

**A GOVERNANÇA DE DADOS E O REUSO DE DADOS NA ADMINISTRAÇÃO PÚBLICA:  
UMA REVISÃO SISTEMÁTICA DE LITERATURA**

**LUCAS DE MAMAN BENZ**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

**ANTONIO CARLOS GASTAUD MAÇADA**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

Agradecimento à orgão de fomento:

Agradecimento à CAPES pelo apoio.

# A GOVERNANÇA DE DADOS E O REUSO DE DADOS NA ADMINISTRAÇÃO PÚBLICA: UMA REVISÃO SISTEMÁTICA DE LITERATURA

## 1. INTRODUÇÃO

Após a passagem da recente pandemia de COVID-19, o estudo realizado por Moser-Plautz & Schmidhuber (2023) mostra que este fator externo acelerou a transformação digital das organizações governamentais. Kitsios *et al.* (2023) apontam que estas organizações ao utilizarem dados entregam aos cidadãos produtos e serviços digitais mais eficientes. Da mesma forma, Dremel (2017), afirma que os dados ao serem entendidos como recursos valiosos podem ajudar as organizações a se tornarem digitais e alcançar vantagens competitivas. A utilização de dados também é citada por Mergel *et al.* (2019) juntamente com a abordagem *data-driven*, e estas são apontadas como recursos importantes para a transformação digital da administração pública. Portanto, devido ao relevante papel do uso dados de dados nas organizações, é necessário que as entidades governamentais entendam que uma vez permitido o acesso a estes ativos, os usuários não o farão somente com o propósito de conhecer seu conteúdo, mas também o reutilizarão para outras finalidades (Janssen, 2011).

A empresa de aconselhamento e pesquisa Gartner (2023) aponta que o reuso de dados aumenta o valor operacional do dado. A empresa de consultoria Deloitte (2021), segue o mesmo caminho, afirmando que os governos estão buscando novas formas de melhorar o compartilhamento de dados, aumentando o valor dos dados para ganhar tempo, economizar dinheiro e inclusive salvar vidas. Nesta busca, uma das tendências apontadas por esta consultoria é que as entidades governamentais têm adotado princípios de facilitação de busca, acesso, interoperabilidade e reusabilidade dos dados. Como exemplo, a McKinsey & Company (2020) apresenta o caso de uma instituição bancária que ao habilitar o reuso de dados, junto de outras ações, pôde eliminar até 20% dos custos de obtenção de dados sobre o risco de crédito e dados de mercado, o que foi estimado em cerca de 100 milhões de dólares anuais. Zhu *et al.* (2008), indicam que em um ambiente de competição entre fornecedores de dados, o reuso do dado e a cópia do conteúdo de bases de dados traz maior benefício quando se criam diferentes propósitos de utilização.

Conforme Gunther *et al.* (2022), o processo de criação de novas aplicações e geração de maior valor sobre um mesmo dado leva tempo e depende das iterações de reconstrução e redefinição de objetivos, pelos quais se ajustam os dados. Para garantir que os dados sejam utilizados de forma correta, transformando-os em um ativo da organização, é necessário que as organizações invistam em tecnologias digitais para facilitar o acesso às informações e tenham capacidade de governança de seus dados, o que potencializa a coordenação e comunicação entre diferentes times das organizações (Zhang *et al.*, 2022). Segundo Paparova *et al.* (2023), o dado transita dentro das organizações, o papel da governança de dados se desdobra, e em algumas situações o tratamento de dados estará subordinado às regras e propósitos de uso, estes chamados por Schmitt (2023) de custo diante dos dados. Outros tipos de estruturas são relatados por Ballantyne & Style (2017) que aborda a existência e a participação de comitês como camadas de governança de dados, já Cerrillo-Martínez & Casadesús-De-mingo (2021) descreve três tipos de modelos de governança de dados dentro da administração pública: centralizado, coordenado e descentralizado, assim como discorre sobre o papel do chefe de dados, porém em

sua pesquisa fica evidenciada a falta de outros trabalhos falando sobre governança de dados na administração pública.

Wang (2021) afirma que os processos de reuso de dados são pouco discutidos de maneira sistemática. Na mesma linha o estudo de Koesten *et al.* (2020), revela que ainda se sabe pouco sobre o que faz uma base de dados mais reusável e que há um gap de pesquisa entre as normativas e os detalhes operacionais para dados. Segundo Alvarez-Romero *et al.* (2023) entender como os dados são gerenciados por estruturas dedicadas é importante para facilitar o compartilhamento de dados e seu reuso, assim, o presente trabalho traz as seguintes questões de pesquisa: (1) Quais fatores estão ligados ao reuso de dados dentro da administração pública?; (2) De que forma a governança de dados vem abordando essa questão nas organizações governamentais?; e (3) Quais as principais lacunas na literatura sobre o reuso de dados no setor público?.

Para responder as questões supracitadas este estudo se divide em cinco seções, iniciando por esta introdução, após trazendo um referencial teórico sobre o reuso de dados e seu contexto. Na terceira seção temos a descrição da metodologia que subsidia a criação do conteúdo da quarta seção, que traz, a partir de 19 artigos incluídos, os temas identificados dentro da revisão sistemática de literatura. Por fim este estudo é finalizado com a discussão sobre os resultados e conclusões. É importante alertar que os assuntos pesquisados ficam dentro dos limites referentes a governança de dados ligadas a processos comerciais ou governamentais, logo o reuso de dados e sua gestão em ambientes estritamente científicos ficam excluídos da pesquisa, pois apresentam características e funcionamentos diversos ao objetivo pretendido por este trabalho. Temáticas relacionadas estritamente a questões de operacionalização de dados abertos pelo governo sem tocar os aspectos de governança da informação, também. não foram consideradas, pois acabam tendo cunho técnico.

## **2.CONCEITOS CHAVE**

Nesta seção são apresentados a fundamentação teórica e os conceitos básicos que apoiam o presente estudo.

### **2.1.REUSO DE DADOS**

O reuso de dado é caracterizado quando se utiliza os dados para propósitos distintos daquele empregado no momento de recebimento do dado de sua fonte original, que serviu para completar uma tarefa ou transação (Black *et al.*, 2023). A literatura também apresenta como sinônimo, para abordar o reuso, o termo "uso secundário do dado" (Black *et al.*, 2023; Green *et al.*, 2023; Varhol *et al.*, 2022; Ballantyne & Style, 2017; Alvarez-Romero *et al.*, 2023). Em seu trabalho, Wang *et al.* (2021) acrescenta que o reuso de dados pode ser dividido em três etapas: iniciação, coleta e exploração e criação de novo propósito. Alertado por Aleksynska & Cazes (2016), os dados que foram construídos a partir de combinações de outros dados podem carregar consigo erros de interpretação e classificação, assim sugerem que os dados podem ser utilizados para análises específicas desde que se compreendam claramente o que pode ou não ser feito com ele.

