



## **Projeto Arquitetura Digital Inteligência de Negócio do MCTIC**

**Relatório da modernização da arquitetura digital de inteligência de negócio do MCTIC, contemplando oito painéis temáticos**

**Relatório Preliminar**

# **Projeto Arquitetura Digital Inteligência de Negócio do MCTIC**

**Relatório da modernização da arquitetura digital de inteligência  
de negócio do MCTIC, contemplando oito painéis temáticos:  
Relatório preliminar**



Brasília, DF  
Dezembro, 2020

## Centro de Gestão e Estudos Estratégicos

### Presidente

*Marcio de Miranda Santos*

### Diretores

*Luiz Arnaldo Pereira da Cunha*

*Regina Silvério*

Relatório da modernização da arquitetura digital de inteligência de negócio do MCTIC, contemplando oito painéis temáticos: Relatório preliminar. Brasília: Centro de Gestão e Estudos Estratégicos, 2020.

85 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 – 21º 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 Arquitetura Digital Inteligência de Negócio do MCTIC**

**Relatório da modernização da arquitetura digital de inteligência de negócio do MCTIC, contemplando oito painéis temáticos:  
Relatório preliminar**

## **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

<b>1.</b>	<b>INTRODUÇÃO .....</b>	<b>1</b>
<b>2.</b>	<b>OBJETIVOS.....</b>	<b>4</b>
<b>3.</b>	<b>METODOLOGIA .....</b>	<b>5</b>
<b>4.</b>	<b>A ARQUITETURA DIGITAL DE INTELIGÊNCIA DE NEGÓCIOS DO MCTI 10</b>	
<b>5.</b>	<b>O EXERCÍCIO DE VALIDAÇÃO DA ARQUITETURA DIGITAL.....</b>	<b>20</b>
<b>5.1.</b>	<b>PROCESSO DE INTELIGÊNCIA DE DADOS E LINHAGEM DE DADOS....</b>	<b>22</b>
<b>5.2.</b>	<b>TEMAS ESTRATÉGICOS IMPLEMENTADOS - PAINÉIS DE INFORMAÇÃO .....</b>	<b>24</b>
<b>5.3.</b>	<b>O CASO ESPECÍFICO DO TEMA ESTRATÉGICO INDICADORES DA COICT<sup>26</sup></b>	
<b>6.</b>	<b>CONCLUSÕES E PRÓXIMOS PASSOS .....</b>	<b>28</b>
<b>7.</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>31</b>
	<b>ANEXO I – AJUSTE DE CRONOGRAMA DO PROJETO .....</b>	<b>33</b>
	<b>ANEXO II – TEMA ESTRATÉGICO LEI DO BEM.....</b>	<b>35</b>
	<b>ANEXO III – TEMA ESTRATÉGICO FNDCT / FUNDOS SETORIAIS .....</b>	<b>54</b>
	<b>ANEXO V – TEMA ESTRATÉGICO INDICADORES DA COICT .....</b>	<b>81</b>

## 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.

Ressalta-se dentre as atividades do MCTI o papel proeminente na gestão do Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI). Considerando a complexidade desse Sistema, a orquestração das ações buscando maximizar resultados de interesse social e econômico para o Brasil requer a integração da informação distribuída nos atores sistêmicos relevantes.

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 do texto que consiste em solução metodológica e de software sobre hardware pré-existente, que objetiva 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 dos Sistemas Nacional de Ciência, Tecnologia e Inovação e Sistema de Comunicações.

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.

Para o alcance desses objetivos, a condução do projeto foi estruturada em três linhas de ação principais, conforme abaixo relacionadas:

- Arquitetura: Atividades de pesquisa, conceituação e avaliação de alternativas tecnológicas para a proposição da arquitetura digital de inteligência de negócio.
- Modelagem: Atividades de desenvolvimento experimental e implementação voltados para a arquitetura tecnológica e modelo de processos de trabalho.
- Experimentação: Atividades de desenvolvimento experimental sobre a arquitetura digital elaborada com foco no atendimento de temas estratégicos estabelecidos pelo MCTI.

Neste relatório são apresentados os resultados alcançados na trilha de Experimentação e contempla a implementação dos temas estratégicos estabelecidos como prioridades para o MCTIC, materializadas sobre a

Arquitetura Digital que foi implantada nas instalações físicas de Tecnologia da Informação do MCTI.

Esse trabalho, realizado em parceria com os Departamentos de Governança Institucional (DGI) e de Tecnologia da Informação (DTI), ambos da SEXEC-MCTI, terá continuidade no decorrer de 2021, tendo em vista as decisões de governança adotadas em reunião de 10 de dezembro de 2020. No decorrer do ano de 2020, os seguintes fatos relevantes e de alto impacto sobre o desenvolvimento dos trabalhos no CGEE e no MCTI motivaram o ajuste de cronograma:

- A pandemia do Covid-19;
- Ajustes do Termo de Referência do Contrato de Gestão no primeiro semestre de 2020;
- A Medida Provisória Nº 980, DE 10 DE JUNHO DE 2020, que altera a Lei nº 13.844, de 18 de junho de 2019, para criar o Ministério da Ciência, Tecnologia e Inovações e o Ministério das Comunicações;
- Definição dos temas estratégicos realizado em 06/08/2020;
- Definição de prorrogação do prazo do projeto conforme 23º Termo Aditivo em processo de assinatura; e
- As tratativas vigentes sobre o novo Contrato de Gestão, que trará ajustes no prazo final para o projeto.

No Anexo I é apresentado o cronograma final de trabalho.



## 2. Objetivos

O objetivo deste relatório é apresentar os resultados da realização desses desenvolvimentos experimentais de produtos de informação com o uso da Arquitetura Digital de Inteligência de Negócio do MCTI. Essa atividade é prevista na linha de ação “Experimentação” da abordagem metodológica do projeto.

Essa linha de ação tem como objetivo a disponibilização de painéis temáticos, onde serão desenhados e implementados, em caráter piloto, pelo menos 8 (oito) painéis selecionados, incluindo seus elementos estruturantes padronizados para uso posterior pelo MCTI.

### 3. Metodologia

O detalhamento completo do projeto com as descrições de suas fases e distribuição no tempo estão registradas em (CGEE, 2019 e 2020), onde é descrita a abordagem metodológica na forma de linhas de ação, conforme mostrado na Figura 1.

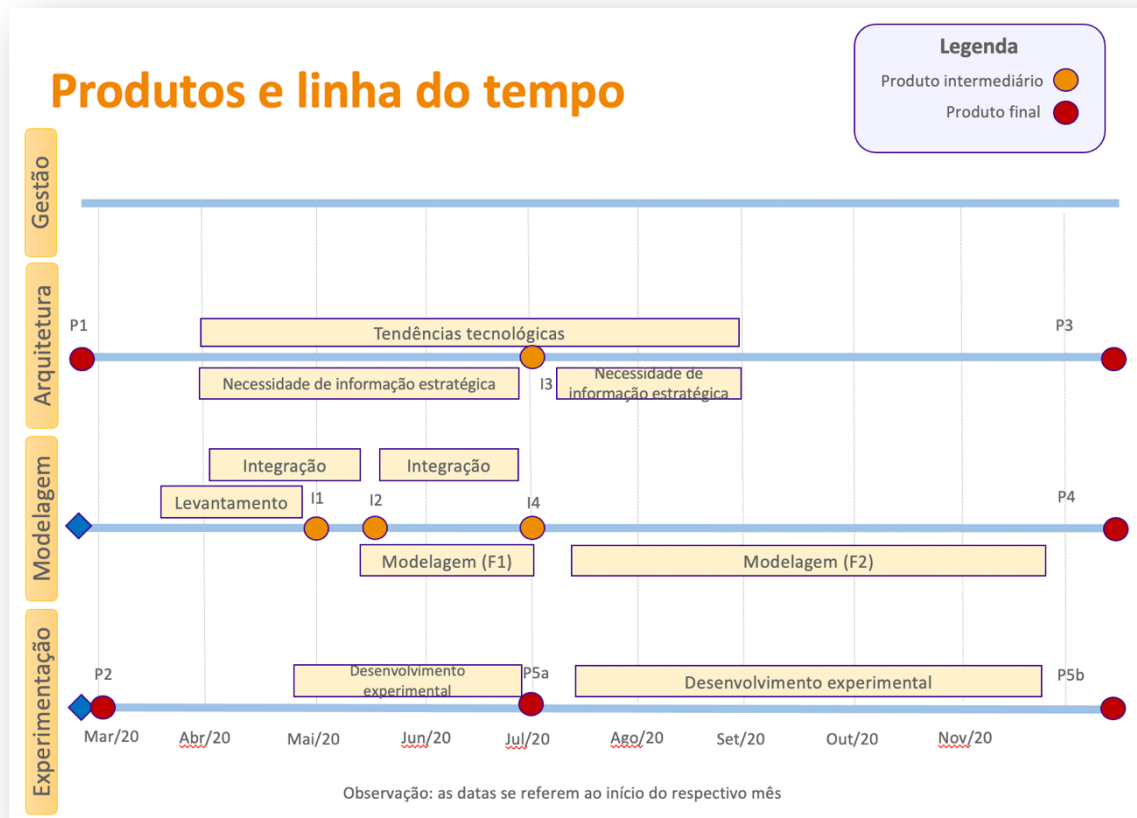


Figura 1 - Plano de trabalho do projeto. Fonte: (CGEE, 2019; 2020).

Em termos gerais, estudos realizados na linha de trabalho “Arquitetura” são utilizados na linha de “Modelagem”. O modelo arquitetural tem como objetivo organizar, em consonância com as práticas mais modernas, processos de trabalho, requisitos de competências e habilidades de atores envolvidos e instrumentação para construção de mecanismos de coleta e tratamento de dados e produção de informação estratégica. Tanto o modelo arquitetural quanto o modelo de processo de trabalho constituem dois dos objetivos específicos deste projeto.

Na linha de ação de Experimentação, objeto deste relatório, ambos os modelos construídos nas linhas anteriores e suas implementações práticas são exercitados. Esses experimentos têm como matéria prima assuntos priorizados pelos gestores do MCTI, denominados **temas estratégicos**. O Departamento de Governança Institucional indicou oito temas estratégicos a serem implementados sobre a Arquitetura Digital, mostrados na Tabela 1.

Tabela 1 - Temas estratégicos definidos pelo MCTI. Fonte: elaboração própria a partir de (MCTI, 2020).

<b>Tema estratégico</b>
<b>Indicadores COICT</b>
<b>Lei do Bem</b>
<b>FNDCT / Fundos Setoriais</b>
<b>Incentivos Fiscais para o Setor de Tecnologias da Informação e Comunicação</b>
<b>UP – Unidades de Pesquisa</b>
<b>Finep – Financiamentos</b>
<b>CNPq – Bolsas</b>
<b>Tema estratégico a ser definido</b>

A abordagem metodológica voltada para a implementação dos temas estratégicos consiste no Ciclo de Inteligência em CTI do CGEE, representado graficamente na Figura 1.



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

A primeira etapa do ciclo é considerada a mais importante. É nesse momento que se buscará compreender as necessidades de informações para a tomada de decisão, ou seja, buscar-se-á 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 criar 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.

As seções seguintes apresentam os resultados das implementações dos temas estratégicos indicados pelo MCTI realizadas no decorrer de 2020, cujo foco foi a implementação de visualizações de informação. Nesse sentido, o processo de inteligência é complementado, quando necessário, com métodos de construção de sistemas analíticos e metodologias ágeis de desenvolvimento de software.

#### **4. A Arquitetura Digital de Inteligência de Negócios do MCTI**

A Arquitetura Digital de Inteligência de Negócios do MCTI tem como conceito central a constituição de um ambiente digital que suporte o armazenamento de fontes de informações heterogêneas, que permita a aplicação de metodologias de tratamento e análise de dados a partir de conjuntos de dados brutos e finalmente, que viabilize a produção e visualização de informação com alto valor agregado para subsídio à tomada de decisão dos gestores e analistas do MCTI. É parte essencial desse ambiente digital a promoção e contribuição para o alcance dos pilares da moderna gestão de dados, a saber, integridade, confiabilidade, disponibilidade, autenticidade.

O cerne desse ambiente digital de informação é um modelo arquitetural que estrutura o processo geral de tratamento de dados em camadas. A aplicação de tratamentos, de qualquer natureza, sobre dados configura agregação de valor ao dado e culmina com a elaboração de produtos de informação orientados para um determinado objetivo. Esse processo pode ser, também, expresso pelo encadeamento de transformações de dados realizadas manualmente ou com auxílio informatizado (ferramentas). Cada transformação aplicada consiste na aplicação de regras de negócio para alcance de objetivos específicos, e contemplam também aspectos de qualidade de dados (integridade, confiabilidade, disponibilidade, autenticidade).

A cadeia de transformações sobre dados é organizada em um modelo em camadas. Cada uma com características marcantes relativas à semântica, tipo de transformação aplicável, tipo de ferramenta utilizada, e em consequência, o grau de valor agregado dos objetos e sua proximidade com objetivos pré-estabelecidos.

O modelo arquitetural da Arquitetura Digital se organiza em quatro camadas referenciais alinhadas com o grau de valor agregado ao dado em sua trajetória

para um determinado produto de informação. A Figura 3 apresenta essas quatro camadas e seus objetivos principais.

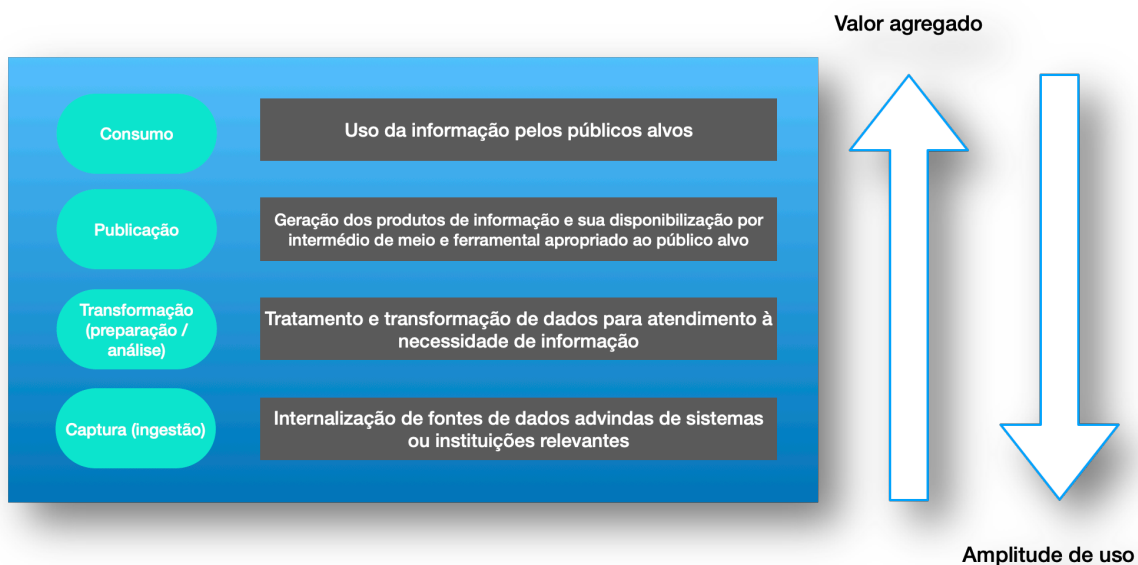


Figura 3 - Camadas do modelo arquitetural e características principais. Fonte: elaboração própria.

À estrutura em camada da arquitetura digital são definidos conceitos arquiteturais associados, mostrados na Figura 4. A figura mostra também as principais fronteiras conceituais que são observadas na construção da arquitetura digital objeto deste projeto.



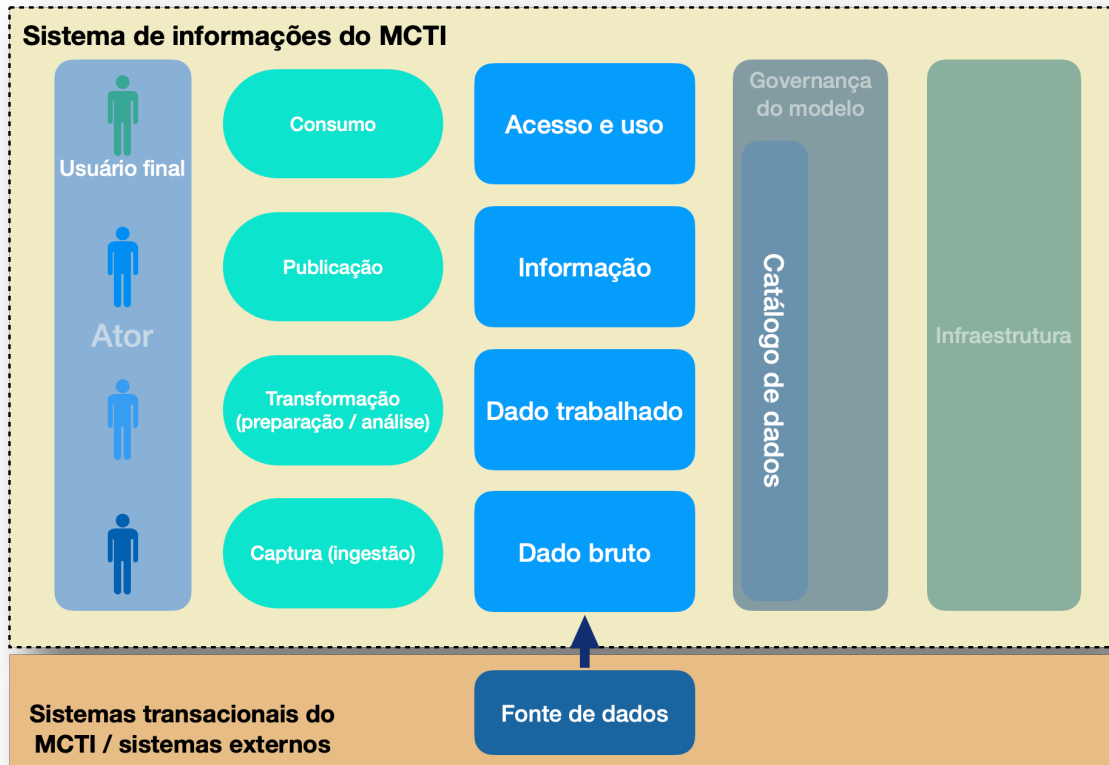


Figura 4 - Visão geral de componentes e fronteiras do modelo arquitetural. Fonte: elaboração própria.

É importante esclarecer neste relatório que o detalhamento conceitual e operacional apresentados na figura anterior e na descrição que segue é objeto dos documentos “Relatório da modernização da arquitetura de ambiente digital” e “Relatório detalhando o processo de gestão de ambiente digital de informação estratégica”, produtos previstos para entrega posterior. Entretanto, uma versão preliminar do detalhamento conceitual e operacional foi disponibilizado ao MCTI (DGI e DTI) por intermédio de um relatório preliminar entregue ao MCTI (DGI e DTI) em 23 de outubro de 2020, intitulado “Projeto Arquitetura Digital de Inteligência de Negócio do MCTIC - Relatório parcial – Outubro 2020”.

