

LLMs and other language tools in technical writing

The fundamentals of technical writing | MUNI 2024

The Red Hat Customer Content Services team



Dominika Borges

Technical writer at Red Hat
Former journalist



Kalyani Desai

Technical Writer at Red Hat, Open source
contributor,

Fun fact: I am obsessed with Frootloops(cereal)
I tried some strawberry shampoo.
It doesn't taste as good as it smells :P



Jan Fiala

Senior Technical Writer, at Red Hat for 4+ years.
Previously, a translator to and from English for
10 years.

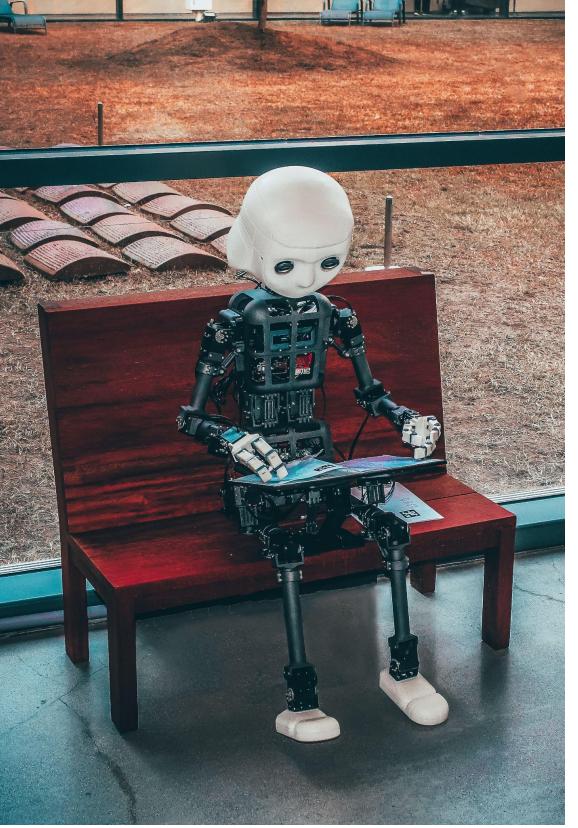
What we'll discuss today

- Writing aids
- How do LLMs work?
- How to use LLMs and other tools in technical writing
- Risks and shortcomings
- Final project



DRESDENCODAK.COM

SS



Which writing aids do you use?

Writing aids

Linters

- [Vale](#)

Other tools

- Spellcheck
- [Grammarly](#)
- [Hemingwayapp](#)

Large language models

- GPT-3.5, 4, 5?
- ~~Bard~~ Gemini
- LLaMA 2
- Alpaca
- ~~Bing AI~~ Copilot



How do large language models work?

- LLMs use the **transformer architecture** for sequence processing.
- Transformers use **self-attention** to weigh word importance efficiently.
- Models are **pre-trained** on vast text data (billions to trillions of tokens).
- Then, they are **fine-tuned** for specific tasks.
- They have billions of **parameters** learned during training.
- They iteratively **predict the next token** based on context.



How do you use AI tools?



LLM use cases in TW

- Brainstorming: “Give me 5 headings for this paragraph”
- Research, explanations: “How does SELinux work?”
- Description of code: “What does this yaml file do?”
- Code snippets or regex: “Create a script that...”
- Formatting: “Convert this csv into a markdown table”
- Ideas for visual elements (runwayml.com)

Neil Perry, Megha Srivastava, Deepak Kumar, Dan Boneh
 We conduct the first large-scale user study examining how users interact with an AI Code assistant to solve a variety of security related tasks across different programming languages. Overall, we find that participants who had access to an AI assistant based on OpenAI's codex-davinci-002 model wrote significantly less secure code than those without access. Additionally, participants with access to an AI assistant were more likely to believe they wrote secure code than those without access to the AI assistant. Furthermore, we find that participants less and engaged more with the language and its features. Finally, we find that participants with access to the AI assistant were more likely to believe they wrote secure code than those without access to the AI assistant. Furthermore, we find that participants less and engaged more with the language and its features.

