

Aluno(s): Evaldo Felipe Lima, Talles Eduardo do Amaral Salles

Residência Técnica em Sistemas de Informação  
2023-2

## **Disponibilizar Ferramentas de Apoio à Gestão Acadêmica**

O respectivo projeto tem como objetivo arquitetar um modelo de sistemas de informação focado em unidades (operacionais) de ensino e suas ramificações para receber informações de diferentes fontes/sistemas (bases de dados internas e externas, documentos em formato textual e tabular, entre outros), tratar e integrar essas informações, além de permitir a consumação e visualização das mesmas de forma dinâmica e interativa, de modo a:

- Apoiar as operações e decisões gerenciais;
- Apoiar as políticas e decisões estratégicas;
- Gerar relatórios que possibilitem a conformidade com órgãos regulatórios;
- Fornecer facilidades para análises centradas no comportamento institucional;
- Fornecer facilidades para análises que revelem as dinâmicas, desempenho e resultados com foco nas áreas de conhecimento, cursos, instituição de ensino, localidades, corpos docentes e discentes;
- Fornecer facilidades para análises preditivas e prescritivas;
- Fornecer facilidades para maior transparência dos atos e fatos administrativos e acadêmicos;
- Possibilitar maior alinhamento com as métricas/objetivos estabelecidos pelos Órgãos Superiores.

Além disso, o acesso e análise de dados em um sistema de apoio acadêmico apresenta um desafio significativo para a instituição, abrangendo diferentes aspectos e impactando diversas áreas. O escopo pode variar de acordo com a abrangência dos dados e a importância estratégica dessas informações para a instituição. Dessa forma, é necessário dedicar esforços para extrair os dados do sistema acadêmico, transformá-los em um formato acessível e disponibilizá-los para análise. Além disso, pode ser necessário adquirir ou desenvolver ferramentas e capacitar a equipe para lidar com a complexidade do sistema acadêmico.

# Sprint 1, Semana 1: Imersão Inicial no problema

A fim de facilitar o acesso e análise de dados no sistema acadêmico e disponibilizar as informações de forma a auxiliar na tomada de decisão pode ser um desafio, mas existem algumas estratégias que podem ajudar. Não obstante, deve-se salientar que cada sistema possui suas particularidades, portanto, é importante avaliar a melhor abordagem com base nas características específicas do sistema e das necessidades da organização. Será inerente ao projeto a aplicação de ferramentas e conceitos de Business Analytics e do processo de ETL (*Extract, Transform and Load*).

Para tal, haja vista a necessidade de cumprimento do roteiro, deve-se realizar a análise do cenário de prática, conforme citado em [1], para pontuarmos as percepções iniciais do time quanto à “Imersão Inicial” (**Situação-Problema, Ambiente e Negócio**) no projeto proposto.

**Situação-problema:** O sistema acadêmico da instituição, que armazena informações sobre alunos, cursos, professores e unidades acadêmicas, possui uma estrutura complexa e limitações de acesso aos dados.

Assim, os responsáveis pela análise de dados enfrentam dificuldades para acessar, extrair, analisar e disponibilizar as informações necessárias para tomada de decisões estratégicas. Posto isso, os principais problemas identificados são:

- **Acesso restrito:** o acesso aos dados do sistema acadêmico é limitado e requer permissões especiais (i.e. isso dificulta a obtenção rápida e fácil das informações necessárias para análise).
- **Grande número de fontes de dados** e suas diferentes organizações (dados estruturados e não estruturados, diferentes bancos de dados, planilhas, arquivos externos, etc).
- **Estrutura de dados complexa:** o sistema acadêmico possui uma estrutura de dados complexa, com várias tabelas relacionadas e campos que não estão claramente documentados (i.e. dificulta o entendimento da organização dos dados e encontrar as informações necessárias para análise).
- **Dificuldade na extração de dados:** a extração dos dados do sistema acadêmico é um processo complexo e demorado (i.e. não há uma única forma eficiente de exportar os dados para ferramentas de análise mais populares, como planilhas ou bancos de dados relacionais).
- **Qualidade e integridade dos dados:** validar se os dados armazenados no sistema acadêmico apresentam problemas de qualidade e integridade, como registros duplicados, valores inconsistentes ou ausência de informações importantes (i.e. isso compromete a confiabilidade das análises realizadas).
- **Limitações de análise:** a complexidade da estrutura dos dados e as limitações de acesso dificultam a realização de análises avançadas, como segmentação de

períodos de tempo em relação à outros dados requeridos, identificação de problemas dentro da instituição e não conformidades com os Órgãos Superiores, análises detalhadas e completas sobre os alunos, professores e cursos.