O foco neste relatório é o desenvolvimento experimental conduzido sobre este modelo arquitetural. Assim, na descrição que segue serão apresentados brevemente as principais estruturas que orientam, organizam e proveem

instrumentação para o desenvolvimento dos painéis de informação que implementa cada tema estratégico.

O foco da modelagem se encontra no “Sistema de informações do MCTI” (MCTI, 2020). O aparato digital e o processo de trabalho, que deriva do modelo arquitetural proposto, têm 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). Esses sistemas transacionais ou externos se comunicam com a arquitetura digital por meio de fontes de dados. Foi pela definição de sistemas transacionais pelo MCTI que se estabeleceu a temática estratégica para os desenvolvimentos experimentais de painéis de informação na Arquitetura Digital, objeto deste relatório.

A arquitetura tecnológica implementada é apresentada na Figura 5. As ferramentas de software componentes da arquitetura, descrita resumidamente, na Tabela 2, é resultante dos requisitos tecnológicos e operacionais descritos no capítulo anterior. Chama-se a atenção para o importante requisito de alinhamento com a infraestrutura de TIC já disponível no MCTI, que conforma o cerne da arquitetura apresentada e discutida a seguir.

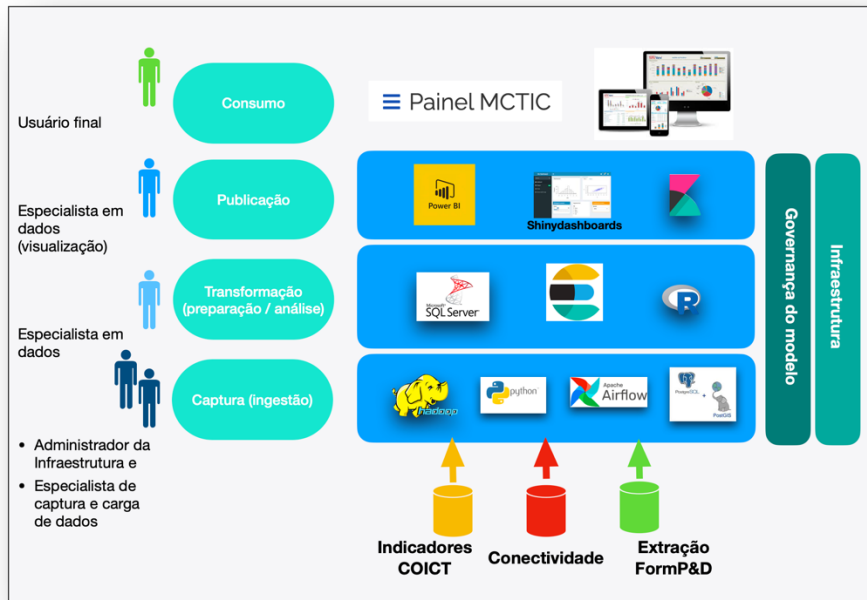


Figura 5 - Visão geral de componentes arquiteturais com respectivas implementações técnicas. Fonte: elaboração própria.

Tabela 2 - Detalhamento dos elementos tecnológicos da implementação inicial do modelo arquitetural. Fonte: elaboração própria.

Componente tecnológico	Função principal
<b>Hadoop</b>	Armazenamento e acesso de grandes volumes de dados, com alta capacidade de processamento com uso intenso de paralelismo
<b>Apache AirFlow</b>	Automação de processos de ETL ( <i>Extract, Transform and Load</i> ).
<b>PostgreSQL / PosGIS</b>	Gerenciador de banco de dados relacional. Armazenamento e acesso de dados relacionais e georreferenciados.
<b>MS SQL Server Big Data Clusters</b>	Disponibilização do Hadoop de forma integrada com o, e abstraída pelo, MS SQL Server, o qual é um outro gerenciador de banco de dados relacional.
<b>ELK ElasticSearch e Kibana</b>	Sistema integrado de coleta de dados e provimento de mecanismos de busca e visualização, incluindo visualização georreferenciada.

<b>R Studio</b>	Ambiente de desenvolvimento integrado para a ferramenta estatística R. R é uma linguagem de computação estatística e gráficos.
<b>ShinyDashboards</b>	Framework web para criação de aplicativos interativos com linguagem R.
<b>Power BI</b>	Ferramenta e serviço na internet para elaboração e disponibilização de visualizações interativas com recursos de business intelligence.
<b>Site <a href="https://paineis.mctic.gov.br/">https://paineis.mctic.gov.br/</a></b>	Página web do MCTI cuja intenção é concentrar a disponibilização de informação sobre atividades e resultados de Ciência, Tecnologia e Inovação para público externo.

Cada tema estratégico trabalhado está hospedado nessa plataforma digital. Os desenvolvimentos experimentais realizados nos temas estratégicos utilizaram os componentes da arquitetura para sua elaboração, bem como produziram como consequência demandas técnico operacionais específicas que exercitaram o modelo arquitetural e promoveram, em alguns casos, integração de novas ferramentas tecnológicas à estrutura tecnológica principal.

Um aspecto importante do desenvolvimento experimental dos temas estratégicos foi a validação e refinamento da estrutura de governança do modelo, que é endereçada por um conjunto de propostas de padrões apresentados na Tabela 3 e por um processo de trabalho, cujo modelo será apresentado em seguida.

Tabela 3 - Proposições de padrões por camada do modelo. Fonte: elaboração própria.

	<b>Camada de Captura</b>	<b>Camada de Transformação</b>	<b>Camada de Publicação</b>	<b>Camada de Consumo</b>
<b>Organização dos objetos típicos</b>	Classe / Subclasse	Classe / Subclasse	Classe	Por tecnologia de acesso

<b>Padrões</b>	Formato de arquivo e conteúdo	Descritivo metodológico	Plano de Dados Abertos - MCTI	Padrão de identidade visual do MCTI Regras de navegação e acessibilidade
<b>Nomenclatura</b>	Padrão para nomenclatura de objetos em <i>datalake</i>	Padrão para nomenclatura de objetos em <i>datalake</i>	Padrão para nomenclatura de objetos em <i>datalake</i>	Padrão para nomenclatura de objetos em <i>datalake</i>

Sob o aspecto de atividades sobre o modelo arquitetural, foi definido um **processo de inteligência de negócio alinhado com a arquitetura de inteligência estratégica do MCTI**, materializada na plataforma digital de informação.

Esse processo, cuja visão geral é apresentada na Figura 1, de trabalho integra ações dos departamentos DGI e DTI, e contempla os demais atores relevantes no contexto de gestão de dados em um fluxo de tarefas que:

- Contempla as atividades necessárias para a produção de informação com valor agregado 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.

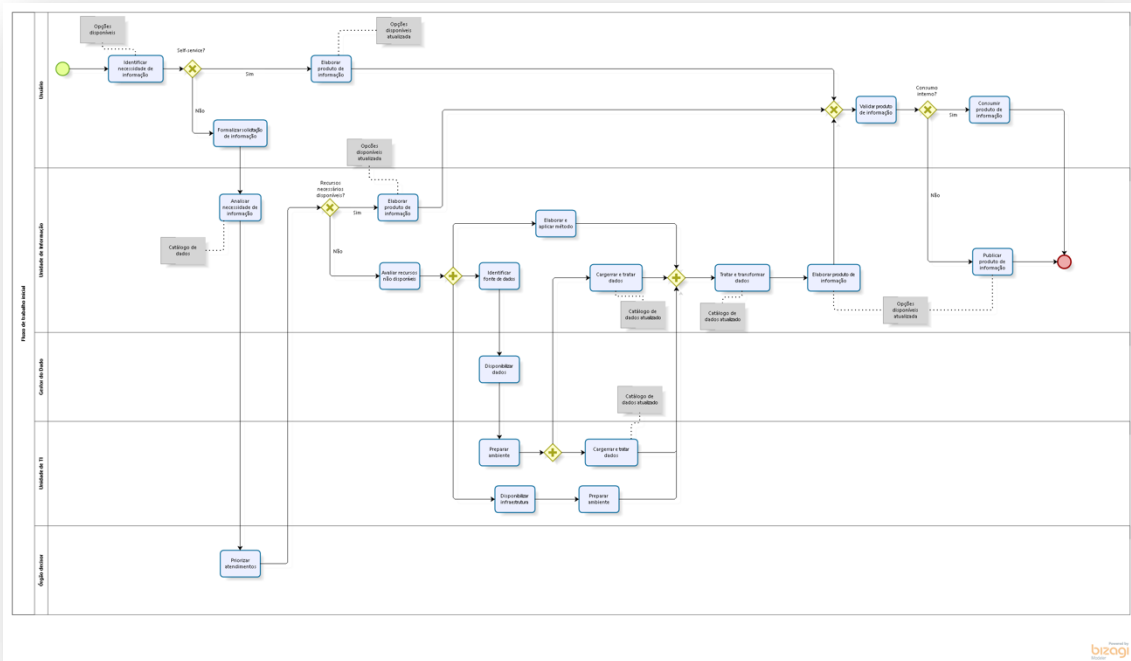


Figura 6 - Modelo de processo de trabalho sobre a Arquitetura Digital. Fonte: elaboração própria.

Sob o aspecto de implementação física da Arquitetura Digital, a Figura 7 apresenta os componentes de software centrais.

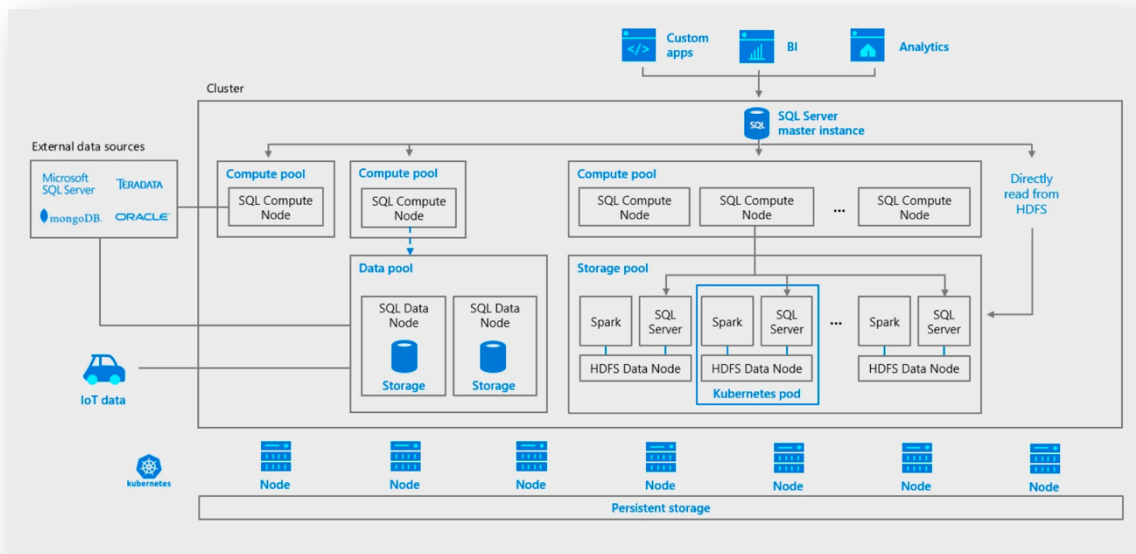


Figura 7 - Módulos da implementação física do Microsoft Big Data SQL Server. Fonte: (WRIGHT, 2018).

A infraestrutura tem como base o Clusters de Big Data SQL Server da Microsoft, isso em estrutura de cluster de contêineres Linux orquestrados pelo *Kubernetes*.

Dentre outros componentes, o cluster Big Data SQL Server possui componentes do Hadoop e seu ecossistema, tal como HDFS para armazenamento distribuído, o Spark e seus componentes para o processamento distribuído e o Apache Kafka para o barramento de mensagens, isso além do próprio SQL Server e os componentes de gerenciamento do cluster.

Essa é uma solução relativamente nova da Microsoft, que apresentou resultados até o momento satisfatórios, dado a facilidade de implementação, gerenciamento, integração dos componentes internos e externos e segurança. Além disso, o MCTI já possui um relacionamento comercial com a fabricante da solução, facilitando assim a aquisição e a evolução da arquitetura com os produtos do fabricante.

A solução permite a implementação de um importante conceito no mundo *big data*, denominado virtualização de dados (*data virtualization*). Virtualização de dados (ROUSE, 2019) é uma abordagem de gestão de dados que permite que uma aplicação acesse e manipule dados sem necessidade de conhecer ou implementar os detalhes técnicos de acesso ao dado, como por exemplo, qual o formato do dado e onde o dado está localizado. Uma camada de abstração é provida, na forma de um servidor de virtualização de dados (DMBOK, 2020, p.294), de tal forma que a aplicação acessa diretamente o dado sem a necessidade de remanejamentos de dados ou cópias. Essa abordagem facilita e, em alguns casos, dispensa a realização de ETL.

No caso do Big Data SQL Server, essa facilidade é disponibilizada por meio do SQL Server (vide Figura 8) executa a função de abstração do ambiente Hadoop e outros gerenciadores de bases de dados, realizando as operações de ETL e integração de dados internamente e provendo a figura de virtualização de dados para o usuário. Assim, o desenvolvedor de uma

aplicação de BI pode consultar tabelas de dados de diferentes formatos de origem ou tipos e fabricantes de gerenciadores de dados, sem se preocupar se a implementação física está em uma base de dados SQL Server, Postgres, NoSQL, Hadoop HDFS, arquivos CSV etc.

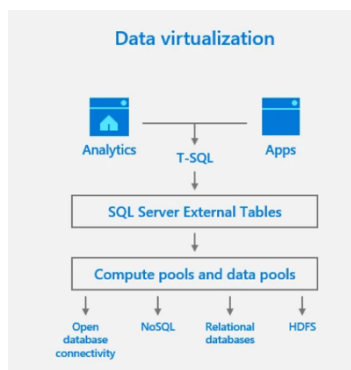


Figura 8 - Estrutura lógica da virtualização de dados com uso do MS Big Data SQL Server. Fonte: (WRIGHT, 2018).



## 5. O exercício de validação da Arquitetura Digital

O elemento central dos exercícios é a experimentação da capacidade da proposta de Arquitetura Digital para hospedar processos de inteligência de dados de natureza e requisitos diversos, em uma estrutura que seja capaz de acomodar essa diversidade e prover mecanismos padronizados para sua gestão observando quesitos de qualidade, desempenho e produtividade.

A Arquitetura Digital, apresentada na seção anterior, é o *locus* e instrumental para o processo de inteligência de dados e produção de informação estratégica. Conforme seus componentes sem mostram prontos para o uso, os exercícios permitem tanto a validação quanto a melhoria da Arquitetura Digital, promovendo, além disso, a disseminação da informação e aprendizados sobre a proposta de trabalho e seus resultados.

Os exercícios realizados tomaram como guia o processo de inteligência de negócio elaborado e descrito resumidamente na seção anterior. Os aspectos de alinhamento exercitados incorporam, além do processo de trabalho, a arquitetura em camadas (modelo arquitetural), as prescrições e os padrões de curadoria de dados e os principais requisitos de segurança da informação. São esses componentes da Arquitetura Digital que foram colocados à prova. São esses os componentes que foram objeto de ajustes e melhorias em decorrência dos exercícios.

Para a realização dos exercícios foi necessário a definição de matéria prima para trabalhar. Para tanto, foram definidos oito temas estratégicos que constituiriam os conteúdos a serem tratados nos exercícios. O processo de seleção desses temas foi conduzido pelo DGI/SEXEC com contribuições da equipe técnica do projeto (DGI, DTI e CGEE). Os temas estratégicos apontados são apresentados na Tabela 4 juntamente com seus interlocutores e natureza de informação esperada.

Tabela 4 - Detalhamento dos temas estratégicos indicados pelo MCTI. Fonte: elaboração própria a partir de (MCTI, 2020).

<b>Tema estratégico</b>	<b>Interlocutor</b>	<b>Decisões</b>
<b>Indicadores COICT</b>	Departamento de Governança Institucional – DGI / Coordenação de Indicadores de Ciência, Tecnologia e Inovação - COICT	Internalizar os indicadores (variáveis) já disponibilizadas no site “Indicadores Nacionais de Ciência, Tecnologia Inovação” <sup>1</sup>
<b>Lei do Bem</b>	Secretaria de Empreendedorismo e Inovação – SEMPI e Grupo de Trabalho de Aperfeiçoamento da Aplicação da Lei do Bem (MCTI)	Implementação dos indicadores resultantes do estudo “Uma análise dos resultados da Lei do Bem: com base nos dados do FormP&D” (CGEE, 2018)
<b>FNDCT / Fundos Setoriais</b>	Departamento de Governança Institucional – DGI / Coordenação-Geral de Governança de Fundos - CGGF	Painéis de informação sobre arrecadação, execução orçamentária, gestão de demanda do FNDCT e panorama de projetos financiados
<b>Incentivos Fiscais para o Setor de Tecnologias da Informação e Comunicação</b>	Secretaria de Empreendedorismo e Inovação - SEMPI	Incentivos fiscais
<b>UP – Unidades de Pesquisa</b>	Secretaria Executiva – SEXEC / Subsecretaria de Unidades Vinculadas - SUV	Unidades de pesquisa vinculadas
<b>Finep Financiamentos</b>	– Departamento de Governança Institucional – DGI	Linhas de financiamento – Valor contratado Linhas de financiamento – Valor desembolsado
<b>CNPq - Bolsas</b>	Departamento de Governança Institucional – DGI	Bolsas

A seguir são abordados os pontos relevantes advindos da experiência realizada assim como principais achados.

<sup>1</sup> [https://antigo.mctic.gov.br/mctic/opencms/indicadores/indicadores\\_cti.html](https://antigo.mctic.gov.br/mctic/opencms/indicadores/indicadores_cti.html)

## 5.1. Processo de inteligência de dados e linhagem de dados

O desenvolvimento e implantação de processos de inteligência de dados relativos aos temas estratégicos concretizaram fluxos de captura, transformação e tratamento de dados e produção de visualizações em painéis. No modelo de processo de trabalho, o principal sub-processo do relacionado com inteligência de dados é "Elaborar produto de informação", cujas tarefas são:

- Elaborar produto de informação,
- Elaborar e aplicar método,
- Internalizar fonte de dados.

Residem nas tarefas destacadas especial acoplamento com a arquitetura da informação. A Figura 9 faz a correspondência aproximada entre essas tarefas e as camadas do modelo arquitetural.