Limitations of LLMs

- No real time information
- No fact checking
- Not for general tasks
- No informed decisions
- Legal limitations
- Security concerns



Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models

Priyan Vaithilingam
 pvaithilingam@g.harvard.edu
 Harvard University
 USA

Tianyi Zhang
 tianyi@purdue.edu
 Purdue University
 USA

Elena L. Glassman
 glassman@seas.harvard.edu
 Harvard University
 USA

ABSTRACT

Recent advances in Large Language Models (LLM) have made automatic code generation possible for real-world programming tasks in general-purpose programming languages such as Python. However, there are few human studies on the usability of these tools and how they fit the programming workflow. In this work, we conducted a within-subjects user study with 24 participants to understand how programmers use and perceive Copilot, a LLM-based code generation tool. We found that, while Copilot did not necessarily improve the task completion time or success rate, most participants preferred to use Copilot in daily programming tasks, since Copilot often provided a useful starting point and saved the effort

on two different kinds of approaches: (1) program synthesis algorithms that search over a large program space defined by a domain-specific language (DSL) [2, 7, 10, 12, 14, 19, 24, 25, 30, 31, 34, 40] and (2) deep learning models that are trained on a large amount of existing code and can generate new code given some form of specifications such as natural language descriptions or incomplete code [5, 16, 17, 22, 38, 39, 48, 49]. Both kinds of approaches have clear drawbacks. On the one hand, existing program synthesis techniques are constrained to pre-defined DSLs and cannot scale to general-purpose programming languages [15]. On the other hand, existing generative models have a hard time learning sophisticated programming patterns from code corpora and often generate co

Retrieving information about TW style?

Here is a comparison of the advice on using contractions across Google Developer Documentation Style Guide, Apple Style Guide, and the IBM Style guide:

Style Guide	Recommendation
Google Developer Documentation Style Guide	Use most types of contractions. Negation contractions, such as isn't, don't, and can't, can be used to make writing more informal and conversational. Noun + verb contractions, such as the browser's, should be avoided because they can be confusing for readers. Double contractions, such as mightn't've, should also be avoided. Its and it's: Don't confuse its (possessive) with it's (noun + verb).
Apple Style Guide	Do not use contractions in documentation. Contractions can make writing informal and unprofessional, and they can be confusing for readers.
IBM Style Guide	Use contractions sparingly. Contractions can make writing less formal and more difficult, and they are necessary to make writing more clear and readable.

Contractions

[Send feedback](#)

In general, we write our documentation in an informal [tone](#), so we recommend using most types of contractions.

Negation contractions

In particular, it's fine to use -n't contractions, such as *isn't*, *don't*, and *can't*.

One reason that such contractions are useful is that it's sometimes easy for a reader to miss the word *not*, whereas it's harder to misread *don't* as *do*. If you need to emphasize the negative, use text formatting such as "is *not*" which renders as "is *not*".

Which text is written by a human TW?

A	B	C
<p>Red Hat Enterprise Linux 9: GIMP 2.99.8 Technology Preview Available</p> <p>This release introduces GNU Image Manipulation Program (GIMP) version 2.99.8 as a Technology Preview in Red Hat Enterprise Linux 9. GIMP 2.99.8 delivers a comprehensive set of image editing and manipulation capabilities, enabling users to:</p> <ul style="list-style-type: none">• Perform basic image editing tasks: Crop, resize, adjust color and brightness, apply filters, and more.• Create and manipulate artwork: Draw vector graphics, paint with various brushes, and work with layers for complex compositions.• Automate repetitive tasks: Utilize scripting capabilities to automate workflows and save time. <p>Important Note: GIMP 2.99.8 is provided as a Technology Preview. This means it is still under development and may contain bugs or stability issues. It is not intended for production use.</p> <p>This Technology Preview allows users to explore the capabilities of GIMP 2.99.8 before its official introduction in Red Hat Enterprise Linux 9. Feedback from users will be crucial in shaping the final offering.</p> <p>Installation:</p> <p>GIMP 2.99.8 is available through the Technology Preview repository. You can install it using the following command:</p> <pre>sudo dnf install gimp-techpreview</pre> <p>Disclaimer:</p> <p>Red Hat does not provide support for Technology Preview features. Use them at your own risk and discretion.</p>	<p>Feature: GNU Image Manipulation Program (GIMP) 2.99.8</p> <p>Description: The GIMP package version 2.99.8 is now available as a Technology Preview in Red Hat Enterprise Linux 9. This pre-release version provides a glimpse into the upcoming GIMP 3 release. Users can explore and test this version using the RPM package for installation.</p> <p>Key Points:</p> <ul style="list-style-type: none">• Technology Preview Status:<ul style="list-style-type: none">◦ GIMP 2.99.8 is presented as a Technology Preview, showcasing features and enhancements that will be fully integrated into the official GIMP 3 release in future versions of RHEL 9.• Pre-Release Version:<ul style="list-style-type: none">◦ This package serves as a pre-release version, allowing users to experience and provide feedback on the latest developments in GIMP. <p>Recommended Action:</p> <ul style="list-style-type: none">• Users interested in previewing the upcoming features in GIMP 3 can install the GIMP 2.99.8 package.• As a Technology Preview, this version is not intended for production use but offers users an opportunity to familiarize themselves with the upcoming GIMP release. <p>Note: GIMP 2.99.8 is a forward-looking Technology Preview, giving users early access to the evolving capabilities of GIMP 3. Feedback from the community during this preview phase is valuable in refining and enhancing the user experience before the official release.</p>	<p>GIMP available as a Technology Preview in RHEL 9</p> <p>GNU Image Manipulation Program (GIMP) 2.99.8 is now available in RHEL 9 as a Technology Preview. The <code>gimp</code> package version 2.99.8 is a pre-release version with a set of improvements, but a limited set of features and no guarantee for stability. As soon as the official GIMP 3 is released, it will be introduced into RHEL 9 as an update of this pre-release version.</p> <p>In RHEL 9, you can install <code>gimp</code> easily as an RPM package.</p>

