

Using LLMs to support architectural decision-making in IoT systems

Rodrigo Nascimento¹, Guilherme Avelino², Davi Viana¹

¹Departamento de Informática – Universidade Federal do Maranhão (UFMA) - Avenida dos Portugueses - 1966 - 65085-580 – São Luís – MA – Brasil

²Departamento de Informática – Universidade Federal do Piauí (UFPI) - Campus Ministro Petrônio Portela – Bairro Ininga - 64049-550 – Teresina – PI – Brasil
Doutorado - Início 06/03/2023 - Término 01/02/2027 - Qualificação 02/03/2026

rodrigo.nascimento@discente.ufma.br, gaa@ufpi.edu.br,

davi.viana@ufma.br

Abstract. *IoT connects physical devices, enabling communication and automation across various areas. However, many IoT systems fail to meet expected requirements because they are complex systems. Reference Architectures offer a common basis for the development of these systems, promoting consistency. This article presents a thesis proposal that focuses on the systematized creation of an RA for IoT, using ProSA-AR methods, and on the development and validation of instantiation and updating mechanisms. The objective is to provide support for software architects, increasing efficiency and quality in the development of IoT systems.*

Resumo. *A IoT conecta dispositivos físicos, permitindo comunicação e automação em diversas áreas. No entanto, muitos sistemas de IoT falham em atender aos requisitos esperados, pois são sistemas complexos. Arquiteturas de Referência oferecem uma base comum para o desenvolvimento desses sistemas, promovendo consistência. Este artigo apresenta uma proposta de tese que se concentra na criação sistematizada de uma AR para IoT, utilizando métodos ProSA-AR, e no desenvolvimento e validação de mecanismos de instanciação e mecanismos de atualização. O objetivo é oferecer suporte para arquitetos de software, aumentando a eficiência e a qualidade no desenvolvimento de sistemas de IoT.*

Link vídeo: <https://youtu.be/sPb1OUn3Wzo>

1. Introdução

A Internet das Coisas (IoT) permite que um grande número de dispositivos físicos se conectem para realizar comunicação e interação de dados sem fio usando a Internet como um ambiente de comunicação global [Zimmermann et al. 2017, Nascimento et al. 2023]. As tecnologias de IoT estão entre as tecnologias transformáveis que estão entrando no mundo físico dos humanos [Ystgaard et al. 2023], e com isso, estão se tornando cada vez mais onipresentes em todos os aspectos da vida humana: em transporte [Moura et al. 2023], na saúde [Bakiris et al. 2024], em ambientes inteligentes (por exemplo, casas inteligentes [Sisodia et al. 2024]).

Porém, um problema comum no desenvolvimento de sistemas de IoT surge quando esses sistemas não atendem às especificações necessárias, o que evidencia a necessidade de esforços de pesquisa, para garantir uma aplicação eficiente das tecnologias de IoT e para compreender melhor como essas tecnologias podem ser implementadas em contextos específicos [Nižetić et al. 2020, Tiwari et al. 2025]. Sistemas de IoT apresentam uma série de características, como conectividade e comunicação entre dispositivos, gerenciamento eficiente de equipamentos, coleta e análise de dados, atuação automatizada, escalabilidade e segurança. Essas características são determinantes para definir como o sistema interage com os usuários finais e como se comporta em diferentes ambientes operacionais.

Diante disso, a identificação e o entendimento dessas características são essenciais para orientar o desenvolvimento de soluções mais eficazes. As Arquiteturas de Referência podem contribuir para enfrentar esse desafio, proporcionando maior consistência e interoperabilidade ao longo de todo o processo de desenvolvimento. Arquiteturas de Referência são modelos que representam uma família de arquiteturas de software para um domínio específico; a partir delas, diversas arquiteturas concretas podem ser instanciadas [Graciano Neto et al. 2015]. O uso de Arquiteturas de Referência permite estabelecer uma base conceitual e técnica que orienta o projeto de sistemas, considerando requisitos, restrições e padrões reutilizáveis. Contudo, ainda não há um método amplamente reconhecido e padronizado para orientar a construção dessas arquiteturas, o que leva a inconsistências em sua implementação [Nakagawa and Antonino 2023], além disso, criar Arquiteturas de Referência custam esforço e conhecimento que muitos desenvolvedores não possuem. Criar Arquiteturas de Referência para sistemas de IoT envolve diversos desafios específicos, cada um impactando a eficácia, segurança, escalabilidade e manutenção da solução final.