Encontram-se relatos, sobre o reuso de dados e seus benefícios em diversas áreas, como exemplo, Fu *et al.* (2021) descreve o reuso de dados facilitando processos de identificação de hipóteses, otimização de biomarcadores e estratificação de pacientes na pesquisa imunooncológica. Utilizando dados de transações financeiras, De Caigny *et al.* (2020) relata o reuso para prever eventos da vida de clientes, como nascimento de filhos, separação, assim como, Kitchens *et al.* (2018), traz o exemplo do trabalho de identificação de diferentes dados, que através do uso secundário poderiam fazer a organização entender melhor seus clientes e obter vantagens competitivas. No turismo, Miah *et al.* (2017) traz o exemplo do uso secundário de dados de geolocalização de fotos para prever o comportamento dos turistas.

Focado em melhorar a reusabilidade dos dados e aumentar o cuidado à longo prazo deste ativo, uma comunidade formada pela academia, indústria, agências de fomento e editores acadêmicos, propôs o princípio de dados FAIR, que respectivamente, conforme sigla, abordam características sobre a facilidade de encontrar o dado, acessibilidade, interoperabilidade e reusabilidade. Estas características devem ser tratadas não só em relação aos dados, mas também seus metadados e infraestruturas utilizadas durante o ciclo de vida do dado (Wilkinson *et al.*, 2016). Complementarmente ao FAIR, Carroll *et al.* (2021) descreve outro princípio, o CARE, sigla que aborda os benefícios coletivos do uso do dado, a autoridade para controlar este ativo, a responsabilidade e a ética, originariamente criado pelo Grupo Internacional de Interesse em Soberania de Dados Indígenas da *Research Data Alliance* (2019). Carroll *et al.* (2021) entende que o princípio FAIR trata de questões técnicas sobre o dado, e o CARE acaba o complementando pois aproxima a discussão com as pessoas sobre os possíveis propósitos de uso do dado levando em consideração questões culturais, éticas, sociais, entre outras.

## **2.2.GOVERNANÇA DE DADOS**

A governança de dados segundo a Organização para a Cooperação e Desenvolvimento Econômico, OECD (2022), é responsável por criar políticas que incentivem a gestão de riscos e transparência no ecossistema de dados. Além disso, ela deve trabalhar para identificar a contribuição de múltiplos atores durante o ciclo de vida dos dados e gerar investimentos para tornar o reuso de dados efetivo, estimulando assim investimentos em tecnologias para abertura e controle de dados. Conforme definição do *Data Governance Institute*, s.d., a Governança de Dados é um sistema de responsabilidades e direitos de decisão para processos relacionados à informação, executado de acordo com modelos acordados que descrevem quem pode tomar quais ações com quais informações e quando, sob quais circunstâncias e usando quais métodos. Resumindo, Paprica *et al.* (2023) acrescentam que governança de dados é o caminho para garantir que o uso dos dados esteja em linha com a legalidade e com a concordância social.

Quanto pior o dado a tendência é que pouco se use a informação, logo a principal função da governança dados é definir e implementar regras e ações que assegurem a qualidade dos dados (Cerrillo-Martínez & Casadesús-De-mingo, 2021). Esta qualidade deve perdurar, e neste sentido, Jarvenpaa & Essén (2023) complementam que questões de curto prazo sobre o dado devem ser observadas, mas não se deve deixar de garantir que a governança de dados, também, observe o uso do dado a longo prazo e a considere que no futuro existirá o reuso do dado com propósitos que são desconhecidos no presente.

A qualidade do dado, considerada uma dimensão da governança de dados (Goel *et al.*, 2024), é fundamental para que as organizações criem valor a partir do dado (Vidgen *et al.*, 2017). Assim, conforme Valencia-Parra *et al.* (2021), a usabilidade do dado está justamente ligada a ela e ao contexto em que ele é utilizado. Por isso, importa para a qualidade do dado, o nível de ajuste da informação ao uso atribuído a ela, sendo assim Côrte-Real *et al.* (2020) relatam, por exemplo, que a qualidade do dado impacta diretamente as capacidades de *big data analytics* e *IoT*, consequentemente influenciando a competitividade das organizações, o que a torna um recurso de conhecimento organizacional.

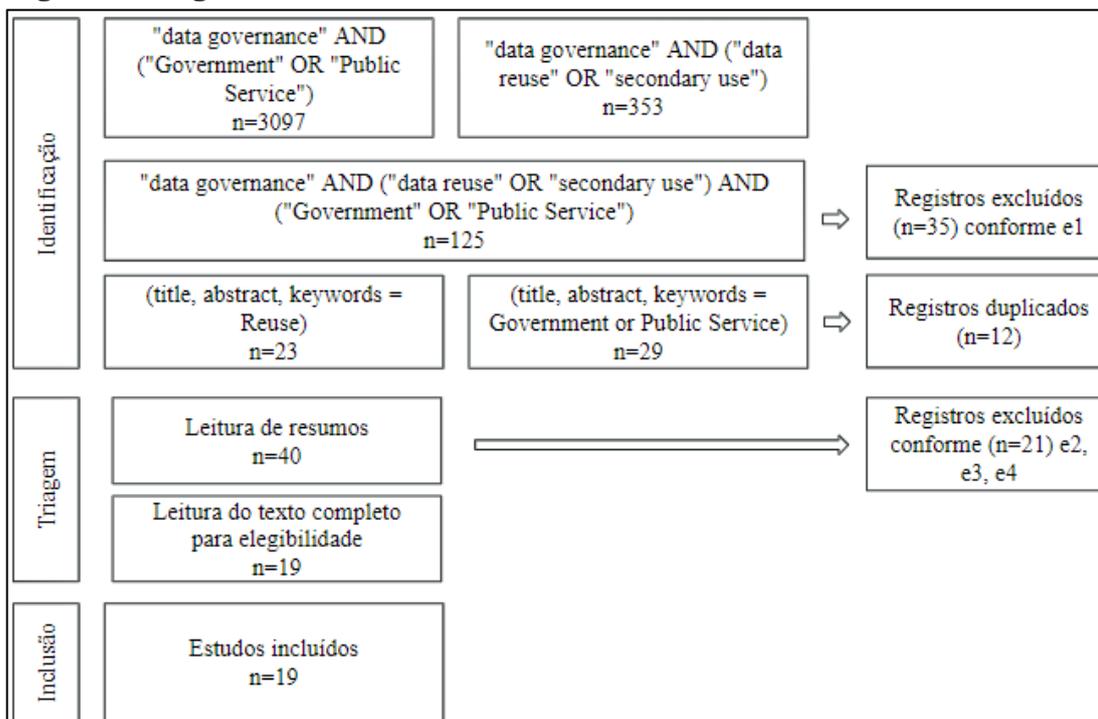
As necessidades das organizações com relação a qualidade dos dados são distintas (Woodall *et al.* 2013). No contexto da análise de risco de crédito, Moges *et al.* (2013) realizaram uma pesquisa com usuários e descobriu que a acurácia é a dimensão mais importante da qualidade dos dados. Além desta, eles identificaram outras dezessete dimensões de qualidade a partir da literatura, como relevância, facilidade de entendimento, facilidade de acesso e rastreabilidade. Ridzuan *et al.* (2024), encontraram na literatura, vinte dimensões, e diversas coincidem com as de Moges *et al.* (2013), entre elas a acurácia, a completude, a atualidade do dado e a consistência, que inclusive são citadas como as mais comuns.