Which text is written by a human TW?

A	B
<p>Deprecated Functionality: Change in Default Settings for SSSD Local User Caching</p> <p>Description: In Red Hat Enterprise Linux 9.4, a change in default settings has been made regarding the caching of local users by the System Security Services Daemon (SSSD). As part of this adjustment, SSSD will no longer cache local users by default.</p> <p>Changes:</p> <ul style="list-style-type: none">• Default Settings Update:<ul style="list-style-type: none">◦ Local user caching, previously enabled by default, is deprecated in this release.◦ SSSD will no longer cache local users by default, impacting the default behavior of the system. <p>Impact:</p> <ul style="list-style-type: none">• Users relying on the previous default behavior of local user caching may experience a change in SSSD's behavior.• This deprecation aligns with evolving security and performance considerations. <p>Recommended Action:</p> <ul style="list-style-type: none">• Administrators relying on local user caching in SSSD are advised to review their configurations.• If local user caching is still required, administrators should explicitly configure SSSD to enable this functionality. <p>Note: This deprecation is a proactive step toward optimizing default settings in SSSD. While the default behavior has changed, users have the flexibility to tailor SSSD configurations based on their specific requirements. The deprecation aims to align with evolving best practices and ensure a more streamlined experience for administrators managing user caching in SSSD.</p>	<p>The SSSD implicit files provider domain is disabled by default</p> <p>The SSSD implicit <code>files</code> provider domain, which retrieves user information from local files such as <code>/etc/shadow</code> and group information from <code>/etc/groups</code>, is now disabled by default.</p> <p>To retrieve user and group information from local files with SSSD:</p> <p>Configure SSSD. Choose one of the following options:</p> <ol style="list-style-type: none">1. Explicitly configure a local domain with the <code>id_provider=files</code> option in the <code>sssd.conf</code> configuration file.<pre>[domain/local] id_provider=files ...</pre>2. Enable the files provider by setting <code>enable_files_domain=true</code> in the <code>sssd.conf</code> configuration file.<pre>[sssd] enable_files_domain = true</pre> <p>Configure the name services switch.</p> <pre># authselect enable-feature with-files-provider</pre>

Which text is written by a human TW?