Por outro lado, a aplicação de técnicas de aprendizado de máquina no apoio para guiar o desenvolvimento de novas arquiteturas concretas, a partir de Arquiteturas de Referência, é promissora para desenvolvedores de sistemas IoT, pois permite a análise eficiente de grandes volumes de dados e a automação de processos complexos. O aprendizado de máquina é uma subárea da Inteligência Artificial focada no desenvolvimento de algoritmos e modelos que permitem aos computadores aprender a partir de dados, isso permite a automação de tarefas complexas, aprimorando a eficiência. Os avanços em aprendizado de máquina têm impulsionado inovações em diversas áreas, como visão computacional, *Large Language Models* (LLM) e sistemas de recomendação. Além disso, modelos de aprendizado de máquina podem melhorar continuamente com o tempo, à medida que são expostos a mais dados, resultando em previsões e decisões cada vez mais precisas [Rincy and Gupta 2020].

A criação de arquiteturas de IoT requer conhecimento multidisciplinar e apresenta desafios como heterogeneidade, interoperabilidade e requisitos de múltiplos domínios, como hospitais, residências. Como resultado, muitos profissionais têm dificuldade em sistematizar esse processo. As LLMs demonstraram capacidade de apoiar processos criativos e técnicos em diversas áreas [Shinde and Shah 2018], incluindo engenharia de software. Sua capacidade de entender e gerar textos técnicos pode ser explorada para dar suporte a decisões arquitetônicas.

Nesta direção, o objetivo da pesquisa de doutorado é explorar o uso do ProSA-RA,

que é um processo que sistematiza o projeto, a representação e a avaliação de arquiteturas de referência [Nakagawa et al. 2014], para a criação de uma Arquitetura de Referência para sistemas IoT, seguido pelo uso e avaliação de Mecanismos de Instanciação, com suporte de LLM, para gerar arquiteturas concretas e Mecanismos de Atualização para garantir a evolução contínua das ARs. Este trabalho visa fornecer ferramentas concretas para que arquitetos de software aumentem a eficiência e a qualidade no desenvolvimento de sistemas de IoT. A proposta visa empregar o potencial desses modelos para analisar um conjunto de arquiteturas concretas e, a partir dessa análise, extrair uma estrutura generalizável que represente os elementos comuns e recorrentes entre elas. Dessa forma, busca-se oferecer suporte aos arquitetos de software no contexto de sistemas IoT, promovendo não apenas a consistência e reusabilidade, mas também a melhoria explícita da qualidade do software, com soluções mais alinhadas às necessidades reais desses ambientes. Como objetivos específicos, destacam-se:

- Aplicar o método ProSA-AR [Nakagawa et al. 2014] para a criação de uma Arquitetura de Referência (AR) para sistemas IoT, baseada na análise de múltiplas arquiteturas concretas e na elicitação de requisitos do domínio.
- Definir uma abordagem de uso de LLMs para a elicitação e o refinamento de recursos arquiteturais com base em arquiteturas existentes.
- Desenvolver Mecanismos de Instanciação que permitam aos arquitetos derivar arquiteturas de sistemas IoT concretas de forma consistente a partir da AR criada.
- Propor Mecanismos de Atualização da AR para garantir a sua adaptabilidade e manutenção ao longo do tempo, em resposta a novas tecnologias e requisitos.
- Avaliar a qualidade e a eficácia da AR, dos Mecanismos de Instanciação e dos Mecanismos de Atualização, usando critérios de qualidade de software e a validação de especialistas.

