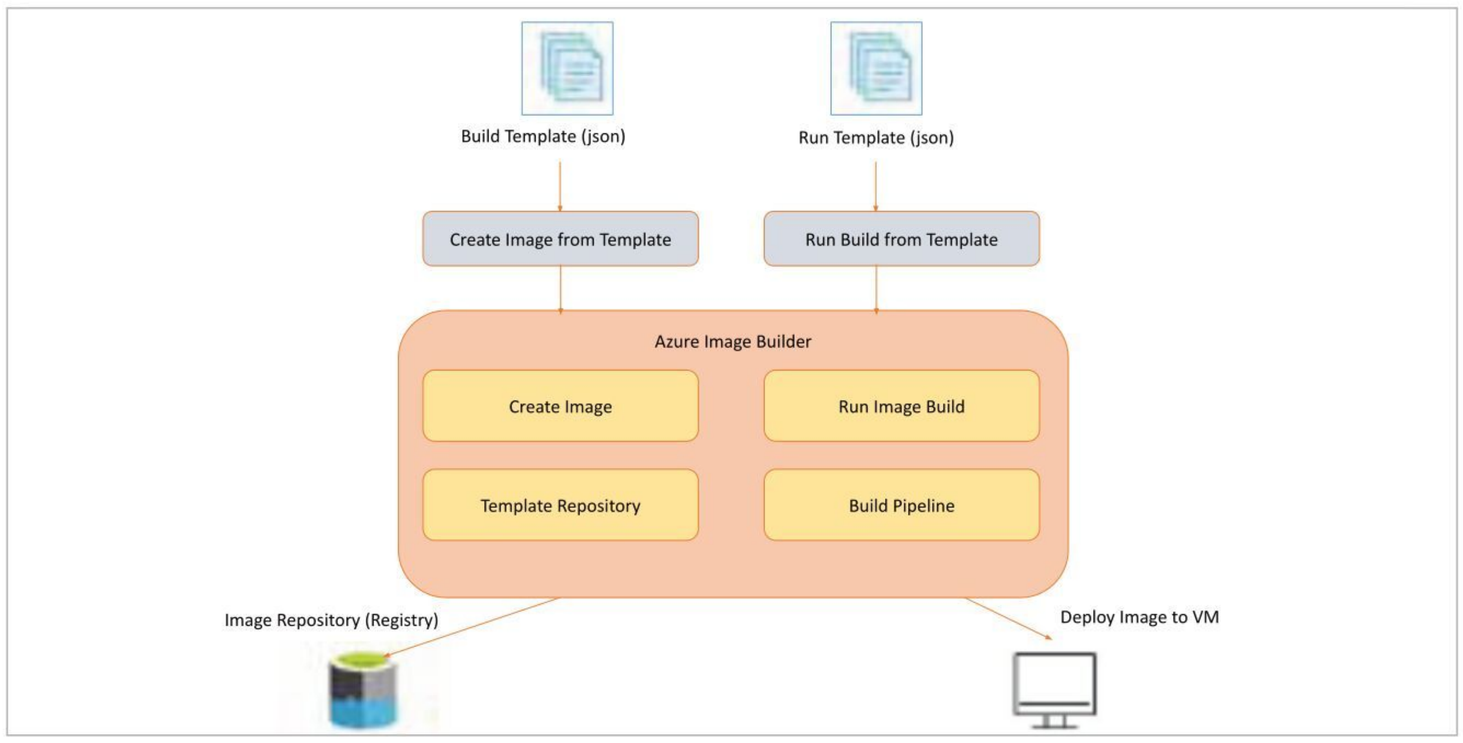


Figure 3: Azure Image Builder

Infrastructure as Code (IaC) templates can be classified as cloud native templates (like AWS CloudFormation, Azure Resource Manager and Google Deployment Manager templates), and cloud-agnostic templates (like Terraform templates).

Cloud native templates are flexible and customisable to the native platform, and can be integrated with any native service easily. Cloud-agnostic templates are suitable for multi-cloud application services in order to have uniform infrastructure scripts across cloud service providers to reduce rework in IaC script activities. For example, if you want to deploy a Web application and API services with disaster recovery services (active-passive deployment) in the AWS platform, it would be better to choose cloud-agnostic templates.

Google Cloud Platform (GCP) has a native template service called Cloud Foundation Toolkit (CFT), which contains a set of reference templates that can be used as is or customised to your IaC needs. The reference templates available in CFT are deployment manager templates (GCP native) or Terraform templates (cloud-agnostic). They are simple Python scripts that can be customised easily.

Other popular cloud native IaC facilities

Infrastructure as Code as well as cloud automation for provisioning and virtualisation requirements are very helpful to accelerate build, deployment, templatizing and cloud management activities. They are even more helpful when it comes to multi-cloud management, as preparing a provisioning script for each cloud is inconsistent and inefficient. Infrastructure as Code helps to automate this across different cloud environments with ease.

Though cloud native facilities like CloudFormation in AWS, ARM templates in Azure and CDM templates in GCP help to achieve this Infrastructure as Code, they are mostly native to the cloud platform and don't have the agility and flexibility for multi-cloud management. On the other hand, Google Anthos, HashiCorp Terraform, and Red Hat Ansible templates can help in multi-cloud management, enabling the development of zero-touch deployment scripts and provisioning templates for cloud tower setup.

There are various VM image building tools like Packer and Terraform, but Azure Image Builder