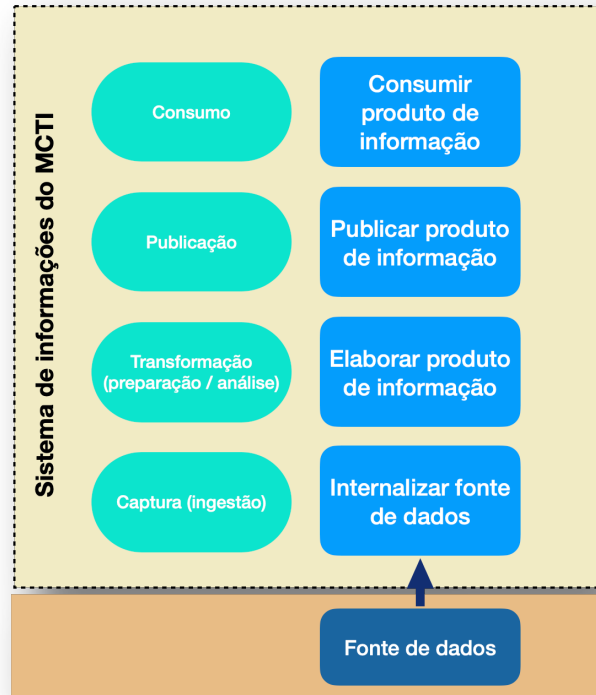


Figura 9 - Tarefas do processo de inteligência de dados e relação com as camadas do modelo arquitetural. Fonte: elaboração própria.

De forma geral esse é o cerne do processo de ingestão e tratamento de dados, e visualização de informação. Cada instância de processo é aplicada à demanda específica formalizada, no caso, os temas estratégicos definidos. Em cada instância de execução desse processo, componentes metodológicos são utilizados ou concretizados. Ou seja, quais dados, quais técnicas ou métodos de tratamento de dados, quais variáveis, métricas e indicadores, quais opções de visualização ou disponibilização da informação, incluindo quais ferramentas eletrônicas, todos esses componentes são arregimentados no bojo da instância específica de execução desse processo.

Com objetivo de descrever e documentar esses fluxos de dados, foram elaboradas **linhagens de dados**. A linhagem de dados (*data lineage*) consiste na descrição de processos de transformação de dados desde o ponto de sua origem até o ponto de seu uso (STEENBEEK, 2019; DAMA, 2020). Este produto é associado a cada instância de execução do cerne do processo de inteligência de dados, de forma a manter uma trilha de informação sobre como um dado é utilizado, assim como por quais transformações passou um dado que faz parte de um painel.

Outro aspecto relevante do exercício sobre a estrutura de camadas foi a objetividade e efetividade na escolha de metodologias de trabalho. Cada camada, com seu objetivo pré-estabelecido, propõe de forma natural um leque de métodos e ferramentas mais adequadas. Na camada de “Ingestão”, as tarefas são dominadas por de métodos de extração e movimento de dados, levando a metodologias de *web scraping*, interoperabilidade entre sistemas, desenvolvimento de artefatos de software de ETL (extração, tratamento e carga de dados).

Já na camada de “Transformação”, as tarefas são mais sofisticadas, de pensar como responder à necessidade de informação do usuário final. Portanto levam a necessidade de levantamento de requisitos, projetos de sistemas (analíticos), modelagem de dados, projeto de interfaces e implementação dos artefatos de softwares produção de informação com valor agregado. Entram em cena

metodologias também sofisticadas de desenvolvimento de sistemas analíticos, *business intelligence*, aprendizado de máquina, dentre outras. No exercício conduzido neste projeto, foi adotado a metodologia o Ciclo de Inteligência Estratégica (CGEE, 2017) para na camada de “Transformação”.

Na camada de “Publicação” o foco é a visualização de informação, com seu repertório de alternativas gráficas de visualização de informação complexa, e ferramentas atuais que facilitam a criação de visualizações ricas com integração de diferentes formas de prover informação (gráficos, mapas, tabelas etc.). Foi utilizada a ferramenta Microsoft PowerBI em consonância com escolhas já realizada pelo MCTI para essa função.

## 5.2. Temas estratégicos implementados - Painéis de informação

Os detalhes operacionais e os resultados da realização do desenvolvimento experimental estão descritos na documentação contida nos anexos deste relatório.

A descrição é feita por intermédio de uma adaptação do método de engenharia de software denominado **histórias de usuários** (*user stories*<sup>2</sup>), o qual foi integrado com o Ciclo de Inteligência Estratégica. Consiste em uma proposição nascida no projeto com objetivo de documentar os seguintes elementos do processo de inteligência de dados implementado em painéis digitais de informação:

- Especificação de requisitos (perguntas norteadoras e suas narrativas de respostas);
- Linhagem de dados;
- Dicionário de variáveis ou indicadores;

---

<sup>2</sup> Descrição de características de sistemas de software em linguagem natural, na perspectiva do usuário final, utilizada em especial em metodologias ágeis de desenvolvimento de software e gestão de produtos (tradução livre a partir de [https://en.wikipedia.org/wiki/User\\_story](https://en.wikipedia.org/wiki/User_story)).

- Projeto gráficos de visualizações (*mock ups*); e
- Modelos de dados.

Essa proposta de padrão mostrou características de facilidade de uso em especial na interação com os usuários finais (em especial as perguntas norteadoras, narrativas de resposta e *mock ups*). A conversão das especificações em artefatos documentais para orientação para desenvolvimento de software (*mock ups*, painéis, bases de dados) também obteve efetividade com produtividade. Além disso, a reunião desses elementos em um documento organizado permite uma visão integral do objeto final elaborado, que no caso dos temas estratégicos foram painéis, a exceção do tema estratégico Indicadores do COICT.

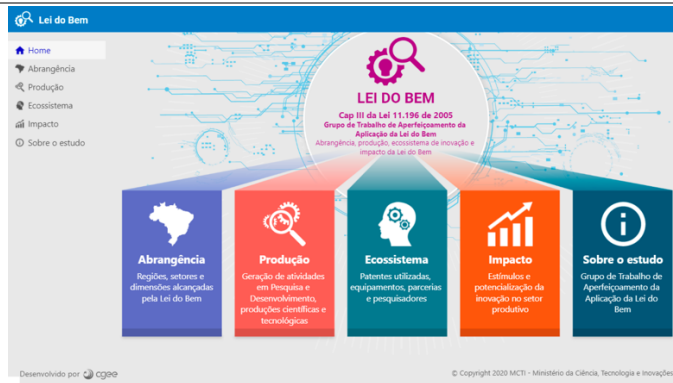
Várias ferramentas automatizadas podem ser utilizadas para registrar os elementos integrados no método documental proposto. Não foi identificada uma ferramenta em uso corrente no MCTI com esse propósito. Entretanto, na medida em que o Ministério adote uma ou mais ferramentas para a documentação, o material produzido pode ser incorporado em seus repositórios. Ressalta-se que os conteúdos dessa documentação são a matéria prima dos conceitos de dicionários de dados ou variáveis (em nível de aplicação e de negócio), catálogos de dados, e a própria linhagem de dados.

Os temas estratégicos trabalhados até o momento têm sua documentação apresentada conforme Tabela 5.

Tabela 5 - Localização da documentação complementar dos temas estratégicos nos anexos deste relatório. Fonte: elaboração própria.

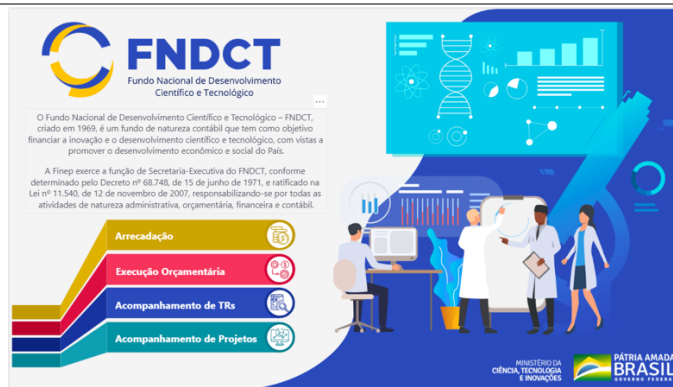
Tema estratégico	Tela de entrada	Anexo
------------------	-----------------	-------

Lei do Bem



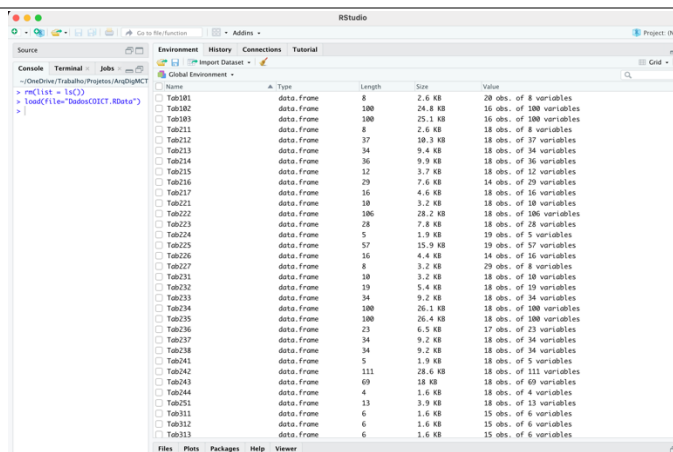
Anexo II

FNDCT



Anexo III

Indicadores da COICT



Anexo IV

### 5.3. O caso específico do tema estratégico Indicadores da COICT

O tema estratégico Indicadores da COICT, 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"<sup>3</sup>. A expectativa do demandante era “apoio do CGEE para incorporar a base de dados da COIND 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.
- 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*<sup>4</sup>), 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 JSON<sup>5</sup> 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).

---

<sup>3</sup> [https://antigo.mctic.gov.br/mctic/opencms/indicadores/indicadores\\_cti.html](https://antigo.mctic.gov.br/mctic/opencms/indicadores/indicadores_cti.html) .

<sup>4</sup> 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).

<sup>5</sup> 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>)



A documentação desse tema estratégico consiste em cadernos Jupyter (*Jupyter Notebooks*) que descrevem as transformações e a documentação das variáveis em formato JSON. Esse material se encontra no Anexo V deste relatório.

## 6. Conclusões e próximos passos

A disponibilização de painéis temáticos teve como objetivo exercitar o ambiente digital proposto neste projeto, ou seja, a Arquitetura Digital de Inteligência Estratégica do MCTI (**Arquitetura Digital**).

O exercício consistiu na elaboração, em caráter piloto de oito painéis sobre temas estratégicos definidos pelo MCTI. Os resultados alcançam objetivos maiores do projeto na medida que experimenta o suporte e o armazenamento a fontes de informações heterogêneas, proporciona contexto administrado para aplicação de metodologias de tratamento e análise de dados de brutos, de variados formatos.

O exercício também permitiu 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<sup>6</sup> e a relação com a Arquitetura Digital, assim como os conceitos, componentes da arquitetura, modelo de trabalho e responsabilidades.

Esse último ponto destacado foi objeto de verbalização por parte do Departamento de Governança Institucional (DGI), em reunião de acompanhamento de projeto, expressado na forma de reconhecimento da importância dos aprendizados alcançados no decorrer dos trabalhos realizados ressaltados abaixo:

- A importância do papel articulador do DGI, internamente no MCTI, para que as equipes técnicas afeitas aos temas estratégicos estejam atuantes durante a construção dos painéis temáticos;
- A importância da proximidade do Departamento de Tecnologia da Informação (DTI) para apoio e fluidez das atividades;
- A importância de se organizar uma base de dados no *datalake* para os temas estratégicos tratados.

Na mesma reunião, foi estabelecido o ajuste de cronograma de atividades para o projeto. Durante o ano de 2020, um conjunto de fatos de alto impacto no projeto (relatados na introdução deste relatório) foram vivenciados e impactaram na fluidez dos trabalhos. Em vista disso, o Ministério apresentou proposta de revisão de cronograma, com extensão do prazo para conclusão dos oito temas estratégicos para meados de 2021.

---

<sup>6</sup> Na acepção de (DAMA, 2020), ciclo de vida de dados incorpora todo o aparato de tarefas relativa a vida do dado na organização, partindo do planejamento de sua aquisição até sua descarte final. Essa abordagem remete à vida do dado também em sistemas transacionais e sistemáticas manuais ou semi-automatizadas, que estão além da fronteira estabelecida na Arquitetura Digital.

---

A continuidade dos trabalhos no primeiro semestre de 2021 envolve a continuação do exercício de desenvolvimento experimental para os demais temas estratégicos e elaboração versão final de relatório.

## 7. Referências Bibliográficas

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, 2018 CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. Uma análise dos resultados da Lei do Bem: com base nos dados do FormP&D. Resumo Executivo. Brasília: Centro de Gestão e Estudos Estratégicos, 2018.

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.

MCTI, 2020 MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES - MCTI. Sistema de informações do MCTI. Brasília, DF: Departamento de Governança Institucional (DGI). Apresentação PPT. Sistema\_de\_Informações\_do\_MCTI(2).pptx. Agosto de 2020.

ROUSE, 2019 Rouse, Margaret. "Data Virtualization". TechTarget/WhatIs.com. Página disponível na Internet em <https://searchdatamanagement.techtarget.com/definition/data->

virtualization. Acessado em 10/11/2020. Última atualização em maio/2019.

STEENBEEK, 2019 STEENBEEK, Irina. Data Lineage 101. Data Crossroads. Disponível em <https://datacrossroads.nl/category/series/data-lineage-101/>. Último acesso em 17/10/2020. Março de 2019.

## Anexo I – Ajuste de cronograma do projeto

# Projeto Arquitetura Digital e Inteligência de Negócio do MCTI

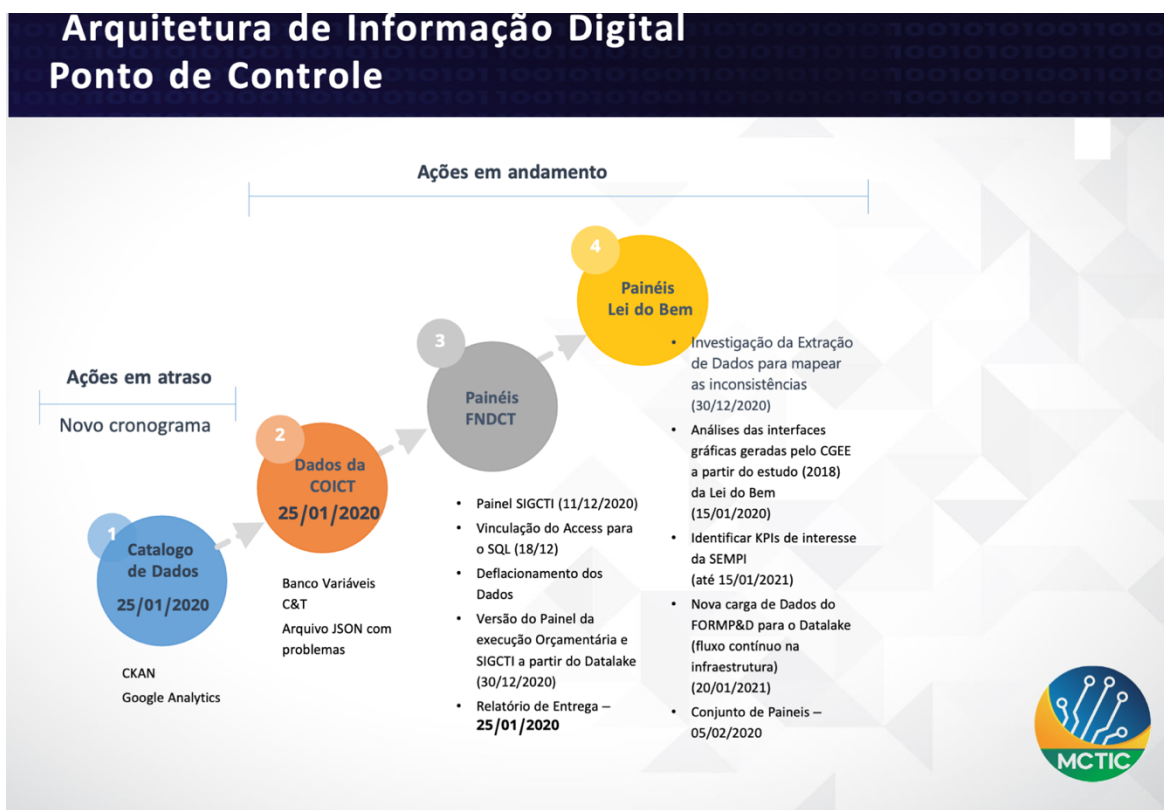
Departamento de Governança Digital  
Departamento de Tecnologia da Informação

Dezembro, 2020

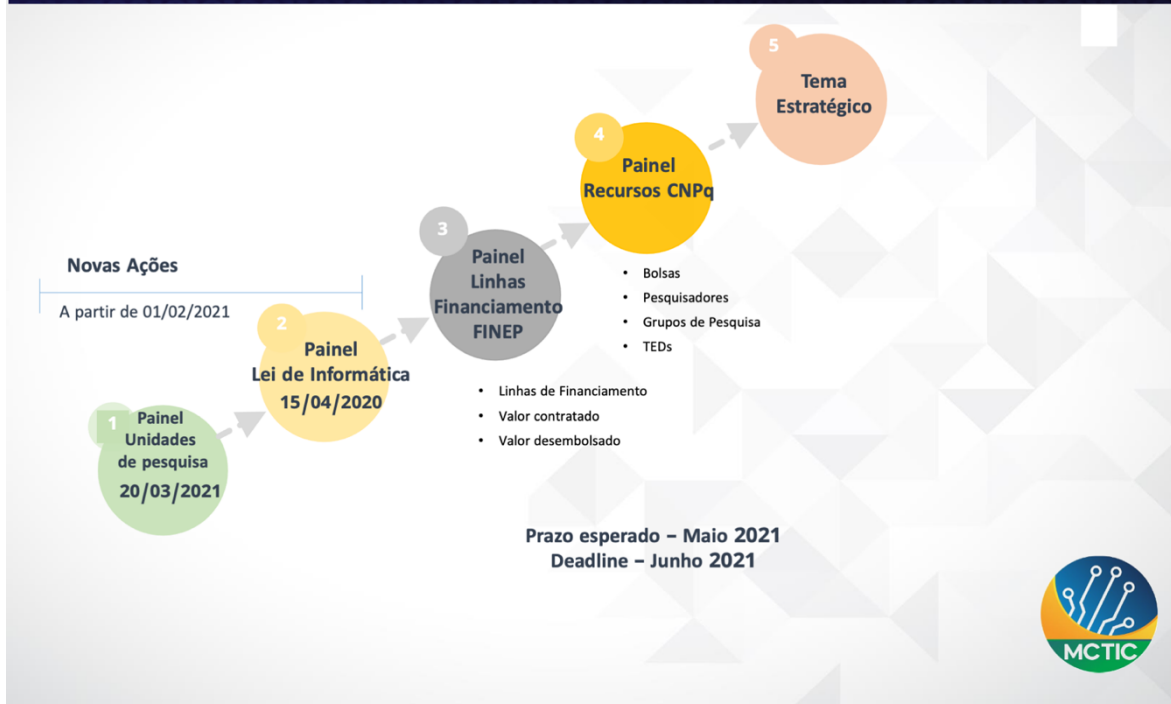


MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA,  
INOVAÇÕES E COMUNICAÇÕES