A	B
<p>Simplified testing with infinispan-dev-mode.xml file</p> <p>Infinispan Server distribution now includes <code>infinispan-dev-mode.xml</code> file that provides you with necessary configuration to simplify development experience with Infinispan and Quarkus projects. The <code>infinispan-dev-mode.xml</code> configuration file enables BASIC client authentication for the Hot Rod client and cross-site configuration.</p> <p>To start the Infinispan Server with this configuration, complete the following steps:</p> <ol style="list-style-type: none">1. Copy the server/ directory and rename the directory to <code>server2/</code>.2. Start the first server using the <code>infinispan-dev-mode.xml</code> file. <code>./bin/server.sh -c infinispan-dev-mode.xml</code>3. In another terminal, start the second server using the same configuration. <code>./bin/server.sh -c infinispan-dev-mode.xml -s server2 -Dinfinispan.site.name=site2 -o 100 -Djgroups.mcast_port=46656</code>	<p>Infinispan Development Mode Configuration File (<code>infinispan-dev-mode.xml</code>)</p> <p>This document provides a brief overview of the <code>infinispan-dev-mode.xml</code> file and how to work with it.</p> <p>What is it?</p> <p>The <code>infinispan-dev-mode.xml</code> file is an XML configuration file used with Infinispan, an open-source data grid platform. This specific file is designed for development environments and configures various aspects of Infinispan, including:</p> <ul style="list-style-type: none">• Networking: How Infinispan nodes communicate with each other.• Caching: Defines cache configuration and behavior.• Security: Sets up authentication and authorization for accessing the cache.• Server: Configures server interfaces, ports, and endpoints. <p>How to Use It:</p> <ol style="list-style-type: none">1. Location: This file is typically located in the Infinispan configuration directory.2. Editing: You can edit this file using a text editor to modify various settings for your development environment.3. Comments: Comments within the file (denoted by ``) explain specific configurations and provide instructions for running the configuration in a multi-server setup. <p>Important Note:</p> <p>This is a development configuration file and is not intended for use in production environments. For production environments, a different configuration with stricter security measures should be used.</p>

EXERCISE: Now you try it!

Cause: Too restricted checks in openssl

Consequence: random tls connection failures

Workaround (if any): just reconnect

Result: handshake succeeds

gemini.google.com

bing.com/chat

chat.openai.com

How did you do compared to basic LLMs?

Naive Gemini	GPT 3.5 with a custom “style guide”	Naive ChatGPT
<p>OpenSSL Strictness Causing TLS Connection Failures (Resolved)</p> <p>Issue: Recent updates to OpenSSL have introduced strictness that can lead to random TLS connection failures in some environments.</p> <p>Impact: Users might experience unexpected connection failures when establishing secure connections (TLS).</p> <p>Resolution: This issue has been addressed in the latest OpenSSL update. Upgrading to the latest version is recommended to ensure secure connections.</p> <p>Workaround (Temporary): If updating OpenSSL is not immediately possible, a temporary workaround is available. Simply reconnecting after encountering a failed connection attempt may allow the handshake to succeed.</p> <p>Note: This workaround is not a long-term solution and should only be used until the system can be updated with the latest OpenSSL version.</p> <p>Recommendation: We strongly recommend upgrading to the latest OpenSSL version to benefit from the improved security measures and avoid potential connection issues.</p>	<p>Release Note: Improved OpenSSL Connection Checks</p> <p>Issue: OpenSSL sometimes fails to connect with TLS when FFDHE is used</p> <p>TLS connections that use finite-field Diffie-Hellman key exchange (FFDHE) sometimes fail. This is due to overly restrictive checks in OpenSSL. To work around this problem, reconnect and the handshake succeeds.</p>	<p>Release Note: OpenSSL Update (Version X.X.X)</p> <p>Issue Addressed: This update addresses the issue where OpenSSL sometimes fails to connect with TLS when FFDHE is used. It improves the connection process by eliminating overly restrictive checks. Now, connections will consistently succeed without requiring users to reconnect.</p> <p>This update enhances the stability and reliability of TLS connections, providing a seamless experience for users.</p>

Exercise: Generating technical documentation with ChatGPT 3.5

Scenario:

"Document how to manage and configure passkey authentication devices in the Identity Management (IdM) environment."

1. Go to ChatGPT 3.5 <https://chat.openai.com/>
2. Copy and paste the scenario above.
2. Save the first response generated by the AI in a gdoc.
3. Experiment with prompting to enhance the AI-generated documentation.
4. Save the best response in the same gdoc.
5. Prepare for discussion to share the reasoning behind selecting the best response and also the strategies you used to generate the documentation.

Exercise: Tips for prompting

"Document how to manage and configure passkey authentication devices in the Identity Management (IdM) environment."

- Provide contextual details.
- Be clear and specific in your prompts.
- Focus on one aspect at the time (content format, structure, language).
- Work one section at a time (heading, abstract, procedural steps, verification...).
- Provide examples of the style or tone you're aiming for.
- Refine the prompts using different wording to get better results.
- Feed ChatGPT with content or reference material.