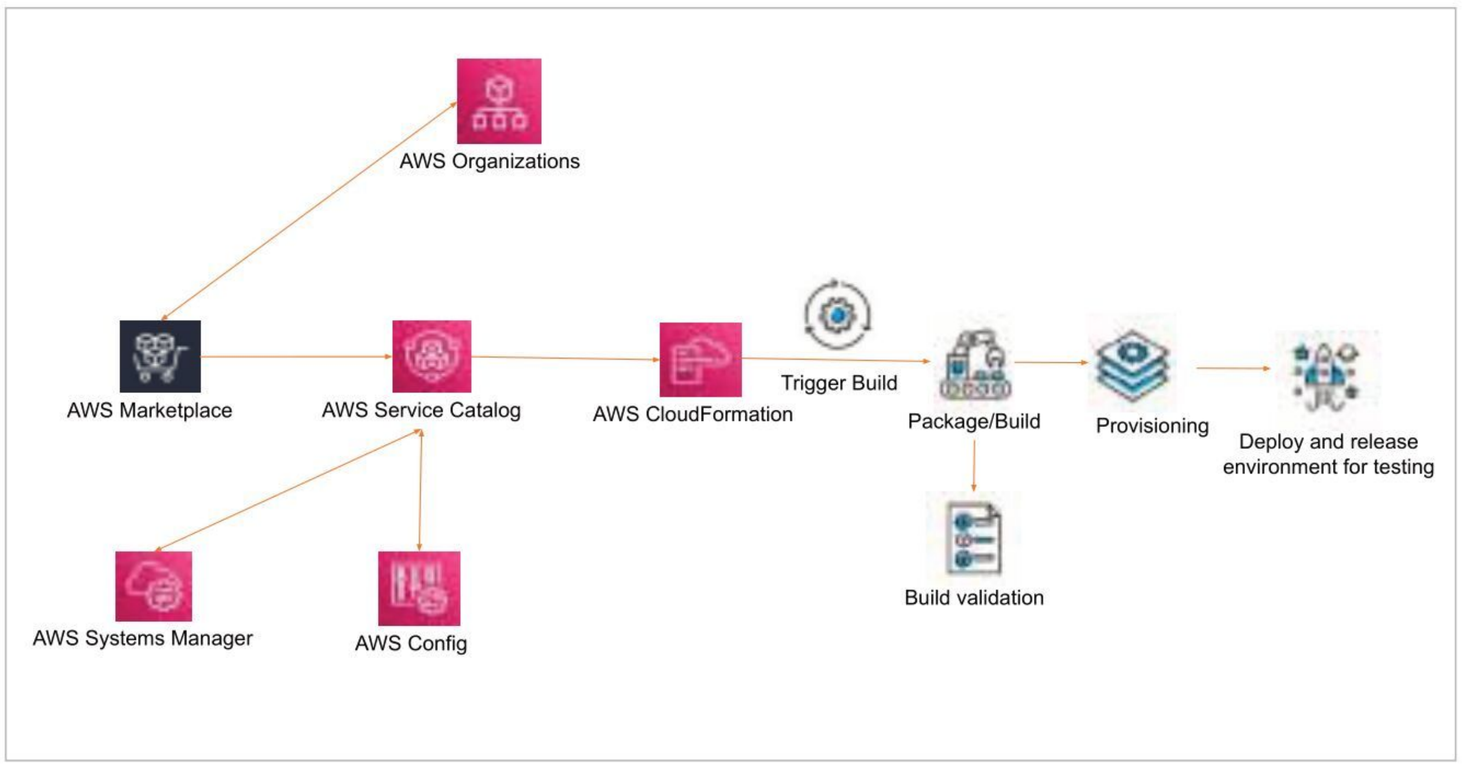


Figure 4: Infra as Code using AWS CloudFormation, Config and Systems Manager

is a comprehensive tool that can be used in combination with Azure Resource Manager (ARM) templates or with Terraform templates and create VHD (virtual hard disk) files of a virtual machine (VM). This can be imported to any cloud platform like AWS EC2 using the VM import tool in AWS or the GCP VM import tool in Google Cloud.

Azure Image Builder has two major functions — to create images from the template builder, and to run the build image with an integrated CI pipeline to deploy it to the target VM in Azure. The image build from this Image Builder service can be a Windows or Linux image, and hence suitable for importing by other cloud service providers as well as for using the VM import tool.

Integrating a CI pipeline with Azure Image Builder is very simple. Integrate it with Azure DevOps or any other customised DevOps pipeline with Ansible, Bamboo or Jenkins. Azure Image Builder also provides role based access control (RBAC) to access the image repository and template repository in a secure way.

AWS CDK has many pre-configured templates available, which can be reused to build a provisioning script. It usually works hand-in-hand with the AWS CloudFormation service. One has to develop an IaC script using AWS CDK and provisions using CloudFormation.

If you are familiar with Terraform, you can use the aws-tf construct library, and if you are familiar with Kubernetes and want to develop a container deployment script, you can use the aws-cdk8s construct library and deploy it in the AWS Elastic Kubernetes Engine (EKS).

Nowadays, Low-code/No-code and Infrastructure as Code (IaC) are getting popular as they are less complex and more efficient for operations. The Systems Manager Service of AWS can be used to group cloud resources, get operational insights and automate operational tasks. It helps to reduce the time and effort used in troubleshooting operational problems. This service can handle orphaned cloud resources and group cloud resources, get monitoring and management activities onto a single UI dashboard quickly, and simplify the operation and management of cloud infrastructure at scale. When you are managing large applications and infrastructure estate in a cloud platform, it is tedious to manage the server provisioning and deployment across different stages. Infrastructure as Code (IaC) simplifies this with template based scripts.

Google Anthos, HashiCorp Terraform and Red Hat CloudForms are leading this space equally for creating, changing and provisioning infrastructure for any cloud platform including service management, security, and compliance and optimisation. When you work with more applications and services during cloud adoption, you may end up with repeated work in infrastructure provisioning, platform setup, and application development and deployment. If you can automate many of these jobs, it will ease the software development journey particularly when you need to repeat activities at different stages like development, testing, pre-production and production. END

References

[1] *https://www.terraform.io/intro/use-cases.html*
[2] *https://www.cycloid.io/open-source/terracognita*
[3] *https://www.linkedin.com/pulse/shorticle-555-reverse-engineering-cloud-using-cycloid-kasthuri*
[4] *https://www.terraform.io/docs/language/modules/syntax.html*

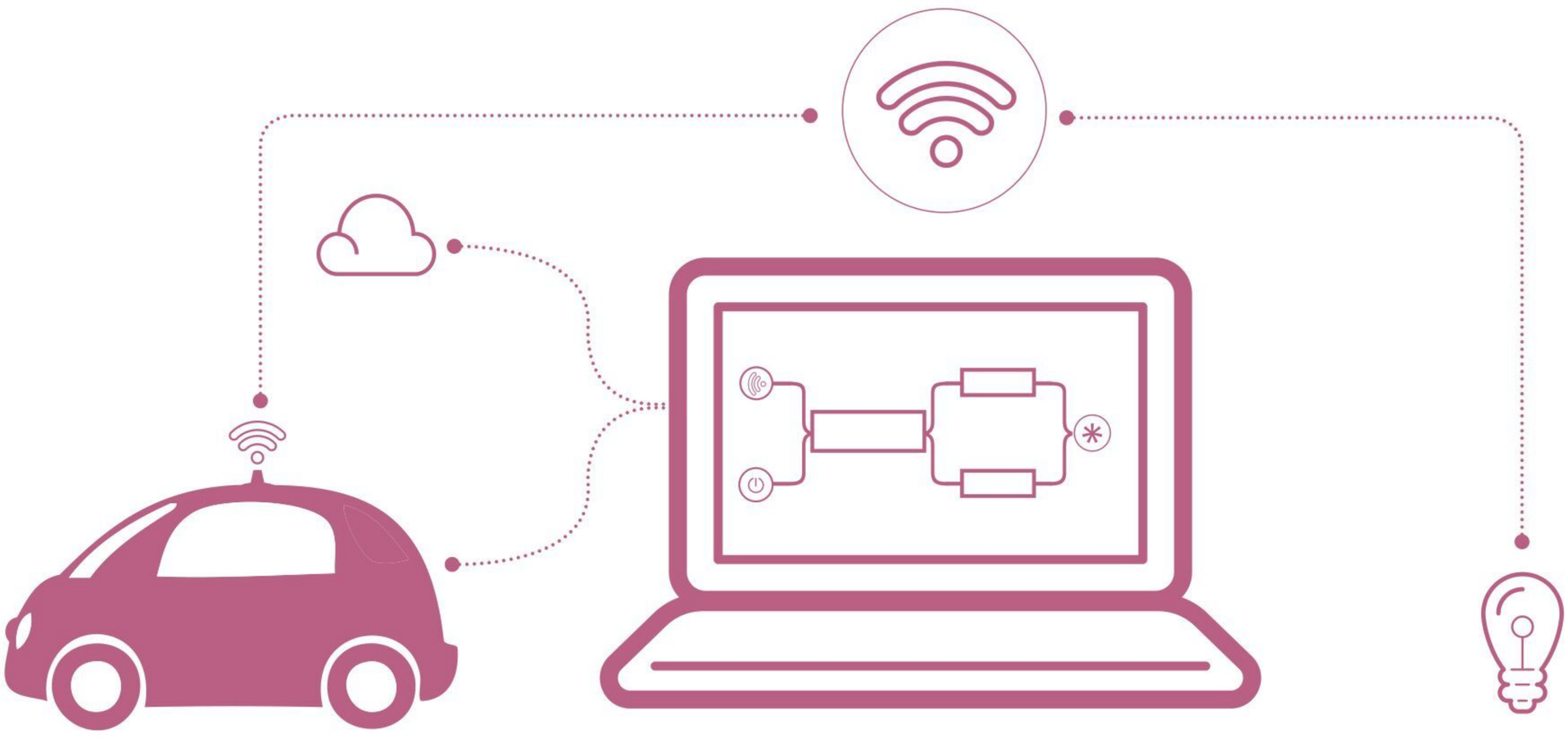
By: Dr Anand Nayyar and Dr Magesh Kasthuri