### 3.MÉTODOS

Para realização de uma revisão sistemática de literatura, foi utilizado o PRISMA (Moher *et al.*, 2009) para definir os artigos relevantes para o estudo, representado pela Figura 1. Como principais bases de pesquisa foram eleitas a Scopus e Web of Science (WoS) que contêm artigos de maior impacto, em comparação ao Google Scholar (Martín-Martín *et al.*, 2018). Nestas bases, primeiramente iniciou-se a exploração de modo mais amplo, em todos os campos das bases, com os seguintes termos de busca "*data governance*" AND ("*government*" OR "*public Service*"), onde obteve-se 3097 resultados. Relacionando governança de dados com o reuso de dados, sem inserir o contexto de governo, em todos os campos das bases, com a seguinte busca "*data governance*" AND ("*data reuse*" OR "*secondary use*") obteve-se um retorno de 353 trabalhos.

Como o foco dado diretamente ao reuso, resultou em poucos artigos relacionando com a administração pública, para buscar aumentar a base de triagem, buscou-se dessa vez o foco naqueles 125 sobre títulos, resumos e palavras chaves que envolvessem o governo de forma mais explícita. Este movimentou apresentou 29 artigos para serem triados, dos quais 11 já haviam sido recuperados. Para os 18 restantes analisou-se seus resumos, direcionando para inclusão mais 10 artigos. Sendo assim, a Figura 1 apresenta detalhadamente os processos realizados até a inclusão dos 19 artigos, trazendo também os critérios de inclusão e exclusão adotados durante o processo na tabela 1.

**Figura 1: Diagrama Prisma**



Fonte: elaborado pelo autor (2024)

**Tabela 1: Critérios de Inclusão e exclusão de artigos**

Critérios de inclusão	Critérios de exclusão
i1: Existência de atividade ou participação governamental ou comercial ligada ao relato de governança, refletindo a busca pelo reuso de dados	e1: O trabalho não é um artigo e2: Aborda somente práticas de dados abertos sem profundidade ou foco no reuso de dados e3: Contexto apenas de gestão e manipulação de dados científicos ou acadêmicos e4: Não existência de atividade ou participação governamental ou comercial ligada ao relato de governança de dados e reuso de dados

Fonte: elaborado pelo autor (2024)

## 4.RESULTADOS E DISCUSSÕES

Nesta seção são apresentados o detalhamento de oito dimensões que influenciam o reuso de dados. As dimensões encontradas foram divididas em duas categorias, uma ligada aos atributos diretamente ligados aos dados, que influenciam sua reutilização, e outra ligada aos atributos de governança de dados que influenciam o reuso na administração pública.

### 4.1.AS DIMENSÕES LIGADAS AO DADO PARA REUSO

#### 4.1.1.INFRAESTRUTURA PARA O REUSO DO DADO

As infraestruturas digitais tangibilizam a implementação de plataformas para o uso e reuso de dados (Schmitt, 2023), logo deve-se revisá-las e ajustá-las para que os usuários consigam enxergar todo escopo de dados, pois isso garante um maior engajamento dos usuários

(Bowers & Choi, 2023). Porém, sem os dados, não há motivos para se utilizar uma plataforma, então, a infraestrutura deve contemplar as necessidades dos produtores do dado, pois estes esperam colher os benefícios do trabalho administrativo de registro dos dados. Assim durante a concepção da infraestrutura deve-se garantir que as necessidades de reuso do dado não ofusquem as de uso primário, sob pena, em caso de não equilíbrio dessas questões, de gerar tensões entre os produtores e os interessados pelo reuso, como relatado por Green et al. (2023), ao narrarem o desconforto dos médicos em relação a dificuldade que tinham para usar encontrar as informações nas plataformas de dados de saúde que eles mesmos ajudavam a alimentar. Além de garantir o atendimento ao propósito primário dos dados, complementarmente, segundo Bowers & Choi (2023), devem ser evitadas as infraestruturas focadas em apenas um propósito de uso, pois como exposto, considerando o contexto de dados educacionais, a infraestrutura de disponibilização de dados foi pensada para focar em questões de responsabilização, e isso acabou por dificultar o reuso das informações para outros propósitos. E neste caso, quando a infraestrutura é incompleta, o valor dos dados deixa de ser obtido (Kempeneer & Heylen, 2023).

De acordo com Alvarez-Romero *et al.* (2023), é o governo que usualmente financia infraestrutura digitais que reúnem e controlam dados de múltiplas fontes e permitem o compartilhamento para seu reuso, fornecendo serviços que facilitam a descoberta e acessibilidade de diferentes conjuntos de dado. Da mesma forma Martin *et al.* (2017) dizem que o governo deve melhorar as estratégias de disponibilização dos dados e das plataformas, e para isso devem utilizar tecnologias que permitam um acesso simples e direto ao dado, a exemplos de links e API, FTP, páginas web entre outros (Koesten *et al.*, 2020; Alvarez-Romero *et al.*, 2023).

