

Resenha do Artigo Thoughtworks Technology Radar

Júlia Moreira Nascimento
Projeto de Software

September 9, 2024

1 Temas

1.1 Licenças de código (quase) aberto

Código open source é um tipo de software cujo código-fonte é disponibilizado publicamente, permitindo que qualquer pessoa possa visualizar, modificar e compartilhar. Ele promove a colaboração e a transparência, já que programadores podem contribuir para melhorar o software, corrigir erros e adicionar funcionalidades. Isso também facilita a adaptação do software às necessidades específicas de usuários ou organizações.

Atualmente, várias ferramentas que mudaram de uma licença de código aberto para um modelo comercial, o que por si só não seria um problema, mas pode ser considerado problemático uma vez que uma funcionalidade central de uma ferramenta amplamente utilizada passa a ser paga. Outra questão é que estamos em um momento de ascensão de ferramentas que se autoproclamam como open source, mas grande parte de suas funcionalidades só estão disponíveis na sua versão paga, como em ferramentas de IA.

1.2 Times de desenvolvimento de software assistidos por IA

A Inteligência Artificial (IA) tem sido amplamente discutida, com foco em ferramentas para desenvolvedores como GitHub Copilot e CodiumAI. Foram analisadas várias ferramentas, inclusive terminais assistidos por IA e ChatOps. Acredita-se que a IA pode beneficiar todos os aspectos do desenvolvimento de software, embora novos riscos à qualidade e segurança também surjam, exigindo vigilância das equipes.

1.3 O surgimento de padrões de arquitetura para LLMs

Com o uso crescente de modelos de linguagem de grande porte (LLMs), estão se popularizando padrões de arquitetura para suportar contextos comuns. A geração aumentada por recuperação (RAG) é o padrão preferido dos autores para melhorar a qualidade dos LLM em sistemas corporativos. Padrões formam um vocabulário importante para tecnologia e

é mapeado uma explosão de padrões junto com o uso das AIs generativas no uso do desenvolvimento de software.

1.4 A proximando pull requests de um CI adequado

A integração contínua, embora fundamental para ciclos de feedback rápidos, enfrenta desafios na prática. A obrigatoriedade de pull requests, por exemplo, pode gerar gargalos no fluxo de trabalho. Ferramentas como gitStream e Github merge queue têm sido propostas como soluções para esses problemas. No entanto, a ascensão da IA generativa introduz uma nova complexidade, com a criação de pull requests maiores e mais difíceis de revisar. Neste contexto, a Thoughtworks, pioneira em CI, recomenda que as equipes busquem estratégias para otimizar seus processos de feedback, considerando as particularidades de cada projeto e as limitações de suas ferramentas.

2 Técnicas

2.1 Geração Aumentada por Recuperação (RAG)

A geração aumentada por recuperação é um padrão que está sendo utilizado para melhorar a qualidade das respostas geradas por LLM. A RAG utiliza informações sobre documentos relevantes e confiáveis, os armazena em um banco de dados que suporta formas de pesquisa eficiente e resulta em saídas de alta qualidade e diminuição de respostas irrelevantes. A janela de contexto (quantidade de informações que um sistema pode processar ou considerar de uma vez) é limitada, o que faz com que seja mais relevante a qualidade do conteúdo. Melhoramos a relevância do conteúdo com uma nova classificação. Documentos grandes precisam ser divididos em partes menores para cálculo de embedding (conversão de dados em vetores numéricos para análise), e uma solução é fazer com que essas partes se sobreponham.

3 Plataformas

3.1 CludEvents

Eventos são mecanismos comuns em arquiteturas baseadas em eventos ou aplicações sem servidor. No entanto, diferentes produtores e provedores de nuvem oferecem suporte a eventos em formatos variados, dificultando a interoperabilidade entre plataformas. O CludEvents é uma especificação que padroniza a estrutura de descrição de dados em formatos comuns, visando proporcionar interoperabilidade entre serviços. O CludEvents é hospedado pela Cloud Native Computing Foundation (CNCF).

4 Ferramentas

4.1 Conan

Conan é uma ferramenta de código aberto para o gerenciamento de dependências para aplicações C/C++. Oferecendo uma interface intuitiva e facilitando a integração de bibliotecas de terceiros, o Conan reduz o tempo de compilação, especialmente para dependências pesadas, fazendo com que as bases de códigos resultantes fiquem mais limpas e fáceis de manter.

4.2 Kaniko

Em outubro de 2022, adicionamos o Kaniko ao Radar após o Kubernetes remover o suporte ao Docker. Desde então, expandimos nosso uso do Kaniko e, devido à sua flexibilidade e desempenho, estamos promovendo-o para a categoria "Adote" como ferramenta padrão para construção de imagens de container.

4.3 Karper

O autoscaling horizontal do Kubernetes permite adicionar ou remover pods conforme a demanda, mas só funciona se os nós já existirem. O Cluster Autoscaler expande o cluster de forma básica, mas o Karpenter, um autoscalador de nós mais avançado e de código aberto, oferece maior flexibilidade. Ele ajusta o tipo e a quantidade de instâncias conforme as necessidades dos workloads e agendamento de pods. Desenvolvido inicialmente pela AWS para o EKS, o Karpenter está se tornando padrão entre provedores de Kubernetes, com o Azure recentemente adicionando suporte ao Karpenter no AKS.

5 Crítica

A crítica ao artigo "Thoughtworks Technology Radar" destaca a relevância dos temas abordados, como a comercialização de ferramentas open source, o impacto da IA no desenvolvimento de software, e o surgimento de padrões de arquitetura para LLMs. Embora essas inovações tragam avanços significativos, também levantam questões sobre acessibilidade, segurança e a necessidade de adaptação das equipes às novas ferramentas e fluxos de trabalho. O artigo ressalta a importância de equilibrar a inovação com governança, otimização de processos e colaboração contínua.