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta a fundamentação teórica; a Seção 3 apresenta os desafios das Arquiteturas de Referência para IoT; a Seção 4 apresenta o método de pesquisa; a Seção 5 apresenta conclusões sobre este trabalho e os trabalhos futuros.

2. Fundamentação Teórica

Esta seção apresenta os conceitos fundamentais de sistemas IoT, Arquiteturas de Referência e LLM. Esses conceitos são essenciais para o desenvolvimento da proposta apresentada neste trabalho.

2.1. IoT

IoT é um paradigma que permite compor sistemas a partir de objetos endereçáveis exclusivamente (coisas) equipados com comportamentos de identificação, detecção ou atuação e recursos de processamento que podem se comunicar e cooperar para alcançar um objetivo [Motta et al. 2018]. Esse paradigma encontra aplicação em muitos domínios diferentes, como por exemplo, na saúde, agricultura e até mesmo cidades inteligentes.

Os sistemas de IoT conectam dispositivos e sensores à internet, permitindo a coleta e a troca de dados em tempo real, o que resulta em automação avançada e eficiência aprimorada em diversas áreas, como saúde, agricultura e indústrias. Porém, devido à

heterogeneidade dos dispositivos e à natureza distribuída dos sistemas, a modelagem arquitetural com o uso de padrões e Arquiteturas de Referência aplica níveis de abstração que ajudam a reduzir a complexidade dos sistemas de IoT, ao definir seus componentes de forma mais estruturada e em um nível mais alto [Saemaldahr et al. 2021].

2.2. LLM

LLMs são modelos de Inteligência Artificial desenvolvidos para compreender, gerar e manipular textos em linguagem natural. Eles aprendem padrões, estruturas gramaticais e contextos a partir de grandes volumes de dados textuais. O uso da LLMs aplicados à engenharia de software tem crescido rapidamente pelos mais diversos domínios [Hou et al. 2024].

Esses modelos têm se mostrado promissores para apoiar e, em alguns casos, transformar a pesquisa em Engenharia de Software [Trinkenreich et al. 2025]. Um exemplo de aplicação é a geração de requisitos, como demonstrado por [Gräßler et al. 2022], que utilizou LLMs para aumentar conjuntos de dados de treinamento com requisitos oriundos de diferentes domínios de software. Essa abordagem contribuiu para a melhoria na tarefa de extração de requisitos, comprovando sua eficácia. Eles já vêm sendo utilizados para diversas outras tarefas, como a geração de dados representativos [Steinmacher et al. 2024], o apoio à análise de dados e à pesquisa qualitativa [Barros et al. 2024], além da automação de atividades repetitivas ou tediosas.

Um benefício dos LLMs é que eles foram treinados com grandes quantidades de informações, incluindo documentos relacionados a design arquitetônico e catálogos de padrões de design e outras soluções para problemas de design [Cervantes et al. 2025]. As LLMs podem apoiar a criação de arquiteturas IoT, apoiando a criação dessas arquiteturas para sistemas de IoT.

2.3. Arquiteturas de Referência

Arquitetura de Referência reduzem a complexidade ao segmentar sistemas complexos em partes gerenciáveis, facilitando o entendimento e a implementação. Além disso, orienta a tomada de decisões, fornecendo um framework claro para avaliar e comparar soluções tecnológicas, resultando em um desenvolvimento mais eficiente e de alta qualidade. Portanto, as orientações fornecidas por uma Arquitetura de Referência são elementos úteis para orientar e facilitar o desenvolvimento de sistemas IoT [Cavalcante et al. 2015, Benotmane et al. 2022].

A Arquitetura de Referência abrange componentes essenciais, suas interações, padrões de design e restrições, permitindo uma abordagem sistemática para resolver problemas recorrentes de engenharia de software. Essa abordagem é interessante em ambientes complexos, como sistemas de sistemas onde as características e complexidades específicas podem ser consideradas um grande desafio [Paes and Neto 2023], incluindo IoT, onde a interoperabilidade e a escalabilidade são cruciais. Além disso, uma arquitetura de IoT precisa suportar atualizações gerenciadas automaticamente e dispositivos gerenciados remotamente. Frequentemente, dispositivos conectados coletam e analisam dados pessoais ou relevantes para a segurança [Zimmermann et al. 2015].

