

MANUAL DE BOAS PRÁTICAS NA CATALOGAÇÃO DE DADOS

Gabriel Nascimento
Yago Rusef

DEZEMBRO de

2024

GABRIEL NASCIMENTO

YAGO RUSEF

Manual de boas práticas na catalogação de dados

Projeto de conclusão de disciplina

ACH-3778 Governo Aberto

Docente: Gisele da Silva Craveiro

São Paulo – SP

2024

Sumário

| | |
|---|-----------|
| INTRODUÇÃO | 4 |
| ETAPA PRÉVIA..... | 5 |
| 1. Compreensão dos Problemas e Desafios | 5 |
| ETAPA INICIAL..... | 7 |
| 2. Identificação e Levantamento de Fontes de Dados | 7 |
| 3. Padronização dos Dados | 9 |
| 4. Privacidade e Proteção de Dados Pessoais | 11 |
| 5. Estruturação do Catálogo | 12 |
| 6. Acessibilidade e Disponibilidade..... | 13 |
| 7. Não discriminação | 14 |
| ETAPA INTERMEDIÁRIA | 15 |
| 8. Atualização Contínua | 15 |
| 9. Reutilização de Dados | 17 |
| ETAPA AVANÇADA..... | 19 |
| 10. Ferramentas de Visualização e Interatividade | 19 |
| REFERENCIAS BIBLIOGRAFICAS | 22 |

INTRODUÇÃO

A transparência de dados públicos é um pilar essencial para fortalecer a confiança entre a administração pública e a sociedade. Ao disponibilizar informações de forma acessível, organizada e compreensível, governos promovem uma cultura de responsabilidade, possibilitando que cidadãos, organizações e instituições monitorem ações públicas, colaborem na formulação de políticas e impulsionam inovações sociais e econômicas.

No entanto, para que a transparência seja eficaz, é necessário superar desafios relacionados à coleta, padronização, proteção e disponibilização de dados. Um catálogo de dados bem estruturado não apenas facilita o acesso às informações, mas também garante sua qualidade, confiabilidade e relevância.

O presente documento aborda dez pontos fundamentais para a catalogação de dados, organizados de acordo com boas práticas classificadas em etapas de prioridade: prévia, inicial, intermediária e avançada. Dessa forma, a estrutura segue uma progressão linear, permitindo que, após a implementação das recomendações de uma etapa, seja possível avançar de forma consistente para as etapas seguintes, garantindo uma abordagem sólida e eficiente.

ETAPA PRÉVIA

1. Compreensão dos Problemas e Desafios

A construção de um catálogo de dados públicos requer a compreensão de como os dados são gerados, armazenados e gerenciados. Esse processo pode ser conduzido por meio da designação de representantes de todas as secretarias, que serão responsáveis por realizar um levantamento interno dessas informações em colaboração com os demais integrantes. Esses representantes, por sua vez, podem debater as metas e desafios do projeto se baseando em metodologias de projeto, como o Design Thinking. Com base nesse trabalho, será possível conduzir uma análise abrangente e integrada de todos os setores.

No contexto da prefeitura de São Vicente, vemos uma prefeitura em fase bastante embrionária no que tange à disponibilidade de dados. Grande parte dos dados estão em formato físico, e os que estão digitalizados ainda estão bastante desorganizados e não normalizados, fazendo com que as informações sejam de difícil acesso e controle. Conceitos de estruturas de gestão de dados, como Data Warehouses, Data Lakes e Data Mesh ajudam a organizar e facilitar o acesso aos dados, além de garantir uma base sólida para futuras integrações.

- **Data Warehouse:** Acordando às definições da AWS (Amazon Web Services), um dos maiores provedores de informações sobre dados do mundo, o Data Warehouse é um repositório centralizado que organiza dados estruturados para análises rápidas e consistentes. É útil para consolidar informações financeiras e tributárias.
- **Data Lake:** Um repositório mais flexível que armazena dados estruturados e não estruturados, permitindo que os dados originais sejam mantidos para processamento futuro.

