



Arquitetura Digital Inteligência de Negócio do MCTI

Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando os painéis temáticos elaborados

Projeto Temático: Arquitetura Digital Inteligência de Negócio do MCTIC

Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando os painéis temáticos elaborados

Projeto Temático: Arquitetura Digital Inteligência de Negócio do MCTIC

Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando os painéis temáticos elaborados



Brasília, DF

Junho, 2022

Centro de Gestão e Estudos Estratégicos

Presidente

Fernando Cosme Rizzo Assunção

Diretores

Ary Mergulhão Filho

Luiz Arnaldo Pereira da Cunha

Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando os painéis temáticos elaborados. Arquitetura Digital Inteligência de Negócio do MCTIC. Brasília: Centro de Gestão e Estudos Estratégicos, 2022.

38 p. : il.

1. Ciência, tecnologia e inovação. 2. Plataforma digital de informações. 3. Sistema de apoio à decisão. 4. Inteligência de dados. Título. II. CGEE.

Centro de Gestão e Estudos Estratégicos - CGEE
SCS Quadra 9 – Torre C – 4º andar – salas 401 a 405
Edifício Parque Cidade Corporate
70308-200 - Brasília, DF
Telefone: (61) 3424.9600
<http://www.cg ee.org.br>

Este relatório é parte integrante das atividades desenvolvidas no âmbito do 2º Contrato de Gestão CGEE – 36º Termo Aditivo, Linha de Ação: Apoio Técnico à Gestão Estratégica do SNCTI / Projeto: Arquitetura digital de inteligência de negócios do MCTIC – 8.10.53.05.01.03/MCTIC/2019.

Todos os direitos reservados pelo Centro de Gestão e Estudos Estratégicos (CGEE). Os textos contidos neste relatório poderão ser reproduzidos, armazenados ou transmitidos, desde que citada à fonte.

Projeto Temático: Arquitetura Digital Inteligência de Negócio do MCTIC

Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando os painéis temáticos elaborados

Supervisão

Luiz Arnaldo Pereira da Cunha

Equipe técnica interna

Alberto Akira Okata

Carlos Duarte de Oliveira Junior

Carlson B. de Oliveira (Coordenador)

Marco Antônio Andrade Dias

Marcus Vinícius T. da Cunha

Wagner Alberto Soares Junior

Equipe técnica externa

Adriano Albernaz Golebiowski

Jorge Millis

Raissa Rondon

Sebastião Gonella

Victor Neves Martorelli

SUMÁRIO

1. INTRODUÇÃO.....	1
2. OBJETIVOS	3
3. PERCURSO METODOLÓGICO	4
4. PRODUTOS RESULTANTES.....	10
4.1. AMBIENTE DE INTELIGÊNCIA DE NEGÓCIO DO MCTI	14
4.1.1. <i>Ambiente tecnológico.....</i>	<i>14</i>
4.1.2. <i>Ambiente de informação.....</i>	<i>16</i>
4.2. EXPERIMENTOS - TEMAS ESTRATÉGICOS	19
4.2.1. <i>Painéis temáticos.....</i>	<i>20</i>
4.2.2. <i>Indicadores da COICT.....</i>	<i>21</i>
5. RESULTADOS ALCANÇADOS	23
5.1. FONTES DE DADOS INTERNAS	23
5.2. REALIZAÇÕES DERIVADAS	26
6. CONCLUSÃO	28
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	32
ANEXO I – GLOSSÁRIO.....	35

1. Introdução

O Ministério da Ciência, Tecnologia e Inovações (MCTI) tem como competências o planejamento, coordenação, supervisão e controle das atividades de ciência, tecnologia e inovação, dentre outras. À Secretaria Executiva – SEEXEC, por sua vez, compete supervisionar e coordenar as atividades de formulação e proposição de políticas, diretrizes, objetivos e metas relativas às áreas de atuação do Ministério, atividades naturalmente demandantes de informação de alto valor agregado.

Dentre as atividades do MCTI deve-se ressaltar o papel proeminente no Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI). Na orquestração das ações buscando maximizar resultados de interesse social e econômico para o Brasil, em meio à complexidade desse Sistema, é fundamental a capacidade de integração da informação distribuída nos atores sistêmicos relevantes para geração de conhecimento e apoio à tomada de decisão.

Este projeto visa, explorar a área de sistemas analíticos modernos (que incluem conceitos tais como *data lake*, *big data*, *business intelligence*) e contribuir para a qualidade da gestão das ações governamentais no SNCTI, aportando conhecimento para a produção de informação estratégica ao MCTI. Como ponto focal do projeto está o desenvolvimento experimental de estruturas tecnológicas e técnicas para a produção de informações e sua apresentação em ambientes virtuais inovadores criados para o apoio à tomada de decisão relacionada a políticas públicas e programas de natureza estratégica.

O objetivo geral do projeto é elaborar e disponibilizar uma Arquitetura digital de inteligência de negócio do MCTI, referenciada como **Arquitetura Digital** no decorrer deste relatório, que consiste em solução metodológica e instrumental (software e hardware) para dar suporte e prover recursos para armazenamento e tratamento de fontes de dados heterogêneas, visualização de informação e gestão de características de qualidade e segurança sobre o conjunto de dados e informações administradas.

O ambiente deve ser interoperável com sistemas de informação legados do MCTI e fontes de informação externas ao Ministério que sejam consideradas relevantes para o monitoramento e avaliação do desempenho do Sistema Nacional de Ciência, Tecnologia e Inovação.

Esse objetivo geral se decompõe nos seguintes objetivos específicos:

- Evoluir a pesquisa e o desenvolvimento de ambientes digitais de acordo com as especificações feitas pelo MCTI e atores relevantes do SNCTI.
- Disponibilizar um modelo integrado de trabalho sobre ambientes informacionais, interoperáveis, que promova a construção e manutenção de catálogo de fontes de dados e informações do MCTI.
- Disponibilizar meios para construção de análises, produção de dados agregados e indicadores com capacidade para conexão com dispositivos móveis e mobilidade em nuvem, de modo a permitir a expansão da arquitetura de informação.

Neste relatório são apresentados os resultados alcançados no projeto contemplando a implementação dos temas estratégicos estabelecidos pelo MCTI para os quais estavam presentes os insumos necessários, materializadas sobre a Arquitetura Digital implantada nas instalações físicas de Tecnologia da Informação do Ministério.

2. Objetivos

O objetivo deste relatório é apresentar os resultados do projeto Arquitetura Digital de Inteligência de Negócio do MCTI no ano de 2021, contemplando os painéis temáticos elaborados.

Incorpora, também, o registro dos resultados relativos à arquitetura da informação e tecnológica proposta, processo de trabalho e os demais elementos estruturantes padronizados de gestão de dados, temas centrais da missão conferida ao projeto.

O objetivo geral do projeto é disponibilizar ambiente digital que suporte o armazenamento de fontes de informações heterogêneas e permita a aplicação de metodologias de análise a partir de conjuntos de dados de brutos, dados estruturados, dados parcialmente estruturados e dados não estruturados disponíveis em diferentes formatos. O ambiente deve ser interoperável com sistemas de informação legados do MCTIC e fontes de informação externas ao ministério que sejam consideradas relevantes para o monitoramento e avaliação do desempenho de ambos os sistemas.

Esse objetivo se desdobra nos seguintes objetivos específicos:

- Evoluir a pesquisa e o desenvolvimento de ambientes digitais de acordo com as especificações feitas pelo MCTI e atores relevantes do SNCTI e do sistema de Comunicações.
- Disponibilizar um modelo integrado de trabalho sobre ambientes informacionais, interoperáveis, que promova a construção e manutenção de catálogo de fontes de dados e informações do MCTI.
- Disponibilizar meios para construção de análises, produção de dados agregados e indicadores com capacidade para conexão com dispositivos móveis e mobilidade em nuvem, de modo a permitir a expansão da arquitetura de informação.

3. Percurso metodológico

Para o alcance desses objetivos, a condução do projeto foi estruturada em três linhas de ação principais, conforme apresentado na Figura 1. O detalhamento completo do projeto, contendo descrições das fases e distribuição no tempo, está registrado em (CGEE, 2019 e 2020).



Figura 1 - Plano de trabalho do projeto. Fonte: (CGEE, 2019; 2020).

No eixo de ação Arquitetura são realizadas atividades de pesquisa, conceituação e avaliação de alternativas tecnológicas para a proposição da arquitetura digital de inteligência de negócio. Para tanto, é necessário conhecer e explorar as capacidades a serem disponibilizadas pela arquitetura digital para atender objetivos de negócio do Ministério. A arquitetura digital tem o papel estruturador de ambiente onde soluções analíticas serão desenvolvidas, integradas, implantadas e mantidas.