3. Desafios das Arquiteturas de Referência para IoT

Apesar da relevância das Arquiteturas de Referência, vários desafios continuam afetando sua ampla adoção. Arquiteturas concretas são projetadas por arquitetos de software, que gastam uma quantidade considerável de tempo tentando garantir a conformidade de arquiteturas concretas com seus Arquiteturas de Referência correspondentes [Angelov et al. 2012].

Conforme apresentado em 2.3, a definição de Arquiteturas de Referência para apoiar o desenvolvimento de sistemas IoT enfrenta desafios como a heterogeneidade dos dispositivos, a necessidade de interoperabilidade, escalabilidade e confiabilidade. Esses fatores se tornam ainda mais complexos devido à diversidade de partes interessadas e requisitos envolvidos em sistemas IoT. Embora a literatura reconheça a importância de Arquiteturas de Referência para padronizar e guiar o desenvolvimento [Nakagawa and Antonino 2023], ainda há lacunas significativas, especialmente em métodos inteligentes e dinâmicos que adaptem arquiteturas às necessidades específicas do usuário. Assim, uma LLM como ferramenta para auxiliar essa especificação pode não apenas ajudar a sistematizar o conhecimento já disponível na literatura, mas também a personalizar e a automatizar processos, melhorando assim a eficiência.

4. Método de Pesquisa

O objetivo principal deste trabalho é propor e validar um método para apoiar o desenvolvimento de uma Arquitetura de Referência para IoT: sua criação, instanciação e atualização, considerando as necessidades específicas de arquitetos de sistemas IoT, como as camadas, padrões de integração, componentes, restrições de domínio e estilo arquitetural, de modo a ajudá-los no desenvolvimento de software. Para tanto, será necessário realizar uma série de etapas. Na Figura 1 são apresentadas as etapas da proposta. O método proposto tem as seguintes etapas principais:

- Condução de um mapeamento sistemático da literatura para identificar soluções existentes e mapear as principais características em arquiteturas de IoT, com o objetivo de extrair e categorizar os principais recursos, padrões e desafios arquitetônicos presentes na literatura.
- Aplicação de um Método sistematizado: Utilização do ProSA-AR para guiar formalmente a concepção da AR a partir dos dados do mapeamento. O foco será na definição formal das camadas, componentes, padrões de integração e restrições de domínio da AR.
- Desenvolvimento dos Mecanismos de Instanciação: Desenvolvimento de artefatos práticos que orientem o arquiteto passo a passo na derivação de uma arquitetura específica a partir da AR, garantindo a conformidade arquitetural.
- Desenvolvimento de Ferramentas: Implementação de ferramentas de software para automatizar parte do processo de instanciação.
- Desenvolvimento dos Mecanismos de Atualização: Definição de um processo que permita a evolução e a manutenção da AR de forma controlada, adaptando-a a novos requisitos ou tecnologias emergentes.
- Execução de Estudos de Caso para demonstrar a eficácia e a aplicabilidade da estrutura, serão realizados estudos de caso em 2-3 domínios de IoT distintos, como casa inteligente e saúde.

- Avaliação e análise de resultados com questionários e entrevistas com especialistas e a análise qualitativa e quantitativa dos resultados.

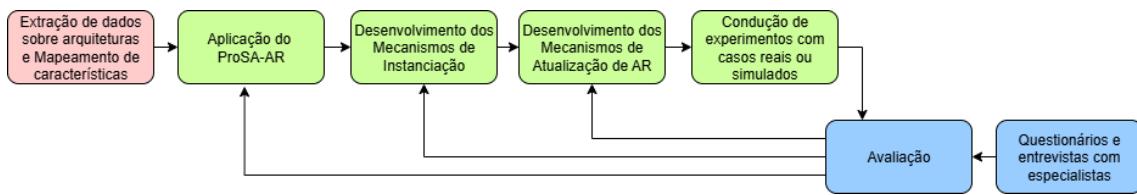


Figura 1. Metodologia Proposta.