Dr Anand Nayyar is a PhD in wireless sensor networks and swarms intelligence. He works at Duy Tan University, Vietnam, and loves to explore open source technologies, IoT, cloud computing, deep learning and cyber security.

Dr Magesh Kasthuri is a senior distinguished member of the technical staff and principal consultant at Wipro Ltd. This article expresses his views and not that of Wipro.

Enhancing the Design of Autonomous Cars Using Node-RED

Autonomous cars are the future. A lot of work goes into designing them to minimise the risk of machine failure and accidents. This article outlines how Node-RED can be used to create flows for developing animations of car movements.



Autonomous cars are set to revolutionise the travel experience for all of us. As per a Gartner report, there were 508,622 autonomous cars on the road in 2021. This figure is expected to go up to 745,705 in 2023. Autonomous cars reduce the possibility of accidents caused by driver distraction, reduce traffic congestion, and optimise the use of fuel. Cameras, radars, and ultrasonic and touch sensors used in these cars replace human decision-making. This article outlines how to enhance the capability of these cars by creating Node-RED flows that develop animations for car movements. The flows are implemented and the movements are observed with various input and output sequences, using the animations developed.

Node-RED

Node-RED is open source software used for the Internet of Things (IoT) and for making Web applications. Unlike normal software that use complex programs and structures, Node-RED uses GUI operations to create 'flows' that are easy to understand. This is done by integrating various 'nodes' without using any complex code. Similar to the successful serverless architecture, the zero coding techniques of Node-RED are gaining popularity for IoT operations. Node-RED allows us to develop real-time applications in IoT, online services and projects with minimal effort. This article will help you create a pleasant graphical representation of blocks named 'nodes' in Node-RED, which provide programmers a fantastic way to develop applications. It can also function as a wiring tool that can connect a large number of both hardware and software devices.

Node-RED development environments can also connect hardware devices to Web services. Users can share their flows with each other, making group based activities convenient. Node-RED can be used

both for simulation and real-time applications. In this article, the simulation mode is explored for the animation of autonomous cars.

Installation of Node-RED and associated components

The latest version of Node.js can be installed from the URL *https://nodejs.org/en/download/*. The checkup wizard for the same can be completed with full administration permissions sanctioned.

To check if Node.js has been installed successfully, without any issues, the following command is used. This also checks the version installed.

```
node-v
```

After its setup, the node package manager (NPM) is installed by typing the command given below:

```
npm install -g node-red
```

To check the installation status or version of Node-RED, give the following command. This command shows the verbose output while the server is being started, as shown in Figure 1.

```
node-red -v
```

After the installation, to view the default Node-RED screen, the website must be opened with the URL *http://127.0.0.1:1880* using the browser. This URL address notifies that we have the Node-RED server running on the local machine on port 1880.

To install the SVG graphics node, the following command is to be given:

```
npm install node-red-contrib-ui-svg
```

Architecture for animation in autonomous cars

A new flow is created first, and an SVG graphics node is added. Later, a car is drawn in the SVG editor using elements like lines and circles. Next, to add animation, there should be specific parameters like direction of motion, time delay between movements, number of times the movement has to be repeated, coordinates of movement and the duration of movement. After creating the animated car, configure the 'button', 'audio out' and the 'inject' nodes. A connection is made between the SVG graphics, inject button and audio out nodes in the Node-RED flow. A stationary car will be visible in the Node-RED dashboard (output window) and animation is triggered once the button is pressed. The entire workflow is depicted in Figure 2.

