

AI30 - Implementar a Governança de Dados na CAPES

Área Demandante: NDI/DTI e NADC/DTI

Responsável pelas informações:

Hermógenes Ramos Batista Correia - Chefe de Divisão

Descrição:

Implementar as boas práticas da Governança de dados na CAPES, conforme melhores práticas adotadas no mercado, por meio da implementação das seguintes ações:

Definir políticas, procedimentos, estruturas, responsabilidades, glossários e metadados:

Definir políticas, procedimentos, estruturas, responsabilidades, glossários, metadados e monitoramento que permitam melhorar a qualidade, descoberta e entendimento dos dados resultando em melhores análises, garantindo o gerenciamento adequado dos ativos de dado.

Implementar gestão, controle e proteção de dados:

Adotar features de proteção de dados da solução Oracle, definindo procedimentos de controle e gestão.

Manter e organizar os modelos de dados:

Promover, organizar, padronizar, unificar os modelos de dados buscando a integridade dos dados, armazenando-os em repositório de acesso às Equipes de Sistemas.

Implementar os Dados-Mestres:

- a. ADD PROCESSO – Acervo de Dados Digitais do Processo de bolsa e programas (inscrição, concessão, acompanhamento e egresso de bolsas).
- b. ADD AVALIAÇÃO – Acervo de Dados Digitais do Processo de Avaliação da Pós-Graduação (APCN e SNPG);
- c. Disponibilizar dados históricos.

Implementar a Qualidade dos dados.

Implementar a gestão de arquitetura dos dados.

Implementar a gestão de configuração e mudança.

Implementar a gestão de metadados.

Implementar a gestão Data Warehouses.

Implementar a gestão do ciclo de vida dos dados.

Requisitos Preliminares:

1. Definir políticas, procedimentos, estruturas, responsabilidades, glossários e metadados.

- Entender os princípios que norteiam os dados na CAPES.
 - Determinar a política que permitirá o cumprimento dos princípios definidos.
 - Identificar papéis e responsabilidades no cuidado dos ativos de dados.
 - Construir o glossário e metadados que definam e eliminem ambiguidades e unifiquem vocabulário.
 - Definir a forma de implantação da Governança de Dados.
2. Implementar gestão, controle e proteção de dados.
- Definir procedimentos de controle e gestão.
 - Estudar features oferecidas pela Oracle.
 - Definir features a serem empregadas.
 - Realizar treinamento nas features selecionadas.
 - Implantar produtos de proteção de dados.
 - Monitorar resultado dos produtos de proteção.
3. Manter e organizar os modelos de dados
- Definir o repositório para armazenamento dos modelos.
 - Disseminar as normas para geração de modelo de dados.
 - Treinar as Equipes de Sistemas na utilização da ferramenta de modelagem de dados.
 - Definir os processos para criação de modelo de dados.
 - Manter modelos atualizados no repositório.
4. Implementar o Dado-Mestre ADD PROCESSO e ADD AVALIAÇÃO.
- Disponibilizar os ADDs para todas as diretorias da CAPES, contendo dados higienizados referentes a inscrição, concessão, acompanhamento e egresso de bolsas.
 - Assegurar que esta seja a única fonte de dados para todas as consultas referentes a este assunto.
5. Implementar a Qualidade de dados dos sistemas da CAPES (AE4 PDTIC 2017-2019).
- Disponibilizar dados precisos, corretos, consistentes, completos e atualizados.
 - “Realizar a limpeza da base e a revisão das entradas de diversos sistemas da CAPES, permitindo às diretorias que trabalhem com informações gerenciais e estatísticas confiáveis”. (AE4 PDTIC 2017-2019).
 - Executar processos que busquem garantir que os dados armazenados sejam corretos, precisos, consistentes, completos e integrados.
 - Gerar indicadores de qualidade de dados com métricas a serem definidas pelos Cientistas de Dados (estatísticos) com o objetivo de avaliar o grau de correção, precisão, consistência, completude, integração e aderência às regras de negócio e aos domínios estabelecidos.