ETAPA PRÉVIA

- Data Mesh: Uma abordagem descentralizada onde os dados são organizados por domínios (como saúde, educação ou infraestrutura), permitindo que diferentes departamentos mantenham controle sobre suas informações enquanto seguem padrões comuns.

O objetivo dessa etapa é preparar a base tecnológica e conceitual para que os dados da cidade de São Vicente sejam digitalizados, organizados, estruturados e estejam prontos para integração, reduzindo problemas de fragmentação e redundância.

Ademais, outro conceito bastante importante no que tange à acessibilidade de dados são as API's (Interface de Programação de Aplicação), que são, acordando à AWS, artifícios que permitem a comunicação entre dois softwares usando um conjunto de protocolos, sendo feito, normalmente, de forma automatizada.

ETAPA INICIAL

2. Identificação e Levantamento de Fontes de Dados

O primeiro desafio para a Prefeitura de São Vicente é mapear todas as fontes de dados existentes, que ainda são predominantemente físicas. A digitalização dessas fontes é crucial para garantir a coleta de dados de forma eficaz. Neste contexto, a governança de dados desempenha um papel fundamental. A gestão dos dados deve ser descentralizada, mas com padrões claros de organização, a fim de garantir a integridade e o controle da informação à medida que ela é transferida para formatos digitais. Para isso, ferramentas como o AWS Glue, ou outros softwares que trabalham ETL e ELT, como o Alteryx Designer e o Apache NiFi, como uma opção aberta e bastante intuitiva por conta da sua interatividade gráfica, podem ser úteis, pois permitem a integração de dados de diferentes fontes, mesmo que estejam em formatos variados, além de ajudar a categorizar esses dados à medida que são digitalizados. A catalogação dos dados, portanto, começa com a identificação dos repositórios e a digitalização de documentos físicos, muitas vezes em um formato não estruturado.

Nessa etapa, é essencial entender conceitos relativos a ETL (Extract Transform Load) e ELT (Extract Load Transform). No ETL, os dados são primeiro extraídos das fontes, transformados para um formato adequado e, em seguida, carregados no repositório de destino. Esse processo é comum quando há necessidade de limpar ou formatar os dados antes de armazená-los. Já no ELT, a extração e o carregamento são feitos primeiro, e a transformação ocorre depois, diretamente no repositório de dados, o que permite que grandes volumes de dados sejam processados mais rapidamente e com mais flexibilidade. Ambos os métodos têm suas vantagens, sendo o ETL mais utilizado em sistemas tradicionais e o ELT mais adequado para ambientes de Big Data, onde o poder de processamento do repositório é aproveitado para a transformação dos dados.

Dado que temos o domínio dos conceitos necessários para o início da digitalização dos dados, uma boa maneira de iniciar o projeto é utilizando do sistema de ETL com interfaces do Apache NiFi, que é uma plataforma aberta e gratuita, com uma documentação precisa e uma grande comunidade. Com seu sistema de interfaces, o Apache NiFi facilita bastante o manuseio do fluxo de dados, permitindo fazer transformações na formatação dos dados e manipulá-los sem a necessidade de um conhecimento em programação, o que facilita bastante a manutenção e o desenvolvimento dos dados. Com isso, teremos uma boa base para seguir com as próximas etapas, como a normalização dos dados, que se baseiam fortemente no processo de transformação de dados.

ETAPA INICIAL

3. Padronização dos Dados

Com os dados digitais mapeados, o próximo desafio para a prefeitura é a padronização e normalização dessas informações. Muitos dados podem estar em formatos incompatíveis ou sem uniformidade em relação a valores, datas e unidades de medida. A normalização de dados é necessária para garantir que os dados sigam um padrão único, tornando-os utilizáveis e consistentes. Por exemplo, se as informações de contato telefônico estiverem em formatos diferentes ((XX) XXXXX-XXXX ou XXXXXXXXXXXX), é fundamental padronizá-las para o mesmo formato. Além disso, a prefeitura pode adotar formatos abertos, como CSV ou JSON, para garantir que os dados sejam facilmente acessíveis e reutilizáveis. O uso de plataformas como Data Lakes pode ser uma solução interessante, pois elas permitem armazenar dados brutos e não estruturados antes de serem processados e organizados.