5. Conclusão e trabalhos futuros

Como resultados parciais, foi realizado um mapeamento sistemático da literatura para identificar, analisar e caracterizar o estado da arte das arquiteturas para sistemas IoT. Para isso, foi realizada uma busca em seis bibliotecas digitais. No total, 648 artigos foram recuperados e analisados, sendo filtrados de acordo com critérios pré-estabelecidos. Desse, 104 estudos foram selecionados com base nos critérios definidos. Nas arquiteturas identificadas em cada artigo, foi feita uma análise e categorização da base teórica e da metodologia utilizadas, bem como das características observadas, tais como as camadas apresentadas e as tecnologias envolvidas.

Este trabalho resultará em uma Arquitetura de Referência para IoT formalmente criada, complementada por mecanismos de instanciação e mecanismos de atualização validados, que servirão como ferramentas para o desenvolvimento de sistemas IoT. Com esta pesquisa, espera-se a redução de erros e promover a adoção de melhores práticas na área, com um método que acelere o desenvolvimento de aplicações de IoT. A solução proposta auxiliará os desenvolvedores na criação de Arquiteturas de Referência eficazes, consistentes e interoperáveis, adaptadas a diversos cenários de aplicação, aumentando a eficiência no desenvolvimento de sistemas IoT. Como próximos passos, estão previstos a finalização da arquitetura gerada pelo ProSA-AR e a criação do guidelines e a ferramenta para apoiar o arquitetura de software.

Referências

- Angelov, S., Grefen, P., and Greefhorst, D. (2012). A framework for analysis and design of software reference architectures. *Information and Software Technology*, 54(4):417–431.
- Bakiris, E., Papadakis, N., Katsarou, V., and Alampasis, N. (2024). Septon toolkit application: An overview of the security techniques used from wearable medical devices to physician's healthcare platform. In *Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '24, page 582–586, New York, NY, USA. Association for Computing Machinery.
- Barros, C., Azevedo, B., Graciano Neto, V., Kassab, M., Kalinowski, M., Nascimento, H., and Bandeira, M. (2024). Large language model for qualitative research – a systematic mapping study.
- Benotmane, M., Elhari, K., and Kabbaj, A. (2022). Meta-comparison of iot reference architectures and models. In Kacprzyk, J., Balas, V. E., and Ezziyyani, M., editors,

Advanced Intelligent Systems for Sustainable Development (AI2SD'2020), pages 337–357, Cham. Springer International Publishing.

Cavalcante, E., Alves, M. P., Batista, T., Delicato, F. C., and Pires, P. F. (2015). An analysis of reference architectures for the internet of things. In *Proceedings of the 1st International Workshop on Exploring Component-Based Techniques for Constructing Reference Architectures*, CobRA '15, page 13–16, New York, NY, USA. Association for Computing Machinery.

Cervantes, H., Kazman, R., and Cai, Y. (2025). An llm-assisted approach to designing software architectures using add.

Graciano Neto, V. V., Garcés, L., Guessi, M., de Oliveira, L. B. R., and Oquendo, F. (2015). On the equivalence between reference architectures and metamodels. In *Proceedings of the 1st International Workshop on Exploring Component-Based Techniques for Constructing Reference Architectures*, CobRA '15, page 21–24, New York, NY, USA. Association for Computing Machinery.

Gräßler, I., Preuß, D., Brandt, L., and Mohr, M. (2022). Efficient extraction of technical requirements applying data augmentation. In *2022 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–8.

Hou, X., Zhao, Y., Liu, Y., Yang, Z., Wang, K., Li, L., Luo, X., Lo, D., Grundy, J., and Wang, H. (2024). Large language models for software engineering: A systematic literature review.

Motta, R. C., de Oliveira, K. M., and Travassos, G. H. (2018). On challenges in engineering iot software systems. In *Proceedings of the XXXII Brazilian Symposium on Software Engineering*, SBES '18, page 42–51, New York, NY, USA. Association for Computing Machinery.