#### **4.1.2.DOCUMENTAÇÃO**

As questões ligadas a facilitação do entendimento do dataset deveriam ser priorizadas, pois isso é crítico para o reuso (Koesten *et al.*, 2020). Entre os pré-requisitos para reutilização de dados a documentação de qualidade é um item relevante (Riley *et al.*, 2024; Miletic *et al.*, 2023; Koesten *et al.*, 2020). Esta documentação deve conter a descrição do contexto em que o dado foi criado (Koesten *et al.*, 2020; Riley *et al.*, 2024), deve também, permitir que as limitações do dado estejam claras, de forma a esclarecer como o dado pode ser utilizado e para que propósitos ele pode não ser útil (Riley *et al.*, 2024; Martin *et al.*, 2017).

O processo real de acesso ao dado também deve constar na documentação. Deve-se garantir que o usuário consiga seguir corretamente os passos para obter os dados e que este processo seja simples e rápido (Riley *et al.*, 2024). Neste caso, segundo Koesten *et al.*, (2020), uma boa estratégia é fornecer junto a documentação exemplos dos dados, ou conjuntos menores, assim os usuários podem testar as informações a serem trabalhadas.

#### **4.1.3.ACOMPANHAMENTO DO CICLO DE VIDA**

Enquanto os dados continuam sendo disponibilizados em Excel ou CSV, sem estarem contemplados por estruturas que acompanhem o seu ciclo de vida e proporcionem sua capacidade de integração, a organização continuará sofrendo com problemas de interoperabilidade (Bowers & Choi, 2023). A não disponibilidade das informações é um

problema para o reuso, porém a sua disponibilidade por um período muito longo também pode se tornar um. Dados obsoletos e não utilizados implicam em custos desnecessários e acabam impactando a busca pela informação, o que dificulta a reutilização dos dados (Cerrillo-Martínez & Casadesús-De-mingo, 2021). Com intuito de entender a utilidade dos dados, deve-se monitorar o acesso dos usuários, e com estas informações é possível não somente trabalhar na melhoria dos padrões dos conjuntos de dados, mas também qualificar os metadados e documentação (Koesten *et al.*, 2020).

#### **4.1.4.QUALIDADE DOS DADOS**

A qualidade dos dados garante a criação de mais valor, assim os governos devem criar padrões para gerenciar o dado (Cerrillo-Martínez & Casadesús-De-mingo, 2021) de forma sistemática (Bowers & Choi, 2023). Para isso, a Administração pública deve manter informações transparentes sobre a qualidade dos dados em sua custódia (Riley *et al.*, 2023). Por exemplo, fornecendo diagramas para explicar o controle e rotinas realizadas para manter a qualidade dos dados, e publicando relatórios sobre essas informações (Riley *et al.*, 2024). Segundo Cerrillo-Martínez & Casadesús-De-mingo (2021) sistemas de gestão dos dados, também, ajudam os dados a manterem sua qualidade ao longo do tempo.

Outros atributos influenciadores da qualidade dos dados providos pelo governo são: a existência de links que identifiquem a ligação entre conjuntos de dados diversos, o que fornece um universo de trabalho mais completo aos usuários (Riley *et al.*, 2024; Koesten *et al.*, 2020); a existência de processos de validação dos dados (Riley *et al.*, 2024); e a granularidade do dado, pois quanto mais desagregado maior a possibilidade de reuso (Alvarez-Romero *et al.*, 2023).

## **4.2.DIMENSÕES LIGADAS A GOVERNANÇA DE DADOS PARA O REUSO**

### **4.2.1.ALINHAMENTO ENTRE INTERESSADOS**

A administração pública possui diversas necessidades para o uso secundário do dado e isso se reflete nas plataformas tecnológicas. Essa condição gera tensões com quem utiliza o dado de forma primária, pois estes têm expectativas diferentes em relação a usabilidade dos sistemas (Green *et al.*, 2023). No caso relatado por (Aaen *et al.*, 2022), o crescimento na base de interessados externos para utilizar os dados de forma secundária, tirou o foco do uso primário pelos médicos e pacientes, e visou atender propósitos comerciais e de pesquisa, fazendo com que os produtores das informações começassem a ficar insatisfeitos com esta situação. Como uma possível solução, a aplicação dos princípios CARE, complementarmente aos princípios FAIR, facilitariam o equilíbrio entre os anseios de onde o dado foi originado e o dos utilizadores, pois reforçam os princípios de autoridade sobre o dado e trabalham o benefício coletivo através da participação comunidade ligada aos dados (Bowers & Choi, 2023; Carroll *et al.*, 2020).

Por outro lado, a governança de dados não pode focar somente nas interações internas as organizações, tendo função de encontrar ligação com atores externos, inclusive de fora das esferas governamentais (Cerrillo-Martínez & Casadesús-De-mingo, 2021). Para manter o equilíbrio do projeto de compartilhamento e reuso, é importante compreender a interação entre todas as partes interessadas, as funcionalidades e a própria informação (Aaen *et al.*, 2022). Por

exemplo, o governo chinês, visando garantir boa governança de dados, entende necessário criar relação com organizações externas ao estado e o público (Li & Guo, 2023). Carroll *et al.* (2020) e Paprica *et al.* (2023) também apontam que os cidadãos, de quem os dados foram originados, devem participar da governança dos dados, e que isso a torna mais confiável. Em outra frente, dentro do contexto de dados sobre agricultura e agricultores, entende-se que o governo é uma parte responsável para colocar os interessados corretos como meios para gerar confiança no processo de compartilhamento de dados e reuso (Kaur *et al.*, 2022).

Do ponto de vista do uso, de forma geral, segundo Paprica *et al.* (2023) usuários de dados devem ser engajados de forma contínua pois são capazes de impulsionar a inovação e a melhoria contínua da qualidade. Para isso, Miletić *et al.* (2023) sugere que os usuários tenham formas de interagir com os dados online, pontuando a importância dos datasets, e Koesten *et al.* (2020) adiciona que se deve viabilizar a colaboração através de discussão sobre mais pessoas. Além disso é importante que os utilizadores ou os interessados sejam consultados para saber que tipos de ontologias, metadados, e interface de sistema eles imaginam ter para consumir as informações (Martin *et al.*, 2017). O governo, também, deveria se engajar com os interessados pelos dados para promover o conhecimento sobre as informações disponibilizadas e incentivar o reuso destes (Martin *et al.*, 2017).

No trabalho de dados deve-se ter atenção na relação entre os cientistas de dados e os tomadores de decisão, pois estas podem apresentar vários problemas, desde falta de entendimento sobre a informação gerada de forma técnica, mas as vezes mesmo havendo entendimento, por interesses políticos a informação não é utilizada. A falta de alinhamento organizacional pode causar o reuso ilegítimo ou ilegal dos dados, assim o governo deveria refletir sobre a necessidade de usar dados e criar uma estratégia de dados que compreendesse toda organização (Kempeneer & Heylen, 2023). Além disso é importante que se trabalhe a cultura sobre o tema de reuso e gestão para essa mudança pois estas são mais importantes que a própria infraestrutura (Schmitt, 2023).