Essa situação-problema reflete um cenário comum em instituições acadêmicas que possuem sistemas acadêmicos antigos ou de grande dimensão, e que foram desenvolvidos sem a percepção das demandas atuais de acesso e análise de dados. Solucionar esses problemas é essencial para melhorar a eficiência das análises e permitir uma tomada de decisão embasada, ágil e assertiva.

**Ambiente da estruturação do problema:** o ambiente da estruturação do problema no acesso e análise de dados em um sistema acadêmico refere-se ao contexto em que o problema ocorre, incluindo os elementos que influenciam e afetam a situação. Esse ambiente pode ser definido considerando os seguintes aspectos:

- **Instituição de Ensino:** o problema ocorre dentro de uma determinada Instituição de Ensino, que possui suas próprias características, gestão, regras e objetivos. A estrutura dessa instituição, as políticas de dados e as práticas de gestão podem influenciar a forma como o acesso e análise de dados são abordados.
- **Tecnologia:** o ambiente tecnológico desempenha um papel fundamental no acesso e análise de dados em um sistema acadêmico. Isso inclui o sistema acadêmico em si, suas ramificações e sistemas adjacentes, bem como outras tecnologias utilizadas para extrair, transformar e analisar os dados. A infraestrutura de hardware, software e redes também faz parte desse ambiente. Além disso, não se pode desprezar as informações que são capturadas e armazenadas de forma manual e física.
- **Regulamentações e conformidade:** do ponto de vista das Instituições de Ensino, sobretudo públicas, existem regulamentações e normas específicas que devem ser consideradas ao acessar e analisar dados (i.e. isso inclui aspectos relacionados à privacidade, segurança e conformidade legal).
- **Objetivos e necessidades da Instituição Acadêmica:** o ambiente também é influenciado pelos objetivos e necessidades específicas da Instituição em relação ao acesso e análise de dados (i.e. isso pode incluir a necessidade de melhorar as tomadas de decisões, aumentar a precisão das previsões, identificar problemas ou identificar possíveis pontos de melhoria).
- **Restrições e limitações:** o ambiente também é influenciado por restrições e limitações, como restrições orçamentárias, prazos apertados, recursos limitados ou a resistência à mudança. Essas restrições podem afetar a abordagem e a viabilidade das soluções propostas neste projeto.

Ao considerar o ambiente da estruturação do problema no acesso e análise de dados em um sistema acadêmico, é possível identificar os fatores que influenciam e afetam a situação, permitindo uma abordagem mais precisa e efetiva na busca de soluções.

**Natureza do negócio na estruturação do problema:** A instituição possui um amplo banco de dados, com informações sensíveis de estudantes, cursos, professores e unidades

acadêmicas, e realiza ofertas, controle e execução de disciplinas e cursos, por meio de seu Sistema Integrado de Gestão Acadêmica (SIGAA).

- **Objetivos do negócio:** os principais objetivos desse sistema para a instituição incluem extrair, refinar e disponibilizar as informações a fim de melhorar a tomada de decisão no ambiente acadêmico, além de prover uma visão micro e macro sobre como o ambiente está funcionando e identificar tendências através de indicadores da educação superior, como citado em [2]. A análise de dados é fundamental para alcançar esses objetivos, permitindo a identificação acadêmica e previsões de demanda e tomada de decisões estratégicas.
- **Necessidades do negócio:** a instituição precisa ter acesso fácil e rápido às informações sobre dados acadêmicos e relatórios para tomar decisões informadas. (e.g. isso inclui dados sobre quantidade de alunos formados em um ano, quantidade de alunos com média global maior que 6,0 que tenham reprovado na disciplina Análise e Projeto de Algoritmos). A disponibilidade dessas informações é essencial para melhorar a eficiência operacional, planejar oferta de turmas e otimizar a gestão de recursos da unidade acadêmica.
- **Restrições do negócio:** a instituição enfrenta restrições de legislação vigente para resolver o problema de acesso e análise de dados do sistema acadêmico. Além disso, a instituição pode ter limitações técnicas em termos de recursos de TI ou falta de conhecimento especializado para lidar com a complexidade do sistema acadêmico.