Assim, os passos iniciais identificaram esses objetivos maiores e os principais interessados, o que resultou no levantamento da situação atual e necessidades que são endereçadas pelos desenvolvimentos de soluções analíticas. A captura de necessidades e preocupações tem como objetivo a definição de princípios, diretrizes e fronteiras (escopo) da arquitetura digital.

Esses levantamentos e o debate realizado com o MCTI permitiram a modelagem da arquitetura digital, definindo sua forma e conteúdo, que são representados por: elementos arquiteturais e seus inter-relacionamentos, descritivo de seus elementos, instrumentos tecnológicos (software, hardware e serviços) que os implementam, padrões e processos de trabalho para realização das soluções analíticas que residirão na arquitetura digital.

Essa abordagem resultou na avaliação conceitual e tecnológica para arquiteturas e a proposição de um modelo arquitetural que constitua um arcabouço de trabalho para atendimento de necessidades de informação estratégica do MCTI.

No eixo Modelagem o foco das atividades foi a proposição de padrões e modelo de trabalho contextualizados na proposta arquitetural. A experimentação também foi utilizada para refinamento da arquitetura digital e elaboração de modelo referencial de processo de trabalho com integração dos diversos atores identificados na proposta da arquitetura. A natureza das atividades estudadas manteve, durante todo o percurso, proximidade com as metodologias e tecnologias de desenvolvimento de sistemas analíticos avançados (*bigdata*, *data lake*, ciência de dados etc.).

Essa abordagem metodológica foi operacionalizada por meio de estudo e análise do ambiente, reuniões de trabalho, experimentação de desenvolvimento de temas estratégicos, apoiada pela modelagem de processos de negócio (Business Process Modeling - BPM), adotado como instrumento de captura de fluxos de atividades e suas características.

Na linha de ação de Experimentação, ambos os modelos construídos nas linhas anteriores e suas implementações práticas são exercitados e implementados. Esses experimentos têm como matéria prima assuntos indicados

pelo Departamento Governança Institucional (DGI) da Secretaria Executiva (SEXEC/MCTI), denominados **temas estratégicos**, para os quais foi possível a interlocução com as áreas finais demandantes ou a disponibilização dos dados brutos. Cada tema estratégico exemplifica uma necessidade de informação estratégica do MCTI e dá origem a um produto de informação.

Os temas estratégicos trabalhados até o momento, seus respectivos produtos de informação e situação ao final de 2021 são mostrados na Tabela 1.

Tabela 1 - Temas estratégicos definidos pelo MCTI. Fonte: elaboração própria.

Tema estratégico	Produto de informação	Situação
Indicadores COICT	Banco de variáveis	Implementado
Lei do Bem	Painéis de informações (dashboards)	Implementadas versões 1 e 2.
FNDCT / Fundos Setoriais	Painéis de informações (dashboards)	Implementado

Consistente com as atuais metodologias ágeis, foi explorada a natureza iterativa de elaboração da arquitetura digital e sua implementação. A implantação de fluxos de dados, a identificação e internalização de ferramentas de software, a elaboração e validação de padrões e processos de trabalho, foram ações que se desenvolveram de forma interativa e cíclica com as equipes do Ministério. Isto se traduz em implantações de versões da arquitetura e seu contínuo aprimoramento e evolução à medida em que se desenvolvem as atividades em cada linha de ação da abordagem metodológica.

A abordagem metodológica central para os experimentos previstos no projeto (desenvolvimento de artefatos de inteligência de dados) foi o Ciclo de

Inteligência em CTI do CGEE (CGEE, 2017), representado graficamente na Figura 2.



Figura 2: Ciclo de inteligência em CTI. Fonte: (CGEE, 2017).

Na primeira etapa do ciclo procura-se compreender as necessidades de informação para a tomada de decisão, ou seja, expressar as incertezas e as dificuldades da organização em relação ao seu processo decisório. Estas incertezas são desdobradas em tópicos e questões chave de inteligência conhecidos que, ao serem respondidos, construirão uma estrutura orientadora para a coleta de informações relevantes. Na abordagem do CGEE, essas questões chave são denominadas perguntas norteadoras.

Se inicia nessa primeira etapa ações a elaboração da estrutura orientadora para responder à necessidade de informação, envolvendo a coleta e tratamento de dados e geração de informação, denominada de narrativas de respostas. Essas narrativas expressam alternativas metodológicas, potenciais fontes de dados, necessidades técnicas previstas inicialmente, até as proposições iniciais do “como” atender às perguntas norteadoras. Além disso, faz parte da primeira etapa do ciclo

a definição e objetivos, desafios e outras características associadas à tecnologia que poderá ser empregada para atender à necessidade de informação.

Na etapa seguinte, denominada Coleta e Armazenamento de Dados, são executadas as seguintes tarefas:

- obtenção de informações a partir de fontes primárias e secundárias;
- definição dos processos de coleta de informações;
- definição dos modelos analíticos que serão posteriormente utilizados, para planejar a organização do ambiente de armazenamento das informações.

Todos os dados e informações coletados são considerados inteligência bruta e, necessitam ser trabalhados para que o seu valor possa emergir na etapa de análise. Por mais qualidade que tenha uma informação, é muito mais o modo como ela será analisada e utilizada do que apenas a sua captura e disponibilização que determinará sua valia. Essas ações da segunda etapa podem ajustar, refinar e gerar novas perguntas norteadoras e correspondentes narrativas de respostas.

Na etapa de Análise dos Dados, transformam-se as informações coletadas em um produto de inteligência. O objetivo é definir o melhor ou os melhores métodos de análise das informações para a geração dos produtos de inteligência que se pretende. Neste momento as narrativas de respostas têm suas implementações realizadas com as tecnologias pensadas (ou preparadas) para o projeto em caráter piloto ou experimental. Por meio das informações reunidas, esta etapa visa a identificação de tendências e padrões significativos, ou seja, percepções exclusivas e conexões até então não relacionadas entre os dados. A condução do CGEE incorpora, também, nesta etapa a experimentação de mecanismos de visualização (por exemplo rascunhos visuais, ou *mock ups*) final sempre que possível.

A etapa da Produção de Resultados e Avaliações envolve a entrega do produto de inteligência, em um formato coerente, claro, objetivo aos clientes finais. Para que o uso ou disseminação dos resultados seja eficiente alguns aspectos precisam ser observados, como por exemplo, o melhor formato do documento a ser entregue

pelos profissionais de inteligência para os responsáveis pela tomada de decisão na organização.

A etapa da Avaliação da Informação tem dois objetivos:

- avaliar se o processo desenhado foi eficiente do ponto de vista da elaboração do produto de inteligência. Diz respeito ao desempenho de cada uma das etapas que compõem o ciclo de inteligência, isto é, se o melhor método de análise foi escolhido, se a escolha das fontes de informação poderia ter sido mais bem direcionada, se o formato do produto foi o mais adequado e assim por diante;
- avaliar a eficiência deste produto para o cliente final, ou seja, verificar os resultados práticos obtidos com o uso dos produtos gerados para o cliente de inteligência.

Estas duas avaliações são imprescindíveis tanto para o aprimoramento do processo quanto para a sua sobrevivência. A consolidação e o reconhecimento da utilidade deste processo só são possíveis a partir dos resultados de seus produtos na tomada de decisão. Caso as atividades do processo terminem na produção de resultados, a organização terá somente adquirido informação, uma vez que a inteligência somente ocorre quando os resultados do processo são utilizados na definição das ações organizacionais.

O percurso metodológico realizado observou em todos os seus componentes e passos as iniciativas de gestão de dados já conduzidas no MCTI e a integração dessas iniciativas como linhas condutoras para elaboração e implementação dos modelos arquitetural e de processo de trabalho.

4. Produtos resultantes

O resultado central do projeto foi o modelo denominado Arquitetura Digital de Inteligência de Negócios do MCTI cujo conceito central a constituição de um ambiente digital para gestão de dados que:

- suporta o armazenamento de fontes de informações heterogêneas,
- permite a aplicação de metodologias de tratamento e análise de dados a partir de conjuntos de dados de brutos,
- viabiliza a produção e visualização de informação com alto valor agregado para subsídio à tomada de decisão dos gestores e analistas do MCTI, e
- contribui para o alcance dos pilares de gestão de dados - integridade, confiabilidade, disponibilidade, autenticidade.

O cerne desse ambiente digital de informação é o modelo arquitetural em camadas, mostrado na Figura 3, que estrutura um modelo de processo de trabalho aplicável ao **tratamento de dados** e **produção de informação**, alinhado com os requisitos de gestão de dados, em especial sua governança.