#### **4.2.2.CONFIANÇA DA SOCIEDADE**

Em um cenário de uso de dados para saúde, em Singapura, identificou-se que o cidadão acha plausível o uso de dados anonimizados pelo governo para pesquisa em saúde. Por exemplo, a população tem mais confiança no que o governo vai fazer com os dados do que na iniciativa privada (Lysaght *et al.*, 2021). Corroborando com encontrado pelo autor anterior, no estudo realizado na Austrália, observou-se que o quando um órgão governamental confiável está no centro do tratamento do dado as barreiras para o reuso são reduzidas (Varhol *et al.*, 2022). Conforme Riley *et al.*, (2023) e Kempeneer & Heylen, (2023) para garantir a confiança do público, estratégias de transparência e publicação sobre os propósitos de reuso de dado devem ser praticadas pelos órgãos governamentais.

O Governo deve reforçar regulamentações para prevenir o uso inadequado dos dados, pois as empresas tendem a tentar contornar as regras criadas para desviar o uso do dado, principalmente as menores que querem evitar custos altos com processamento de dados (Li & Guo, 2023). Se o ecossistema de dados mantiver suas bases acordadas estáveis, ele tende a ter sucesso (Aaen *et al.*, 2022). E a administração pública é justamente quem cria bases para responsabilização e adequação sobre o uso do dado, e adicionalmente deve criar padrões e

procedimentos que evidenciem a posse correta das informações, com o objetivo de garantir uma governança equilibrada (Kaur *et al.*, 2022). Quando a posse do dado não é clara, conforme Varhol *et al.* (2022), criam-se dificuldades para o compartilhamento do dado e seu uso secundário. Complementarmente as políticas e processos também devem atender requerimentos legais estabelecidos externamente (Paprica *et al.*, 2023). Porém, conforme relato de Aaen *et al.* (2022) não basta que exista bases e procedimentos legais, mas estes devem ser claros o suficiente, pois caso não sejam o lado obscuro do reuso dos dados pode emergir.

#### **4.2.3.GOVERNANÇA DO PROPÓSITO DE USO**

O sistema de compartilhamento de dados públicos de saúde da Dinamarca, que permitia que os médicos acessem históricos dos pacientes e que estes acompanhassem seus dados de saúde, colapsou durante a expansão do uso para propósitos secundários, quando se perdeu o controle sobre o reuso do dado. Isso gerou tensões e levou o parlamento a ordenar a exclusão de toda a base de dados (Aaen *et al.*, 2022). Para que estas iniciativas não sejam descontinuadas, e os investimentos não sejam perdidos, conforme Lysaght *et al.* (2021) o governo deve priorizar os projetos em que os dados estão gerando valor para a administração pública e que esse reuso seja transparente. Ballantyne & Style (2017), também em um cenário sobre dados públicos de saúde, entende que o interesse público é uma justificativa plausível para o uso secundário dos dados, porém deve-se criar regras claras para definir este tipo de situação, acompanhando seus reais benefícios, os riscos envolvidos no reuso e garantindo que aquilo que for realizado por esta motivação se acessível publicamente.

#### **4.2.4.ESTRUTURA DE GOVERNANÇA**

Existe uma necessidade de conceber mecanismos para melhorar as escolhas estratégicas dos diferentes participantes e promover a circulação e utilização de dados para diversos assuntos (Li & Guo, 2023). A governança de dados pode estar desenhada de forma hierárquica, onde existirão mecanismos de comando e controle e relação vertical entre os participantes, e pode ser trabalhada de forma colaborativa onde há maior interação entre os participantes (Cerrillo-Martínez & Casadesús-De-mingo, 2021). Porém, Varhol *et al.* (2022) indica que esta infraestrutura da coleta até a disponibilização do dado deve ser realizada de forma centralizada pelo governo para garantir uma governança transparente. Alvarez-Romero *et al.* (2023) colabora observando que organizações centralizadas são mais comuns, do que as descentralizadas na gestão de infraestruturas de compartilhamento de dados, e cita o governo como provedor de acesso à informação, o qual realiza a avaliação de propósito de uso secundário.

A gestão não sistemática dos dados torna eles menos úteis (Bowers & Choi, 2023). Sugerida por (Ballantyne & Style, 2017) como uma estrutura de governança de dados, um comitê composto por profissionais de negócio e de dados, interessados e cidadãos deveria ser criado para avaliar o uso secundário do dado. Esta configuração poderia se encaixar em dois dos tipos de modelos (Cerrillo-Martínez & Casadesús-De-mingo, 2021) : o coordenado, onde cita-se um corpo colegiado de diferentes áreas atuando internamente, que podem focar na construção das políticas e ou na implementação destas; e o modelo descentralizado, também

formado por um colegiado, tendo este, uma característica mais estratégica, com olhar de futuro, realizando contato com entes externos a organização, e reportando-se a autoridade da organização. Schmitt (2023) ainda acrescenta na estrutura o papel dos custodiantes dos dados, os quais são relevantes para contribuir com o uso e reuso dos dados públicos e Cerrillo-Martínez & Casadesús-De-mingo (2021) narra a existência do escritório do chefe de dados, posicionado como autoridade central da gestão de dados para garantir qualidade, confiabilidade, não duplicidade, precisão, consistência e completude dos dados.

### 4.3. CATEGORIZAÇÃO

Observando-se os trabalhos selecionados, conseguiu-se identificar dois categorias de dimensões diferentes em relação ao reuso do dado, aquela que compete intrinsecamente a atributos do dado que vão levar a sua reutilização, conforme Tabela 2, e aquelas ligadas atributos de governança de dados em relação ao uso secundário do dado conforme Tabela 3.