Exercise: Tips for prompting

Bonus: Custom instructions (system prompt)

Go to **Your profile > Customize ChatGPT** to customize the responses.

For example, create your own style guide:

1. **Be Clear:** Use simple and direct language. Avoid unnecessary jargon unless it's needed.
2. **Stay Formal:** Keep a formal and neutral tone. Don't add personal opinions or emotions.
3. **Talk to the Reader:** Address the reader directly with "you".
4. **Use Present Tense:** Describe actions or states in the present tense, not future tense.
5. **Use Asciidoc:** Format text using Asciidoc markup, not Markdown.
6. **Give Actions:** Provide clear instructions or guidance on what the reader should do.



You need TW skills even for AI tools

- User focus
- Minimalism
- Stylistics, applying a style guide



Technical writing is not just writing

- Research of new topics
- Planning, content strategy, user/content journey
- Verification and testing
- User experience



Final project

Week 5

Final project schedule

- Week 1, March 26 (after class 6, Tools II): Select a project, sign up for issues, fork the repository, and join the communication channel to introduce yourself.
- Week 2, April 2 (after class 7, Tools III): Conduct thorough research on the subject. Ask questions to your SMEs.
- Week 3, April 9 (after class 8, Usability): Create a draft.
- Week 4, April 16 (after class 9, Soft Skills): Create a pull request and undergo SME review.
- **Week 5, April 23 (after class 10, LLMs): Close the SME review.**
- Week 6, April 30: Peer review.
- **Week 7, May 7 at 9:59**: Deadline to submit your content.

Present the project.

Conclude the review process with the subject matter experts (SME)

- Review the pull request thoroughly, including all threads, questions, and comments from SMEs.
- Address all necessary feedback provided by the SME and other reviewers.
- Finalize your documentation based on the feedback received.



If you're unsure about any aspect, ask clarifying questions to ensure there are no remaining action items and that your pull request adequately addresses the issue.

Next class: Crowd-sourced peer review of your final project!

Prerequisites

- Have a final project assigned and a PR created.
- Create at least the first draft of a part of the content.
- Think about any questions or issues you might have at this point. The teachers (as well as your classmates) will be able to help you make progress if you have blockers.
- Optional: Submit your PR in advance for class-wide review, and specify whether you'd like it to be anonymized.
 - If you'd like to volunteer, contact Jirka Herrmann on Discord

Need help?

Discord channels:

General queries, organization, and grading: #general

Final project queries

- Ansible #ansible-final-projects
- Foreman #foreman-final-projects
- GNOME User docs #gnome-final-projects

Technical and project specific queries

Use the community forums, alternatively reach out to SMEs directly:

- Ansible dnaro@redhat.com
- Foreman apetrova@redhat.com
- GNOME Settings feborges@redhat.com



Red Hat

Opportunities

for university students

Masaryk University
April 2024 Brno University of Technology



Red Hat

Courses at MUNI Spring semester 2024

MORE DETAILS



red.ht/RHCoursesSpring24

A woman with dark hair, wearing a white top and black pants, sits cross-legged in a red bean bag chair. She is looking at a laptop on her lap. To her right is a small round table with a red book and a white coffee cup. A tall red floor lamp stands behind her. The background is a large pink circle.

Trh práce v IT

A woman in a red top and black pants stands next to a large, stylized brain diagram. The brain is composed of red lines and dots representing neurons and data flow. Binary code (010101) is shown near the bottom left. The background is a grey circle.

Applied Machine Learning
(FF MU)

A woman in a red top and white pants stands holding a laptop. Behind her are several red server racks. A white cloud icon is positioned above the racks. The background is a pink circle.

Virtualizace od A do Z

A man with a beard, wearing a white t-shirt, sits at a desk working on a laptop. To his left is a red shield icon. Above him is a circular diagram with two headphones and a play button symbol. The background is a pink circle.