# Arquitetura de Informação Digital Ponto de Controle



**Arquitetura Digital de Inteligência de Negócios do MCTI (ArqDigMCTI)**  
Operacionalização de atividades a partir do segundo semestre considerando a proposta do Sistema de Informações do MCTI (DGI)  
Encaminhamentos da reunião de 10/12/2020

Item	Descrição	Resultados (parciais / finais)	dez/20	jan/21	fev/21	mar/21	abr/21	mai/21	jun/21
1	Catálogo de dados	Instalação do CKAN na Infraestrutura	15	31					
2	Dados da COICT	Geração de .Rdata	15	21					
3	Painéis FNDCT	Vinculação do Access para o SQL (1) Inclusão do deflacionamento dos dados (2) Painéis a partir do Datalake (versão 1) (3)	??	31					
4	Painéis Lei do Bem (Homologação)	Relatório de entrega Investigação da extração de dados Apoio na análise das interfaces gráficas (4) Identificação de variáveis de interesse Nova carga de dados (fluxo contínuo) (5) Painéis a partir do Datalake (versão 2)	15	26	15	15	31	13	
5	Painel Unidades de Pesquisa						30		
6	Painel Lei de Informática						15		
7	Painel Linhas Financiamento FINEP							15	
8	Painel Recursos CNPq								31
9	Painel 8								15
10	Relatório final								30

**Observações**

- (a) Dependência do acesso a Rede MCTI; (b) Pádua tem que disponibilizar a última versão da base Access, e todas as atividades sobre a base deverão ser realizadas sobre o novo ambiente.
- (a) Depende da data de entrega do componente e (b) conhecermos o seu funcionamento.
- (a) Dependência do acesso a Rede MCTI
- A análise é conduzida pelo MCTI, o CGEE faz apoio
- A extração dos dados é desenvolvida pelo MCTI, o CGEE faz apoio

**Anexo II – Tema estratégico Lei do Bem**

# **ESPECIFICAÇÃO TÉCNICA**

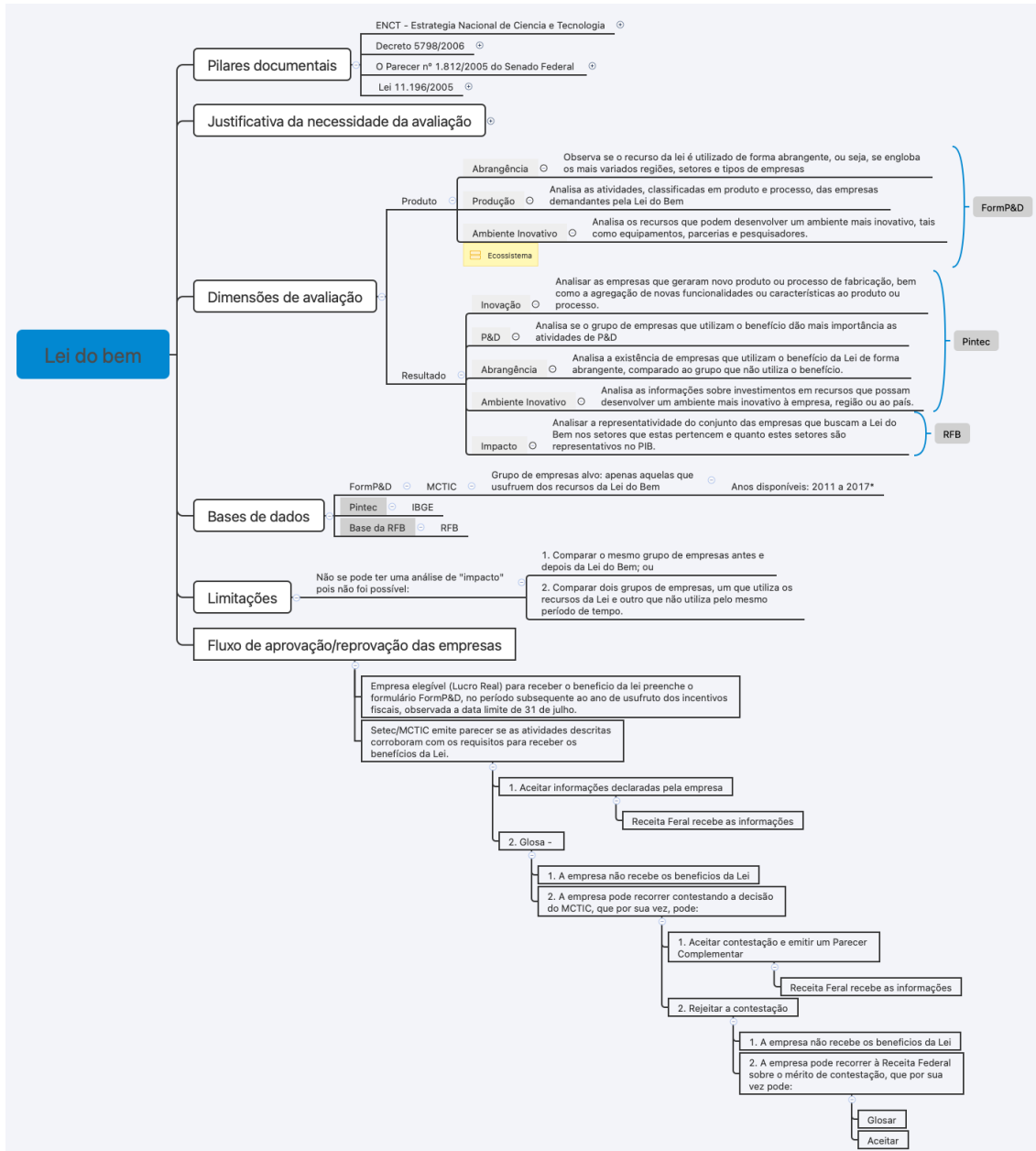
LEI DO BEM



# 1. INTRODUÇÃO

Este documento tem como objetivo especificar todos os requisitos referentes ao projeto Lei do Bem.

## 1.1. CONTEXTUALIZAÇÃO



## 2. HISTÓRIA DE USUÁRIO

### 2.1. Visualizar Painel Home

ID	NOME
HST001	Visualizar Painel Home

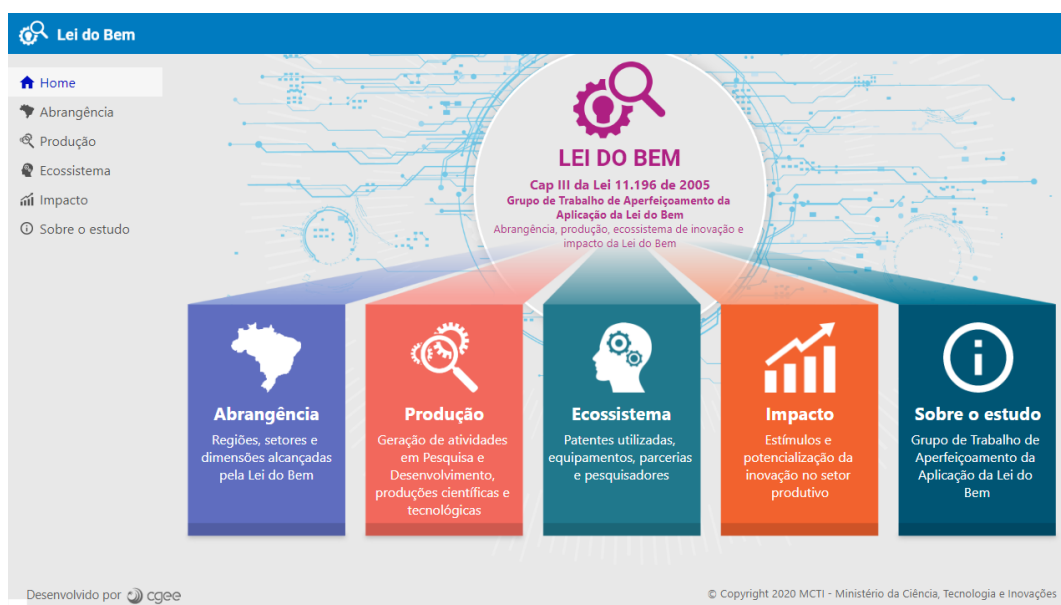
**COMO** Gestor Público  
**QUERO** visualizar todos painéis referentes a Lei do Bem  
**PARA** que eu possa escolher qual painel desejo acessar

#### 2.1.1. Critérios de Aceitação

<b>CA001.</b>	A tela inicial deverá exibir atalhos para acessar os 5 painéis disponíveis: <b>Abrangência, Produção, Ecosistema, Impacto e Sobre o estudo.</b>
---------------	---

#### 2.1.2. Protótipo

##### 2.1.2.1. Tela Home



### 2.2. Visualizar Painel Abrangência

ID	NOME
HST002	Visualizar Painel Abrangência

**COMO** Gestor Público  
**QUERO** ter acesso ao painel Abrangência  
**PARA** que eu possa visualizar a evolução anual do número de empresas demandantes

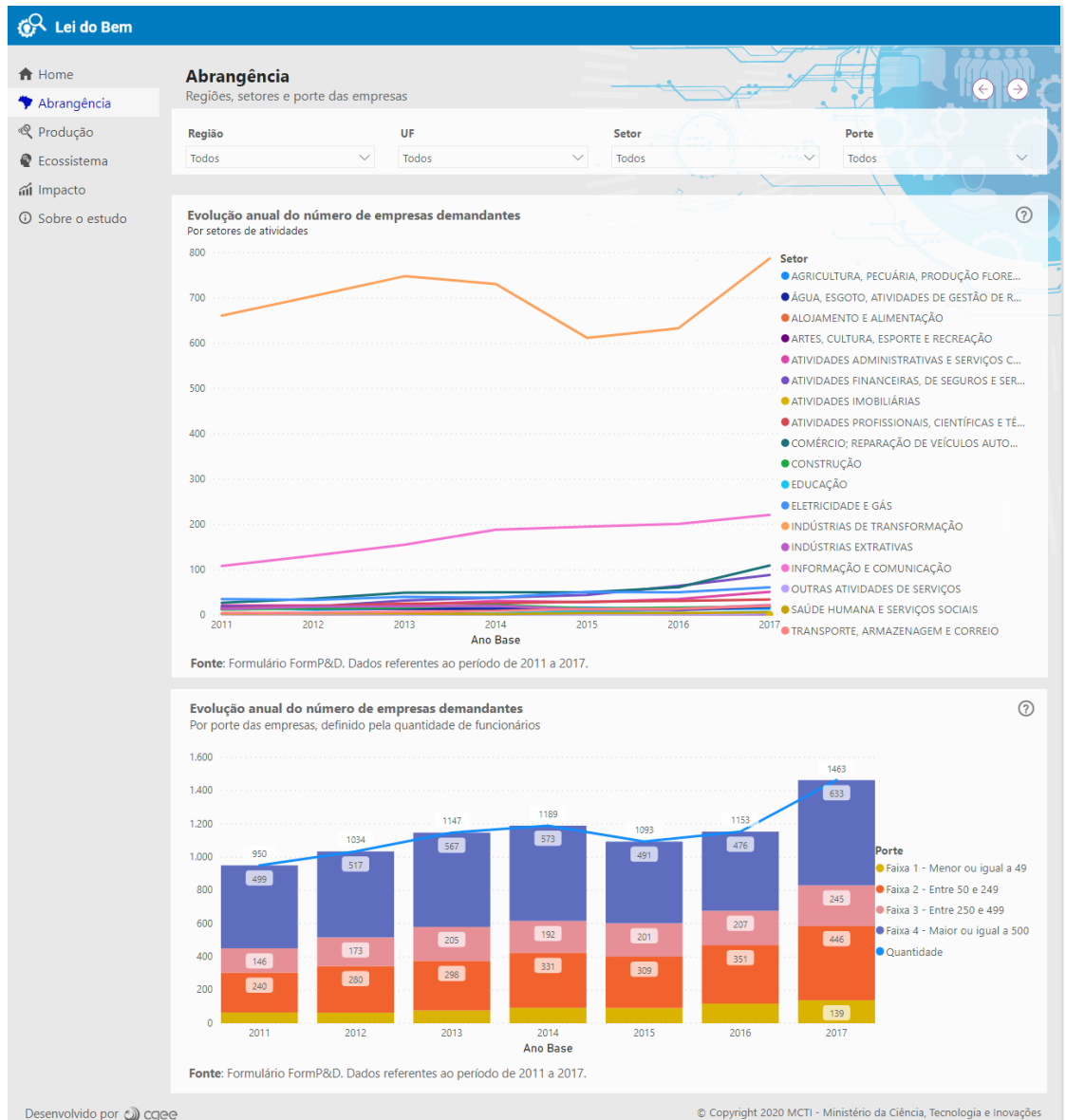
#### 2.2.1. Critérios de Aceitação

<b>CA002.</b>	O painel abrangência deverá apresentar 2 (dois) gráficos referentes a evolução anual do número de empresas demandantes, são eles: <b>Por setores de atividades e Por porte das empresas.</b>
<b>CA003.</b>	Deverá permitir filtrar os gráficos por: <ul style="list-style-type: none"> <li>▪ Região: deve exibir as 5 (cinco) regiões do Brasil;</li> <li>▪ UF: deve exibir a sigla de todas as unidades federativas do Brasil;</li> <li>▪ Setor: deve exibir todos os setores de atividades [Erro! Fonte de referência não encontrada.]; e</li> <li>▪ Porte: deve exibir todas as faixas de porte das empresas [Erro! Fonte de referência não encontrada.].</li> </ul>
<b>CA004.</b>	Serão 21 (vinte e um) setores de atividades, são eles: <ul style="list-style-type: none"> <li>▪ Administração pública, defesa e seguridade social;</li> <li>▪ Agricultura, pecuária, produção florestal, pesca e aquicultura;</li> <li>▪ Água, esgoto, atividades de gestão de resíduos e descontaminação;</li> <li>▪ Alojamento e alimentação;</li> <li>▪ Artes, cultura, esporte e recreação;</li> <li>▪ Atividades administrativas e serviços complementares;</li> <li>▪ Atividades financeiras, de seguros e serviços relacionados;</li> <li>▪ Atividades imobiliárias;</li> <li>▪ Atividades profissionais, científicas e técnicas;</li> <li>▪ Comércio; reparação de veículos automotores e motocicletas;</li> <li>▪ Construção;</li> <li>▪ Educação;</li> <li>▪ Eletricidade e gás;</li> <li>▪ Industrias de transformação;</li> <li>▪ Indústrias extrativas;</li> <li>▪ Informação e comunicação;</li> <li>▪ Organismos internacionais e outras instituições extraterritoriais;</li> <li>▪ Outras atividades de serviços;</li> <li>▪ Saúde humana e serviços sociais;</li> <li>▪ Serviços domésticos; e</li> <li>▪ Transporte, armazenagem e correio.</li> </ul>
<b>CA005.</b>	Serão 4 (quatro) faixas de porte por número de empregados, são elas: <ul style="list-style-type: none"> <li>• Faixa 1 – Menor ou igual a 49;</li> <li>• Faixa 2 – Entre 50 e 249;</li> <li>• Faixa 3 – Entre 250 e 499; e</li> <li>• Faixa 4 – Maior ou igual a 500.</li> </ul>
<b>CA006.</b>	Para o gráfico <b>por setores de atividades</b> , deverá apresentar um gráfico de linhas com a evolução anual das empresas em cada setor de atividade. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda com todos os setores disponíveis e suas respectivas cores; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes</li> </ul>

	informações: A cor e nome do setor e quantidade de empresas em cada setor.
<b>CA007.</b>	<p>Para o gráfico <b>por porte das empresas</b>, deve apresentar um gráfico de colunas e linhas com a evolução anual das empresas por porte. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda com todos os setores disponíveis e suas respectivas cores; e</li> <li>▪ Um hint nas colunas gráfico com as seguintes informações: Ano base, porte e quantidade de empresas do porte em questão; e</li> <li>▪ Um hint em cada ano do gráfico com as seguintes informações: Ano base e a quantidade total de empresas no ano.</li> </ul>
<b>CA008.</b>	Deverá apresentar a fonte dos dados em cada gráfico com o período de referência.

## 2.2.2. Protótipo

### 2.2.2.1. Painel Abrangência



### 2.3. Visualizar Painel Produção

ID	NOME
HST003	Visualizar Painel Produção

**COMO** Gestor Público  
**QUERO** ter acesso ao painel Produção  
**PARA** que eu possa visualizar informações referentes a geração de atividades em Pesquisa e Desenvolvimento, produções científicas e tecnológicas.