## Sprint 1, Semana 2 e 3: Método de Estruturação de Problema

### Aplicação do Método SODA (Structured Objective Decision Analysis)

O método **SODA** (Structured Objective Decision Analysis) é uma metodologia que auxilia na estruturação e resolução de problemas complexos. No contexto de extrair e disponibilizar dados de um sistema de gestão acadêmica, podemos seguir os seguintes passos usando o método SODA:

- **Definir o objetivo:** o primeiro passo é identificar claramente o objetivo da extração e disponibilização dos dados do sistema de gestão acadêmica (e.g. o objetivo pode ser melhorar a eficiência na análise de dados acadêmicos ou fornecer informações para facilitar a análise e o uso dos dados acadêmicos).
- **Identificar os critérios relevantes:** em seguida, é necessário identificar os critérios que são relevantes para alcançar o objetivo definido (e.g. os critérios podem ser a precisão dos dados extraídos, a facilidade de acesso aos dados ou a segurança das informações).
- **Identificar as alternativas:** neste passo, é necessário identificar as diferentes alternativas disponíveis para extrair e disponibilizar os dados do

sistema de gestão acadêmica (i.e. pode incluir o desenvolvimento de uma nova API, a integração com um sistema de business intelligence, a implementação de um data warehouse ou a adoção de ferramentas de “Extract, Transform, Load”).

- **Estruturar as relações de influência:** agora é necessário identificar as relações de influência entre os critérios e as alternativas (e.g. a escalabilidade da arquitetura pode ser influenciada pela utilização de uma arquitetura em nuvem ou pela adoção de ferramentas de ETL eficientes).
- **Avaliar as alternativas:** com as relações de influência estabelecidas, é possível avaliar cada alternativa em relação aos critérios definidos. Isso pode ser feito atribuindo pesos para cada critério e avaliando cada alternativa em relação a esses pesos (e.g. se a precisão dos dados extraídos for considerada o critério mais importante, as alternativas serão avaliadas com base nesse critério).
- **Tomar a decisão:** com base na avaliação das alternativas, é possível tomar uma decisão informada sobre qual a melhor abordagem para extrair e disponibilizar os dados do sistema de gestão acadêmica (i.e. a decisão pode ser baseada na alternativa que obteve a maior pontuação geral ou na alternativa que melhor atende aos critérios mais relevantes).
- **Implementar a solução:** após tomar a decisão, é necessário implementar a solução escolhida. Isso pode envolver o desenvolvimento da API, a integração com o sistema de business intelligence, a configuração das ferramentas selecionadas ou a integração com outros sistemas.
- **Monitorar e ajustar:** por fim, é importante monitorar a implementação da solução e realizar ajustes conforme necessário. Isso pode envolver a verificação da precisão dos dados extraídos, a análise do feedback dos usuários ou a resolução de eventuais problemas que surgirem.

Seguindo esses passos, é possível aplicar o método SODA para extrair e disponibilizar dados de um sistema de gestão acadêmica de forma estruturada e eficiente.

## Aplicação do Canvas de Hipótese de Negócio

O Canvas de Hipótese de Negócio [4] é uma ferramenta flexível que pode ser adaptada e iterada à medida que o empreendimento evolui e novas informações são obtidas. Ele ajuda a testar e validar as suposições subjacentes ao negócio, permitindo uma melhor compreensão dos riscos e oportunidades envolvidos.

- **Segmento de clientes:** Instituições de ensino que utilizam sistemas de gestão acadêmica, que não possuem integração dos sistemas internos e/ou não possuem ferramentas de análise de dados.
- **Proposta de valor:** Facilitar a extração e disponibilização de dados do sistema de gestão acadêmica, permitindo uma melhor análise dos dados acadêmicos e sua correlação com o ambiente e tomada de decisão.

- **Canais:** Plataforma online que permite aos interessados acessem e visualizem os dados do sistema de gestão acadêmica com relatórios e análises personalizadas.
- **Relacionamento com clientes:** Suporte técnico para auxiliar os clientes na extração e disponibilização dos dados, além de fornecer atualizações e melhorias contínuas baseadas nos interesses dos stakeholders.
- **Fontes de receita:** O projeto terá suporte financeiro da união e da instituição de ensino que fizerem uso do sistema.
- **Recursos chave:** Equipe de análise e desenvolvimento para manter e aprimorar a plataforma, infraestrutura de servidores para armazenamento e processamento dos dados.
- **Atividades chave:** Desenvolvimento e manutenção da plataforma, suporte técnico aos stakeholders, atualização e melhoria contínua da plataforma com base em novas demandas.
- **Parcerias chave:** Parcerias com instituições de ensino e unidades acadêmicas visando o desenvolvimento assertivo com base nos interesses dos stakeholders, a validação da eficácia da plataforma, além da obtenção de feedback dos usuários.
- **Estrutura de custos:** Custos de desenvolvimento e manutenção da plataforma, custos de infraestrutura de servidores, custos de suporte técnico e treinamento.
- **Métricas chave:** Número de usuários, taxa de adesão à plataforma, satisfação do cliente, tempo médio de resultado de uma consulta aos dados, satisfação com os filtros e relatórios disponibilizados.