**Tabela 2 - Dimensões ligadas ao dado**

Dimensões ligadas ao dado	Atributos relacionados	Citações
Infraestrutura	Adequação a uso primário e secundário, nível de especialização da plataforma, serviços de descoberta, estratégias de disponibilização de dados, papel do governo em relação as plataformas	(Kempeneer & Heylen, 2023), (Bowers & Choi, 2023), (Green <i>et al.</i> , 2023), (Alvarez-Romero <i>et al.</i> , 2023), (Koesten <i>et al.</i> , 2020), (Martin <i>et al.</i> , 2017), (Schmitt, 2023)
Documentação	Descrição do contexto de criação dos dados, descrição do escopo e das limitações dos dados	(Koesten <i>et al.</i> , 2020) (Riley <i>et al.</i> , 2024), (Martin <i>et al.</i> , 2017), (Miletic <i>et al.</i> , 2023)
Ciclo de vida do dado	Monitoramento do acesso dos usuários aos dados, acompanhamento do uso dos dados disponibilizados, descontinuação de dados não utilizados	(Bowers & Choi, 2023), (Cerrillo-Martínez & Casadesús-De-mingo, 2021), (Koesten <i>et al.</i> , 2020)
Qualidade dos dados	Fornecimento de diagramas e transparência das rotinas de qualidade, existência de padrões de gerenciamento, links entre bases de dados, granularidade da informação, informações sobre a propriedade do dado	(Cerrillo-Martínez & Casadesús-De-mingo, 2021), (Bowers & Choi, 2023), (Riley <i>et al.</i> , 2024), Riley 2023, (Koesten <i>et al.</i> , 2023), (Alvarez-Romero <i>et al.</i> , 2023)

Fonte: elaborado pelo autor (2024)

A dimensão mais citada entre os trabalhos pesquisados, trata das ações de alinhamento entre os interessados, aparecendo em quatorze dos dezoito trabalhos, e está em linha com o que a OECD (2022) identifica como trabalho da governança de dados, a qual deve diagnosticar a contribuição de diversos tipos de atores durante a jornada do dado. Paprica *et al.* (2023) cita a concordância social como um dos atributos de alinhamento, assim evidenciando a necessidade da governança de dados trabalhar também com atores externos a administração pública. Além disso encontram-se indícios que capacidade de alinhamento com os interessados impacte diretamente, pelo menos em outras três dimensões identificadas no presente estudo.

**Tabela 3 - Dimensões ligadas a governança do dado**

Dimensões ligadas a Governança do Dado	Atributos relacionados	Citações
Alinhamento entre interessados	Gerenciamento de expectativas entre produtores e utilizadores, utilização de abordagens CARE e FAIR, relacionamento com público interno e externo, papel do governo na governança e engajamento dos atores, colaboração dos interessados internos e externos ao governo, alinhamento organizacional entre atores	(Aaen <i>et al.</i> , 2022), (Kempeneer & Heylen, 2023), (Bowers & Choi, 2023), (Carroll <i>et al.</i> , 2020), (Martin <i>et al.</i> , 2017), (Green <i>et al.</i> , 2023), (Cerrillo-Martínez & Casadesús-De-mingo, 2021), (Li & Guo, 2023), (Paprica <i>et al.</i> , 2023), (Kaur <i>et al.</i> , 2022), (Miletic <i>et al.</i> , 2023), (Koesten <i>et al.</i> , 2023), (Martin <i>et al.</i> , 2017), (Schmitt, 2023)
Confiabilidade da sociedade	Nível de confiabilidade do órgão de governo, estratégias de transparência e publicidade dos propósitos de uso, clareza sobre a a responsabilidade pelo dado, estabilidade das regras, adequação a regramentos externos ao governo, clareza dos procedimentos	(Li & Guo, 2023), (Kaur <i>et al.</i> , 2022), (Aaen <i>et al.</i> , 2022), (Paprica <i>et al.</i> , 2023), (Varhol <i>et al.</i> , 2022), (Lysaght <i>et al.</i> , 2021), (Riley <i>et al.</i> , 2023) (Kempeneer & Heylen, 2023)
Governança do propósito de uso	Priorização de uso, acompanhamento do uso em relação a benefício e riscos, disponibilização dos propósitos de uso, identificação da relevância do dado	(Green <i>et al.</i> , 2023), (Aaen <i>et al.</i> , 2022), (Lysaght <i>et al.</i> , 2021), (Riley <i>et al.</i> , 2023), (Ballantyne & Style, 2017), (Kempeneer & Heylen, 2023), (Miletic <i>et al.</i> , 2023)
Estrutura de governança	Estruturas centralizadas e descentralizadas e seus papéis em relação a criação de políticas e alinhamento com alta gestão; Formação e atuação de comitês.	(Ballantyne & Style, 2017), (Cerrillo-Martínez & Casadesús-De-mingo, 2021), (Alvarez-Romero <i>et al.</i> , 2023), (Varhol <i>et al.</i> , 2022), (Li & Guo, 2023), (Schmitt, 2023), (Bowers & Choi, 2023)

Fonte: elaborado pelo autor (2024)

Considerando que para se construir uma infraestrutura adequada e completa é necessário estar alinhado com os produtores do dado, os usuários da informação primária e usuários que exercem o uso secundário (Bowers & Choi, 2023), identifica-se a dependência da dimensão de alinhamento das organizações públicas com os interessados. A necessidade deste alinhamento sugere ligação como nível de confiabilidade da sociedade sobre o reuso dos dados, quando busca-se dar transparência ao uso dos dados (Riley *et al.*, 2023; Kempeneer & Heylen, 2023). Por último observa-se conexão com a dimensão de estruturas de governança, justamente quando precisa-se discutir a relação das estruturas com os interessados, a exemplo da existência de comitês que podem ter diversos participantes de dentro e de fora das organizações, e de estruturas centrais que prestam contas a alta gestão (Cerrillo-Martínez & Casadesús-De-mingo, 2021).

Quando se abordam os aspectos ligados ao dado e sua reutilização com participação governamental, a literatura trata de forma mais igual as dimensões encontradas. A infraestrutura e a qualidade de dados apresentam-se como as mais abordadas. Os atributos da dimensão de qualidade, encontrados na revisão de literatura, não coincidem diretamente com as dimensões pesquisadas por Moges *et al.* (2013) ou Ridzuan *et al.* (2024), por um lado isso fica de acordo com a afirmação que cada organização terá um tipo de necessidade com relação a qualidade do dado Woodall *et al.* (2013), por outro lado o fato de não se encontrar alinhamento nem das

dimensões mais comuns como a acurácia, a completude, a consistência e atualidade, o tratamento superficial desse tema ligado ao reuso de dados na administração pública.

## 5. CONSIDERAÇÕES FINAIS

Através da revisão sistemática da literatura foi possível identificar que há poucos estudos que mostram o reuso de dados para propósitos internos às organizações governamentais e a interação da governança de dados neste processo, representando a mesma dificuldade que Cerrillo-Martínez & Casadesús-De-mingo (2021) enfrentaram em seu estudo. Foram identificadas duas principais abordagens sobre o tema, a primeira, ligada a primeira questão de pesquisa, salientando aspectos mais ligados a atributos do dado e infraestrutura que facilitam seu uso, tocando temas, ao menos de forma superficial, como documentação e qualidade, e a segunda, respondendo a segunda questão de pesquisa, abrange a governança de dados, alcançando temas como alinhamento entre os interessados pelo uso e reuso do dado, regras e procedimentos e estruturas de governança.