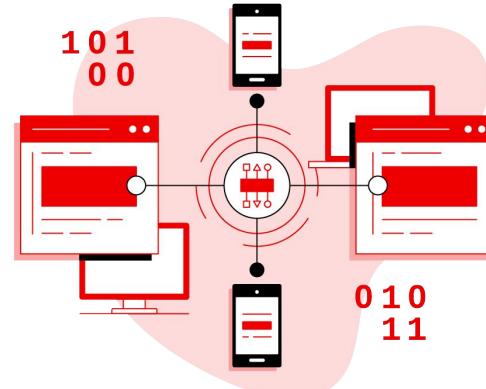
Technical Writing



Red Hat

Sneak Peek: Courses at MUNI Autumn semester 2024

- ▶ Tentative list of courses taught by Red Hat engineers in the upcoming semester
- ▶ All courses are single seminar group (20 places)
- ▶ Registration starts 3. 6. 2024



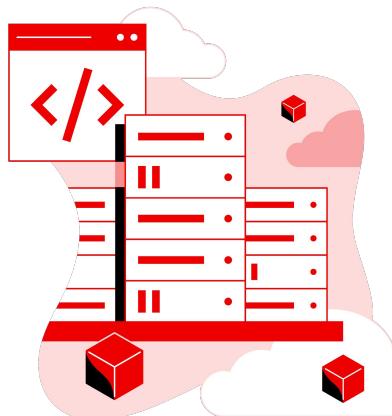
PV278 Vývoj intuitivních uživatelských rozhraní



PB173 Programování v C++ a frameworku Qt



PB173 Coding in Go



PB173 Kernel Development Learning Pipeline



Red Hat

Courses at BUT Summer + Winter 2024

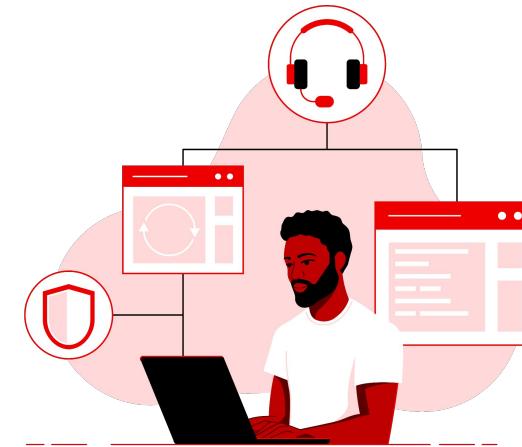
- ▶ IAN Binary Code Analysis and the Technical Writing workshop are taught in the summer semester
- ▶ ILI Advanced Topics of Linux Administration is in winter semester
- ▶ All courses are single seminar group (20 places)



IAN Binary Code Analysis



ILI Advanced Topics of
Linux Administration



Technical writing block workshop
(as a part of Professional Practice at FEEC BUT)



Red Hat Theses

- ▶ Red Hat welcomes students interested to participate on final theses (Bc/Mgr/Ing) with us
- ▶ From 2012, we have participated in over 450 theses (>350 at FI MUNI)
- ▶ Website with open topics will be available from Autumn 2024
 - In the meantime, feel free to come up with your own topic idea (related to open source, linux or anything from our portfolio)
- ▶ If interested, send your idea to thesis@redhat.com





Red Hat Research Fellowships

- ▶ Participate in cutting-edge research where Czech universities cooperate with the industry
 - **Sec-certs** (data mining from security certifications)
 - **Perun** (versioned code performance profiles)
 - **DiffKemp** (checking semantic API equivalence)
 - Multiple other projects ...
- ▶ Long-term cooperation, flexible working hours, financially supported by Red Hat
- ▶ Explore all projects at research.redhat.com
- ▶ If you are interested to join research with Red Hat, contact Martin Ukrop at mukrop@redhat.com





Red Hat Internship program

- ▶ Participate on developing Red Hat products, such as Red Hat® Enterprise Linux®, Middleware, Ansible® OpenShift®, and much more.
- ▶ Flexible working hours, 20 hours/week
- ▶ Paid positions for 12 months+
- ▶ Openings in SW Engineering, Quality Engineering, IT, HR, Sales, and more
- ▶ Positions will be open from **April 15th** to **April 28th**

Follow Red Hat's career page red.ht/RHjobs



DEVCONF.cz

open source komunitní konference

Chceš být v obrazu a kontaktu?

Přijd' se podívat na DevConf.CZ, komunitní konferenci o moderních technologiích a open source.