Para realizar as transformações dos dados, podemos utilizar a ferramenta supracitada do Apache NiFi, onde o fluxo de dados pode ser facilmente interpretado pela interatividade por interfaces gráficas, onde o usuário “clica e arrasta” os blocos de operações para cada parte específica do fluxo de dados, por exemplo, permitindo que números de CPF’s sejam padronizados em um formato como 123.456.789-10, ao invés de 12345678910.

Além dos dados relativos a telefone e CPF, outros dados comumente problemáticos no contexto de dados são os de:

- **Endereço:** Comumente ocorrem diferenças de escritas para a mesma referência, como “Av. Rebouças, 123” e “Avenida Rebouças, 123”. Esses dados não são problemáticos em contexto de uso humano, mas quando aplicados a uma análise automatizada, a diferenciação entre os dados é danosa.

- Nomes: Ocorre em diversos casos a abreviação de alguns sobrenomes ou mesmo a exclusão de partes de nomes, como “João Pinto da Silva” sendo tratado como “João P. Silva”. Novamente, os dados representam a mesma pessoa, mas num contexto de automatização são problemáticos.
- Datas: Num contexto de dados, normalmente vemos diversas opções diferentes para formatação de datas, como DD-MM-YYYY, DD/MM/YYYY, YYYY-MM-DD, entre diversas outras. O uso da ferramenta de padronização do Apache NiFi seria de grande importância para a manutenção do projeto.

Vale lembrar que dados em XLXS, o formato padrão do Microsoft Excel, não são os mais adequados no que tange à disponibilidade de dados, pois não é um formato aberto, como o JSON, o CSV ou mesmo arquivos de planilha ODS, que pertence ao software aberto Libreoffice. Ademais, o uso de Data Lakes é interessante num primeiro momento, mas no longo prazo, pode se mostrar uma escolha caótica, principalmente, por conta do grande volume de dados presentes na prefeitura. Com isso, uma possível recomendação seria a implementação temporária de Data Lakes, para que os dados sejam organizados dentro de Data Meshes, sendo uma opção mais robusta, escalável e organizada

ETAPA INICIAL

4. Privacidade e Proteção de Dados Pessoais

A Lei Geral de Proteção de Dados (LGPD) exige que a prefeitura tome medidas rigorosas para garantir que dados pessoais sejam protegidos e não utilizados de forma indevida. No contexto de digitalização, isso envolve garantir que dados sensíveis, como informações de contato, saúde ou educação, sejam anonimizados sempre que possível. A anonimização é uma prática que transforma dados pessoais em informações irreconhecíveis, assegurando a privacidade dos cidadãos. A utilização de AWS Macie, ou do software Amnesia, desenvolvida pelo projeto openAIRE, como uma alternativa aberta e gratuita, por exemplo, pode ajudar a detectar dados sensíveis durante a digitalização e aplicar técnicas de anonimização automaticamente. A conformidade legal também inclui a implementação de controle de acessos rigorosos, para garantir que apenas pessoas autorizadas possam acessar informações pessoais.

ETAPA INICIAL

5. Estruturação do Catálogo

A criação de um catálogo de dados eficaz é vital para garantir que os dados públicos possam ser facilmente acessados e utilizados. Isso requer a modelagem de dados e uma organização coerente que permita aos usuários encontrarem as informações de forma intuitiva. Para a Prefeitura de São Vicente, a categorização dos dados pode ser feita com base em temas como:

- saúde
- educação
- segurança pública
- infraestrutura
- Meio-Ambiente

A governança de dados aqui entra como um conceito chave para garantir que diferentes departamentos (saúde, educação, etc.) mantenham a responsabilidade e o controle sobre suas próprias informações, mas dentro de um esquema comum de catalogação e metadados. O uso de Data Mesh, por exemplo, pode ser uma abordagem interessante, onde cada setor tem sua própria governança, mas todos seguem padrões que permitem integração.