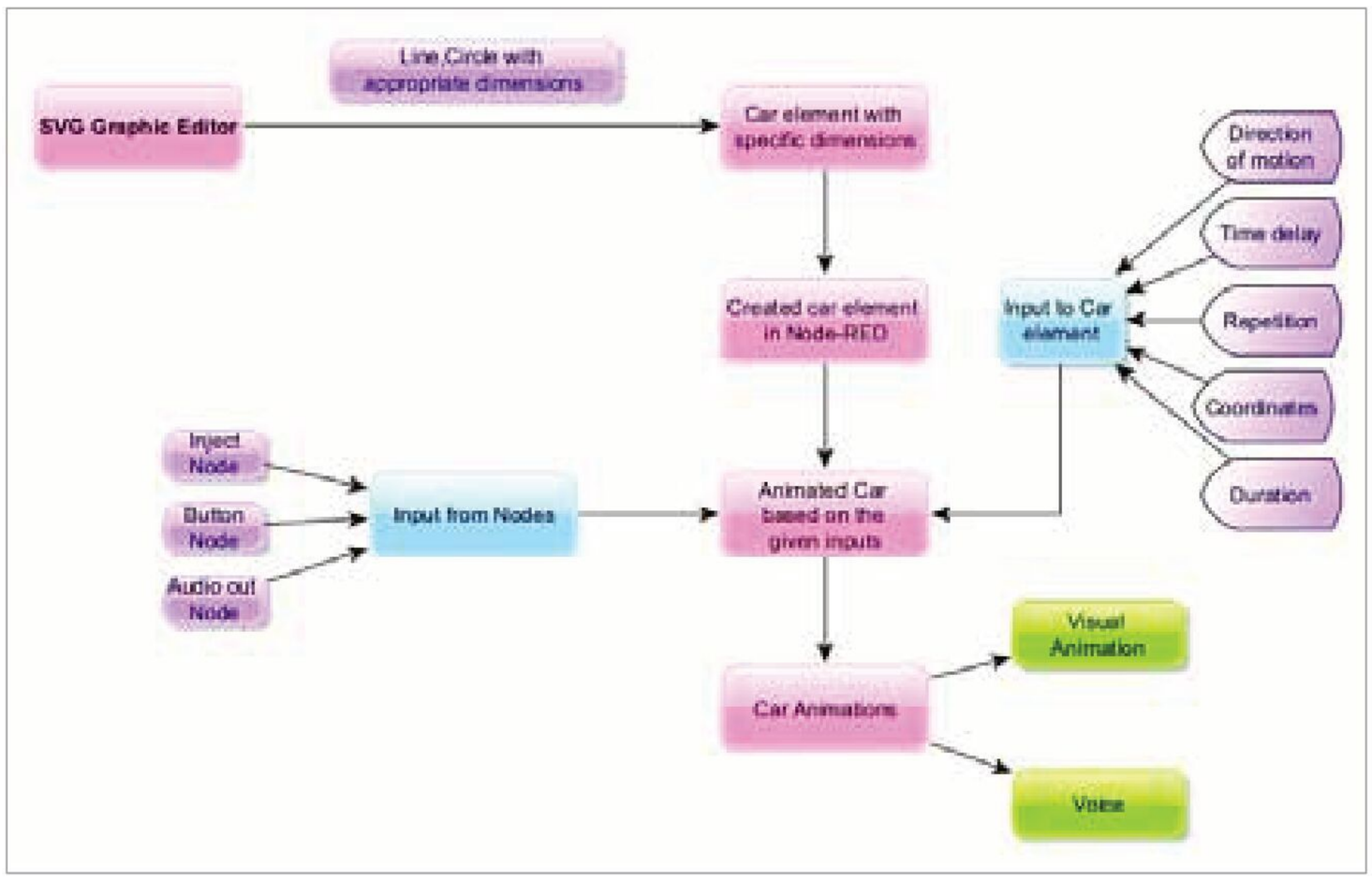


Figure 1: Command to start the Node-RED server

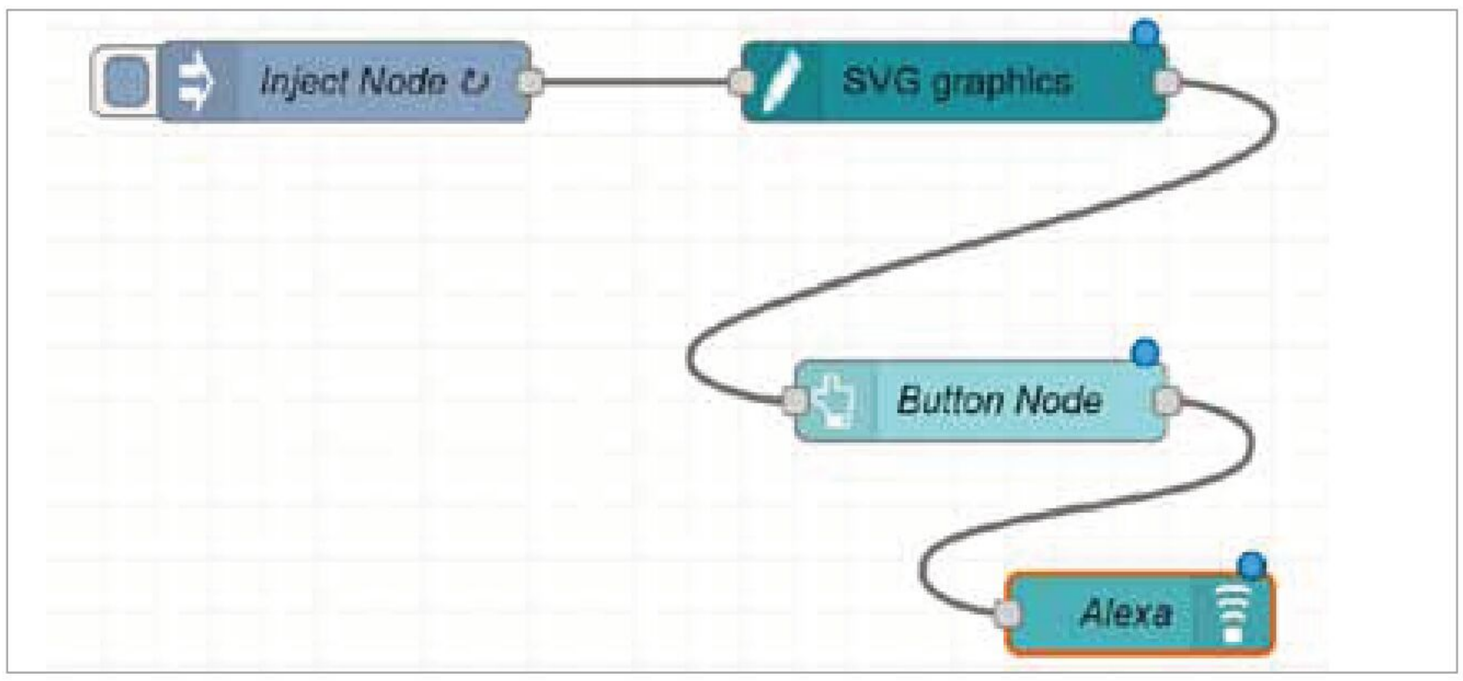


Figure 2: Architecture diagram for autonomous car animation



Figure 3: Node-RED flow for autonomous car design

Design for autonomous car animation

The steps for installing and opening the Node-RED interface are

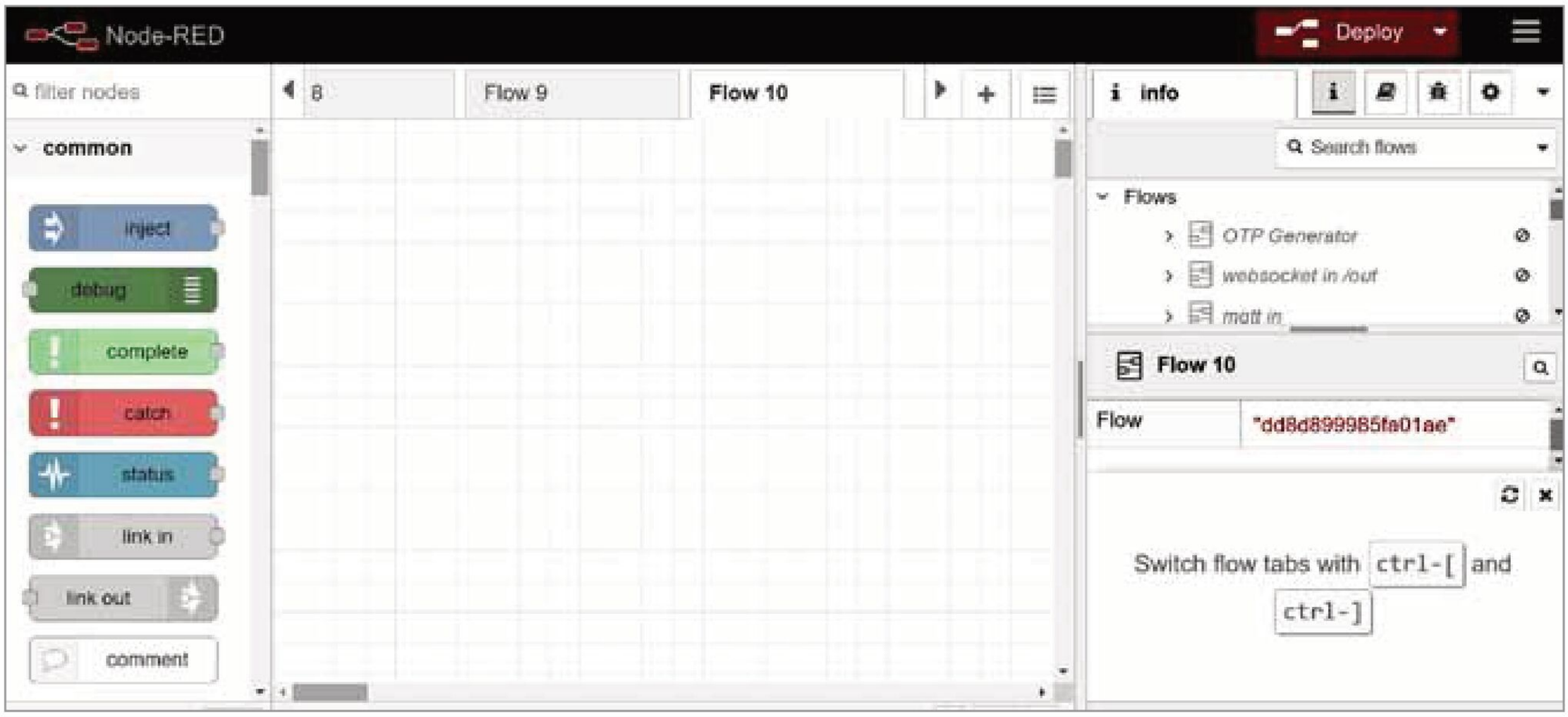


Figure 4: Newly created flow in the Node-RED interface

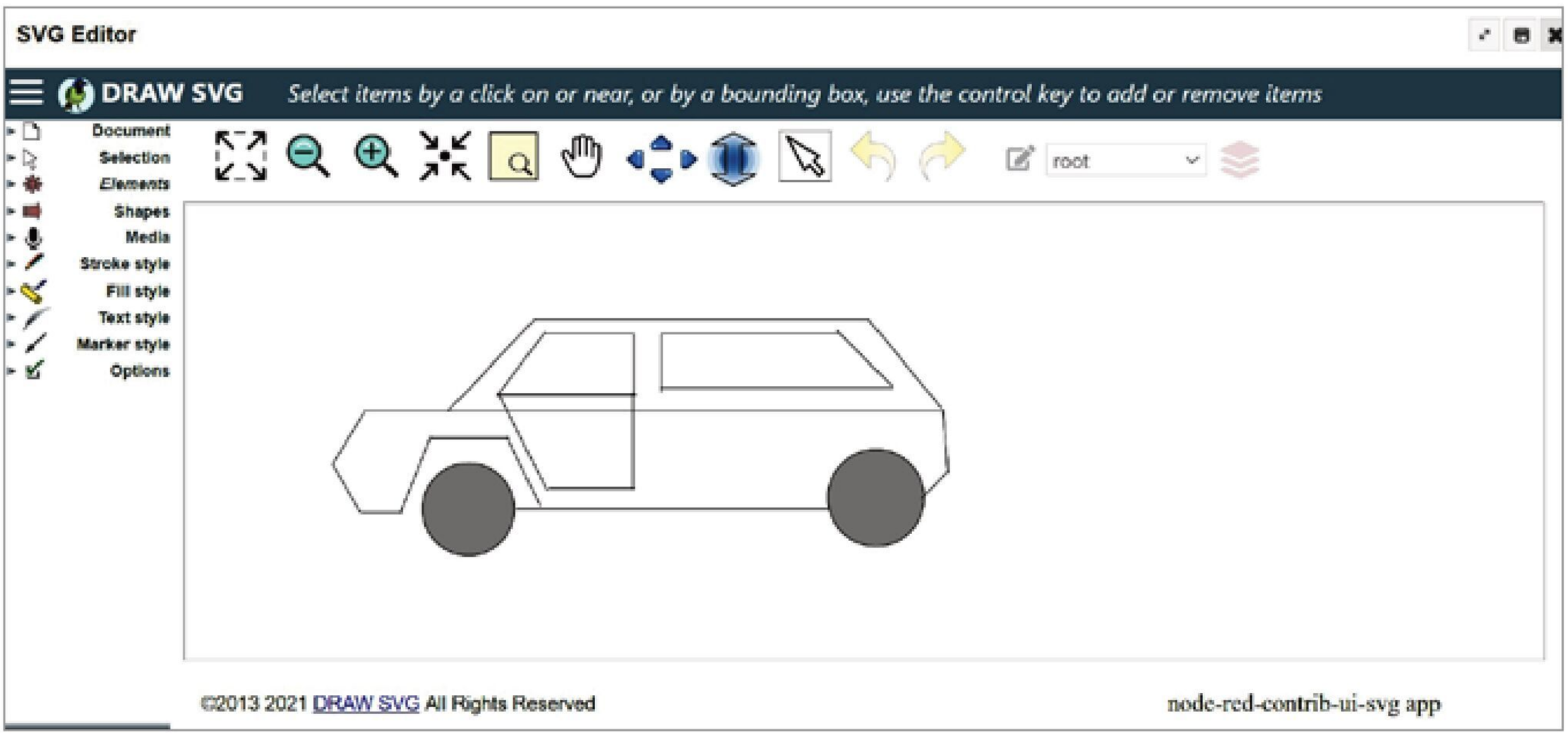


Figure 5: Car drawing in the SVG editor

explained below.

Open a new flow by clicking the '+' symbol on the top right of the interface. This interface generally consists of three columns, as shown in Figure 4. In the first column, all the nodes in the node palette are visible. The second column shows the flows under operation, and the third column consists of tabs for information, debug messages, help and node configuration. The flows, sub-flows and global configuration nodes are visible under the 'information' tab, and the 'help' tab is useful for all the Node-RED libraries. The messages that are given to produce a text output are visible under the 'debug messages' tab. All the uses of nodes are visible under the 'node configuration' tab.

Look for the SVG graphics node in the node palette and insert it into the created flow. When it is opened, the SVG editor appears. The various options required for a drawing, like document, selection, elements, shapes, media, stroke style, fill style, text style, marker style, etc, are visible on the left side of this editor's window.

Select the *Elements* option in the editor window. Shapes, lines, images and other options are visible here. After selecting the line, shape or element to be used, drag it from the *Elements* option on the left side. If you wish