#### 2.3.1. Critérios de Aceitação

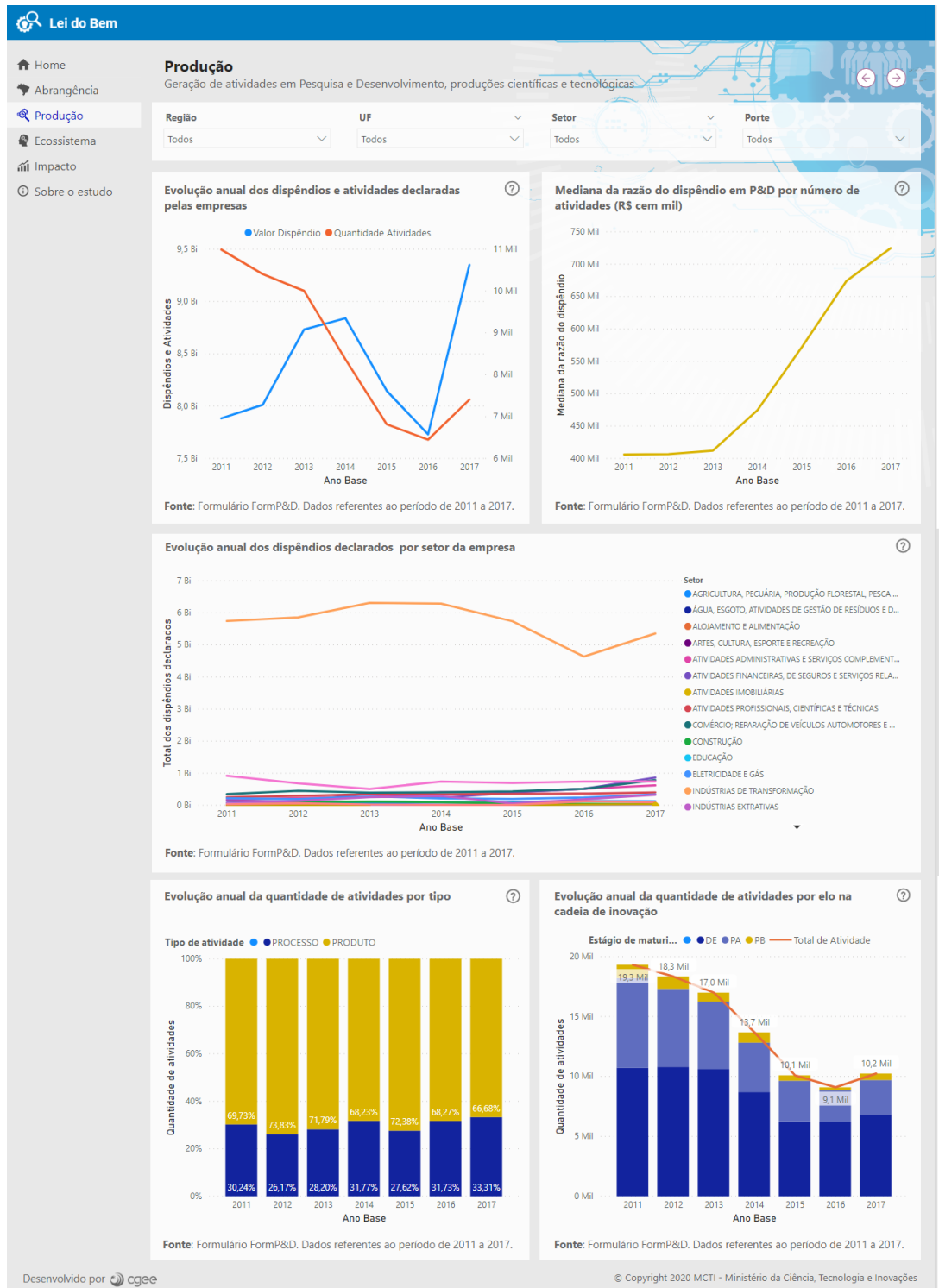
<b>CA009.</b>	O painel Produção deverá apresentar 5 (cinco) gráficos, são eles: <ul style="list-style-type: none"> <li>▪ Evolução anual dos dispêndios e atividades declaradas pelas empresas;</li> <li>▪ Mediana da razão do dispêndio em P&amp;D por número de</li> </ul>
---------------	---

	<p>atividades (R\$ cem mil);</p> <ul style="list-style-type: none"> <li>▪ Evolução anual dos dispêndios declarados por setor da empresa;</li> <li>▪ Evolução anual da quantidade de atividades por tipo; e</li> <li>▪ Evolução anual da quantidade de atividades por elo na cadeia de inovação.</li> </ul>
<b>CA010.</b>	Deverá permitir filtrar os gráficos igual ao critério <b>[CA003]</b>
<b>CA011.</b>	<p>Para o gráfico <b>Evolução anual dos dispêndios e atividades declaradas pelas empresas</b>, deverá apresentar um gráfico de linhas com os valores dos dispêndios e das quantidades de atividades por ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar a cor do valor do dispêndio e a cor da quantidade das atividades; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: Ano, Valor do dispêndio e Quantidade de atividades.</li> </ul>
<b>CA012.</b>	<p>Para o gráfico <b>Mediana da razão do dispêndio em P&amp;D por número de atividades (R\$ cem mil)</b>, deverá apresentar um gráfico de linhas com a evolução da mediana da razão do dispêndio por ano. Também deverá exibir um hint em cada ponto do gráfico com as seguintes informações: Ano e Mediana da razão do dispêndio.</p>
<b>CA013.</b>	<p>A mediana da razão do dispêndio é calculada da seguinte forma:</p> <ol style="list-style-type: none"> <li>1. Primeiro é calculada a razão do dispêndio: <math display="block">\text{Razão Dispêndio} = \frac{\text{Valor dispêndio}}{\text{Quantidade Atividades}}</math> </li> <li>2. Em seguida é realizada a mediana com a razão do dispêndio de todas as empresas.</li> </ol>
<b>CA014.</b>	<p>Para o gráfico <b>Evolução anual dos dispêndios declarados por setor da empresa</b>, deverá apresentar um gráfico de linhas com os valores totais dos dispêndios declarados por ano em cada setor. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada setor [Nome dos setores Erro! Fonte de referência não encontrada.]; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: Ano, Nome do Setor e Valor total dos dispêndios declarados no setor.</li> </ul>
<b>CA015.</b>	<p>Para o gráfico <b>Evolução anual da quantidade de atividades por tipo</b>, deverá apresentar um gráfico de colunas com a quantidade de atividades em cada tipo por ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada tipo de atividade (Processo e Produto); e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ano base, TPResultado, % Atividade e Total Atividades.</li> </ul>
<b>CA016.</b>	Para o gráfico <b>Evolução anual da quantidade de atividades por elo na cadeia de inovação</b> , deverá apresentar um gráfico de colunas e

	<p>linhas com a quantidade de atividades em cada elo por ano. Também deverá exibir:</p> <ul style="list-style-type: none"><li>▪ Uma legenda para indicar as cores de cada elo (DE, PA e PB); e</li><li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ano base, Tipo de atividade e Quantidade de atividades.</li></ul>
<b>CA017.</b>	Deverá apresentar a fonte dos dados em cada gráfico com o período de referência.

### 2.3.2. Protótipo

#### 2.3.2.1. Painel Produção



## 2.4. Visualizar Painel Ecosistema

ID	NOME
HST004	Visualizar Painel Ecosistema

**COMO** Gestor Público  
**QUERO** ter acesso ao painel Ecosistema  
**PARA** que eu possa visualizar informações referentes aos pesquisadores,



equipamentos, parcerias e patentes utilizadas.

#### 2.4.1. Critérios de Aceitação

<p><b>CA018.</b></p>	<p>O painel Produção deverá apresentar 7 (sete) gráficos, são eles:</p> <ul style="list-style-type: none"> <li>▪ Evolução anual da quantidade de pesquisadores por titulação;</li> <li>▪ Evolução anual do total de parcerias;</li> <li>▪ Evolução anual do total de dispêndios com parcerias;</li> <li>▪ Evolução anual do percentual do dispêndio com parcerias por tipo;</li> <li>▪ Evolução anual dos dispêndios com equipamentos;</li> <li>▪ Evolução anual do número de equipamentos adquiridos; e</li> <li>▪ Evolução anual do percentual da quantidade de equipamentos adquiridos para atividades de P&amp;D por porte da empresa.</li> </ul>
<p><b>CA019.</b></p>	<p>Deverá permitir filtrar os gráficos igual ao critério <b>[CA003]</b> mais o filtro por dedicação: Dedicção Exclusiva ou Dedicção Parcial</p>
<p><b>CA020.</b></p>	<p>Para o gráfico <b>Evolução anual da quantidade de pesquisadores por titulação</b>, deverá apresentar um gráfico de colunas e linhas com os valores totais de pesquisadores por titulação em cada ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada titulação [Doutor, Mestre, Pós-Graduação, Graduação, Técnico e Tecnólogo]; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ano base, o nome da titulação e a quantidade de pesquisadores.</li> <li>▪ Um hint em cada ponto da linha do gráfico com as seguintes informações: Ano base e a quantidade total de pesquisadores.</li> </ul>
<p><b>CA021.</b></p>	<p>Para o gráfico <b>Evolução anual do total de parcerias</b>, deverá apresentar um gráfico de colunas com a quantidade de parcerias por ano. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Ano base e o total de parcerias.</p>
<p><b>CA022.</b></p>	<p>Para o gráfico <b>Evolução anual do total de dispêndios com parcerias</b>, deverá apresentar um gráfico de colunas com o valor de dispêndios com parcerias por ano. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Ano base e o valor de dispêndios.</p>
<p><b>CA023.</b></p>	<p>Para o gráfico <b>Evolução anual do percentual do dispêndio com parcerias por tipo</b>, deverá apresentar um gráfico de colunas com o percentual do dispêndio com parcerias para cada tipo por ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada tipo de prestador <b>[CA024]</b>; e</li> </ul>

	<ul style="list-style-type: none"> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ano base, tipo prestador e valor (real e percentual).</li> </ul>
<b>CA024.</b>	<p>Serão 6 (seis) tipos de prestadores, são eles:</p> <ul style="list-style-type: none"> <li>▪ Apoio técnico, tecnologia industrial básica e viagens;</li> <li>▪ Instituição de pesquisa;</li> <li>▪ Inventor independente;</li> <li>▪ Microempresa;</li> <li>▪ Pequeno porte; e</li> <li>▪ Universidade.</li> </ul>
<b>CA025.</b>	<p>Para o gráfico <b>Evolução anual dos dispêndios com equipamentos</b>, deverá apresentar um gráfico de colunas com o valor dos dispêndios com equipamentos por ano. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Ano base e o valor de dispêndios com equipamentos.</p>
<b>CA026.</b>	<p>Para o gráfico <b>Evolução anual do número de equipamentos adquiridos</b>, deverá apresentar um gráfico de colunas com a quantidade de equipamentos adquiridos no ano. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Ano base e a quantidade de equipamentos.</p>
<b>CA027.</b>	<p>Para o gráfico <b>Evolução anual do percentual da quantidade de equipamentos adquiridos para atividades de P&amp;D por porte da empresa</b>, deverá apresentar um gráfico de colunas com o percentual de equipamentos adquiridos em cada porte e em cada ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada porte [Erro! Fonte de referência não encontrada.]; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ano base, porte e percentual total de equipamentos.</li> </ul>
<b>CA028.</b>	<p>Deverá apresentar a fonte dos dados em cada gráfico com o período de referência.</p>

## 2.4.2. Protótipo

### 2.4.2.1. Painel Ecosystema

Lei do Bem

- Home
- Abrangência
- Produção
- Ecosistema
- Impacto
- Sobre o estudo

### Ecosistema de inovação

Pesquisadores, equipamentos, parcerias e patentes utilizadas

Região: Todos

UF: Todos

Setor: Todos

Porte: Todos

Dedicação: Todos

#### Evolução anual da quantidade de pesquisadores por titulação

Ano	Total	Doutor	Mestre	Pós-Graduado	Graduado	Técnico
2011	70300	10847	17338	36430	10847	17338
2012	74591	9373	19193	39325	9373	19193
2013	67695	14429	6007	40054	14429	6007
2014	24910	5264	14570	5264	14570	5264
2015	54464	9682	6350	33384	9682	6350
2016	47460	8309	6401	28502	8309	6401
2017	66795	11452	8479	41899	11452	8479

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual do total de parcerias

Ano	Quantidade
2015	21074
2016	29832
2017	29961

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual do total de dispêndios com parcerias

Ano	Valor (R\$ milhões)
2015	2.485
2016	2.550
2017	2.842

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual do percentual do dispêndio com parcerias por tipo

Ano	UNIVERSIDADE	PEQUENO PORTE	MICRO EMPRESA	INVENTOR INDEPENDENTE	INSTITUIÇÃO DE PESQUISA	APOIO TÉCNICO, TECNOLOGIA INDUSTRIAL B...
2015	46,63%	6,97%	6,01%	10,10%	30,13%	
2016	46,98%	8,10%	5,96%	7,71%	31,07%	
2017	53,00%	7,71%	7,07%	27,62%		

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual dos dispêndios com equipamentos

Ano	Valor (R\$ milhões)
2015	100,78
2016	96,75
2017	188,61

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual do número de equipamentos adquiridos

Ano	Quantidade
2015	1905
2016	1804
2017	2024

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

#### Evolução anual do percentual da quantidade de equipamentos adquiridos para atividades de P&D por porte da empresa

Ano	Faixa 4 - Maior ou igual a 500	Faixa 3 - Entre 250 e 499	Faixa 2 - Entre 50 e 249	Faixa 1 - Menor ou igual a 49	(Em branco)
2015	76,33%	12,97%	9,13%		
2016	78,38%	6,10%	7,43%	8,04%	
2017	69,27%	11,91%	17,89%		

Fonte: Formulário FormP&D. Dados referentes ao período de 2011 a 2017

Desenvolvido por

© Copyright 2020 MCTI - Ministério da Ciência, Tecnologia e Inovações

## 2.5. Visualizar Painel Impacto

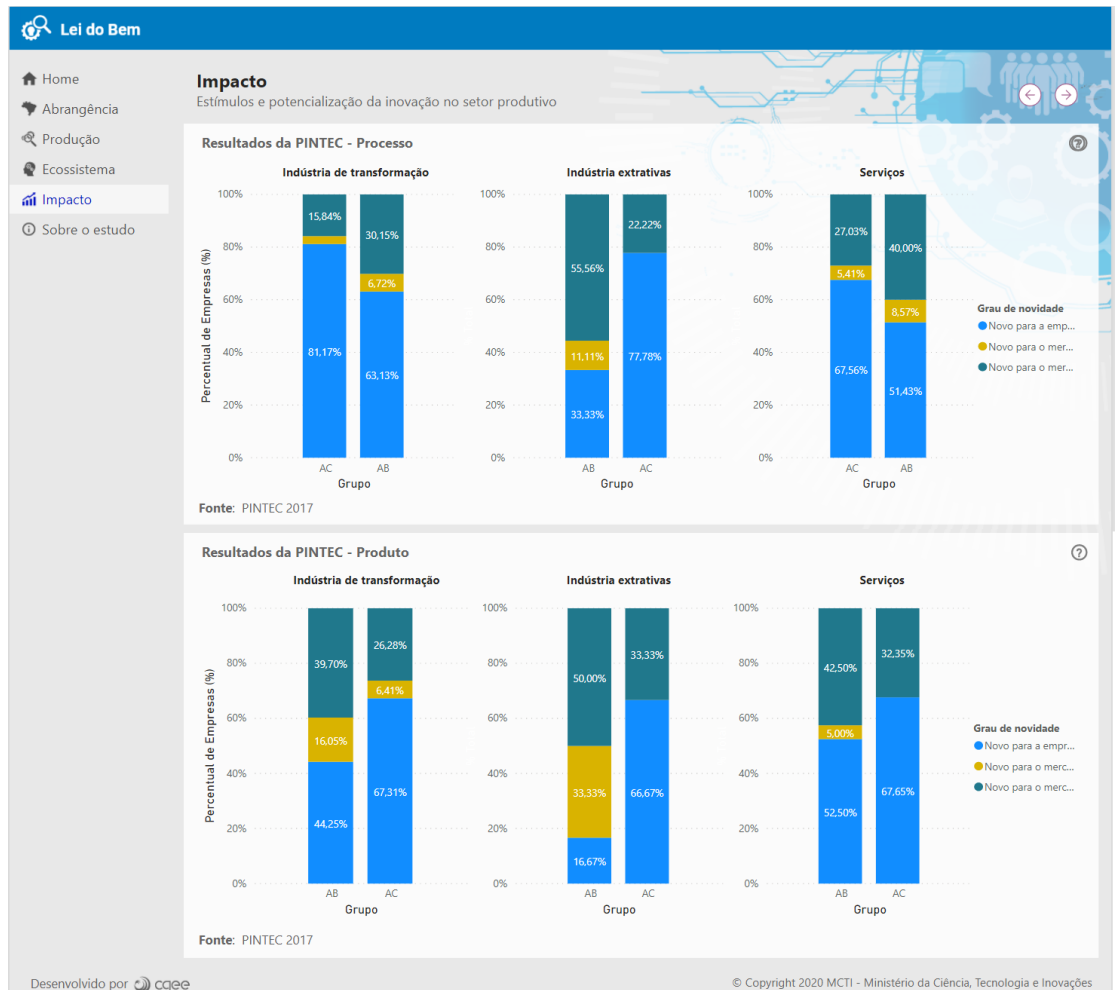
ID	NOME
HST005	Visualizar Painel Impacto
<p><b>COMO</b> Gestor Público  <b>QUERO</b> ter acesso ao painel Impacto  <b>PARA</b> que eu possa visualizar informações referentes aos estímulos e potencialização da inovação no setor produtivo</p>	

### 2.5.1. Critérios de Aceitação

<b>CA029.</b>	<p>O painel Impacto deverá apresentar 2 conjuntos de gráficos:</p> <ul style="list-style-type: none"> <li>- Resultados da PINTEC – Processo: <ul style="list-style-type: none"> <li>▪ Indústria de transformação;</li> <li>▪ Indústria extrativas; e</li> <li>▪ Serviços.</li> </ul> </li> <li>- Resultados da PINTEC – Produto: <ul style="list-style-type: none"> <li>▪ Indústria de transformação;</li> <li>▪ Indústria extrativas; e</li> <li>▪ Serviços.</li> </ul> </li> </ul>
<b>CA030.</b>	<p>Serão 3 (três) tipos de graus de novidades, são eles:</p> <ul style="list-style-type: none"> <li>▪ Novo para a empresa, mas já existente no mercado nacional;</li> <li>▪ Novo para o mercado mundial; e</li> <li>▪ Novo para o mercado nacional, mas já existente no mercado mundial.</li> </ul>
<b>CA031.</b>	<i>[Descrição dos grupos - Monique]</i>
<b>CA032.</b>	<p>Os gráficos deveram apresentar em cada setor (Indústria de transformação, Indústria extrativas e Serviços) o percentual de empresas em cada grupo e grau de novidade. Também deverá apresentar:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para indicar as cores de cada grau de novidade [CA030]; e</li> <li>▪ Um hint em cada coluna dos gráficos com as seguintes informações: Grupo, Grau de novidade e Percentual total.</li> </ul>

### 2.5.2. Protótipo

#### 2.5.2.1. Painel Impacto



## 2.6. Visualizar Painel Sobre o Estudo

ID	NOME
HST006	Visualizar Painel Sobre o Estudo
<p><b>COMO</b> Gestor Público  <b>QUERO</b> ter acesso ao painel Sobre o Estudo  <b>PARA</b> que eu possa saber mais sobre o estudo referente a aplicação da Lei do Bem</p>	

### 2.6.1. Critérios de Aceitação

<b>CA033.</b>	O painel deverá apresentar um texto referente ao estudo da aplicação da Lei do Bem.
<b>CA034.</b>	<a href="#">[Texto Lei do Bem - Gerson]</a>

### 2.6.2. Protótipo

#### 2.6.2.1. Painel Sobre o Estudo

Lei do Bem

- Home
- Abrangência
- Produção
- Ecosistema
- Impacto
- Sobre o estudo

### Sobre o estudo

Grupo de Trabalho de Aperfeiçoamento da Aplicação da Lei do Bem

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum mollis justo sit amet ligula vehicula aliquet. Sed arcu diam, auctor sit amet euismod a, ultricies et lacus. Duis nisi risus, fermentum ut tincidunt non, elementum ut eros. Maecenas metus nibh, viverra eget hendrerit ut, rhoncus non purus. Sed semper lorem maximus convallis imperdiet. In mollis venenatis venenatis. Etiam erat arcu, finibus id mi in, euismod consequat enim. Mauris iaculis ligula non posuere dictum. Fusce sit amet efficitur ipsum. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

Aliquam a porta lorem. Sed dapibus leo nec nulla tincidunt mollis at at quam. Donec venenatis enim sit amet neque gravida, eget sodales lacus feugiat. Aenean at lorem neque. Nam arcu ipsum, facilisis aliquet ipsum et, cursus ullamcorper odio. Etiam hendrerit, sem at finibus elementum, libero lectus iaculis nisi, id imperdiet nisi mauris non ligula. Nulla congue cursus massa, a condimentum nisi convallis hendrerit. Proin facilisis eros eu volutpat vestibulum.