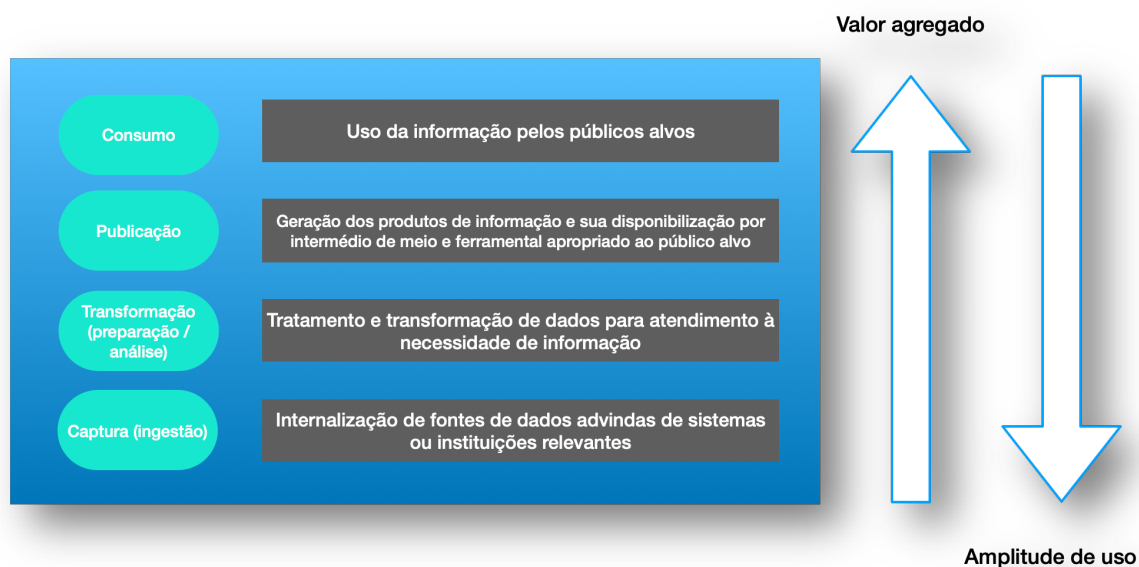


Figura 3 - Camadas do modelo arquitetural e características principais. Fonte: elaboração própria.

Cada camada do modelo organiza elementos estruturantes que compartilham de um determinado nível semântico associado a um objetivo de negócio. Ou seja, as camadas mais baixas agrupam conjuntos de dados e lógicas de tratamento de dados mais gerais com objetivo de receber dados brutos e prepará-los para o reuso. As camadas mais altas incorporam as necessidades específicas de informação, ou seja, objetivos de negócio, que se refletem em tratamentos e conjuntos de dados resultantes semanticamente alinhados com necessidades específicas de tomada de decisão. Por fim, a cada camada do modelo é acrescentado valor agregado ao dado até alcançar o produto de informação. Na Figura 4 são mostrados elementos conceituais e fronteiras da arquitetura digital.

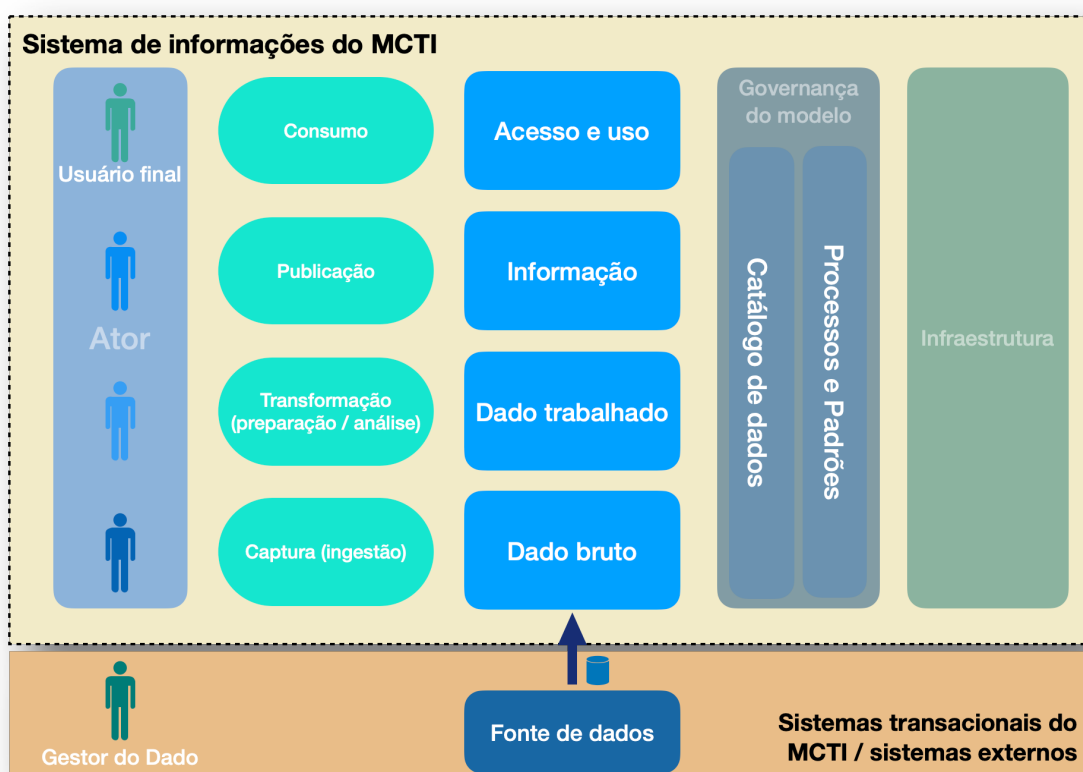


Figura 4 - Visão geral de componentes e fronteiras do modelo arquitetural. Fonte: elaboração própria.

Vale destacar que na camada denominada Governança do Modelo a atenção é dirigida aos itens mostrados na Figura 5. Consistem das disciplinas de governança de dados priorizadas no contexto do projeto.

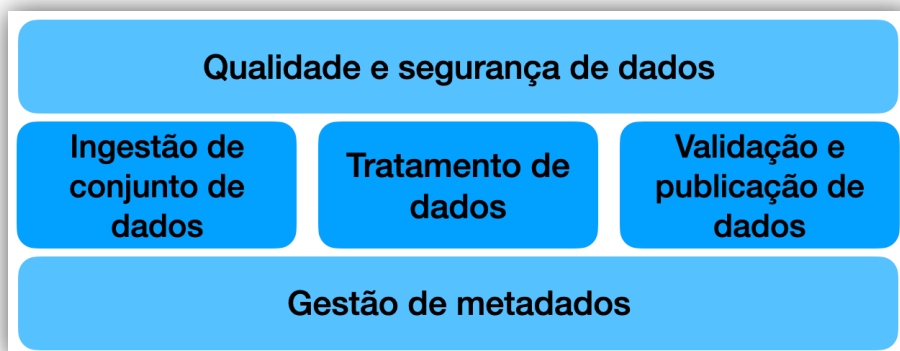


Figura 5 - Atividades em foco na Governança do Modelo. Fonte: elaboração própria.

O atendimento à demanda de informação dos usuários finais com o uso da arquitetura digital materializada no *data lake* do MCTI envolve atividades administrativas e técnicas. O primeiro tipo se refere à interação entre unidades e atores organizacionais na formulação da demanda, autorizações e providências de cunho administrativo. As atividades técnicas, por sua vez, são aquelas realizadas para o atendimento propriamente dito da necessidade de informação apresentada pelo usuário e são conduzidas com o apoio da arquitetura digital.

Para ambos os casos foi modelado um **processo de inteligência de negócio alinhado com a arquitetura de inteligência estratégica do MCTI**. O modelo geral do processo de trabalho é apresentado na Figura 6, onde se percebe integração de ações dos usuários finais e dos departamentos de Governança Institucional (DGI) e de Tecnologia da Informação (DTI) e outros atores relevantes no contexto de gestão de dados. O processo de trabalho organiza e descreve fluxos de tarefas que:

- Contempla as atividades necessárias para a elaboração de produtos de informação para subsídio à tomada de decisão dos gestores do MCTI;

- Promove o atendimento aos pilares de gestão de dados (integridade, confiabilidade, disponibilidade, autenticidade); e
- Está integrado com as características regimentais do Ministério.