## Sprint 2, Semana 4: Seleção dos pontos chave e priorização de problemas junto ao PO

### Aplicação da Matriz G.U.T (Gravidade, Urgência, Tendência)

A **Matriz GUT** é uma ferramenta de análise utilizada para priorizar problemas ou situações que requerem ação imediata. GUT é um acrônimo para Gravidade, Urgência e Tendência, conforme pontuado em [5].

- **Gravidade (G):** avalia o impacto negativo que o problema pode causar caso não seja tratado (i.e. quanto maior a gravidade, maior a pontuação atribuída).
- **Urgência (U):** reflete a necessidade de resolver o problema o mais rápido possível (i.e. quanto maior a urgência, maior a pontuação atribuída).
- **Tendência (T):** indica a tendência de piora do problema caso não seja tratado (i.e. quanto maior a tendência, maior a pontuação atribuída).

Cada critério foi pontuado individualmente em uma escala de 1 a 5, sendo 1 a menor pontuação e 5 a maior. A multiplicação dos valores atribuídos a cada critério resulta em uma

pontuação total para cada problema analisado. Quanto maior a pontuação final, maior a prioridade para ser resolvido.

Dessa forma, a Matriz GUT permitiu uma análise objetiva e sistematizada dos problemas, ***disponível no Anexo 1, Matriz GUT dos apêndices***, priorizando aqueles que apresentam maior impacto, urgência e tendência negativa. Isso auxilia os gestores na tomada de decisões eficientes para a resolução dos problemas identificados.

## **Sprint 2, Semana 4: Validar junto aos Stakeholders e realizar os ajustes necessários para finalizar a estruturação do problema**

### **Plano/Programa Geral do Projeto - PGP**

Um **Programa Geral do Projeto (PGP)** é um documento detalhado que descreve todos os aspectos essenciais de um projeto. Ele serve como um guia abrangente que define o escopo, os objetivos, as metas, os recursos, as atividades e as estratégias do projeto.

Dessa maneira, ***disponível no Anexo 2, Programa Geral do projeto / Programa - PGP dos apêndices***, o PGP é geralmente criado no início do projeto, durante a fase de planejamento, e é atualizado e refinado ao longo do ciclo de vida do projeto, à medida que novas informações e requisitos surgem.

Alguns dos elementos-chave que podem ser encontrados em um PGP incluem:

- ***Objetivos do projeto***: declarações claras e mensuráveis que descrevem o que o projeto espera alcançar.
- ***Escopo do projeto***: uma descrição detalhada do trabalho a ser realizado, incluindo produtos ou entregas específicas.
- ***Cronograma***: um cronograma geral das principais atividades do projeto, estimando a duração e a sequência de cada fase.
- ***Riscos***: uma análise dos riscos potenciais que podem afetar o projeto e as estratégias para mitigá-los.

O PGP fornece uma visão geral do projeto, ajudando todos os envolvidos a compreenderem seus objetivos e planos (i.e. serve como uma referência importante para a gestão do projeto, facilitando o acompanhamento e a revisão do progresso, e garantindo que todas as partes interessadas estejam alinhadas e informadas sobre o projeto).

# **Sprint 3, Semana 5 e 6: PBB, Requisitos Funcionais, não funcionais e Histórias de Usuário junto ao PO**

## **PBB - Product Backlog Building**

O **PBB** (Product Backlog Building) tem como principal objetivo ajudar na construção de um backlog de forma colaborativa, construindo um entendimento compartilhado e alinhado com todas as partes envolvidas no produto, ou seja, todos compreendem o contexto do negócio, conforme pontuado em [6] e **disponível no Anexo 5, Product Backlog Building - PBB dos apêndices**.