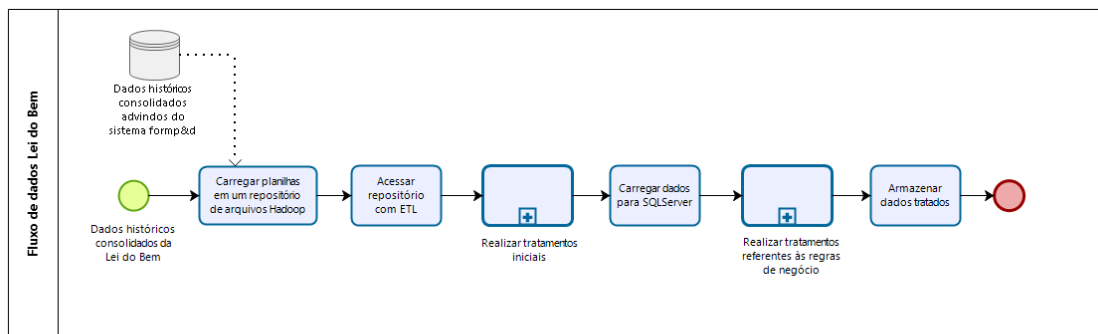
Etiam a tristique mauris, eu porta ipsum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Suspendisse in ipsum ipsum. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Phasellus quis quam pellentesque, tincidunt justo at, consequat urna. Vestibulum facilisis, neque eget placerat vehicula, augue neque rutrum turpis, et imperdiet nibh ligula et nisi. Suspendisse luctus posuere augue ut condimentum. Mauris ligula urna, blandit vitae tincidunt at, ultrices vitae felis. Praesent non odio luctus quam tincidunt sollicitudin sed quis dolor.

Donec tincidunt, quam et tincidunt imperdiet, lacus dolor hendrerit magna, ac bibendum tortor magna hendrerit orci. Quisque mollis consectetur lacinia. Nullam lacinia magna eu dapibus aliquam. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Fusce pulvinar eros in lorem vehicula, quis tempor odio aliquam. Vivamus pharetra tellus non augue luctus euismod. Pellentesque dapibus laoreet elit. Vestibulum eget suscipit massa, a blandit nisi. Fusce quis lacus purus. Morbi massa metus, consequat sed felis nec, auctor semper ante. Sed ac dolor malesuada, posuere purus suscipit, congue quam. Duis rhoncus, nibh nec elementum pharetra, arcu leo aliquam leo, eget sollicitudin lacus libero at dolor. Pellentesque est neque, consequat vitae tempus eu, euismod id risus.

Desenvolvido por cgée
© Copyright 2020 MCTI - Ministério da Ciência, Tecnologia e Inovações

### 3. DOCUMENTAÇÃO DATA PIPELINE

#### Fluxo de dados Lei do Bem



Powered by **bizagi** Modeler

**Realizar tratamentos iniciais:** subprocesso do fluxo de dados Lei do Bem responsável por realizar os tratamentos iniciais que consistem em:

- Quando uma coluna for do tipo decimal e quando o valor for igual a null, então deverá preencher com o valor 0.0. Caso o valor preenchido seja diferente de null, então deverá manter o valor existente.
- Quando uma coluna for do tipo inteiro e quando o valor for igual a null, então deverá preencher com o valor 0. Caso o valor preenchido seja diferente de null, então deverá manter o valor existente.
- Quando uma coluna for do tipo texto e quando o valor for igual a null, então deverá preencher com o valor ""'. Caso o valor preenchido seja diferente de null, então deverá manter o valor existente.
- Quando a planilha vem com uma coluna que não é do tipo inteiro e deveria ser, então deverá ser realizado um tratamento para forçar a coluna se tornar do tipo inteiro.

- Quando a planilha vem com a coluna ano base faltando, então deverá ser criada a coluna e preenchida com o ano base que deverá ser retirado do nome do arquivo; e
- Quando a planilha 01 do ano de 2017 não vier com a coluna PCParcelaFaturamentoInov então deverá ser criado a coluna e ser preenchida com o valor 0.

**Realizar tratamentos referentes às regras de negócio:** subprocesso do fluxo de dados Lei do Bem responsável por realizar os tratamentos referentes às regras de negócio definidas que consistem em:

- Referente à receita líquida, foi definido o limite superior de R\$ 327 bilhões levando em consideração o ranking do valor econômico, 2017 onde a empresa de maior faturamento foi a Petrobrás com aproximadamente 284 bilhões. Com isso, foi aplicado uma taxa de segurança de 15% no valor de 284 bilhões e definiu-se o valor do limite superior.
- Referente à receita líquida, foi definido o limite inferior de R\$ 100 mil/ano levando em consideração que a empresa elegível ao benefício precisa ter lucro no ano fiscal.
  - Neste caso, foram identificadas 45 (quarenta e cinco) ocorrências com receita líquida abaixo de R\$ 100 mil/ano, das quais 2 (duas) ocorrências de receita líquida negativa, 7 (sete) ocorrências com o valor igual a zero e 28 (vinte e oito) ocorrências que apresentam valores inferiores que R\$ 2,00/ano. Os valores identificados foram substituídos por NA.
- Referente ao dispêndio em P&D, foi definido que o limite superior seria o mesmo do limite superior da receita líquida.
- Referente ao dispêndio em P&D, foi definido o limite inferior como número negativo.
- Referente ao dispêndio em P&D, foi definido um limite inferior apenas para análise dos dispêndios de R\$ 10 mil/ano.
  - Neste caso, foi identificado um total de 85 (oitenta e cinco) ocorrências, onde 64 (sessenta e quatro) ocorrências com valor 0 (zero), 3 (três) ocorrências com valor de R\$1 à R\$100 e 18 (dezoito) ocorrências entre R\$1.000,00 e R\$10.000,00. Os valores identificados foram substituídos por NA.
- Referente ao filtro de consistência, foi definido a condição:  $(\text{dispêndio em RH}) + (\text{Dispêndio com Equipamento}) + (\text{Dispêndios com parcerias}) \leq (\text{dispêndio em atividades P\&D})$ .
  - Neste caso, foram identificadas 41 ocorrências falsas e 3716 ocorrências verdadeiras para a regra. Com isso, um pouco mais de 1% da amostra é referente a ocorrências falsas. Estas informações não foram utilizadas na análise de dispêndio porque não há como ter certeza de qual a informação é correta a informação do

dispêndio por insumo (RH, equipamento e parceria) ou o dispêndio total em atividades de P&D.

- Referente ao dispêndio em P&D por titulação, foi definido a condição: *(Total dos dispêndios com pesquisadores) = (soma dos dispêndios com pesquisadores com cada titulação)*.
  - Neste caso, foi identificado que o total não corresponde ao desagregado (por titulação) e das 7.965 ocorrências na base apenas 3.623 são verdadeiras para esta regra. Desta forma, foi criado um novo total com a soma das informações desagregadas e calculado o percentual de cada titulação com base nesse novo total.

Com esse resultado, haveria uma perda de aproximadamente 60% das observações do total da base. Optou-se então por uma validação referente a variável *valortotalrh*, onde tem-se os valores somadas de todas as titulações, e as variáveis que contabilizam os valores de cada titulação. Dessa forma, foram substituídas por NA as observações que na coluna por titulação, diferia em 15% para mais ou para menos da variável *valortotalrh*.
- Referente a apenas para análises de RH, foi definido um limite inferior (por empresa) de acordo com a condição: *quando for igual a zero na variável dispêndio em RH ou quando for igual a zero na variável horas de RH*.
  - Neste caso, foram substituídos por NA as observações que se encontravam nessa condição referente ao dispêndio.
- Referente a variável de horas, ela existe na base apenas contabilizando por titulações, não tendo ela de forma somada. Então, criou-se uma variável com os valores de horas somados.
  - Neste caso, quando a variável somada apresentou valores nulos então foi substituída por NA.
- Referente a apenas para análises de RH, foi definido um limite superior (por empresa) de acordo com a condição: *(dispêndio em RH) ≤ (dispêndio total das atividades de P&D da empresa)*.
- Referente a análise da relação dispêndio em RH e horas de RH, foi aplicado o Limite superior (por empresa) de acordo com a condição: *(dispêndio em RH)/(horas de RH) < 1.000,00*
  - Criou-se uma variável *disp\_horas\_rh* onde fez-se a razão da variável *valortotalrh* por *horassomadasrh* e foram substituídos por NA os valores maiores que 1000. Os valores antigos estão sem filtros da variável.
- Referente a apenas para análises de equipamentos, foi definido um limite inferior (por empresa) de acordo com a condição: *quando for igual a zero na variável dispêndio em Equipamentos ou quando for igual a zero na variável quantidade de equipamentos*.



- Referente a apensas para análises de equipamentos, foi definido um limite superior (por empresa) de acordo com a condição:  $(disp\hat{e}ndio\ em\ Equipamento) \leq (disp\hat{e}ndio\ total\ das\ atividades\ de\ P\&D\ da\ empresa)$ .
  - Foram identificadas 13 (treze) ocorrências equivalente a aproximadamente 12% das ocorrências (106) com dispêndios em equipamentos. (Ainda não aplicado)
- Referente a apensas para análises de parceria, foi definido um limite inferior (por empresa) de acordo com a condição: *quando for igual a zero na variável dispêndio em parcerias ou quando for igual a zero na variável quantidade de parcerias.*
- Referente a apensas para análises de parceria, foi definido um limite superior (por empresa) de acordo com a condição:  $(disp\hat{e}ndio\ em\ parcerias) \leq (disp\hat{e}ndio\ total\ das\ atividades\ de\ P\&D\ da\ empresa)$ .

## 4. DICIONÁRIO DE VARIÁVEIS

O dicionário de variáveis se encontra na planilha **Mapa conceitual Vfinal3.xlsx**

Extrato da planilha:

Dimensão	Indicador	Tipo de indicador		Descrição	Cálculo	Parâmetro		Fonte	Ano do FormP&D					
		Fimto	Desempenho			classe 1	classe 2		2014	2017	2016	2015	2014	
Evolução anual do número absoluto de empresas demandantes por "parâmetro"(a);	Resultado	Eficácia	Evolução anual do número absoluto de empresas demandantes dos benefícios da Lei, por setor;	Evolução anual do número absoluto de empresas demandantes dos benefícios da Lei, por região;	Evolução anual do número absoluto de empresas demandantes dos benefícios da Lei, por UF;	Evolução anual do número absoluto de empresas demandantes dos benefícios da Lei, por porte;	Setor (s);	t é o ano, onde 2011 ≤ t ≤ 2017; α = s, onde 1 ≤ s ≤ 21 (ver planilha ANEXO I dos setores CNAE/2015)	FormP&D, CNAE 2.0/2015	1,2;	1,2;	1,2;	1,2;	1,2;
							Região geográfica (r);	t é o ano, onde 2011 ≤ t ≤ 2017; α = r, onde r ∈ {N; S; NE; SE; CO}	FormP&D	1,10;	1,10;	1,10;	1,10;	1,10;
							UF (u);	t é o ano, onde 2011 ≤ t ≤ 2017; α = u, onde u ∈ {AC; AL; AP; AM; BA; CE; DF; ES; GO; MA; MT; MS; MG; PA; PB; PR; PE; PI; RJ; RN; RS; RO; RR; SC; SP; SE; TO}	FormP&D	1,9;	1,9;	1,9;	1,9;	1,9;
							Porte da empresa (p)	t é o ano, onde 2011 ≤ t ≤ 2017; α = p, onde 1 ≤ p ≤ 5 (ver planilha ANEXO II com a classificação)	FormP&D	2,3;	2,3;	2,3;	2,3;	2,3;
Evolução anual do percentual de empresas demandantes dos benefícios da Lei, por setor;	Resultado	Eficácia	Evolução anual do percentual de empresas demandantes dos benefícios da Lei, por região;	Evolução anual do percentual de empresas demandantes dos benefícios da Lei, por UF;	Evolução anual do percentual de empresas demandantes dos benefícios da Lei, por porte;	Setor (s);	t é o ano, onde 2011 ≤ t ≤ 2017; α = s, onde 1 ≤ s ≤ 21 (ver planilha ANEXO I dos setores CNAE/2015)	FormP&D, CNAE 2.0/2015	1,2;	1,2;	1,2;	1,2;	1,2;	
						Região geográfica (r);	t é o ano, onde 2011 ≤ t ≤ 2017; α = r, onde r ∈ {N; S; NE; SE; CO}	FormP&D	1,10;	1,10;	1,10;	1,10;	1,10;	
						UF (u);	t é o ano, onde 2011 ≤ t ≤ 2017; α = u, onde u ∈ {AC; AL; AP; AM; BA; CE; DF; ES; GO; MA; MT; MS; MG; PA; PB; PR; PE; PI; RJ; RN; RS; RO; RR; SC; SP; SE; TO}	FormP&D	1,9;	1,9;	1,9;	1,9;	1,9;	
						Porte da empresa (p)	t é o ano, onde 2011 ≤ t ≤ 2017; α = p, onde 1 ≤ p ≤ 5 (ver planilha ANEXO II com a classificação)	FormP&D	2,3;	2,3;	2,3;	2,3;	2,3;	
Evolução anual da taxa de concentração anual de empresas demandantes dos benefícios da Lei, por setor;	Resultado	Eficácia	Evolução anual taxa de concentração anual de empresas demandantes dos benefícios da Lei, por região;	Evolução anual taxa de concentração anual de empresas demandantes dos benefícios da Lei, por UF;	Evolução anual taxa de concentração anual de empresas demandantes dos benefícios da Lei, por porte;	Setor (s);	t é o ano, onde 2011 ≤ t ≤ 2017; α = s, onde 1 ≤ s ≤ 21 (ver planilha ANEXO I dos setores CNAE/2015); Como "i" possui muitas (21) classes, então i = α;	FormP&D, CNAE 2.0/2015	1,2;	1,2;	1,2;	1,2;	1,2;	
						Região geográfica (r);	t é o ano, onde 2011 ≤ t ≤ 2017; α = r, onde r ∈ {N; S; NE; SE; CO}. Como "r" possui apenas 5 classes, então i = α;	FormP&D	1,10;	1,10;	1,10;	1,10;	1,10;	
						UF (u);	t é o ano, onde 2011 ≤ t ≤ 2017; α = u, onde u ∈ {AC; AL; AP; AM; BA; CE; DF; ES; GO; MA; MT; MS; MG; PA; PB; PR; PE; PI; RJ; RN; RS; RO; RR; SC; SP; SE; TO}. Como "u" possui muitas classes (27), então i = α;	FormP&D	1,9;	1,9;	1,9;	1,9;	1,9;	
						Porte da empresa (p)	t é o ano, onde 2011 ≤ t ≤ 2017; α = p, onde 1 ≤ p ≤ 5 (ver planilha ANEXO II)	FormP&D	2,3;	2,3;	2,3;	2,3;	2,3;	

## 5. REFERÊNCIAS

- Planilha: Mapa conceitual Vfinal3.xlsx
- Planilha: planilhas\_campos\_cglt.xlsx
- Apresentação: LdB, GT, Indicadores v3.1 Status FINAL\_GT 01set2020.pptx



---

**Anexo III – Tema Estratégico FNDCT / Fundos Setoriais**

# **ESPECIFICAÇÃO TÉCNICA**

FNDCT/Fundos Setoriais

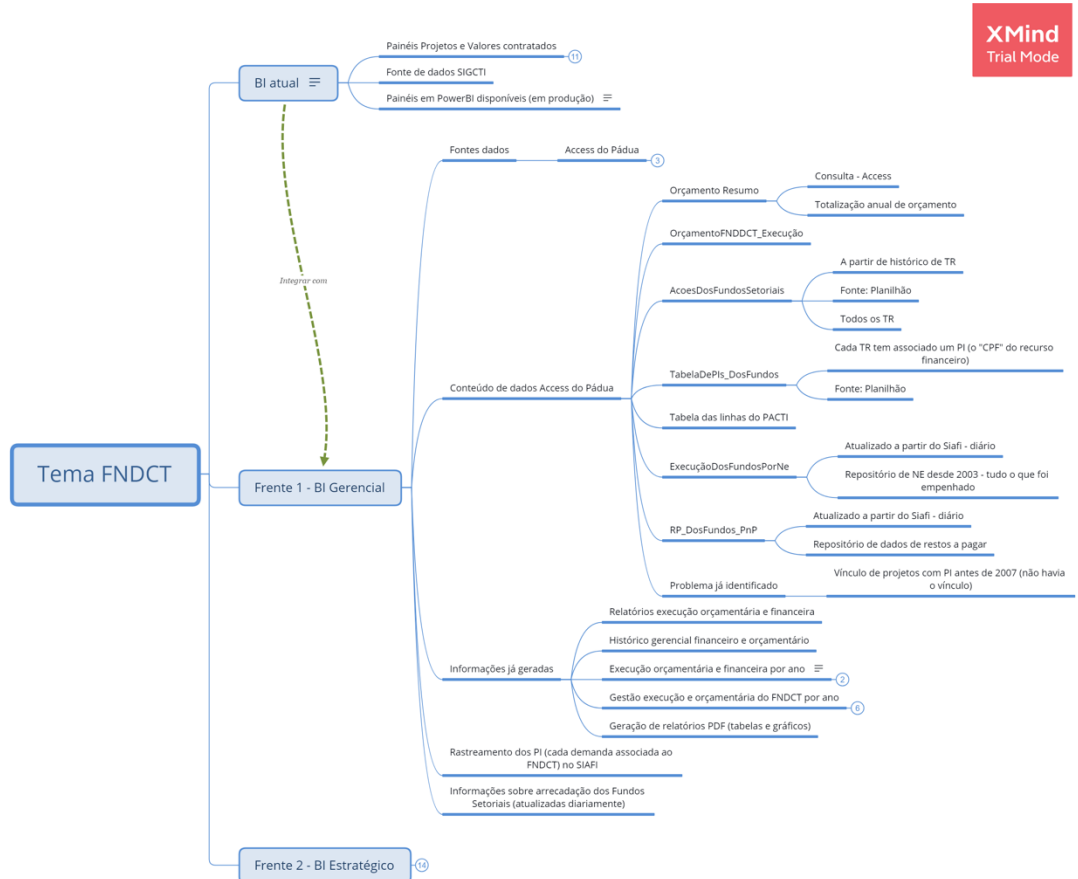
# 1. INTRODUÇÃO

Este documento tem como objetivo especificar todos os requisitos referentes ao projeto FNDCT – Fundo Nacional de Desenvolvimento Científico e Tecnológico.