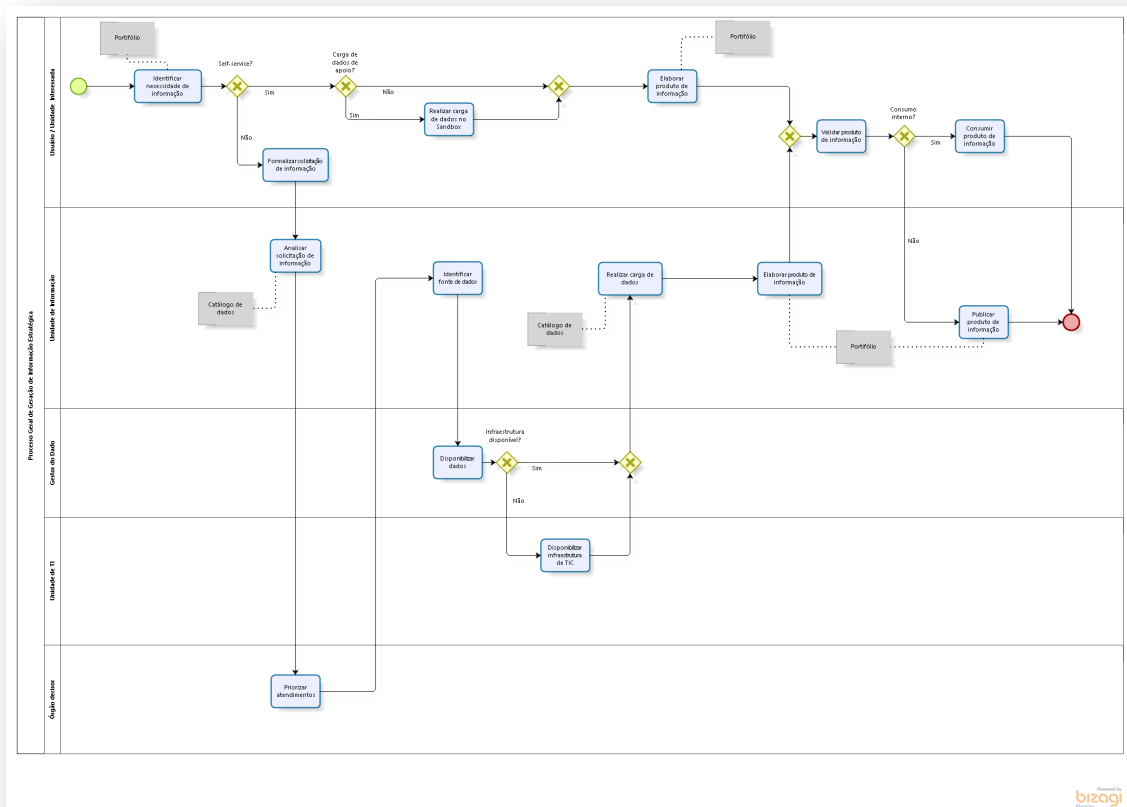


Figura 6 - Modelo de processo de trabalho sobre a Arquitetura Digital. Fonte: elaboração própria.

O objetivo da modelagem contempla, também, o estabelecimento das bases do “Sistema de informações do MCTI” iniciativa interna do Ministério descrita em (MCTI, 2020), que tem como missão a organização dos dados e informações dispersas no Órgão e suas entidades vinculadas e integração com objetivo de prover subsídio à tomada de decisão nos níveis estratégicos. O aparato digital e o processo de trabalho que deriva do modelo arquitetural proposto tem como objetivo subsidiar esse sistema analítico de informações, o qual está assentado sobre o conjunto de sistemas transacionais internos do MCTI ou externos ao MCTI

(incluindo aí aqueles sistemas transacionais das suas unidades vinculadas). Ainda que os sistemas transacionais / externos não façam parte do modelo arquitetural, eles se comunicam com a arquitetura digital por meio de fontes de dados.

A descrição detalhada da arquitetura digital é objeto do relatório (CGEE, 2022). A descrição do processo de trabalho está em (CGEE, 2022a).

4.1. Ambiente de inteligência de negócio do MCTI

4.1.1. Ambiente tecnológico

O modelo arquitetural funciona como um arcabouço lógico para a construção de fluxos de coleta e tratamento de dados e geração de informação com alto valor agregado. A implementação do modelo em termos de ferramentas de software e hardware foi conduzida pelo Departamento de Tecnologia da Informação (DTI) do MCTI com apoio do CGEE. O resultado é mostrado na Figura 7, e os detalhes descritos em (CGEE, 2022a).

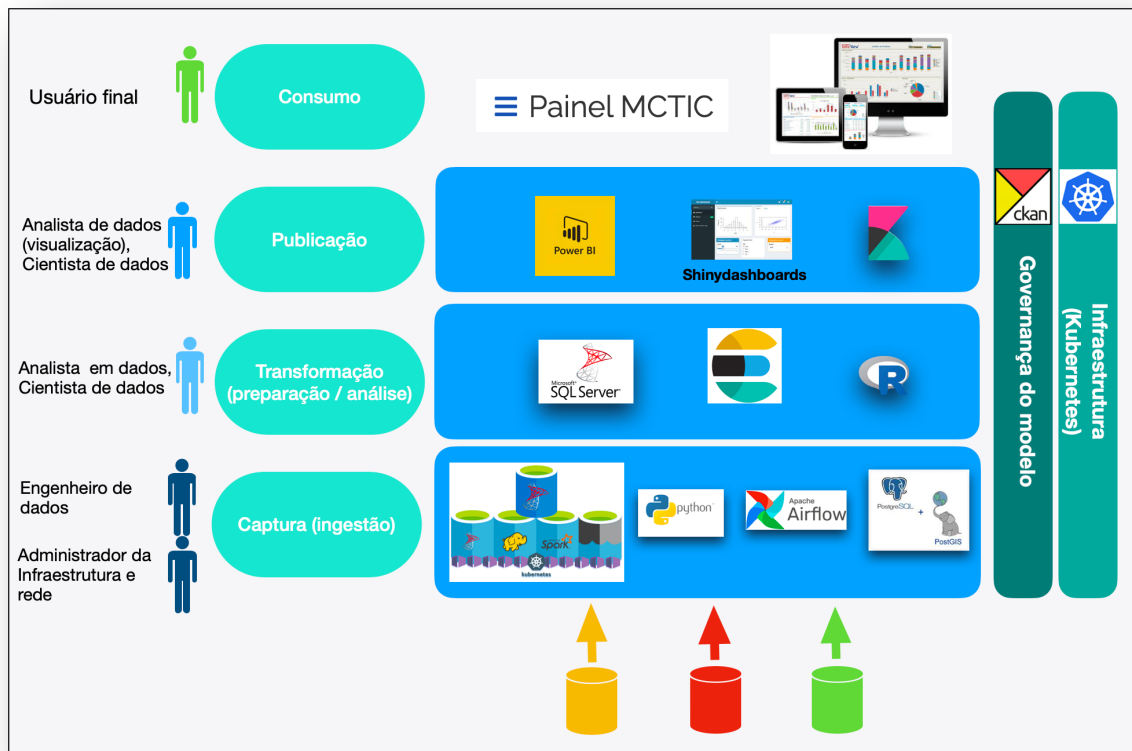


Figura 7 - Visão geral de componentes arquiteturais com respectivas implementações técnicas. Fonte: elaboração própria.

Essa infraestrutura é acessível aos usuários a partir do instrumento *Kit Data Lake*, cuja tela de entrada é mostrada na Figura 8.

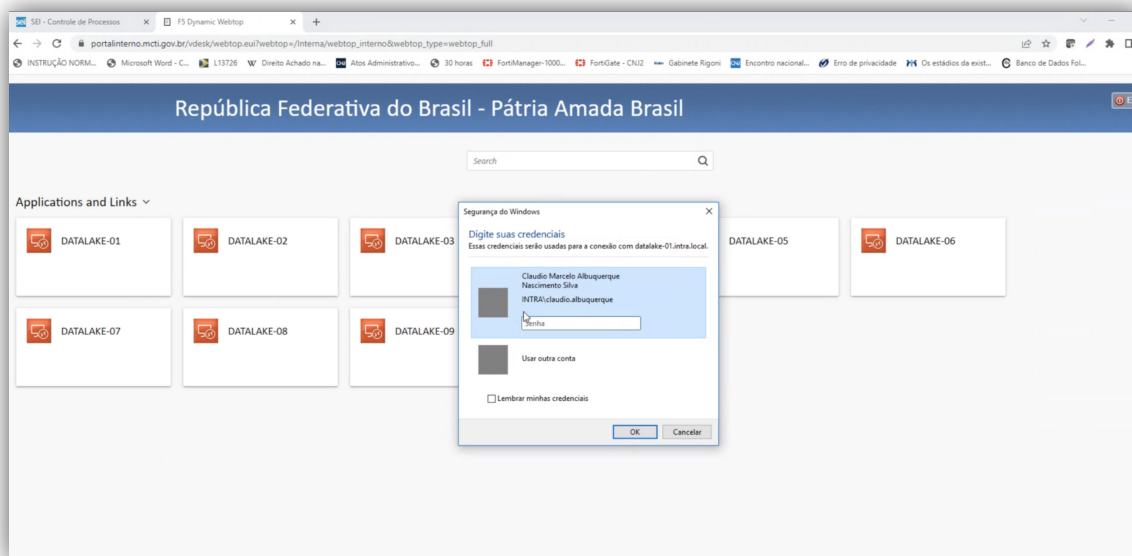


Figura 8 - Kit Data Lake do MCTI – ambiente de apoio à captura e transformação de dados. Fonte: DTI/MCTI.