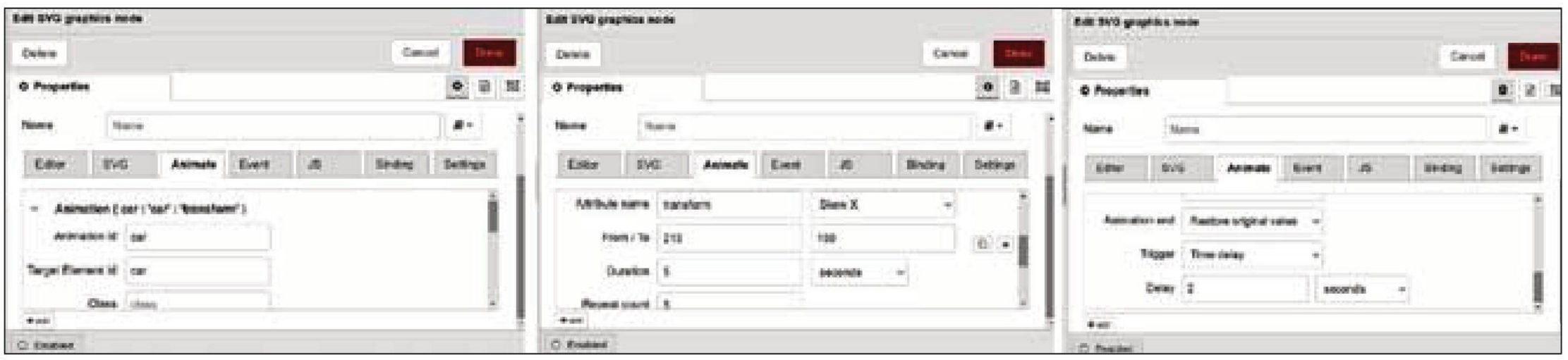


Figure 6: Completing the configuration of the 'animate' tab

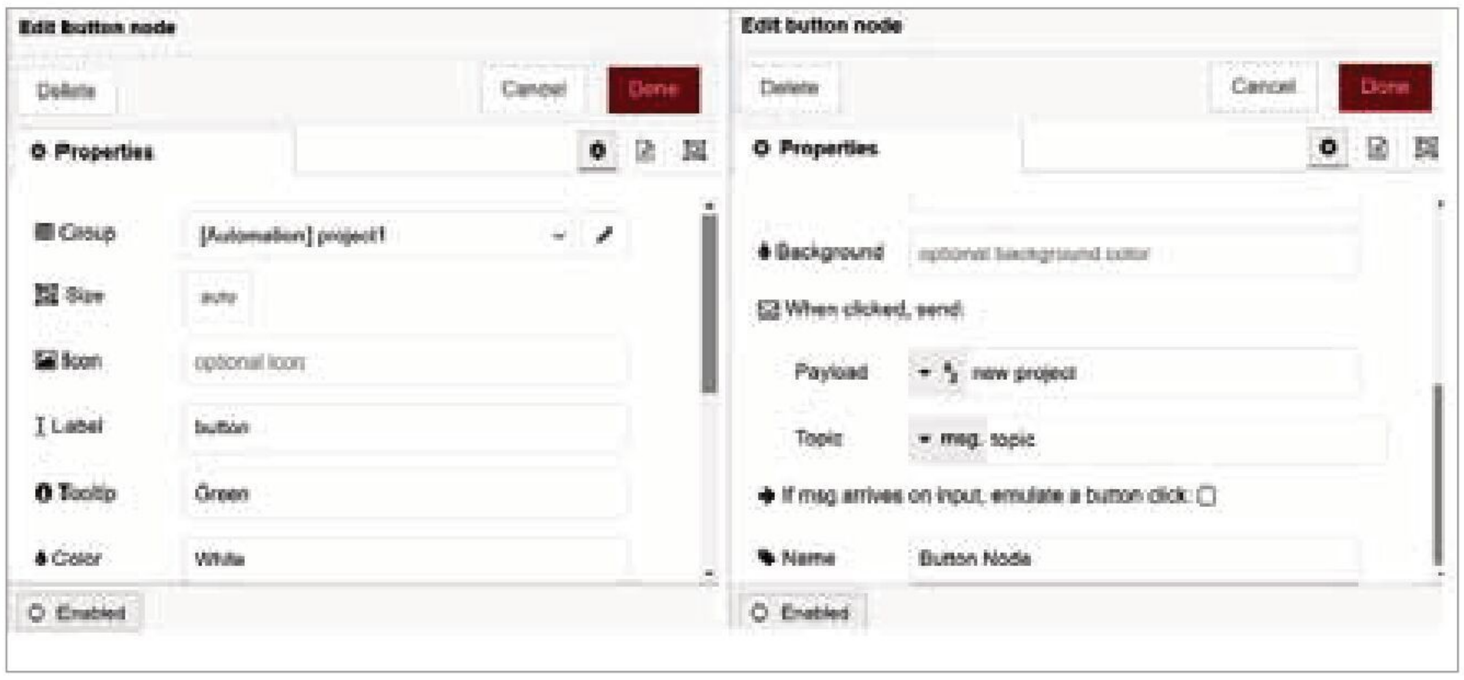


Figure 7: Filling the button node *Properties* dialog box

to add colours to your drawing, then double click the specific element and go to *Properties*. It will have a fill style option. In the 'fill colour' tab, we can change the colours used for filling the closed figure elements. In the 'stroke style' tab, the colour, width and other properties of the drawing pencil can be modified. The 'elements' tab takes care of properties like ID, default style, tag and path. Any shape can be created easily using the SVG editor. Any number of shapes and lines (all sizes) can be used to create the drawing we require. In this case, a car is created with 24 lines and 2 circles, with colours filled in the wheels, as shown in Figure 5.

After drawing the car, save and close the editor. Then select the 'events' tab in the SVG graphics node, and add a new element. The selector is given a name (car) and 'click' is the action for stimulation. The messages 'car has started' and '#car' are given in the *msg.payload* and *msg.topic* options, respectively.

Once the event tab of the SVG graphics node is configured, select the JS tab. The selector ID is given as 'car' and 'click' is given as the action for stimulation.

After configuring the event and JS tabs in the SVG editor, select the 'animate' tab and add a new element. Both the animation and target element IDs are specified as 'car', as shown in Figure 6. Change the attribute to 'transform' and the property to 'Skew

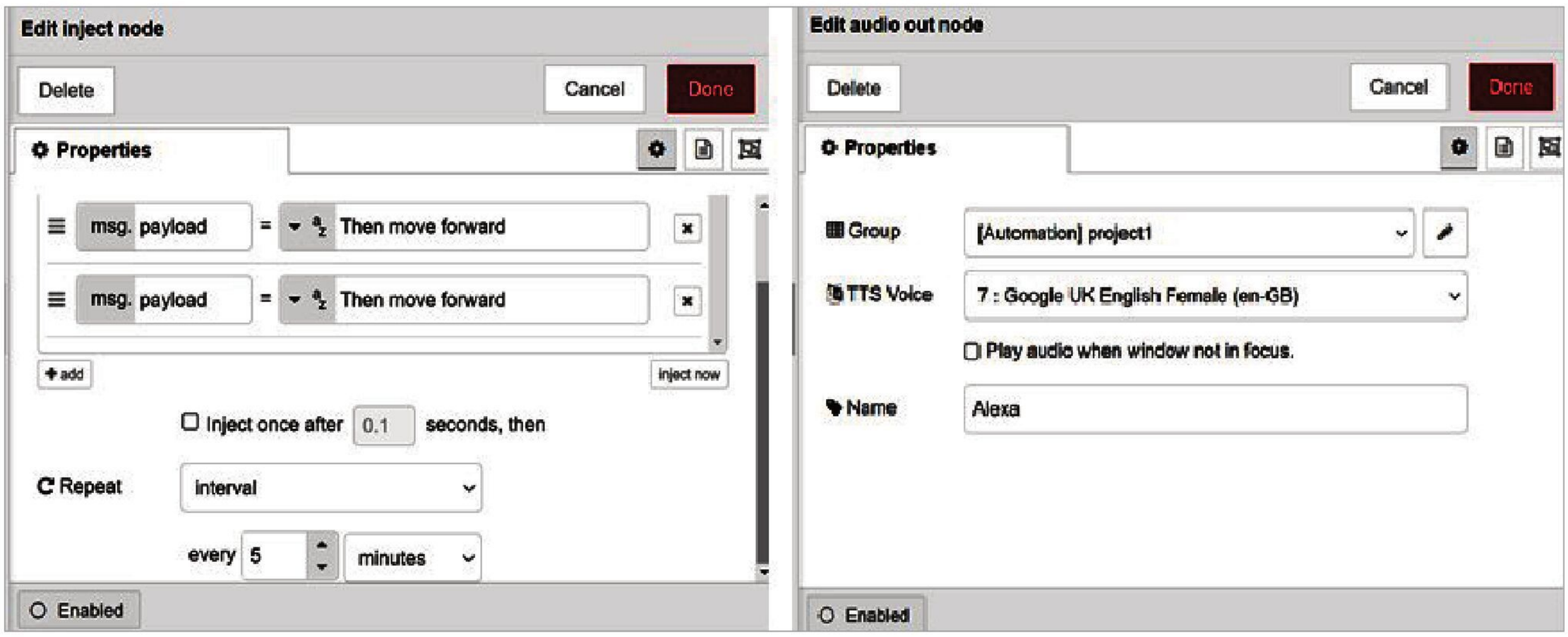


Figure 8: Filling the audio out node properties and inject node properties

X'. Give the coordinates as 215 and 175, with a duration of 5 minutes. Set the repeat count at 5, as shown in Figure 6. Select the option to restore the original value, so that at the end of the animation the car comes to its original position. A time delay can be created in between the movements of the car. The operation in the SVG graphics node is now complete.