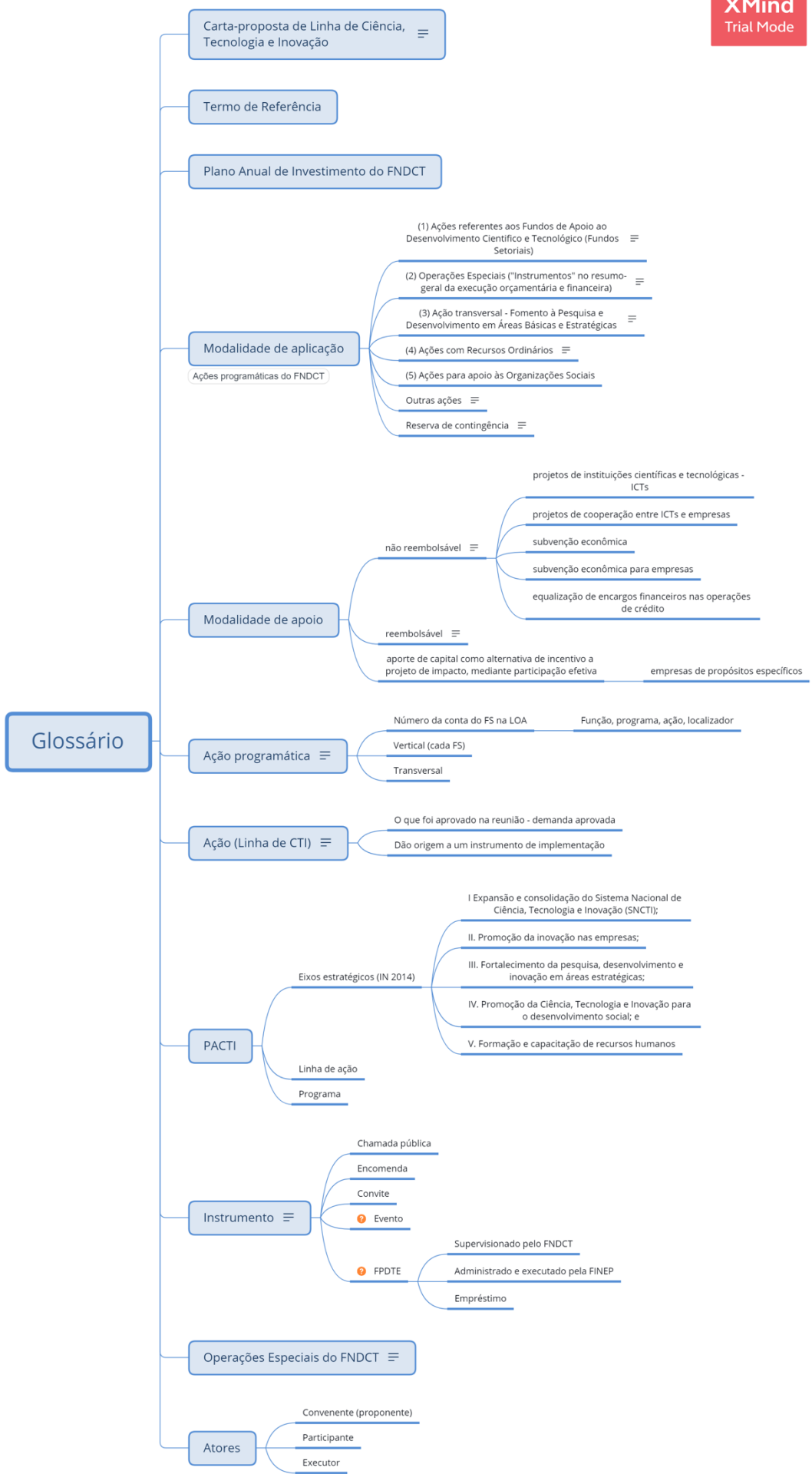
## 1.1. CONTEXTUALIZAÇÃO

Foram gerados 4 mapas mentais para auxiliar na execução do projeto. São eles:

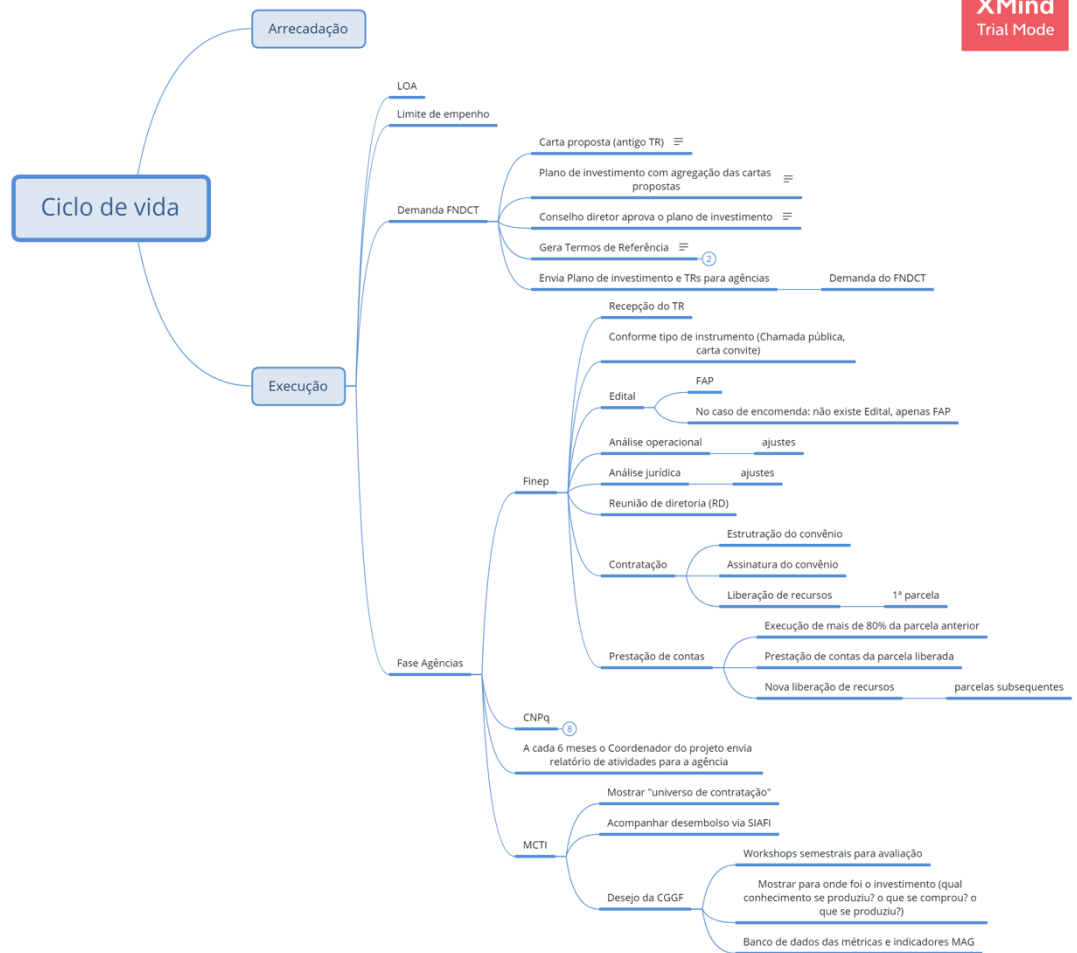
### 1. Ponto de Partida



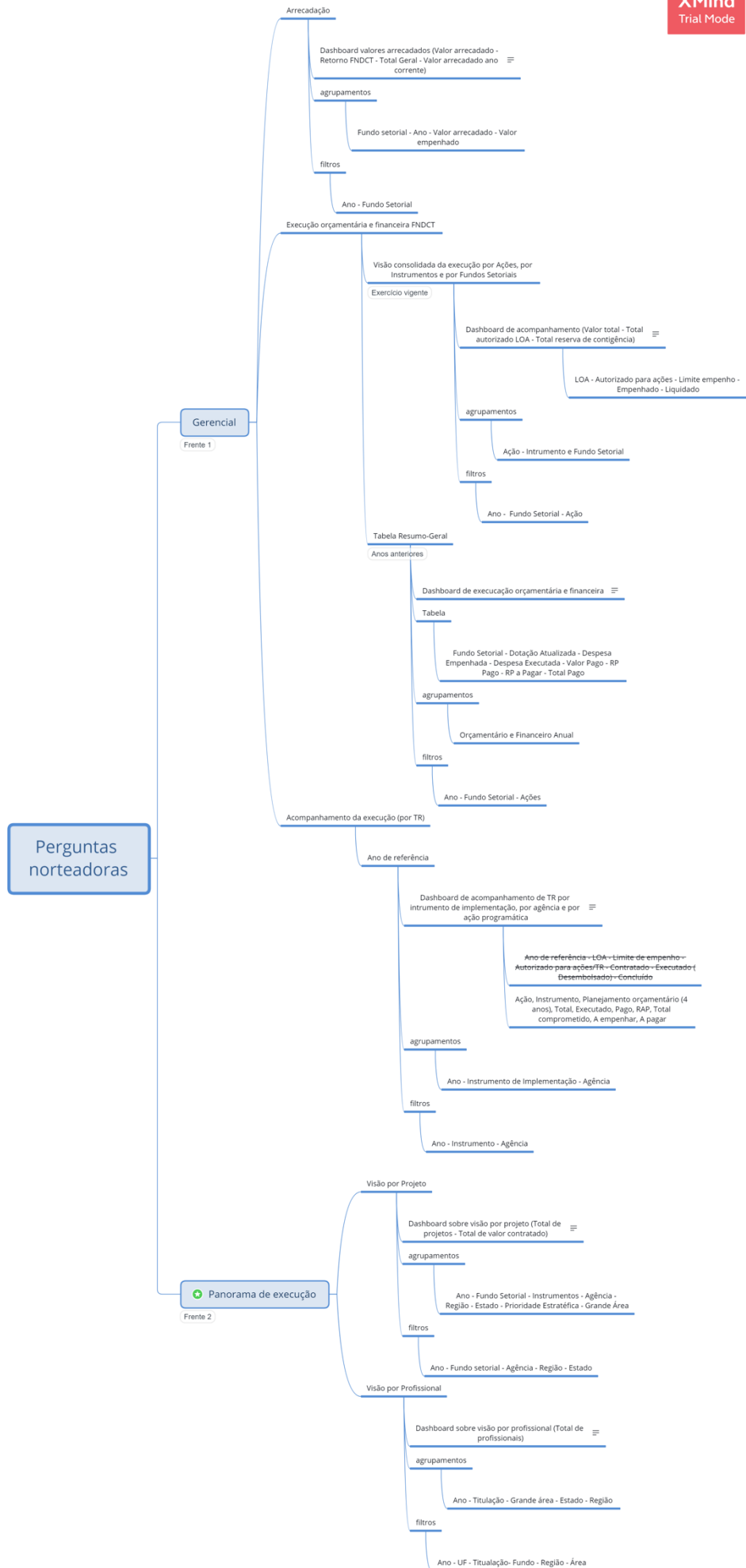
## 2. Glossário



### 3. Ciclo de Vida



### 4. Perguntas Norteadoras



## 2. HISTÓRIA DE USUÁRIO

### 2.1. Visualizar Painel Home FNDCT

ID	NOME
HST001	Visualizar Painel Home FNDCT

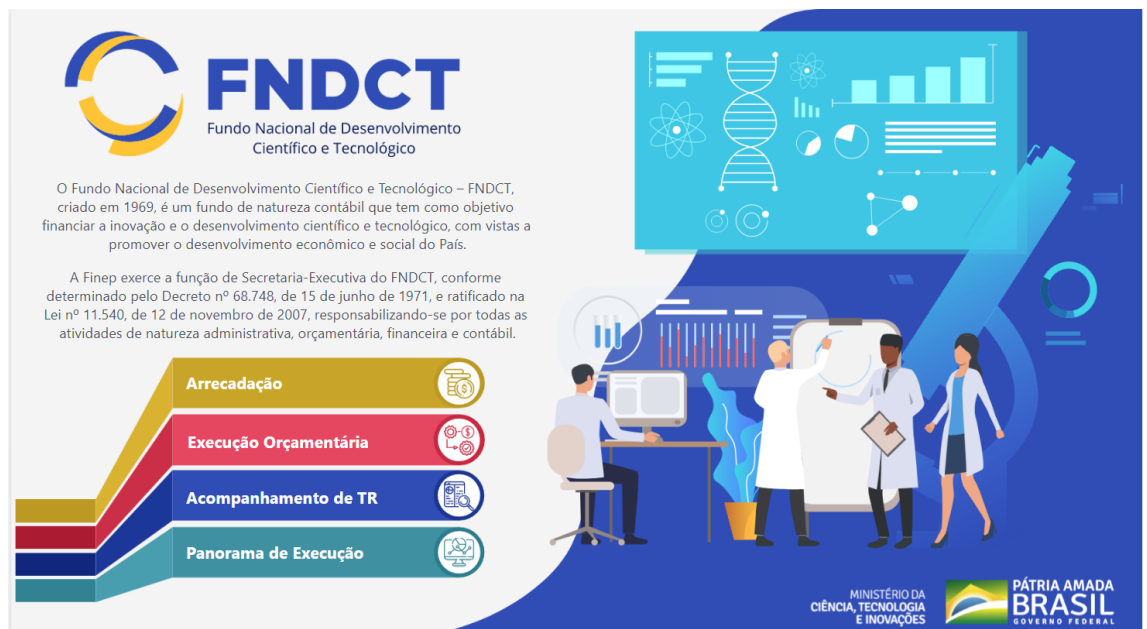
**COMO** Gestor  
**QUERO** visualizar todos painéis do FNDCT  
**PARA** que eu possa escolher qual painel desejo acessar

#### 2.1.1. Critérios de Aceitação

<b>CA035.</b>	A tela inicial deverá exibir atalhos para acessar os 4 (quatro) painéis disponíveis: <b>Arrecadação, Execução Orçamentária, Acompanhamento de TR e Panorama de Execução</b>
---------------	---

#### 2.1.2. Protótipo

##### 2.1.2.1. Tela Home FNDCT



### 2.2. Visualizar Painel Arrecadação

ID	NOME
HST002	Visualizar Painel Arrecadação

**COMO** Gestor  
**QUERO** ter acesso ao painel Arrecadação  
**PARA** que eu possa visualizar informações referentes a arrecadações por fundo e por empenho



### 2.2.1. Critérios de Aceitação

<b>CA036.</b>	O painel Arrecadação deverá apresentar 3 (três) gráficos, são eles: <ul style="list-style-type: none"> <li>▪ Valor arrecadado por Fundo;</li> <li>▪ Valor arrecadado por Ano; e</li> <li>▪ Valor arrecadado e empenhado por Ano.</li> </ul>
<b>CA037.</b>	Deverá permitir filtrar os gráficos por: <ul style="list-style-type: none"> <li>▪ Ano; e</li> <li>▪ Fundo.</li> </ul>
<b>CA038.</b>	No campo <b>Valor Arrecadado</b> , deverá exibir: <ul style="list-style-type: none"> <li>▪ Totalizador Fundo Setoriais: Resultado do somatório dos valores arrecadados em todos os fundos setoriais; e</li> <li>▪ Totalizador Retorno FNDC: Valor de retorno FNDCT.</li> </ul>
<b>CA039.</b>	No campo <b>Valor Arrecadado</b> , deverá exibir um ícone com informações complementares sobre o campo, que quando o usuário passar o mouse, deverá exibir o seguinte texto: <i>Valor arrecadado dos fundos setoriais e valor de retorno FNDCT</i>
<b>CA040.</b>	No totalizador <b>Total Geral</b> , deverá exibir o resultado do somatório dos valores arrecadados dos fundos setoriais com o valor de retorno FNDCT. - Total Geral = [Valor fundos setoriais] + [Retorno FNDCT]
<b>CA041.</b>	No totalizador <b>Total Geral</b> , deverá exibir um ícone com informações complementares sobre o campo, que quando o usuário passar o mouse, deverá exibir o seguinte texto: <i>Valor total arrecadado somando os valores dos fundos setoriais mais o retorno FNDCT</i>
<b>CA042.</b>	No totalizador <b>Arrecadado ano corrente</b> , deverá exibir o resultado do somatório dos valores arrecadados no ano corrente.
<b>CA043.</b>	No totalizador <b>Arrecadado ano corrente</b> , deverá exibir um ícone com informações complementares sobre o campo, que quando o usuário passar o mouse, deverá exibir o seguinte texto: <i>Valor total arrecadado somando os valores dos fundos setoriais em 2020</i>
<b>CA044.</b>	Para o gráfico <b>Valor arrecadado por Fundo</b> , deverá apresentar um gráfico de colunas com o valor total arrecadado em cada fundo setorial e o valor total arrecadado no retorno FNDCT. Também deverá exibir um hint nas colunas do gráfico com as seguintes informações: Fundo setorial e Valor.
<b>CA045.</b>	Para o gráfico <b>Valor arrecadado por Ano</b> , deverá apresentar um gráfico de linhas com o valor arrecadado em cada ano nos fundos, no retorno FNDCT e o valor total da arrecadação (fundos + retorno). Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores arrecadados dos fundos, do retorno FNDCT e do valor total arrecadado; e</li> <li>▪ Um hint em cada ponto das linhas com as seguintes informações: Ano base, valor arrecadado fundos, valor arrecadado retorno e valor total arrecadado.</li> </ul>

<b>CA046.</b>	Para o gráfico <b>Valor arrecadado e empenhado por Ano</b> deverá apresentar um gráfico de colunas e linha com o valor arrecadado em cada ano nos fundos, no retorno FNDCT e o valor total da arrecadação (fundos + retorno). Também deverá exibir: <ul style="list-style-type: none"><li>▪ Uma legenda para representar as cores dos valores arrecadados dos fundos, do retorno FNDCT e do valor total arrecadado; e</li><li>▪ Um hint em cada ponto das linhas com as seguintes informações: Ano base, valor arrecadado fundos, valor arrecadado retorno e valor total arrecadado.</li></ul>
---------------	--