Nessa etapa, sobretudo, teremos em grande parte a atuação do “L” de ETL, onde, após os processos de normalização e anonimização, deveremos começar a organizar os dados em seus devidos locais, garantindo que cada área tenha uma boa governança, trazendo uma maior escalabilidade e eficiência para cada parte do Data Mesh, onde cada servidor tem acesso aos dados relativos à sua área, sem a interferência de dados de outros setores indesejados para aquele dado momento.

ETAPA INICIAL

6. Acessibilidade e Disponibilidade

Para a Prefeitura de São Vicente, garantir a acessibilidade e a disponibilidade dos dados é um passo essencial para promover a transparência. A criação de uma plataforma centralizada, similar ao portal da transparência, onde os cidadãos possam acessar facilmente os dados, é uma necessidade. Utilizando APIs abertas, como a API do IBGE, que será melhor apresentada mais a frente, os dados podem ser disponibilizados em formatos acessíveis como CSV, JSON ou ODS, para garantir que possam ser reutilizados e manipulados por qualquer interessado. A plataforma também precisa ser escalável, para acomodar um grande volume de dados à medida que a cidade avança em seu processo de digitalização. A implementação de Data Meshes pode ser útil aqui, pois esses repositórios centralizados em seu próprio contexto (saúde, educação, etc.) permitem que os dados sejam facilmente consultados e utilizados para análise e visualização.

ETAPA INICIAL

7. Não discriminação

Por fim, a transparência e a abertura dos dados são fundamentais para que a Prefeitura de São Vicente cumpra seu papel de prestar contas à sociedade. Todos os dados públicos devem ser acessíveis sem barreiras, ou seja, sem exigir autenticação ou obstáculos desnecessários. A documentação deve ser clara e fornecer explicações sobre a origem dos dados, a metodologia utilizada na coleta e possíveis limitações. O uso de proveniência de dados (data provenance) ajuda a garantir que os cidadãos compreendam a origem dos dados, a confiabilidade da informação e as limitações do conjunto de dados. A plataforma deve permitir que os cidadãos acessem dados de diferentes áreas, como saúde, educação, transporte e segurança, com total transparência sobre sua origem e utilização. O uso de uma plataforma de governança de dados, explicando cada atributo de seus dados e documentando alguns metadados como período de atualização e como os dados são capturados, é essencial para a correta execução dessa tarefa.

Porém, apesar da necessidade de transparência e abertura, não podemos nos esquecer das grandes necessidades relativas a anonimização dos dados pessoais, visando estar acordando a LGPD.

ETAPA INTERMEDIÁRIA

8. Atualização Contínua

Manter a atualização contínua dos dados é essencial para assegurar a relevância, a confiabilidade e a utilidade das informações para os cidadãos que as consomem. Dados desatualizados podem gerar impactos negativos, como decisões inadequadas, perda de confiança nos sistemas e redução da eficiência organizacional. Para manter essa boa prática, devem ser seguidas as duas diretrizes

Frequência de atualização:

Devem ser estabelecidas e divulgadas as frequências de atualização dos dados, alinhando-a às necessidades dos usuários e aos objetivos organizacionais, para que as informações permaneçam relevantes e úteis para os cidadãos e partes interessadas.

Segundo o Guia para o corpo de conhecimento em gerenciamento de dados (DAMA-DMBOK), uma governança de dados eficiente exige que os dados sejam tratados como ativos estratégicos. Isso inclui definir padrões de atualização baseados na criticidade e na volatilidade das informações. Dados que não são atualizados regularmente perdem sua utilidade, especialmente em setores como saúde, segurança pública ou transporte, onde decisões rápidas e baseadas em dados atualizados são críticas. Estabelecer uma frequência formalizada de atualização promove transparência e permite que os cidadãos saibam quando esperar novas versões dos dados, aumentando a confiança no órgão responsável. Por exemplo, dados ambientais podem ser atualizados mensalmente, enquanto informações de transporte público podem exigir atualizações diárias ou em tempo real. Essa categorização deve ser comunicada claramente aos cidadãos para atender suas expectativas.