Foi possível identificar práticas com relação a governança de dados vinculadas ao reuso, porém não se pode afirmar que os achados representam o cenário completo dos órgãos públicos em relação ao tema de reuso de dados e atuação da governança de dados. Assim, respondendo a última questão de pesquisa, sugere-se um que pesquisas busquem identificar junto aos órgãos públicos como estes tem tratado as dimensões identificadas nessa revisão sistemática de literatura e quais outras práticas têm sido empregadas. De outra forma, mesmo considerando que as bases de pesquisa escolhidas foram as mais relevantes, conforme (Martín-Martín *et al.*, 2018), é possível que exemplos relevantes possam não ter sido incluídos, deixando assim com que outras práticas no contexto governamental emergissem, assim para estudos futuros sugere-se que se analise de forma criteriosa outras bases de conhecimento em busca de complemento ao presente trabalho.

## REFERÊNCIAS

- Aaen, J., Nielsen, J. A., & Carugati, A. (2022). The dark side of data ecosystems: A longitudinal study of the DAMD project. *European Journal of Information Systems*, 31(3), 288–312. <https://doi.org/10.1080/0960085X.2021.1947753>
- Aleksynska, M., & Cazes, S. (2016). Composite indicators of labour market regulations in a comparative perspective. *IZA Journal of Labor Economics*, 5(1). doi:10.1186/s40172-016-0043-y
- Alvarez-Romero, C., Martínez-García, A., Bernabeu-Wittel, M., & Parra-Calderón, C. L. (2023). Health data hubs: an analysis of existing data governance features for research. *Health Research Policy and Systems*, 21(1), 70. <https://doi.org/10.1186/s12961-023-01026-1>
- Ballantyne, A., & Style, R. (2017). Health data research in New Zealand: Updating the ethical governance framework. *New Zealand Medical Journal*, 130(1464), 64–71
- Black, S., Davern, M., Maynard, S. B., & Nasser, H. (2023). Data governance and the secondary use of data: The board influence. *Information and Organization*, 33(2), 100447. ISSN 1471-7727. <https://doi.org/10.1016/j.infoandorg.2023.100447>.

Bowers, A. J., & Choi, Y. (2023). Building School Data Equity, Infrastructure, and Capacity Through FAIR Data Standards: Findable, Accessible, Interoperable, and Reusable. *Educational Researcher*, 52(7), 450–458. <https://doi.org/10.3102/0013189X231181103/FORMAT/EPUB>

Carroll, S.R., Herczog, E., Hudson, M. et al. Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Sci Data* 8, 108 (2021). <https://doi.org/10.1038/s41597-021-00892-0>

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for indigenous data governance. *Data Science Journal*, 19(1), 1–12. <https://doi.org/10.5334/DSJ-2020-043>

Cerrillo-Martínez, A., & Casadesús-De-mingo, A. (2021). Data governance for public transparency. *Profesional de La Informacion*, 30(4), e300402. <https://doi.org/10.3145/EPI.2021.JUL.02>

Côrte-Real, N., Ruivo, P., & Oliveira, T. (2020). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, 57(1), 103141. ISSN 0378-7206. <https://doi.org/10.1016/j.im.2019.01.003>.

Data Governance Institute. Data governance definition. Recuperado de. Acessado em 13 Junho de, 2024 <https://datagovernance.com/the-data-governance-basics/definitions-of-data-governance/>.

De Caigny, A., Coussement, K., & De Bock, K. W. (2020). Leveraging fine-grained transaction data for customer life event predictions. *Decision Support Systems*, 130, 113232. ISSN 0167-9236. <https://doi.org/10.1016/j.dss.2019.113232>.

Deloitte. (2021). Fluid data dynamics: Generating greater public value from data. Recuperado em 07 de abril de 2023, de <https://www2.deloitte.com/uk/en/insights/industry/public-sector/government-trends/2021/fluid-government-data-dynamics.html>

Dremel, C., Herterich, M., Wulf, J., Waizmann, J.-C., & Brenner, W. (2017). How AUDI AG Established Big Data Analytics in Its Digital Transformation. *MIS Quarterly Executive*, 16(2).

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., & Liu, X. S. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Medicine*, 12(1). doi:10.1186/s13073-020-0721-z

Gartner. (2023). 5 Ways to Enhance Your Data Engineering Practices. Recuperado em 07 de abril de 2023, de <https://www.gartner.com/document/4220199?ref=solrResearch&refval=363164487>

Goel, K., Martin, N., & ter Hofstede, A. (2024). Demystifying data governance for process mining: Insights from a Delphi study. *Information & Management*, 61(5), 103973. ISSN 0378-7206. <https://doi.org/10.1016/j.im.2024.103973>.

Green, S., Hillersdal, L., Holt, J., Hoeyer, K., & Wadmann, S. (2023). The practical ethics of repurposing health data: how to acknowledge invisible data work and the need for prioritization. *Medicine, Health Care and Philosophy*, 26(1), 119–132. <https://doi.org/10.1007/s11019-022-10128-6>