Neste processo de qualidade de dados serão executadas as seguintes atividades:

- a) Carregar dados provenientes dos sistemas transacionais para a tabela Source;
 - b) Estudar os dados, utilizando ferramentas SAS / Data Quality;
 - c) Analisar os dados, a partir das estatísticas fornecidas pelas ferramentas;
 - d) Aplicar regras de transformação - ETL (Extração, Carga e Transformação);
 - e) Gerar base de alertas para Sistemas e Gestores de Negócio;
 - f) Gerar relatório de rastreabilidade da transformação (visibilidade das transformações do dado de origem);
 - g) Validar base tratada – Validação da base realizada pela área gestora, e caso haja necessidade de ajustes, retorna à atividade n.4 (Aplicar regras de transformação – ETL).
6. Implementar a gestão de arquitetura dos dados.
 - Trata-se de analisar os requisitos dos dados, implantar o seu modelo de dados, definir como será a manutenção destes modelos de dados, projetar estruturas de bancos de dados para suportar suas necessidades, projetar como será o versionamento e integração de dados e modelo de dados, projetar planos de testes, projetar planos de migração entre outras atividades.
 7. Implementar a gestão de configuração e mudança
 - Disponibilizar, na WIKI, informações sobre os processos utilizados na gestão de configuração e mudança do NDI – Núcleo de Disseminação de Informações.
 8. Implementar a gestão de metadados.
 - Gestão dos dados presentes nos ADDs para reduzir a redundância e garantir melhor qualidade dos dados através da definição padrões e usos dos valores dos dados.
 9. Implementar a gestão Data Warehouses.
 - Projetar modelos de dados que permitam a geração de informações para tomada de decisão sob várias perspectivas (dimensões).
 10. Implementar a gestão do ciclo de vida dos dados
 - Gerenciar o ciclo de vida dos dados, em conformidade com a LGPD – Lei Geral de Proteção de Dados, artigo 5º, assegurando:
 - Coleta dos dados;
 - Processamento dos dados;
 - Análise dos dados;
 - Compartilhamento dos dados;
 - Armazenamento dos dados;
 - Reutilização dos dados;
 - Eliminação dos dados.

Expectativa de Prazo (semestre/ano):

Assinalar uma única opção para INÍCIO e outra para TÉRMINO

| | INÍCIO | | | TÉRMINO | |
|------|-------------------|--------------|--|--------------|-------------------|
| 2020 | 1º SEMESTRE: X | 2º SEMESTRE: | | 1º SEMESTRE: | 2º SEMESTRE: |
| 2021 | 1º SEMESTRE: | 2º SEMESTRE: | | 1º SEMESTRE: | 2º SEMESTRE: |
| 2022 | 1º SEMESTRE: | 2º SEMESTRE: | | 1º SEMESTRE: | 2º SEMESTRE: |
| 2023 | 1º SEMESTRE: | 2º SEMESTRE: | | 1º SEMESTRE: | 2º SEMESTRE: X |

Metas:

| Ano | Meta |
|------|---|
| 2020 | <p><u>1º semestre:</u> Definição de princípios, políticas, glossário e metadados, forma de implantação, procedimentos de controle e gestão, definição do Dado Mestre e disponibilização do ADD PROCESSO para todas as diretorias, definição do repositório para armazenamento dos modelos, disseminação das normas para geração de modelo de dados e treinamento das equipes de sistemas.</p> <p><u>2º semestre:</u> Definição do Dado Mestre e disponibilização do ADD AVALIAÇÃO, incluindo a integração de dados do APCN, para todas as diretorias e definição e treinamento de features a serem adotadas. Definição dos processos para criação dos modelos de dados.</p> |
| 2021 | <p><u>1º semestre:</u> Definição e implementação da gestão de configuração e mudança (WIKI), produtos de proteção de dados implantados, avaliação dos modelos armazenados no repositório de modelos de dados.</p> <p><u>2º semestre:</u> Definição e implementação da gestão de metadados, monitoramento dos produtos de proteção de dados implantados,</p> |



| | |
|------|---|
| | avaliação dos modelos armazenados no repositório de modelos de dados. |
| 2022 | <p><u>1º semestre:</u> Implementação do processo de qualidade de dados, monitoramento dos produtos de proteção de dados implantados, avaliação dos modelos armazenados no repositório de modelos de dados.</p> <p><u>2º semestre:</u> Monitoramento dos produtos de proteção de dados implantados, avaliação dos modelos armazenados no repositório de modelos de dados.</p> |
| 2023 | <p><u>1º semestre:</u> Avaliar os processos de governança de dados implementados, monitoramento dos produtos de proteção de dados implantados, avaliação dos modelos armazenados no repositório de modelos de dados.</p> <p><u>2º semestre:</u> Avaliar os processos de governança de dados implementados, monitoramento dos produtos de proteção de dados implantados, avaliação dos modelos armazenados no repositório de modelos de dados.</p> |

Observações Técnicas, Ações Internas necessárias ou Necessidade de Aquisição/Melhorias:

Unidade DTI Responsável: DTI

Subunidade DTI Responsável: NDI

[NDI] - Equipes compostas de 2 gerentes de projetos, 1 Gerente de Banco de Dados, 2 analistas OLAP, 5 analistas de ETL, 4 DBA(s), 4 AD(s) e 1 Auxiliar Administrativo.