ETAPA INTERMEDIÁRIA

Processos automatizados

Implementar processos automatizados para garantir que os dados sejam atualizados regularmente, reduzindo a possibilidade de falhas humanas e inconsistências nos registros.

Conforme descrito no DAMA-DMBOK, a automação é uma das principais ferramentas para mitigar riscos relacionados à atualização manual, como erros humanos, atrasos e falta de padronização. A automação também é essencial para lidar com volumes crescentes de dados em órgãos públicos, permitindo que os sistemas funcionem com eficiência em tempo real.

Deve se utilizar Ferramentas de integração contínua, como ETL e ELT, descritas no capítulo 1, utilizam um conjunto estruturado de regras para processar dados provenientes de diversas fontes antes de sua integração em um repositório centralizado. Essas ferramentas garantem que os dados sejam filtrados, classificados e limpos, eliminando inconsistências e redundâncias. Como resultado, as informações são transmitidas e processadas de maneira contínua e eficiente, assegurando alta consistência, qualidade e utilidade.

Além disso, processos automatizados permitem monitoramento em tempo real, com alertas para sinalizar eventuais problemas, como falhas de atualização ou inconsistências nos dados, promovendo intervenções rápidas e precisas para tomada de decisão.

Um exemplo prático pode ser encontrado em sistemas de tráfego urbano, onde sensores de trânsito enviam dados automaticamente para servidores, atualizando mapas e sistemas de navegação em questão de segundos.

Dessa forma, a combinação de processos automatizados e frequência definida de atualização atende aos princípios fundamentais de governança de dados: **qualidade, transparência e responsabilidade.**

ETAPA INTERMEDIÁRIA

9. Reutilização de Dados

Para promover a reutilização dos dados, é fundamental aplicar licenças abertas, como as da Creative Commons. Essas licenças permitem que os dados sejam reutilizados, compartilhados e modificados sem restrições, desde que atendidas as condições estabelecidas. Além disso, é importante utilizar meios abertos, conforme descrito no Capítulo 2, para facilitar a manipulação e análise dos dados em diversas aplicações. Para ilustrar os benefícios da reutilização de dados, apresentaremos um exemplo prático utilizando o serviço de dados do IBGE.

O Instituto Brasileiro de Geografia e Estatística (IBGE) disponibiliza uma documentação detalhada para o consumo de suas informações via API. Essa documentação inclui um guia passo a passo que descreve:

- Parâmetros de entrada: especificando cada parâmetro necessário e seu tipo.
- Dados de saída: detalhando o formato e as informações fornecidas.

Essa estrutura facilita o acesso e a reutilização das informações em diversas áreas de interesse, como saúde, educação e economia.

Demonstração de Reutilização de Dados Como exemplo prático, utilizamos a API do IBGE para realizar uma análise simples. O processo incluiu:

1. Consultar a API do IBGE, seguindo as orientações da documentação oficial.
2. Tratar as informações obtidas, conforme necessário.
3. Exportar os dados tratados para uma planilha.