**Problema:** Nos dias atuais não existe um sistema ou ferramenta acadêmica que possa disponibilizar informações, relatórios, gráficos, entre outros, de forma dinâmica e interativa que possibilitem análises e que apoiem as decisões estratégicas e decisões relacionadas às unidades acadêmicas. Além disso, as ferramentas atuais não integram os vários bancos de dados das instituições e outras fontes, como planilhas. Portanto é necessário modelar uma arquitetura para extrair e disponibilizar dados de um sistema de gestão acadêmico, enfrentando desafios como a complexidade da estrutura de dados, os relacionamentos entre os dados, a escalabilidade, o tempo de resposta, a segurança e privacidade, integração com sistemas externos e a manutenção e evolução do sistema.

**Propósito:** O propósito desta arquitetura é fornecer uma solução robusta e eficiente para extrair dados do sistema de gestão acadêmico e disponibilizá-los de forma segura e integrada a sistemas externos, garantindo a integridade dos dados, a escalabilidade e a realização de consultas rápidas e precisas.

### **Benefícios:**

- Acesso facilitado a dados acadêmicos para usuários e sistemas externos.
- Alta disponibilidade e tempo de resposta eficiente para consultas de dados.
- Segurança e privacidade dos dados acadêmicos garantidos.
- Integração simplificada com sistemas externos.
- Capacidade de escalabilidade para lidar com grandes volumes de dados acadêmicos.
- Flexibilidade para manutenção e evolução contínua do sistema.

### **Abordagem:**

- Análise dos requisitos: Realizar uma análise detalhada dos requisitos do sistema de gestão acadêmico, identificando as principais entidades, relacionamentos, dados necessários e requisitos de segurança, escalabilidade e integração.
- Design da estrutura de dados: Projetar uma estrutura de dados adequada para armazenar os dados extraídos do sistema acadêmico. Considerar a criação de um



banco de dados dedicado para armazenar as informações acadêmicas, adotando técnicas eficientes de modelagem de dados para garantir a consistência e integridade dos dados.

- **Definição de integrações:** Identificar os sistemas externos que precisam acessar os dados acadêmicos e definir os mecanismos de integração adequados. Isso pode envolver o uso de APIs, web services, protocolos de comunicação e outros métodos de troca de informações.
- **Desenvolvimento da camada de extração:** Implementar a camada responsável por extrair os dados do sistema acadêmico. Isso pode envolver a criação de scripts para acessar as fontes de dados legadas, a utilização de ferramentas de ETL (Extract, Transform, Load) ou até mesmo a integração direta com o sistema acadêmico (se possível).
- **Implementação da camada de disponibilização:** Construir uma camada de disponibilização dos dados extraídos, que permita consultas eficientes e seguras. Isso pode ser feito através da criação de APIs, serviços web ou outros mecanismos de acesso controlado às informações acadêmicas.
- **Configuração de segurança:** Implementar medidas de segurança adequadas, como autenticação, autorização e criptografia, para garantir a proteção dos dados acadêmicos.
- **Testes e validações:** Realizar testes rigorosos para validar a correção e desempenho da arquitetura. Isso deve incluir a verificação da integridade dos dados extraídos, a eficiência das consultas e a interoperabilidade com os sistemas externos.
- **Implementação e manutenção contínua:** Implantar a arquitetura em ambiente de produção e garantir que exista um processo eficiente de manutenção e evolução contínua do sistema. Isso pode envolver a criação de rotinas de atualização de dados, monitoramento do desempenho e recursos do sistema, bem como ajustes e melhorias regulares.

**Conclusão:** A modelagem de uma arquitetura para extrair e disponibilizar dados de um sistema de gestão acadêmico é um desafio complexo, mas com uma abordagem adequada e a consideração dos requisitos específicos, é possível criar uma solução eficaz que atenda às necessidades dos usuários e sistemas externos. A adoção de práticas de segurança, escalabilidade e integração adequadas garantirá a disponibilidade, integridade e privacidade dos dados acadêmicos, proporcionando uma experiência positiva para os usuários e uma base sólida para tomada de decisões e análise de dados.

## **Requisitos Funcionais, não funcionais e Histórias de Usuário**

**Requisitos Funcionais:** Os requisitos funcionais referem-se às funcionalidades e comportamentos específicos que um sistema deve possuir para atender a determinadas necessidades do usuário. Esses requisitos descrevem as ações que o sistema deve ser capaz de realizar, os inputs e outputs esperados e as condições necessárias para a execução correta das funcionalidades.