13.–15. června, 2024, Brno, FIT VUT

Vice info na www.devconf.cz

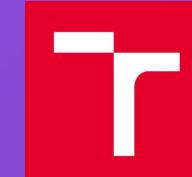
Sponzoři konference:



Red Hat

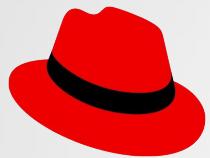


fedora



FACULTY
OF INFORMATION
TECHNOLOGY

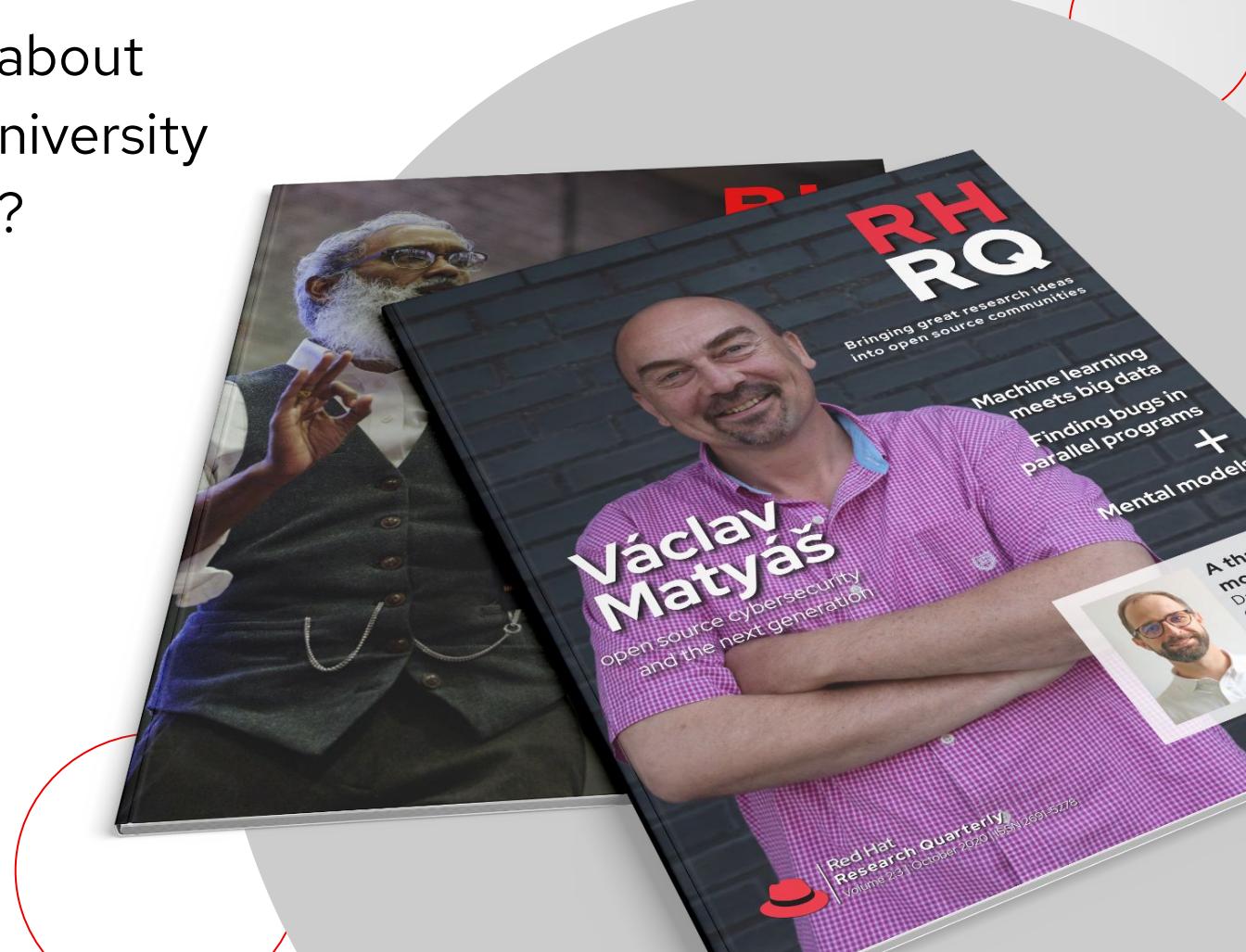




Red Hat Research Quarterly

Are you interested in open source research and innovation? Do you want to learn more about Red Hat's collaboration with Masaryk University and other universities around the world?

Subscribe to the
Red Hat Research
Quarterly in paper
or digital version
for free.



Thank you