4.1.2. Ambiente de informação

O ambiente tecnológico disponibilizado materializa o conceito do *data lake* do MCTI. Institui um ponto central de referência para repositório de dados (*big data*) controlado por padrões e processos de gestão de dados.

Para o usuário conhecer os recursos de dados disponibilizados no *data lake* foi implantada a ferramenta digital de catálogo de dados CKAN, resultado do processo de seleção dentre várias soluções pesquisadas que atendiam aos requisitos funcionais definidos. Dentre os pré-requisitos para a implantação do catálogo de dados foi estabelecido um padrão de metadados alinhado com o modelo arquitetural em camadas. Na Figura 9 é mostrada a tela inicial do CKAN configurado para o MCTI.

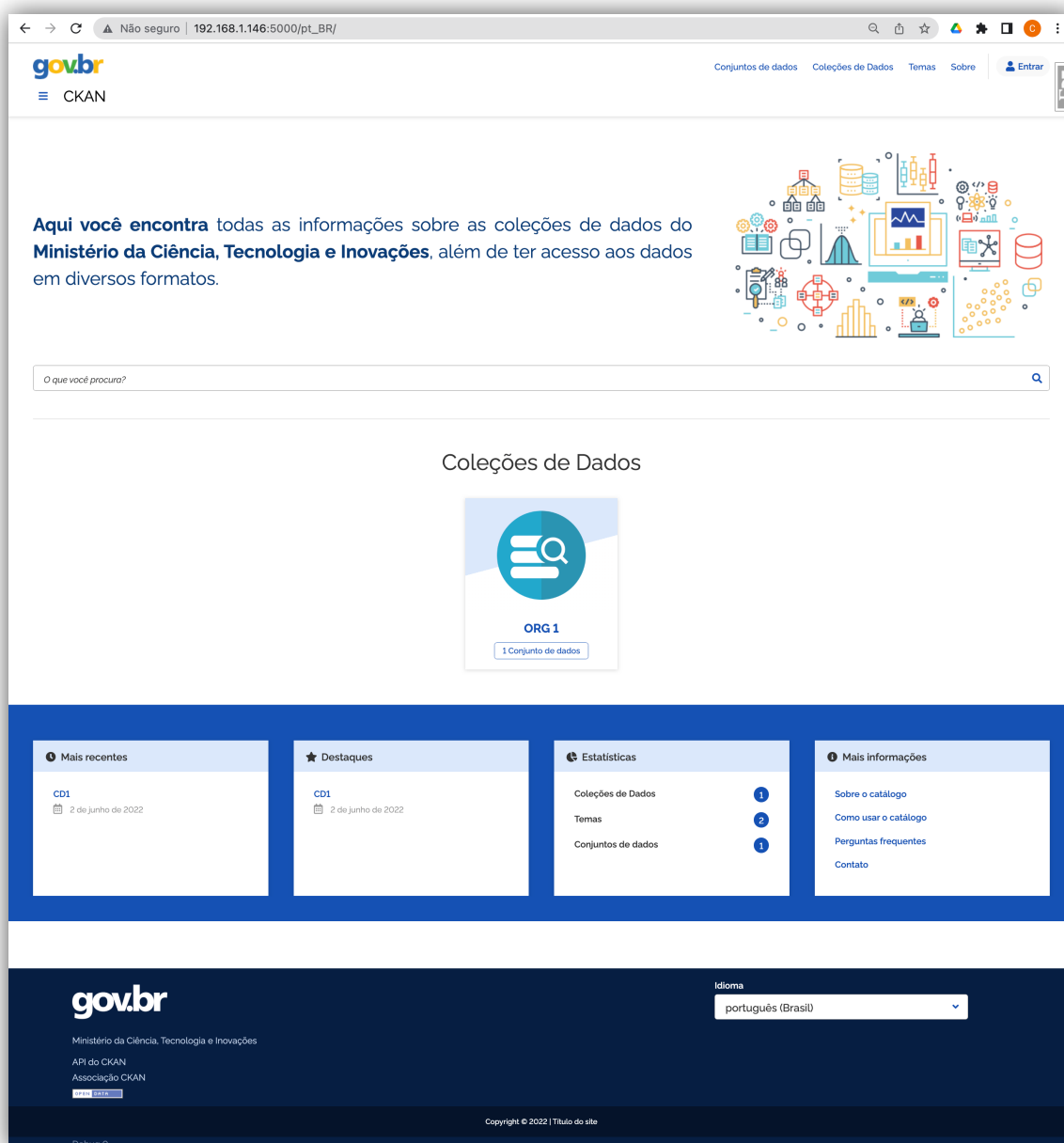


Figura 9 - Tela de entrada no CKAN configurado para o MCTI. Fonte: elaboração própria.

Outro fundamento necessário para facilitar consulta por parte do usuário e para estruturação do processo de governança de dados foi o estudo e explicitação da arquitetura da informação do MCTI. A partir dos estudos e experiências conduzidas foi possível a aplicação de uma arquitetura da informação com a estruturação temática dos ambientes de dados de caráter corporativo que permite

organizar os conjuntos de dados do *data lake* do MCTI. A Tabela 2 mostra a arquitetura da informação adotada.

Tabela 2 - Estrutura do conteúdo da arquitetura da informação inicial. Fonte: (COGCD, 2020).

ASSUNTO	TEMAS
POLÍTICAS SETORIAIS DE CT&I	<ul style="list-style-type: none"> – Política Nuclear – Política Espacial – Política de Energia, Petróleo e Mineração – Política de Inclusão Digital e Tecnologias Assistivas – Política de C&T para o Desenvolvimento Sustentável (Meio Ambiente/ Clima) – Política de Projetos na Fronteira Tecnológica – Política de Bens Sensíveis – Política de Transformação Digital – Bioma – Bioeconomia – Biotecnologia
ASPECTOS TRANSVERSAIS DA POLÍTICA DE CT&I	<ul style="list-style-type: none"> – Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI) – Propriedade Intelectual – Fomento em Ciência e Tecnologia – Extensão e Serviços Tecnológicos (ACTC¹) – Instrumentos de incentivos Fiscais e apoio ao P&D Empresarial – Fundos CTI – Ações de Empreendedorismo
ACOMPANHAMENTO DE ATIVIDADES DO MCTI	<ul style="list-style-type: none"> – Acompanhamento das Unidades de Pesquisa – Acompanhamento das Entidades Vinculadas

¹ Atividade Científica e Técnica Correlata

Outro importante benefício resultante da arquitetura da informação é subsidiar a elaboração de política de governança de dados, endereçando aspectos tais como definição de gestores dos dados, controle de acesso e auxílio na busca e reuso dos dados disponíveis no data lake são otimizados por meio da clara delimitação das categorias de informação inerentes à missão do Ministério.

4.2. Experimentos - Temas estratégicos

Os desenvolvimentos experimentais associados à Arquitetura Digital produziram insumos importantes para enriquecimento dos modelos arquiteturais e de processo. Os exercícios de implementação dos temas estratégicos viabilizaram, também, o estabelecimento da primeira aproximação para a arquitetura da informação subjacente à ação do Ministério, a seleção, configuração e integração do ferramental de software, a criação de conjuntos de dados de alto valor agregado. Outro resultado foi a disponibilização de painéis temáticos, que sob uma óptica mais prática, representa a aplicação e validação dos modelos, padrões processos de trabalho e ferramentas tecnológicas em produtos de informação implantados e disponibilizados para usuários internos.

A matéria prima de dados para os desenvolvimentos experimentais foram assuntos indicados pelo Departamento Governança Institucional (DGI) da Secretaria Executiva (SEXEC/MCTI), denominados temas estratégicos. Cada tema estratégico exemplificou uma necessidade de informação estratégica do MCTI e deram origem a um produto de informação.

Para o alcance dos produtos de informação, descritos na seção subsequente, foi empregado o Ciclo de Inteligência em CTI do CGEE (CGEE, 2017) em uma abordagem interativa e coparticipativa com interlocução constante com as áreas finais demandantes.