**Requisitos Não Funcionais:** Já os requisitos não funcionais referem-se às características e atributos que um sistema deve ter, além das funcionalidades. Esses requisitos descrevem as restrições de desempenho, confiabilidade, segurança, usabilidade, entre outros aspectos, que são essenciais para a satisfação do usuário, mas que não estão diretamente relacionados às funcionalidades do sistema.

**Histórias de Usuário:** são uma técnica utilizada no desenvolvimento de software ágil para capturar requisitos de uma forma mais centrada no usuário. Elas consistem em pequenas descrições de uma funcionalidade ou necessidade do usuário, escritas em linguagem natural, que descrevem o que o usuário deseja alcançar e por quê.

Conforme pontuado em [6] e **disponível no Anexo 6, Requisitos Funcionais, não funcionais e Histórias de Usuários dos apêndices.**

## **Sprint 4, Semana 7: Definição do Estudo de Caso e Indicadores de Educação Superior junto ao PO**

### **Definição do Estudo de Caso para Análise de Desempenho Acadêmico**

É um método detalhado de coleta de informações que envolve a análise aprofundada de um determinado fenômeno, contexto ou evento. No estudo de caso, o pesquisador investiga um único caso ou um conjunto limitado de casos para entender as nuances e complexidades envolvidas (i.e. o caso pode ser uma pessoa, um grupo, uma organização, uma comunidade ou mesmo um evento específico).

Para conduzir um estudo de caso, coletamos informações com base em entrevistas com o PO (Product Owner) do projeto, fizemos observações e redigimos um documento que registra como as informações analisadas de forma sistemática foram interpretadas para identificar padrões, tendências e insights significativos.

Os estudos de caso são valiosos porque permitem uma compreensão aprofundada de um fenômeno específico em seu contexto real, conforme pontuado em [7] e **disponível no Anexo 7, Estudo de caso - Análise de Desempenho dos apêndices.**

# **Sprint 5 e 6, Semana 8, 9 e 10: Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico**

## **Estudos sobre Extração, integração e visualização dos dados para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico**

A extração, integração e visualização de dados desempenham um papel fundamental na modelagem da arquitetura do sistema de análise de dados no ambiente acadêmico. Essas etapas estão relacionadas com a obtenção, organização e representação dos dados, garantindo que as informações sejam utilizadas de forma eficaz para atender às necessidades dos usuários.

Dessa maneira, é importante destacar que a modelagem da arquitetura do sistema de análise de dados deve considerar as necessidades específicas do ambiente acadêmico, como a privacidade dos dados dos alunos, a conformidade com regulamentos e a necessidade de segurança da informação.

Além disso, a colaboração entre profissionais de TI, analistas de dados e usuários finais é fundamental para garantir que a arquitetura seja eficaz e atenda às necessidades de todas as partes interessadas, **disponível no Anexo 8, Estudos sobre Extração, integração e visualização dos dados para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico.**

## **Aplicação da Metodologia CRISP para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico**

A metodologia CRISP-DM (**C**ross-Industry **S**tandard **P**rocess for **D**ata **M**ining) é um modelo de processo amplamente utilizado para a realização de projetos de mineração de dados. Embora seja mais frequentemente aplicada em projetos de mineração de dados, pode ser adaptada e utilizada para a modelagem de arquiteturas de sistemas de análise de dados, como no caso de um sistema que extrai e disponibiliza dados de um sistema de gestão acadêmico. A seguir, as etapas do CRISP-DM e como elas podem ser aplicadas nesse contexto específico:

- **Entendimento do negócio:** Nesta etapa, é fundamental compreender o contexto acadêmico, as necessidades dos usuários e os objetivos que a arquitetura do sistema de análise de dados pretende atingir (i.e. entrevistas com os stakeholders e a análise dos requisitos para garantir que a solução atenda às necessidades identificadas).