Connect the button node from the exit side of the SVG graphics node. On double-clicking, the *Properties* dialog box will appear. Set the group as 'project1' and tab as 'automation', as shown in Figure 7. The colour of the button can be modified. Any message can be given to the *msg. payload* option, as shown in Figure 7.

After setting up the button node, connect the inject node to the SVG graphics node. The desired input for the movement of the car can be given, so that the car moves according to the given instructions. Double-click the inject node to change its properties. The payload messages that can be given are 'start the car', 'move backward' and 'move forward'. The animation can be made to repeat after a certain interval (5 minutes, as shown in Figure 8). Another choice is to trigger the start of the flow at a certain time every day. The audio out node is added to the flow and is wired to the button node. The group and tab are automatically set because of the connection we made between the nodes. An option is provided to choose the desired language and gender of the voice for the output, as shown in Figure 8.

After finishing the connections depicted in Figure 2, the Node-RED dashboard (output window) can be viewed by opening the URL *http://127.0.0.1:1880/ui* in a browser.

In the output window, the car is initially in a stationary position, as shown in Figure 9. Once the button is pressed, the car will move back and forth according to the values specified. This can be seen in Figure 10.

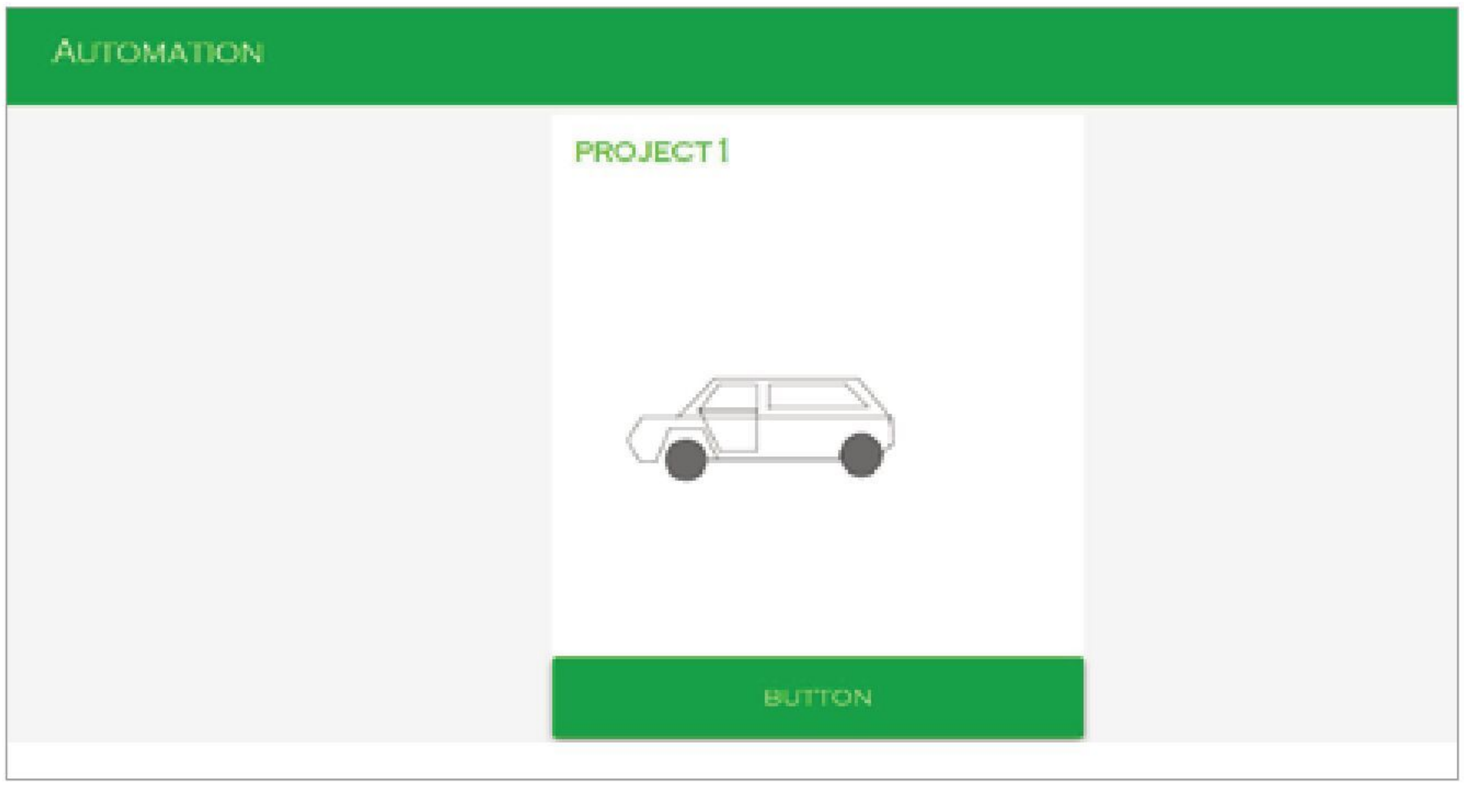


Figure 9: Node-RED dashboard showing the output

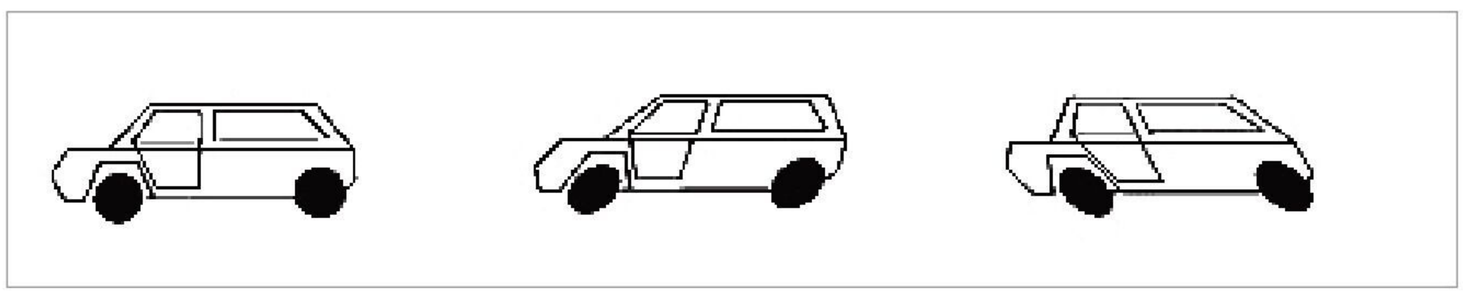


Figure 10.1: Stationary car

Figure 10.2: Forward movement of the car

Figure 10.3: Backward movement of the car

Autonomous cars can be of great help in critical locations like nuclear power plants, atomic centres, huge data centres, industrial applications with extremely high or low temperatures, and space centres. They are a great boon for the elderly and visually impaired. But they need to be designed with the utmost care to ensure there are no accidents. END