ETAPA INTERMEDIÁRIA

Para fins de demonstração, pesquisamos a quantidade de estados por região no Brasil. A partir desse exemplo, fica evidente como os dados disponibilizados podem ser reutilizados para análises mais complexas, como dos dados das Pesquisa de Assistência Médico-Sanitária e Saúde do Escolar.

| | A | B |
|---|--------------|-----------------------|
| 1 | Regiao | Quantidade de Estados |
| 2 | Norte | 7 |
| 3 | Sul | 3 |
| 4 | Nordeste | 9 |
| 5 | Sudeste | 4 |
| 6 | Centro-Oeste | 4 |
| 7 | | |
| 8 | | |

Figura 2: Dados consumidos do IBGE

```
{
  "codigo": "AA",
  "nome": "Pesquisa Nacional de Saúde do Escolar",
  "nome_ingles": null,
  "situacao": "Ativa",
  "categoria": "Especial",
  "periodicidade_coleta": "Eventual",
  "periodicidade_divulgacao": "Eventual",
  "classificacoes_tematicas": [ ... ]
},
{
  "codigo": "AM",
  "nome": "Pesquisa de Assistência Médico-Sanitária",
  "nome_ingles": null,
  "situacao": "Ativa",
  "categoria": "Especial",
  "periodicidade_coleta": "Eventual",
  "periodicidade_divulgacao": "Eventual",
  "classificacoes_tematicas": [ ... ]
}
```

Figura 1: dados de pesquisa disponibilizados em formato aberto (JSON)

```
async function getData() {
  const response = await axios.get('https://servicodados.ibge.gov.br/api/v1/localidades/estados/');
  formatData(response.data);
} catch (error) {
  console.error(error);
}

function formatData(data) { ... }

function csvWriter(data) {
  const createCsvWriter = require('csv-writer').createObjectCsvWriter;
  const csvWriter = createCsvWriter({
    path: 'csv/estados.csv',
    header: [
      {id: 'regiao', title: 'Regiao'},
      {id: 'qtd', title: 'Quantidade de Estados'}
    ]
  });
  const records = [
    {regiao: 'Norte', qtd: data[0]},
    {regiao: 'Sul', qtd: data[3]},
    {regiao: 'Nordeste', qtd: data[1]},
    {regiao: 'Sudeste', qtd: data[2]},
    {regiao: 'Centro-Oeste', qtd: data[4]}
  ];
  csvWriter.writeRecords(records) // returns a promise
    .then(() => {
      console.log('Done');
    });
}
```

Figura 3: Algoritmo (JS) simples de mineração de dados

Com base nessa aplicação prática do serviço do IBGE, observa-se que sua facilidade de uso, a ampla documentação disponível e as diversas possibilidades de análise evidenciam os ganhos em confiabilidade e transparência proporcionados pela adoção da boa prática de reutilização de dados.

ETAPA AVANÇADA

10. Ferramentas de Visualização e Interatividade

A inclusão de ferramentas de visualização e interatividade é uma prática fundamental para democratizar o acesso e a compreensão de dados. De acordo com o DAMA-DMBOK, a disponibilização de dados deve considerar a usabilidade e a acessibilidade para diferentes públicos, promovendo transparência e engajamento. Abaixo, destacam-se os principais pontos relacionados a esta abordagem que devem ser implementados para cumprimento dessa boa prática:

Gráficos e dashboards: Disponibilizar ferramentas de visualização, como gráficos interativos e dashboards, que ajudem os cidadãos a entender e explorar os dados de maneira intuitiva e acessível, trazendo benefícios como:

- Visualizações interativas transformam dados brutos em informações compreensíveis, reduzindo barreiras técnicas para o público em geral.
- Gráficos e dashboards promovem a "inteligibilidade dos dados", que é crucial para apoiar a tomada de decisões e aumentar a transparência pública.
- Ferramentas como essas permitem análises rápidas, destacando tendências e padrões que poderiam passar despercebidos em planilhas ou relatórios estáticos.

Filtros e análises: Incluir filtros que permitam aos cidadãos personalizar a visualização dos dados, possibilitando análises específicas e comparações que atendam a diferentes necessidades informativas.

O DAMA-DMBOK enfatiza que a governança de dados deve priorizar a "flexibilidade no uso dos dados", permitindo que cada usuário extraia insights de acordo com suas necessidades específicas. Ferramentas de filtragem permitem explorar dados em níveis granulares ou agregados, adaptando a visualização

ETAPA AVANÇADA

para diferentes finalidades, como estudos acadêmicos, monitoramento de políticas públicas ou uso empresarial.