- **Entendimento dos dados:** Nesta etapa, é importante realizar uma análise dos dados disponíveis no sistema de gestão acadêmico. Dessa forma, é necessário entender a estrutura dos dados, sua qualidade, quais informações potencialmente relevantes podem ser extraídas e como esses dados se relacionam com outros sistemas acadêmicos.
- **Preparação dos dados:** Com base na análise dos dados realizada na etapa anterior, é necessário preparar os dados para extração e integração no novo sistema (i.e. pode envolver a realização de limpeza, padronização, transformação e até mesmo enriquecimento dos dados, a fim de garantir sua qualidade e consistência).
- **Modelagem de arquitetura:** Nesta etapa, a modelagem da arquitetura do sistema de análise de dados é definida (i.e. importante identificar as tecnologias e ferramentas necessárias para a extração dos dados do sistema, como a criação de APIs ou a realização de consultas diretamente ao banco de dados). Além disso, é necessário planejar a integração dos dados coletados em um novo sistema centralizado, que pode envolver a implementação de uma camada de data warehouse e a utilização de técnicas para garantir a segurança e privacidade dos dados.
- **Implementação:** Após a definição da arquitetura, é realizada a implementação do sistema de extração e disponibilização dos dados (i.e. inclui a codificação das funcionalidades necessárias, a configuração das ferramentas e tecnologias selecionadas e a integração com o sistema de gestão acadêmico).
- **Avaliação:** Nesta etapa, a arquitetura do sistema é testada e avaliada para garantir que atenda aos requisitos definidos no início do projeto (i.e. testes de funcionalidade, desempenho e usabilidade). Dessa forma, os feedbacks dos usuários e os ajustes necessários são incorporados nesta fase.
- **Implantação:** Após a avaliação e ajustes, a arquitetura do sistema é implantada em ambiente de produção. Nesta etapa, é importante realizar o monitoramento do sistema para garantir sua estabilidade e eficácia.
- **Manutenção:** Após a implantação, é necessário garantir a manutenção contínua do sistema, com a realização de atualizações, correção de possíveis erros e evolução conforme novos requisitos e necessidades surgirem.

Aplicar a metodologia CRISP-DM contribui para garantir uma abordagem sistemática e eficiente para a modelagem da arquitetura do sistema de análise de dados no ambiente acadêmico, permitindo a extração e disponibilização de dados de um sistema de gestão acadêmico de forma estruturada, confiável e eficiente, conforme pontuado em [9] e disponível no **Anexo 10, Aplicação da Metodologia CRISP para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico.**

**Modelagem da Arquitetura do Sistema de Análise de Dados no  
Ambiente Acadêmico**

**Sprint 7, Semana 11 e 12: Validação da  
Modelagem da Arquitetura do Sistema de Análise de  
Dados no Ambiente Acadêmico**

**Validação da Modelagem da Arquitetura do Sistema de Análise de  
Dados no Ambiente Acadêmico**

**Sprint 8, Semana 13, 14, 15 e 16: Documentação  
da Arquitetura do Sistema de Análise de Dados no  
Ambiente Acadêmico**