- Günther, W. A., Mehrizi, M. H. R., Huysman, M., Deken, F., & Feldberg, F. (2022). Resourcing with data: Unpacking the process of creating data-driven value propositions. *Journal of Strategic Information Systems*, 31(4), 101744. <https://doi.org/10.1016/j.jsis.2022.101744>
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446-456. doi: 10.1016/j.giq.2011.01.004
- Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. *Information and Organization*, 33(1), 100449. ISSN 1471-7727. <https://doi.org/10.1016/j.infoandorg.2023.100449>.
- Kaur, J., Hazrati Fard, S. M., Amiri-Zarandi, M., & Dara, R. (2022). Protecting farmers' data privacy and confidentiality: Recommendations and considerations. *Frontiers in Sustainable Food Systems*, 6, 903230. <https://doi.org/10.3389/fsufs.2022.903230>
- Kempeneer, S., & Heylen, F. (2023). Virtual state, where are you? A literature review, framework and agenda for failed digital transformation. *Big Data and Society*, 10(1). <https://doi.org/10.1177/20539517231160528>
- Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data. *Journal of Management Information Systems*, 35(2), 540–574. <https://doi.org/10.1080/07421222.2018.1451957>
- Kitsios, F., Kamariotou, M., & Mavromatis, A. (2023). Drivers and outcomes of digital transformation: The case of public sector services. *Information*, 14(1), 43. <https://doi.org/10.3390/info14010043>
- Koesten, L., Vougiouklis, P., Simperl, E., & Groth, P. (2020). Dataset Reuse: Toward Translating Principles to Practice. *Patterns*, 1(8), 100136. <https://doi.org/10.1016/j.patter.2020.100136>
- Li, X., & Guo, X. (2023). A tripartite evolutionary game analysis of sports data rights protection from the perspective of stakeholder. *PLoS ONE*, 18(11), e0292914. <https://doi.org/10.1371/journal.pone.0292914>
- Lysaght, T., Ballantyne, A., Toh, H. J., Lau, A., Ong, S., Schaefer, O., Shiraishi, M., van den Boom, W., Xafis, V., & Tai, E. S. (2021). Trust and trade-offs in sharing data for precision medicine: A national survey of singapore. *Journal of Personalized Medicine*, 11(9), 921. <https://doi.org/10.3390/jpm11090921>
- Martin, E. G., Law, J., Ran, W., Helbig, N., & Birkhead, G. S. (2017). Evaluating the quality and usability of open data for public health research: A systematic review of data offerings on 3 open data platforms. *Journal of Public Health Management and Practice*, 23(4), e5–e13. <https://doi.org/10.1097/PHH.0000000000000388>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160-1177. ISSN 1751-1577. <https://doi.org/10.1016/j.joi.2018.09.002>.
- McKinsey & Company. (2020) Reducing data costs without jeopardizing growth, Recuperado em 07 de abril de 2023, de <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/reducing-data-costs-without-jeopardizing-growth>

- Mergel, I., Edelman, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, 36(4), 101385. <https://doi.org/10.1016/j.giq.2019.06.002>
- Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information & Management*, 54(6), 771-785. ISSN 0378-7206. <https://doi.org/10.1016/j.im.2016.11.011>.
- Miletić, A., Kuveždić Divjak, A., & Welle Donker, F. (2023). Assessment of the Croatian Open Data Portal Using User-Oriented Metrics. *ISPRS International Journal of Geo-Information*, 12(5), 185. <https://doi.org/10.3390/ijgi12050185>
- Moges, H.-T., Dejaeger, K., Lemahieu, W., & Baesens, B. (2013). A multidimensional analysis of data quality for credit risk management: New insights and challenges. *Information & Management*, 50(1), 43-58. ISSN 0378-7206. <https://doi.org/10.1016/j.im.2012.10.001>.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Moser-Plautz, B., & Schmidhuber, L. (2023). Digital government transformation as an organizational response to the COVID-19 pandemic. *Government Information Quarterly*, 101815. <https://doi.org/10.1016/j.giq.2023.101815>
- OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris. <https://doi-org.ez94.periodicos.capes.gov.br/10.1787/40d53904-en>
- Paparova, D., Aanestad, M., Vassilakopoulou, P., & Bahus, M. K. (2023). Data governance spaces: The case of a national digital service for personal health data. *Information and Organization*, 33(1), 100451. <https://doi.org/10.1016/j.infoandorg.2023.100451>
- Paprica, P. A., Crichlow, M., Maillet, D. C., Kesselring, S., Pow, C., Scarnecchia, T. P., Schull, M. J., Cartagena, R. G., Cumyn, A., Dostmohammad, S., Elliston, K. O., Greiver, M., Nelson, A. H., Hill, S. L., Isaranuwachai, W., Loukipoudis, E., McDonald, J. T., McLaughlin, J. R., Rabinowitz, A., McGrail, K. (2023). Essential requirements for the governance and management of data trusts, data repositories, and other data collaborations. *International Journal of Population Data Science*, 8(4), 01. <https://doi.org/10.23889/ijpds.v8i4.2142>
- Research Data Alliance International Indigenous Data Sovereignty Interest Group. (Setembro 2019). “CARE Principles for Indigenous Data Governance.” The Global Indigenous. Data Alliance. [GIDA-global.org](https://gida-global.org)
- Ridzuan, F., & Wan Zainon, W. M. N. (2024). A review on data quality dimensions for big data. *Procedia Computer Science*, 234, 341-348. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2024.03.008>.
- Riley, M., Robinson, K., Kilkenny, M. F., & Leggat, S. G. (2023). The suitability of government health information assets for secondary use in research: A fit-for-purpose analysis. *Health Information Management Journal*, 52(3), 157–166. <https://doi.org/10.1177/18333583221078377>
- Riley, M., Robinson, K., Kilkenny, M. F., & Leggat, S. G. (2024). The knowledge and reuse practices of researchers utilising government health information assets, Victoria, Australia, 2008- 2020. *PLoS ONE*, 19(2), e0297396. <https://doi.org/10.1371/journal.pone.0297396>

Schmitt, T. (2023). Implementing Electronic Health Records in Germany: Lessons (Yet to Be) Learned. *International Journal of Integrated Care*, 23(1), 13. <https://doi.org/10.5334/ijic.6578>

Valencia-Parra, Á., Parody, L., Varela-Vaca, Á. J., Caballero, I., & Gómez-López, M. T. (2021). DMN4DQ: When data quality meets DMN. *Decision Support Systems*, 141, 113450. ISSN 0167-9236. <https://doi.org/10.1016/j.dss.2020.113450>.

Varhol, R. J., Randall, S., Boyd, J. H., & Robinson, S. (2022). Australian general practitioner perceptions to sharing clinical data for secondary use: a mixed method approach. *BMC Primary Care*, 23(1), 167. <https://doi.org/10.1186/s12875-022-01759-y>

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, 261(2), 626-639. ISSN 0377-2217. <https://doi.org/10.1016/j.ejor.2017.02.023>.

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

Woodall, P., Borek, A., & Parlikad, A. K. (2013). Data quality assessment: The Hybrid Approach. *Information & Management*, 50(7), 369-382. ISSN 0378-7206. <https://doi.org/10.1016/j.im.2013.05.009>.

Zhang, Q., Sun, X., & Zhang, M. (2022). Data Matters: A Strategic Action Framework for Data Governance. *Information & Management*, 59(4), 103642. <https://doi.org/10.1016/j.im.2022.103642>

Zhu, H., Madnick, S. E., & Siegel, M. D. (2008). An Economic Analysis of Policies for the Protection and Reuse of Noncopyrightable Database Contents. *Journal of Management Information Systems*, 25(1), 199–232. <http://www.jstor.org/stable/40398931>