Moura, D. L. L., Aquino, A. L. L., and Loureiro, A. A. F. (2023). On the integration of ledger technology and edge computing for intelligent transportation systems. In *Proceedings of the Int'l ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, PE-WASUN '23, page 9–16, New York, NY, USA. Association for Computing Machinery.

Nakagawa, E. Y. and Antonino, P. O. (2023). *An Overview of Reference Architectures*, pages 5–15. Springer International Publishing, Cham.

Nakagawa, E. Y., Guessi, M., Maldonado, J. C., Feitosa, D., and Oquendo, F. (2014). Consolidating a process for the design, representation, and evaluation of reference architectures. In *2014 IEEE/IFIP Conference on Software Architecture*, pages 143–152.

Nascimento, R., Santos, V., Carvalho, B., Correia, J., Rivero, L., Santos, R., Silva, F., Teles, A., and Viana, D. (2023). Sysiotml: A technique for modeling applications in the context of iot. In *Proceedings of the 25th International Conference on Enterprise Information Systems - Volume 2: ICEIS*, pages 187–194. INSTICC, SciTePress.

Nižetić, S., Šolić, P., López-de-Ipiña González-de-Artaza, D., and Patrono, L. (2020). Internet of things (iot): Opportunities, issues and challenges towards a smart and sustainable future. *Journal of Cleaner Production*, 274:122877.

- Paes, C. and Neto, V. G. (2023). Sosml: Rumo a uma linguagem de modelagem arquitetural para sistemas-de-sistemas intensivos em software. In *Anais do V Workshop em Modelagem e Simulação de Sistemas Intensivos em Software*, pages 61–65, Porto Alegre, RS, Brasil. SBC.
- Rincy, T. and Gupta, R. (2020). A survey on machine learning approaches and its techniques:. In *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pages 1–6.
- Saemaldahr, R., Thapa, B., Maikoo, K., and Fernandez, E. B. (2021). Reference architectures for the iot: A survey. In Saeed, F., Mohammed, F., and Al-Nahari, A., editors, *Innovative Systems for Intelligent Health Informatics*, pages 635–646, Cham. Springer International Publishing.
- Shinde, P. P. and Shah, S. (2018). A review of machine learning and deep learning applications. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pages 1–6.
- Sisodia, D., Li, J., Mergendahl, S., and Cam, H. (2024). A two-mode, adaptive security framework for smart home security applications. *ACM Trans. Internet Things*, 5(2).
- Steinmacher, I., Penney, J. M., Felizardo, K. R., Garcia, A. F., and Gerosa, M. A. (2024). Can chatgpt emulate humans in software engineering surveys? In *Proceedings of the 18th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, ESEM '24, page 414–419, New York, NY, USA. Association for Computing Machinery.
- Tiwari, N., Pandey, A., and Kokala, A. (2025). Challenges and solutions for countering emerging security threats to iot devices. *International Journal of Experimental Research and Review*, 47:59–68.
- Trinkenreich, B., Calefato, F., Hanssen, G., Blincoe, K., Kalinowski, M., Pezzè, M., Tell, P., and Storey, M.-A. (2025). Get on the train or be left on the station: Using llms for software engineering research.
- Ystgaard, K., Atzori, L., Palma, D., Heegaard, P., Bertheussen, L., Jensen, M., and De Moor, K. (2023). Review of the theory, principles, and design requirements of human-centric internet of things (iot). *Journal of Ambient Intelligence and Humanized Computing*, 14:1–33.
- Zimmermann, A., Schmidt, R., Sandkuhl, K., Jugel, D., Bogner, J., and Möhring, M. (2017). Decision management for micro-granular digital architecture. In *2017 IEEE 21st International Enterprise Distributed Object Computing Workshop (EDOCW)*, pages 29–38.
- Zimmermann, A., Schmidt, R., Sandkuhl, K., Wißotzki, M., Jugel, D., and Möhring, M. (2015). Digital enterprise architecture - transformation for the internet of things. In *2015 IEEE 19th International Enterprise Distributed Object Computing Workshop*, pages 130–138.