4.2.1. Painéis temáticos

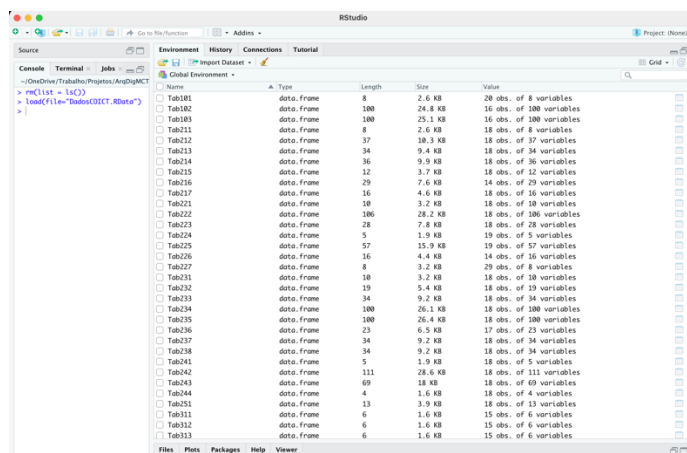
Os produtos de informação resultantes do desenvolvimento experimental foram de dois tipos: (a) painéis de informação, propriamente ditos, para os temas estratégicos selecionados; e (b) base de dados de indicadores em CTI. Esses produtos são mostrados na Tabela 3 e estão disponibilizados no endereço da internet abaixo indicado:

<https://paineis.mcti.gov.br>

Tabela 3 – Produtos de informação resultantes no período 2020/2021. Fonte: elaboração própria.

Tema estratégico	Tela de entrada	Tipo
Lei do Bem		Painéis de informação estratégica
FNDCT		Painéis de informação estratégica

Indicadores da COICT



Name	Type	Length	Size	Value
Tab101	data.frame	8	2.6 KB	20 obs. of 8 variables
Tab102	data.frame	100	24.8 KB	16 obs. of 100 variables
Tab103	data.frame	100	25.1 KB	16 obs. of 100 variables
Tab111	data.frame	8	2.6 KB	18 obs. of 8 variables
Tab112	data.frame	37	10.3 KB	18 obs. of 37 variables
Tab113	data.frame	34	9.4 KB	18 obs. of 34 variables
Tab114	data.frame	36	9.9 KB	18 obs. of 36 variables
Tab115	data.frame	12	3.7 KB	18 obs. of 12 variables
Tab116	data.frame	29	7.6 KB	14 obs. of 29 variables
Tab117	data.frame	16	4.6 KB	18 obs. of 16 variables
Tab121	data.frame	10	3.2 KB	18 obs. of 10 variables
Tab122	data.frame	106	28.2 KB	18 obs. of 106 variables
Tab123	data.frame	28	7.8 KB	18 obs. of 28 variables
Tab124	data.frame	5	1.9 KB	19 obs. of 5 variables
Tab125	data.frame	57	15.9 KB	19 obs. of 57 variables
Tab126	data.frame	16	4.4 KB	14 obs. of 16 variables
Tab127	data.frame	8	3.2 KB	29 obs. of 8 variables
Tab131	data.frame	10	3.2 KB	18 obs. of 10 variables
Tab132	data.frame	19	5.4 KB	18 obs. of 19 variables
Tab133	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab134	data.frame	100	26.1 KB	18 obs. of 100 variables
Tab135	data.frame	100	26.4 KB	18 obs. of 100 variables
Tab136	data.frame	23	6.5 KB	17 obs. of 23 variables
Tab137	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab138	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab141	data.frame	5	1.9 KB	18 obs. of 5 variables
Tab142	data.frame	111	28.8 KB	18 obs. of 111 variables
Tab143	data.frame	69	18 KB	18 obs. of 69 variables
Tab144	data.frame	4	1.6 KB	18 obs. of 4 variables
Tab151	data.frame	13	3.9 KB	18 obs. of 13 variables
Tab111	data.frame	6	1.6 KB	15 obs. of 6 variables
Tab112	data.frame	6	1.6 KB	15 obs. of 6 variables
Tab113	data.frame	6	1.6 KB	15 obs. of 6 variables

Base de variáveis

4.2.2. Indicadores da COICT

O caso específico do tema estratégico Indicadores da COICT – Coordenação de Indicadores de Ciência, Tecnologia e Inovação do Departamento de Governança Institucional, pelas características específicas da demanda, teve condução, resultados e forma de documentação diferenciada.

A demanda, nesse caso, consistiu em captar e internalizar o conjunto de indicadores e variáveis já existentes e disponíveis na página Web do Ministério "Indicadores Nacionais de Ciência, Tecnologia Inovação"². A expectativa do demandante era “apoio do CGEE para incorporar a base de dados da COIND³ e demais bases à Infraestrutura da DTI” (MCTI, 2020).

Essa expectativa resultou nas seguintes especificações e critérios de aceitação:

- Conversão de formato, de arquivos no formato CSV, para uma base de dados de variáveis facilmente manipulável pela ferramenta estatística R / RStudio.

² Sítio Web “Indicadores Nacionais de Ciência, Tecnologia e Inovações” disponível até 2021 em https://antigo.mctic.gov.br/mctic/opencms/indicadores/indicadores_cti.html e a partir de 2022 publicado disponível em <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/indicadores>.

³ Antiga denominação da COICT/DGI.

- Integração das variáveis em um único repositório, conservando ao máximo os metadados existentes na fonte de dados origem, de modo a documentar as variáveis.

Esses objetivos orientaram uma condução de trabalhos distinta dos demais temas estratégicos, uma vez que as variáveis já constituem a informação desejada, não havendo uma necessidade de informação a ser atendida. Assim, a condução se voltou para o trabalho de captura de dados a partir de uma página web (*web scraping*⁴), a ser realizado uma única vez, tratamento (conversão de forma, extração de metadados e outras características, limpeza de caracteres especiais) e ingestão na Arquitetura Digital.

O resultado foi a construção de um repositório com 5.083 variáveis, dispostos nos formatos JSON e RData e documentados no formato JSON⁵ com os seguintes atributos: nome da tabela original, título, subtítulo, descrição, fontes de dados, notas de rodapé, variáveis (metadados: nome para a variável gerado automaticamente, título da coluna da tabela original referente à variável, referências de rodapé específicas da variável).

⁴ Web scraping: raspagem web, é uma forma de mineração que permite a extração de dados de sites da web convertendo-os em informação estruturada para posterior análise. (Wikipedia).

⁵ JSON (JavaScript Object Notation - Notação de Objetos JavaScript) é uma formatação leve de troca de dados. Para seres humanos, é fácil de ler e escrever. Para máquinas, é fácil de interpretar e gerar. (<http://json.org/json-pt.html>)


5. Resultados alcançados

Em termos de resultados alcançados, complementares aos produtos previstos nos objetivos do projeto, estão recursos de dados e informações, e aprendizados organizacionais.

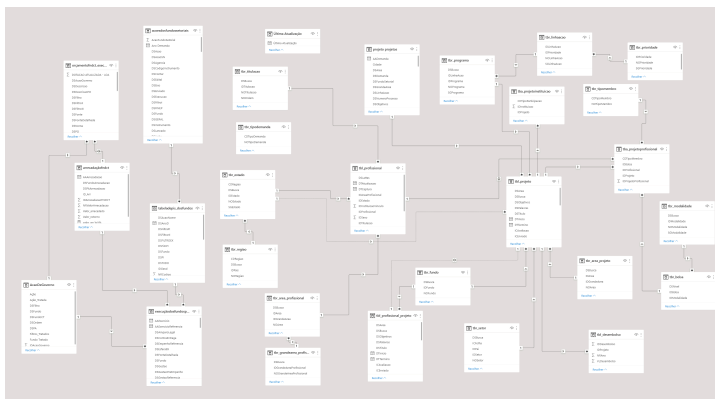
5.1. Fontes de dados internas

Três conjuntos de dados na camada de transformação foram gerados (Tabela 4). Esses produtos intermediários configuram exemplos de dados transformados e preparados para uso geração de produtos de informação. Os formatos de armazenamento escolhidos estão associados aos objetivos finais em cada tema estratégico. O acréscimo continuado de conjuntos de dados como esses no *Data Lake* do MCTI, com sua respectiva documentação no catálogo de dados, produzirá um grande repositório de informações de negócio para o Ministério.

Tabela 4 - Bases de dados gerados na camada de transformação. Fonte: elaboração própria.

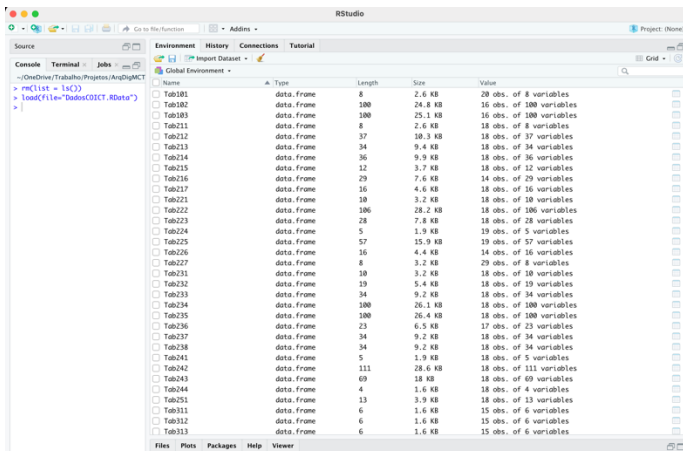
Tema estratégico	Tela de entrada	Formato	Informação de negócio
Lei do Bem		Banco de dados relacional	Base de dados da Lei do Bem extraído dos sistemas transacionais FormP&D e FORMS. Contém o histórico de dados coletados junto aos beneficiários da Lei do Bem para os anos base 2009 à 2021. Os dados foram normalizados e ajustados para um padrão único de cobertura para todos os anos base.