**Documentação da Arquitetura do Sistema de Análise de Dados no  
Ambiente Acadêmico**

# Apêndices

- Anexo 1, Matriz GUT. Disponível em: <https://docs.google.com/spreadsheets/d/1Y2JBq4tRQ-X0HWMChcOZV8yxdXDMXuX9ONOtuhBZNk/edit?usp=sharing>
- Anexo 2, Programa Geral do projeto / Programa - PGP. Disponível em: [https://docs.google.com/document/d/1NRA0N2eHDptkmTOcMQXyhpCZeIL\\_q\\_sY/edit?usp=sharing&ouid=117524354183475428983&rtpof=true&sd=true](https://docs.google.com/document/d/1NRA0N2eHDptkmTOcMQXyhpCZeIL_q_sY/edit?usp=sharing&ouid=117524354183475428983&rtpof=true&sd=true)
- Anexo 3, Estrutura Analítica do Projeto. Disponível em: [https://drive.google.com/file/d/1COIRK60qbV2EvNj3Z--hTbisF\\_uoVmv\\_/view?usp=drive\\_link](https://drive.google.com/file/d/1COIRK60qbV2EvNj3Z--hTbisF_uoVmv_/view?usp=drive_link)
- Anexo 4, Gráfico de Gantt do Cronograma Geral do Projeto. Disponível em: [https://drive.google.com/file/d/14wojtw4Nat9U7VNP2jEs0XuPwENyEwHd/view?usp=drive\\_link](https://drive.google.com/file/d/14wojtw4Nat9U7VNP2jEs0XuPwENyEwHd/view?usp=drive_link)
- Anexo 5, Product Backlog Building - PBB. Disponível em: [https://miro.com/app/board/uXjVNNF-EZU=/?share\\_link\\_id=589692803401](https://miro.com/app/board/uXjVNNF-EZU=/?share_link_id=589692803401)
- Anexo 6, Requisitos Funcionais, não funcionais e Histórias de Usuários. Disponível em: [https://docs.google.com/document/d/1NZHIIVnMve1IPQ06NZ8IHEN\\_x9vdZ1Uzel\\_tNiE6BgC/edit?usp=drive\\_link](https://docs.google.com/document/d/1NZHIIVnMve1IPQ06NZ8IHEN_x9vdZ1Uzel_tNiE6BgC/edit?usp=drive_link)
- Anexo 7, Estudo de caso - Análise de Desempenho. Disponível em: [https://docs.google.com/document/d/1H-ghe1stbspwZkKws2CqOkQ2F0MZ0Xr5vIV5jYx7B2c/edit?usp=drive\\_link](https://docs.google.com/document/d/1H-ghe1stbspwZkKws2CqOkQ2F0MZ0Xr5vIV5jYx7B2c/edit?usp=drive_link)
- Anexo 8, Estudos sobre Extração, integração e visualização dos dados para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico. Disponível em: [https://docs.google.com/document/d/1fw20uuXIAvFCTI5bUIpavisGRd1f6QLkEQzn3ic2kGc/edit?usp=drive\\_link](https://docs.google.com/document/d/1fw20uuXIAvFCTI5bUIpavisGRd1f6QLkEQzn3ic2kGc/edit?usp=drive_link)
- Anexo 9, Detalhamento dos Indicadores de Educação Superior. Disponível em: [https://docs.google.com/document/d/1QdVP0FxaTSXkruF4v48qIJWkSV1rIRMV/edit?usp=drive\\_link&ouid=117524354183475428983&rtpof=true&sd=true](https://docs.google.com/document/d/1QdVP0FxaTSXkruF4v48qIJWkSV1rIRMV/edit?usp=drive_link&ouid=117524354183475428983&rtpof=true&sd=true)
- Anexo 10, Aplicação da Metodologia CRISP para a Modelagem da Arquitetura do Sistema de Análise de Dados no Ambiente Acadêmico. Disponível em: [https://docs.google.com/document/d/1CIAuJvIHbZ\\_iMPnjjc7mgiKdYh5GYMKMxIfJK06rOos/edit?usp=drive\\_link](https://docs.google.com/document/d/1CIAuJvIHbZ_iMPnjjc7mgiKdYh5GYMKMxIfJK06rOos/edit?usp=drive_link)
- Visão geral das atividades (Projects GitHub). Disponível em: <https://github.com/Tallesedu/projects/4/views/1>

# Referências Bibliográficas

- [1] Anexo 2, Imersão Inicial. Disponível em: [https://drive.google.com/file/d/1x9ZaYNt4q0-ltcjCcM2f3r-iOGdKRGxv/view?usp=drive\\_link](https://drive.google.com/file/d/1x9ZaYNt4q0-ltcjCcM2f3r-iOGdKRGxv/view?usp=drive_link)
- [2] Indicadores de Fluxo da Educação Superior, INEP. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadores-educacionais/indicadores-de-fluxo-da-educacao-superior>
- [3] Plataforma Analisa - UFG. Disponível em: <https://analisa.ufg.br/>
- [4] Business Model Canvas: como construir seu modelo de negócio?. SEBRAE, 2021. Disponível em: <https://digital.sebraers.com.br/blog/estrategia/business-model-canvas-como-construir-seu-modelo-de-negocio/>
- [5] Matriz GUT: entenda o que é e como aplicá-la na priorização dos seus projetos. EUAX, 2019. Disponível em: <https://www.euax.com.br/2019/04/matriz-gut/>
- [6] Como construir Backlog com PBB. LAMBDA3, 2022. Disponível em: <https://www.lambda3.com.br/2022/06/como-construir-backlog-com-pbb/>
- [7] PEREIRA, L. DE T. K.; GODOY, D. M. A.; TERÇARIOL, D.. Estudo de caso como procedimento de pesquisa científica: reflexão a partir da clínica fonoaudiológica. Psicologia: Reflexão e Crítica, v. 22, n. 3, p. 422–429, 2009. Disponível em: <https://doi.org/10.1590/S0102-79722009000300013>
- [8] Ronald Lopes Oliveira; Mateus Panizzon. O papel dos sistemas de apoio à decisão para análise dos custos e planejamento de novos cursos: o caso de uma universidade. Disponível em: <https://anaiscbc.emnuvens.com.br/anais/article/view/1004>
- [9] Introdução ao CRISP-DM. IBM, 2023. Disponível em: <https://www.ibm.com/docs/pt-br/spss-modeler/18.4.0?topic=guide-introduction-crisp-dm>