References

[1] Felix Richter, 'Vehicles Are Ready for Autonomy - But Are We?', Statista, 2019, *https://www.statista.com/chart/20091/annual-increase-in-number-of-vehicles-equipped-with-hardware-for-fully-autonomous-driving/*

[2] Varun Totakura, 'Improved Safety of Self-Driving Car using Voice Recognition through CNN', Research gate, 2021, *https://www.researchgate.net/publication/348590701_Improved_Safety_of_Self-Driving_Car_using_Voice_Recognition_through_CNN*

[3] Guest Writer, 'How IoT is Driving the Autonomous Vehicle Revolution', iotforall, 2018, *https://www.iotforall.com/iot-and-autonomous-vehicles*

[4] X.Krasniqi,'Use of IoT Technology to Drive the Automotive Industry from Connected to Full Autonomous Vehicles', Science direct, 2016, *https://www.sciencedirect.com/science/article/pii/S2405896316325162*

[5] Sabrina Sicari, 'Smart transport and logistics: A Node-RED implementation', Wiley Online library, 2018, *https://onlinelibrary.wiley.com/doi/full/10.1002/itl2.88*

By: Josiga S. and Dr Maheswari R.

Josiga S. is a computer science enthusiast and a constant learner.

Dr Maheswari R. is an associate professor in the Vellore Institute of Technology, Tamil Nadu. She has served as a resource person for hundreds of national and international workshops and conferences in the field of IoT.

EFY GROUP
YOURS SINCE 1969

Thank You IIT Madras

It was in your hostel room that the idea was conceived



The first issue in Jan ‘69



The Jan ‘22 Issue

Celebrating 54 years in TECH media

Magazines
- Electronics For You
- Electronics Bazaar
- Open Source For You

Portals
- www.electronicsforu.com
- www.opensourceforu.com
- www.eleb2b.com
- www.electronicsb2b.com

Events
- India Electronics Week
  - IOTshow.in
  - EFY Expo
  - Test & Measurement India
- Open Source India
- Techworldcongress.com

EFY Enterprises Pvt Ltd
D-87/1, Okhla Industrial Area, Phase-1, New Delhi-110020

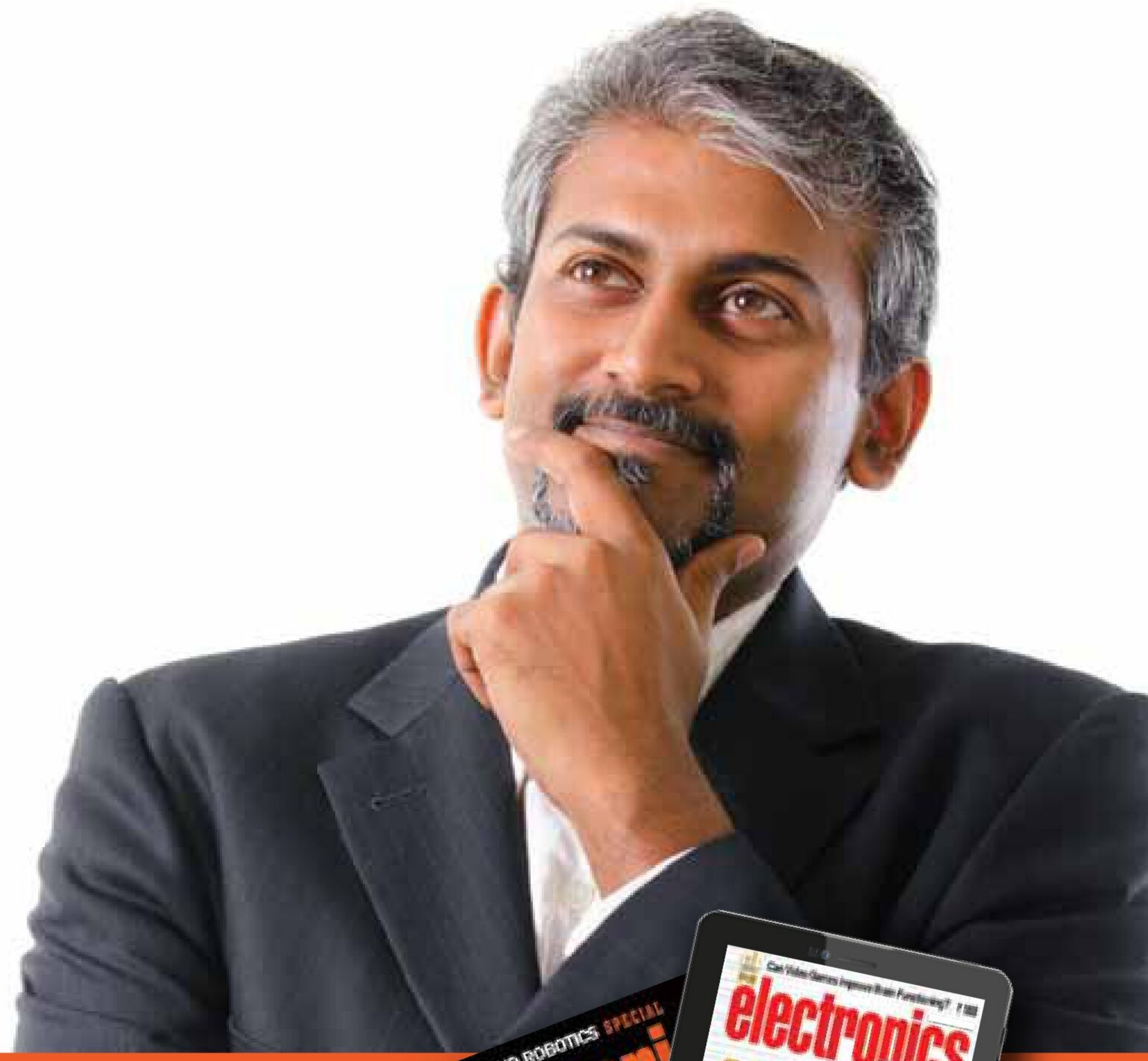
• Phone: +91-98111-55335
• Email: info@efy.in • Website: www.efy.in

R N I No. DELENG/2012/49440, Mailed on 27/28th of Advance month
Published on 27th January 2022
Delhi Postal Regd. No. DL(S)-01/3443/2022-24
Licenced to Post without Pre-Payment Licence No U(S)-56/2019-21
Posted at LPC Delhi RMS Delhi 110006

India's No.1 Electronics Magazine
RANKING AS PER ELCINA-IMRB REPORT

"I started reading it when I was a student... ...and I am still reading it, as a student"

—CEO, Design House



electronics FOR YOU
YOURS SINCE 1969

To Subscribe:
https://subscribe.efy.in

OR

Scan This Code

For any query, call: +91-98111-55335 or email: support@efy.in