FNDCT



Banco de dados relacional Base de dados de execução orçamentária e projetos do FNDCT. Dados extraídos da aplicação operacional em Access utilizada na Coordenação-Geral de Governança de Fundos – CGGF/DGI, integrados com os dados do sistema SIGCTI (<https://sigcti.mcti.gov.br/>).

Indicadores da COICT



Name	Type	Length	Size	Value
Tab181	data.frame	8	2.6 KB	29 obs. of 8 variables
Tab182	data.frame	100	24.8 KB	16 obs. of 100 variables
Tab183	data.frame	100	25.1 KB	16 obs. of 100 variables
Tab111	data.frame	8	2.6 KB	18 obs. of 8 variables
Tab212	data.frame	37	10.3 KB	18 obs. of 37 variables
Tab213	data.frame	34	9.4 KB	18 obs. of 34 variables
Tab214	data.frame	36	9.9 KB	18 obs. of 36 variables
Tab215	data.frame	12	3.7 KB	18 obs. of 12 variables
Tab216	data.frame	29	7.6 KB	14 obs. of 29 variables
Tab217	data.frame	16	4.6 KB	18 obs. of 16 variables
Tab221	data.frame	10	3.2 KB	18 obs. of 10 variables
Tab222	data.frame	106	28.2 KB	18 obs. of 106 variables
Tab223	data.frame	28	7.8 KB	18 obs. of 28 variables
Tab224	data.frame	5	1.9 KB	19 obs. of 5 variables
Tab225	data.frame	57	15.9 KB	19 obs. of 57 variables
Tab226	data.frame	16	4.4 KB	14 obs. of 16 variables
Tab227	data.frame	8	3.2 KB	29 obs. of 8 variables
Tab231	data.frame	10	3.2 KB	18 obs. of 10 variables
Tab232	data.frame	19	5.4 KB	18 obs. of 19 variables
Tab233	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab234	data.frame	100	26.1 KB	18 obs. of 100 variables
Tab235	data.frame	100	26.4 KB	18 obs. of 100 variables
Tab236	data.frame	23	6.5 KB	17 obs. of 23 variables
Tab237	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab238	data.frame	34	9.2 KB	18 obs. of 34 variables
Tab241	data.frame	5	1.9 KB	18 obs. of 5 variables
Tab242	data.frame	111	28.6 KB	18 obs. of 111 variables
Tab243	data.frame	69	18 KB	18 obs. of 69 variables
Tab244	data.frame	4	1.6 KB	18 obs. of 4 variables
Tab251	data.frame	13	3.9 KB	18 obs. of 13 variables
Tab311	data.frame	6	1.6 KB	15 obs. of 6 variables
Tab312	data.frame	6	1.6 KB	15 obs. of 6 variables
Tab313	data.frame	6	1.6 KB	15 obs. of 6 variables

RDATA e JSON Base de variáveis com 5.083 variáveis advindas do sítio Web “Indicadores Nacionais de Ciência, Tecnologia e Inovações”.

5.2. Realizações derivadas

A necessidade de acesso a dados, principal insumo para o desenvolvimento das atividades do projeto, colocou em evidência o papel da atividade de **articulação** interna e externa entre as unidades organizacionais e entre organizações. A boa condução de um projeto de desenvolvimento de produtos de informação é função direta da capacidade de disponibilizar dados e manter canais operacionais de funcionamento da geração, transferência e uso dos dados. E a articulação entre atores envolvidos é o meio pelo qual se torna concreta a disponibilização e atualização continuada da fonte de dados.

As dificuldades de acesso à fonte de dados foi fator de risco preponderante da capacidade de construção de produtos de informação. Temas estratégicos apontados no início do projeto não foram viabilizados pela indisponibilidade da fonte de dados, ainda que a demanda pela informação estratégica estivesse presente.

Esse aprendizado corrobora as iniciativas do DGI e reforça sua intenção em se consolidar como ponto de parceria para construção de articulação com unidades organizacionais internas, unidades vinculadas e demais organizações externas que detem fontes de dados de interesse dos gestores do MCTI.

O MCTI possui um conjunto de sistemas transacionais que apoiam de forma digital atividades das suas atribuições regimentais, tais como a condução de programas e políticas públicas no SNCTI. Esses sistemas transacionais apresentaram dois principais desafios quando da necessidade de extração de dados operacionais neles residentes. O primeiro é exemplificado pelo caso dos sistemas FormP&D e Forms, que implementam o sistema de acompanhamento da Lei do Bem. Esses sistemas se apoiam em uma base de dados voltada para montagem de formulários eletrônicos (base de dados de metamodelo). As consultas sobre essas bases são de extrema complexidade pois tem que remontar o questionário em tempo de consulta. Nesse caso, o projeto implementou duas esteiras de transformação de dados (ETL) que lê o

metamodelo e armazena os dados em um modelo relacional tradicional, voltado para a consultas e construção de relatórios e painéis gráficos.

O segundo desafio está afeito a casos em que os dados dos sistemas transacionais precisam de intervenção humana para complementos e correções. As unidades organizacionais respectivas se apoiam em planilhas de dados complementares para montagem do cenário correto de informação operacional. A decisão de projeto foi não utilizar esses casos como temas estratégicos e introduzir o desenvolvimento de ETL automatizados em projetos subsequentes e focados em cada cenário. Esse encaminhamento se mostra adequado em vista dos padrões de governança de dados que estão em consolidação no MCTI.

6. Conclusão

Na perspectiva operacional do projeto, a Tabela 5 apresenta um resumo dos produtos entregues em suas respectivas categorias.

Tabela 5 - Situação de componentes constitutivos dos produtos previstos no projeto. Fonte: elaboração própria.

Produto previsto	Componente	Sub-componente
Relatório da modernização da arquitetura de ambiente digital	Modelo arquitetural em camadas	
	Documentos de governança de dados	Organização de conjuntos de dados
		Padrões
		Nomenclatura
	Ferramenta de catálogo de dados	Catálogo de dados (CKAN)
Relatório detalhando o processo de gestão de ambiente digital de informação estratégica	Processo de trabalho	Processo geral de geração de informação estratégica e 5 subprocessos correlatos.
Relatório final da modernização da arquitetura digital de inteligência de negócio do MCTI, contemplando painéis temáticos elaborados	Indicadores COICT	Base de indicadores
	Painéis Lei do Bem	Painéis versão 1 e versão 2, base de dados transformados (anos base 2009 a 2021)
	Painéis FNDCT	Painel de informação e base de dados integrada (Access e SIGCTI)

Foram alcançados marcos importantes do projeto, relacionados abaixo:

- Elaboração e revisão de modelo arquitetural para sistemas analíticos com uso de *data lake*, frente à experiência prática nas dependências do MCTI.
- Plataforma digital implementada, pelo DTI/SEXEC/MCT com apoio do CGEE, contemplando a infraestrutura de hardware e as ferramentas de software que implementam as funções de ingestão de dados, fluxos de tratamento de dados, visualização de informação, catálogo de dados e controle de acesso.
- Arquitetura de dados, padrões e processo de trabalho integrado com a plataforma digital implementada.
- Implantação de três (dos oito) temas estratégicos previstos, a saber: Lei do Bem, FNDCT e Indicadores da COICT.
- Estruturação inicial de arquitetura da informação que contextualiza os temas estratégicos implementados e apoia a curadoria de dados para a Arquitetura Digital do MCTI.
- Implementação da segunda versão dos painéis da Lei do Bem, e projeto dos painéis da Lei de Informática e Programa Padis.

Em paralelo foi concluída nova versão do modelo de processos alinhado com a Arquitetura Digital, debatida com o Departamento de Tecnologia da Informação, incluindo o desenho da arquitetura de informação que o *data lake* implementa considerando os quatro temas estratégicos trabalhados. Outro resultado alcançado foi a definição do CKAN como ferramenta de catálogo de dados e sua implantação na infraestrutura de TIC do Ministério com apoio do CGEE. Essa decisão é relevante tendo em vista a importância que um catálogo de dados tem para uma arquitetura digital para *data lake*.

Os resultados alcançam objetivos maiores do projeto na medida que experimenta o suporte e o armazenamento a fontes de informações heterogêneas. Além disso, proporciona arcabouço instrumental organizado e

controlado para aplicação de metodologias de tratamento e análise de dados de brutos, de variados formatos. Os exercícios também permitiram a experimentação da interoperabilidade da Arquitetura Digital com sistemas de informação legados do MCTIC e fontes de informação externas ao ministério.