Exemplos práticos incluem portais como o **Portal da Transparência do Estado de São Paulo**, que disponibiliza análises multivariadas, oferecendo uma visão global dos dados e, ao mesmo tempo, detalhando itens individuais, como despesas por órgão e valores pagos ao longo dos anos. No exemplo, a funcionalidade de filtros possibilita a análise detalhada de cada item de despesa, permitindo o cruzamento de informações, como valores por ano e órgão pagador.



Figura 4: portal da transparência do estado de São Paulo

Canal de feedback: Oferecer meios para que cidadãos e organizações possam sugerir melhorias nos dados e no catálogo, promovendo colaboração e alinhamento às demandas reais.

ETAPA AVANÇADA

Segundo o DAMA-DMBOK, o feedback contínuo dos usuários é uma das formas mais eficazes de melhorar a qualidade dos dados e sua apresentação. Um canal de feedback bem estruturado, como o existente no **Portal da Transparência do Estado de São Paulo**, ajuda a identificar lacunas e atender melhor às expectativas dos usuários. Perguntas como "Você conseguiu encontrar a informação desejada?" e "Tem alguma sugestão de melhoria?" promovem uma cultura de melhoria contínua.

Este mecanismo também fomenta a transparência organizacional, pois demonstra o comprometimento com a evolução e a adequação do sistema às necessidades da população.



PORTAL DA TRANSPARÊNCIA
Poder Executivo do Estado de São Paulo

SOBRE O PORTAL • LEGISLAÇÃO • CANAIS DE COMUNICAÇÃO • LEI DE ACESSO À INFORMAÇÃO • PAINÉIS • LINKS EXTERNOS

MELHORADO PELO Google

Avaliação do Portal da Transparência

☐ Prefiro ficar anônimo

Nome

E-mail

1) Qual a sua avaliação do Portal da Transparência do Estado de São Paulo?

☐ Ótimo

Participação cidadã: Estimular a participação ativa na criação e revisão de dados, por meio de consultas públicas, fóruns de discussão ou enquetes, garantindo que as informações reflitam as necessidades reais da população.

O DAMA-DMBOK aponta que o engajamento das partes interessadas é essencial para alinhar a gestão de dados aos objetivos organizacionais e sociais. A participação cidadã cria um senso de pertencimento, aumentando a confiança nas políticas públicas e a adesão a iniciativas governamentais baseadas em dados. Consultas públicas ou fóruns, além de promoverem inclusão, geram ideias valiosas sobre a percepção e a usabilidade dos dados.

REFERENCIAS BIBLIOGRAFICAS

DAMA. Guia para corpo de conhecimento em gerenciamento de dados. 1 edição. Technics Publications, 2010.

Controladoria geral da união. Guia de transparência ativa: para órgãos e entidades do poder executivo federal. Gov.br, 2022. Disponível em: (<https://www.gov.br/acessoainformacao/pt-br/central-de-conteudo/publicacoes/gta-7-guia-de-transparencia-ativa-final.pdf>). Acesso em: (25/11/2024).

Controladoria geral do município. CMBD - Catálogo Municipal de Bases de Dados. Dados Abertos da prefeitura de São Paulo, 2024. Disponível em: (http://dados.prefeitura.sp.gov.br/dataset/e9b8a70e-aeec-48ee-8da3-641697fd4bde/resource/742e0594-e513-4f3b-8788-7886bde77c28/download/relatorio_cmbd_2024-vfinal.pdf). Acesso em: (25/11/2024).

Ministério da Economia. Guia de padronização de informações das empresas Estatais Federais nos portais da internet. 3 edição. Gov.br, 2022. Disponível em: (https://www.gov.br/gestao/pt-br/assuntos/estatais/central-de-conteudo/guias-e-manuais/guia_padronizacao_informacoes_portais_internet_edicao_3_versao_9.pdf). Acesso em: (25/11/2024).