Os temas estratégicos definidos para o trabalho enriqueceram a Arquitetura Digital com:

- Definição de padrão de documentação para linhagem de dados;
- Ajustes no alinhamento de conteúdo nas camadas arquiteturais;
- Estabelecimento dos primeiros conjuntos de dados, nas camadas de Ingestão e Transformação;
- Validação da integração entre as ferramentas de uso cotidiano do MCTI com a Arquitetura Digital;
- Execução das tarefas previstas no modelo de trabalho, com esclarecimentos sobre objetivos das tarefas, sua natureza e atores envolvidos e responsáveis;
- Homogeneização do conhecimento sobre as dimensões do ciclo de vida de dados e a relação com a Arquitetura Digital, assim como os conceitos, componentes da arquitetura, modelo de trabalho e responsabilidades.

Os resultados diretos (hardware, software, metodologias organizados e integrados por um processo de trabalho), assim como os resultados indiretos (aprendizados, experiências e internalização de conceitos e habilidades) permitem manifestar que o objetivo primeiro do projeto foi realizado. O MCTI hoje possui um ambiente digital que suporta o armazenamento de fontes de informações heterogêneas que permite a aplicação de metodologias de análise inteligente de dados. Esse ambiente se alinha às tendências tecnológicas de mercado que reconhece a diversidade de fontes de dados, coexistência de ferramentas de software, aplicabilidade de técnicas modernas de aprendizado de máquina. A arquitetura digital está alinhada com as boas práticas de governança de dados, quando conta com padrões e processos de trabalho que organizam a entrada e acesso aos dados disponíveis no data lake do MCTI.

Foram exploradas as tendências tecnológicas e alternativas de mercado para estabelecimento desses elementos (arquitetura tecnológica e processos de governança de dados) que reforça e instrumentaliza a capacidade do MCTI na realização de inteligência de dados e produção de informação estratégica.

7. Referências Bibliográficas

- BRASIL, 2019 BRASIL. DECRETO Nº 10.046, DE 9 DE OUTUBRO DE 2019. Diário Oficial da União. Brasília, DF. 10 de outubro de 2019.
- CGEE, 2017 CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. Desenho e detalhamento do primeiro nível do metaprocesso Inteligência Estratégica em CTI. Brasília, DF: Centro de Gestão e Estudos Estratégicos, 2017.
- CGEE, 2019 CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. Plano de trabalho para o desenvolvimento da arquitetura digital de inteligência de negócio do MCTIC. In: Arquitetura digital de inteligência de negócios do MCTIC. Brasília, DF: Centro de Gestão e Estudos Estratégicos, dezembro / 2019.
- CGEE, 2020 CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. Projeto Arquitetura Digital de Inteligência de Negócio do MCTIC - Execução do plano de trabalho. In: Arquitetura digital de inteligência de negócios do MCTIC. Brasília, DF: Centro de Gestão e Estudos Estratégicos, maio / 2020.
- CGEE, 2022 CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. Relatório da modernização da arquitetura de ambiente digital. In: Arquitetura digital de inteligência de negócios do MCTIC. Brasília, DF: Centro de Gestão e Estudos Estratégicos, junho / 2022.

em: <<http://www.ime.usp.br/~vwsetzer>>. Acesso em: 12 jun.
2004. Coleção Ensaio Transversais.

Anexo I – Glossário

Termo	Uso no projeto	Referências complementares
Arquitetura digital	<i>Modelo lógico (desenho) organização dos atores e elementos que compõem a solução tecnológica proposta, bem como os relacionamentos entre esses elementos e atores. Inclui os elementos tecnológicos, atores e processo de trabalho</i>	(DAMA, 2017)
Big data	<i>Conceito amplo que aborda as alterações tecnológicas ocorridas na última década que objetivam habilitar e viabilizar a geração, armazenamento e análise de crescentes volumes de dados. Os critérios mais utilizados para caracterizar a aplicação de big data é endereçar grandes volumes de dados, ampla variedade de dados, velocidade de produção de resultados. Outros critérios utilizados lidam com a veracidade da informação e com o seu valor (impactos resultantes do uso das informações).</i>	Elaboração própria a partir de (DAMA, 2017).
Business intelligence	<i>Categoria de análise de dados que tem como alvo a compreensão das atividades e oportunidades organizacionais por meio de informação sobre produtos, serviços e clientes subsidiando a tomada de decisões para atingir objetivos estratégicos. As ferramentas utilizadas para a realização dessas análises, consideradas ferramentas de business intelligence (BI), implementam suporte à consulta, mineração e visualização de dados. Dentre os instrumentos mais utilizados para BI está os data warehouses, repositórios de grandes quantidades de dados estruturados com foco na análise de negócio, criação de relatórios e painéis de informação. No histórico recente o instrumental de BI foi complementado com a introdução de data</i>	Elaboração própria a partir de (DAMA, 2017).

	<p><i>lakes, que permitem a realização de análises aprofundadas de dados (análise preditiva, analítica sobre dados não estruturados e geração de insight) a partir de aprendizado de máquina sobre dados estruturados ou não estruturados.</i></p>	
Conjunto de dados	<p><i>Um arquivo digital de dados de qualquer formato (estruturado ou não estruturado), armazenado em um meio digital e acessível por intermédio de uma ferramenta de software.</i></p>	
Dado	<p><i>Sequência de símbolos ou valores, representados em algum meio, produzidos como resultado de um processo natural ou artificial. Entende-se que dados são observações ou o resultado de uma medida (por investigação, cálculo ou pesquisa) de aspectos característicos da natureza, estado ou condição de algo de interesse, que são descritos através de representações formais e, ao serem apresentados de forma direta ou indireta à consciência, servem de base ou pressuposto no processo cognitivo.</i></p>	<p>(HOUAISS, 2001; SETZER, 2001)</p>
Data lake	<p><i>Um ambiente digital onde um vasto volume de dados de diferentes tipos e estruturas pode ser ingerido (internalizado), armazenado, acessado e analisado. Faz parte de um Data lake um conjunto variado e integrado de ferramentas de hardware e software com foco em tratamento de dados. Uma atividade crítica associada a data lakes é a administração de metadados para garantir a organização, correção e consistência dos dados ingeridos e disponibilizados aos usuários.</i></p>	<p>Elaboração própria a partir de (DAMA, 2017).</p>
Ferramenta	<p><i>Artefato de software adquirido, internalizado (software livre) ou desenvolvido internamente, que implementa uma funcionalidade.</i></p>	

Fonte de dados	<i>Um provedor de dados externo ao sistema analítico representado pela Arquitetura Digital. Um provedor pode ser um sistema transacional interno ou externo ao MCTI, serviços ou repositórios digitais de dados pagos ou aberto.</i>	Definição do projeto.
Governança de dados	<i>Exercício de autoridade e controle que permite o gerenciamento de dados sob as perspectivas do compartilhamento, da arquitetura, da segurança, da qualidade, da operação e de outros aspectos tecnológicos</i>	(BRASIL, 2019)
Informação	<i>Dados, processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, contidos em qualquer meio, suporte ou formato;</i>	(BRASIL, 2019)
Metadado	<i>Conjunto de informações descritivas sobre os dados, incluindo as características do seu levantamento, produção, qualidade e estrutura de armazenamento, essenciais para promover a sua documentação, integração e disponibilização, bem como possibilitar a sua busca e exploração.</i>	(MPOG, 2018)
Objeto do repositório digital	<i>Um item residente no repositório digital. Pode ser um arquivo, uma pasta digital, uma tabela de banco de dados, um esquema de banco de dados, índices, painéis, gráficos etc. Qualquer um destes itens sobre o qual há interesse em administrar e tornar disponível para uso.</i>	
Plataforma digital	<i>Implementação do arcabouço conceitual da arquitetura digital no que tange aos elementos de Tecnologia da Informação. Conjunto integrado e funcional de hardware e software que implementa fisicamente a Arquitetura Digital do MCTI e hospeda o data lake do Ministério.</i>	

Produto de dados	<i>Conjunto de dados tratados com uma finalidade preestabelecida, disponibilizado via qualquer meio digital, que pode ser incorporado ou usado por outra aplicação digital.</i>
Produto de informação	<i>Conjunto de dados tratados, disponibilizado no seu estado final de valor agregado para consumo humano, e que atende uma ou mais necessidades de informação.</i>
Repositório digital; ou repositório digital de referência	<i>Parte da Plataforma digital que implementa a função de armazenamento de dados brutos (fontes de dados originais) e trabalhados (produtos intermediários ou finais) que requeiram armazenamento. Está sujeito a controle de acesso.</i>