## 2.2.2. Protótipo

### 2.2.2.1. Tela Painel Arrecadação

**FNDCT**  
Fundo Nacional de Desenvolvimento Científico e Tecnológico

ARRECADADAÇÃO

EXECUÇÃO ORÇAMENTÁRIA

ACOMPANHAMENTO DE TR

PANORAMA DE EXECUÇÃO

**ARRECADADAÇÃO**

**Filtros**

Ano: 2020

Fundo: Todos

Valor Arrecadado

**5 Bi**  
Fundo Setorias

**1 Bi**  
Retorno FNDCT

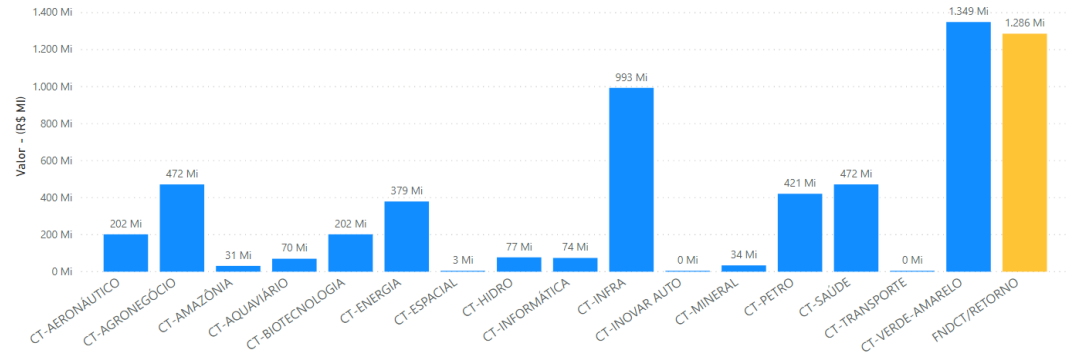
Total Geral

**6 Bi**

Arrecadado ano corrente

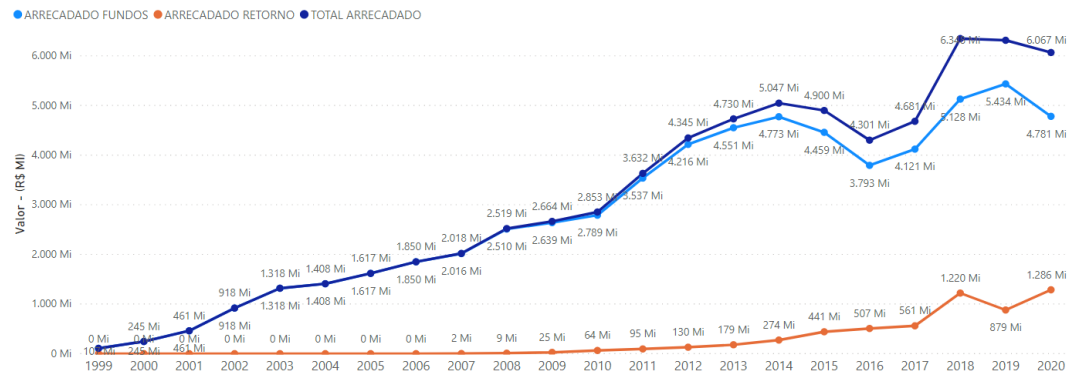
**6 Bi**

**Valor arrecadado por Fundo**



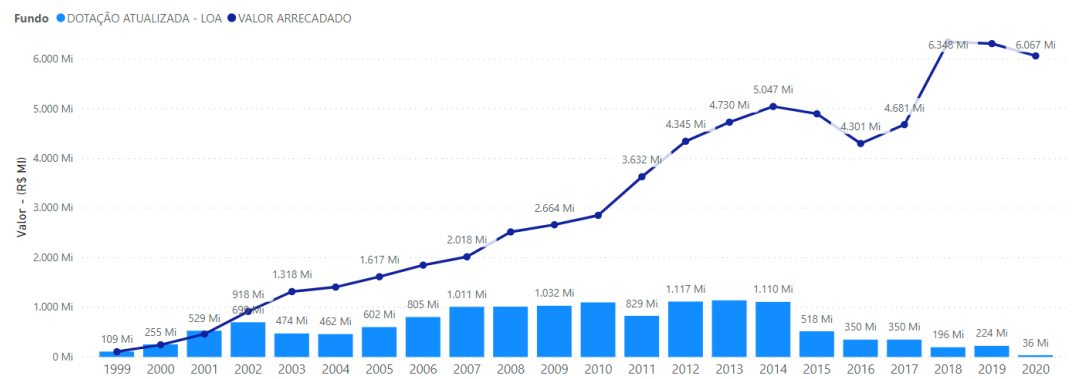
Fonte: Siafi.

**Valor arrecadado por Ano**



Fonte: Siafi.

**Valor arrecadado e empenhado por Ano**



Fonte: Siafi.

## 2.3. Visualizar Painel Execução Orçamentária – Visão Consolidada

ID	NOME
HST003	Visualizar Painel Execução Orçamentária – Visão Consolidada
<p><b>COMO</b> Gestor  <b>QUERO</b> ter acesso ao painel Execução Orçamentária visão Consolidada  <b>PARA</b> que eu possa visualizar informações referentes a execução orçamentária por ação, por instrumentos e por fundos setoriais.</p>	

### 2.3.1. Critérios de Aceitação

<b>CA047.</b>	O painel Execução Orçamentária Visão Consolidada deverá apresentar 4 (quatro) gráficos, são eles: <ul style="list-style-type: none"> <li>▪ Execução Orçamentária por Ação;</li> <li>▪ Execução Orçamentária por Instrumentos;</li> <li>▪ Execução Orçamentária por Fundo Setorial; e</li> <li>▪ Execução Financeira por Fundo Setorial.</li> </ul>
<b>CA048.</b>	Deverá permitir filtrar os gráficos por: <ul style="list-style-type: none"> <li>▪ Ano;</li> <li>▪ Fundo Setoriais; e</li> <li>▪ Ação.</li> </ul>
<b>CA049.</b>	No totalizador <b>Total Autorizado LOA</b> , deverá exibir o resultado do somatório dos valores autorizados LOA.
<b>CA050.</b>	No totalizador <b>Total Autorizado LOA</b> , deverá exibir um ícone com informações complementares sobre o campo, que quando o usuário passar o mouse, deverá exibir o seguinte texto: <i>Valor total LOA</i>
<b>CA051.</b>	No totalizador <b>Total Empenhado</b> , deverá exibir o resultado do somatório dos valores empenhados.
<b>CA052.</b>	No totalizador <b>Total Autorizado LOA</b> , deverá exibir um ícone com informações complementares sobre o campo, que quando o usuário passar o mouse, deverá exibir o seguinte texto: <i>Valor total empenhado</i>
<b>CA053.</b>	No totalizador <b>Total Reserva de Contingência</b> , deverá apresentar o resultado do somatório das reservas de contingência.
<b>CA054.</b>	Para o gráfico <b>Execução Orçamentária por Ação</b> , deverá apresentar um gráfico de colunas com os valores de despesas empenhadas e dotação atualização – LOA por cada ação programática. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores das despesas empenhadas e dotação atualizada – LOA; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ação e Valor Dotação Atualizada ou Valor Despesas Empenhadas.</li> </ul>
<b>CA055.</b>	Para o gráfico <b>Execução Orçamentária por Instrumentos</b> , deverá

	<p>apresentar um gráfico de colunas com os valores de despesas empenhadas e dotação atualização – LOA por cada instrumento. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores das despesas empenhadas e dotação atualizada – LOA; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Ação e Valor Dotação Atualizada ou Valor Despesas Empenhadas.</li> </ul>
<p><b>CA056.</b></p>	<p>Para o gráfico <b>Execução Orçamentária por Fundo Setorial</b>, deverá apresentar um gráfico de colunas com os valores das arrecadações, valores empenhados e valores autorizados LOA para cada fundo Setorial. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores arrecadados, valores autorizados LOA e valores empenhados; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna do valor arrecadado: Fundo Setorial e Valor arrecadado;</li> <li>- Quando for coluna do valor autorizado LOA: Fundo Setorial e Valor Autorizado LOA</li> <li>- Quando for coluna do valor empenhado: Fundo Setorial e Valor Empenhado.</li> </ul> </li> </ul>
<p><b>CA057.</b></p>	<p>Para o gráfico <b>Execução Financeira por Fundo Setorial</b>, deverá apresentar um gráfico de colunas com os valores das despesas pagas, restos a pagar pagos e o valor total pago para cada fundo Setorial. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores das despesas pagas, restos a pagar e do total pago; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna do valor despesas pagas: Fundo Setorial e Valor despesas pagas;</li> <li>- Quando for coluna do valor restos a pagar pagos: Fundo Setorial e Valor restos a pagar pagos; e</li> <li>- Quando for coluna do valor total pago: Fundo Setorial e Valor total pago</li> </ul> </li> </ul>

### 2.3.2. Protótipo

#### 2.3.2.1. Tela Painel Execução Orçamentária Visão Consolidada

**FNDCT**  
Fundo Nacional de Desenvolvimento Científico e Tecnológico

ARRECAÇÃO
EXECUÇÃO ORÇAMENTÁRIA
ACOMPANHAMENTO DE TR
PANORAMA DE EXECUÇÃO

**EXECUÇÃO ORÇAMENTÁRIA**

Exercício Vigente    Resumo geral

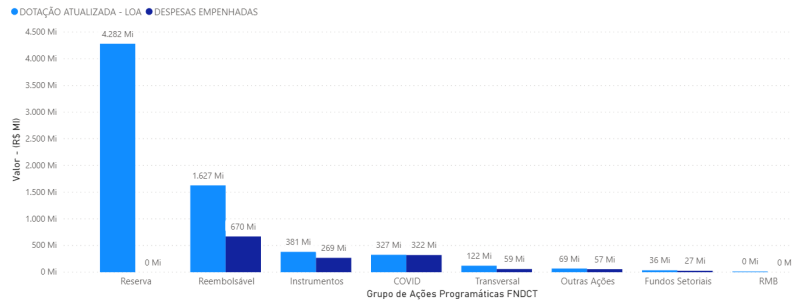
**Filtros**

Ano: 2020    Fundo Setoriais: Todos    Ação: Todos

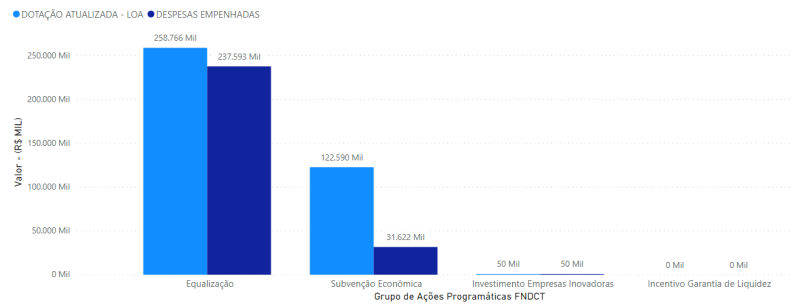
Total Autorizado LOA  
**6,84 Bi**

Total Reserva de Contingência  
**1,40 Bi**  
**4,28 Bi**

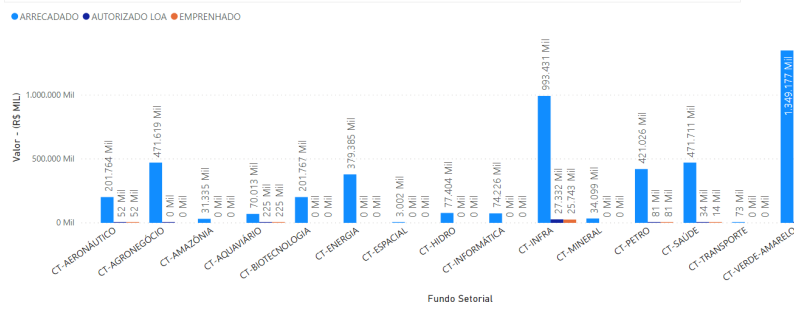
**EXECUÇÃO ORÇAMENTÁRIA POR AÇÃO**  
DOTAÇÃO ATUALIZADA - LOA e DESPESAS EMPENHADAS



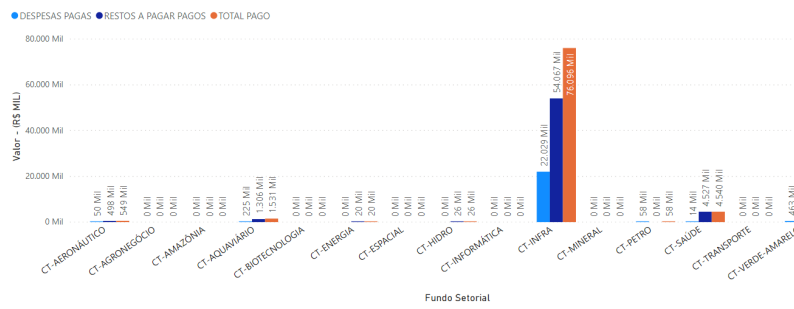
**EXECUÇÃO ORÇAMENTÁRIA POR INSTRUMENTOS**  
DOTAÇÃO ATUALIZADA - LOA e DESPESAS EMPENHADAS



**EXECUÇÃO ORÇAMENTÁRIA POR FUNDO SETORIAL**  
ARRECAÇÃO, DESPESAS EMPENHADAS E LOA



**EXECUÇÃO FINANCEIRA POR FUNDO SETORIAL**  
DESPESAS PAGAS E RESTOS A PAGAR



## 2.4. Visualizar Painel Execução Orçamentária – Visão Resumo Geral

ID	NOME
HST004	Visualizar Painel Execução Orçamentária – Visão Resumo Geral
<p><b>COMO</b> Gestor</p> <p><b>QUERO</b> ter acesso ao painel Execução Orçamentária visão Resumo Geral</p> <p><b>PARA</b> que eu possa visualizar informações gerais referentes a execução orçamentária em cada ano base</p>	

### 2.4.1. Critérios de Aceitação

<b>CA058.</b>	<p>O painel Execução Orçamentária Visão Resumo Geral deverá apresentar 2 (dois) gráficos e uma tabela:</p> <ul style="list-style-type: none"> <li>▪ Execução Orçamentária Gráfico);</li> <li>▪ Execução Financeira (Gráfico); e</li> <li>▪ Resumo Geral da Execução Orçamentária e Financeira (Tabela).</li> </ul>
<b>CA059.</b>	<p>Deverá permitir filtrar os gráficos por:</p> <ul style="list-style-type: none"> <li>▪ Ano;</li> <li>▪ Fundo Setoriais; e</li> <li>▪ Ação.</li> </ul>
<b>CA060.</b>	<p>Para o gráfico <b>Execução Orçamentária</b>, deverá apresentar um gráfico de linhas com os valores das arrecadações, valores empenhados e valores autorizados LOA para cada ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores arrecadados, valores autorizados LOA e valores empenhados; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: Ano, Valor arrecadado, Valor Autorizado Loa e Valor empenhado</li> </ul>
<b>CA061.</b>	<p>Para o gráfico <b>Execução Financeira</b>, deverá apresentar um gráfico de linhas com os valores pago no exercício, valores restos a pagar pagos e valores total pago para cada ano. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores pago no exercício, valores restos a pagar pagos e valores total pago; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: Ano, Valor Pago no exercício, Valor Restos a pagar pagos e Valor pago.</li> </ul>
<b>CA062.</b>	<p>Para a tabela <b>Resumo Geral da Execução Orçamentária e Financeira</b>, deverá apresentar uma tabela com as seguintes informações:</p> <ul style="list-style-type: none"> <li>▪ Fundos;</li> <li>▪ Dotação Atualizada;</li> <li>▪ Despesa Empenhada;</li> </ul>

	<ul style="list-style-type: none"><li>▪ Despesas Executadas;</li><li>▪ Valor Pago;</li><li>▪ RP Pagos;</li><li>▪ RP a Pagar;</li><li>▪ Total Pago; e</li><li>▪ Total.</li></ul>
--	---

## 2.4.2. Protótipo

### 2.4.2.1. Tela Painel Execução Orçamentária – Visão Resumo Geral



ARRECADAÇÃO
EXECUÇÃO ORÇAMENTÁRIA
ACOMPANHAMENTO DE TR
PANORAMA DE EXECUÇÃO

**EXECUÇÃO ORÇAMENTÁRIA**

Visão consolidada
Resumo geral

←
→

**Filtros**

Ano

2020

Fundo Setoriais

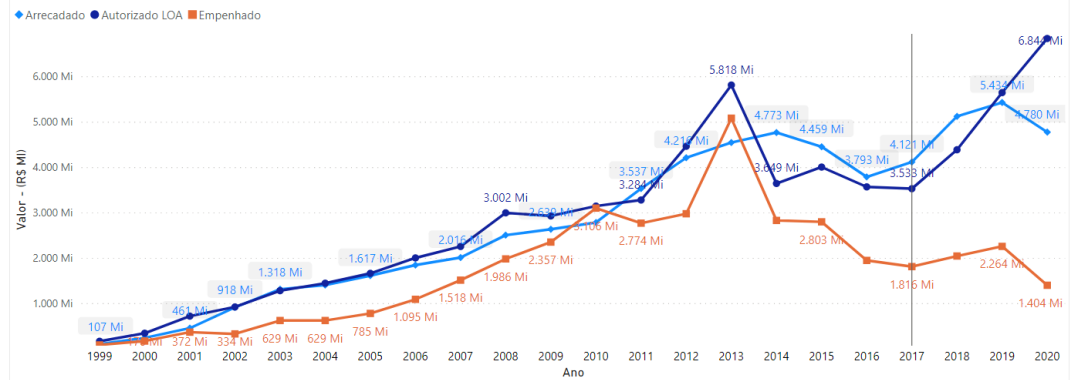
Todos

Ações

Todos

### EXECUÇÃO ORÇAMENTÁRIA

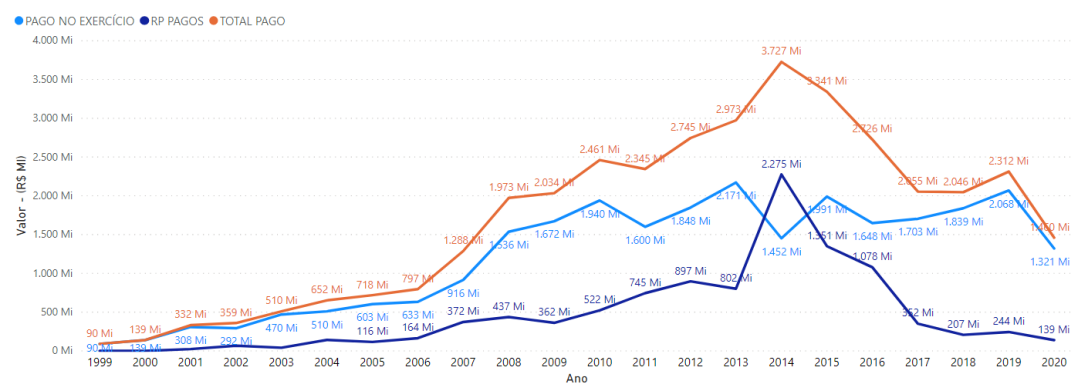
ARRECADAÇÃO, DESPESA EMPENHADA e LOA



Fonte: Siafi.

### EXECUÇÃO FINANCEIRA

DESPESAS PAGAS e RESTOS A PAGAR PAGO



Fonte: Siafi.

### RESUMO-GERAL DA EXECUÇÃO ORÇAMENTÁRIA E FINANCEIRA

Fundos	Dotação Atualizada	Despesa Empenhada	Despesas Executadas	Valor Pago	RP Pagos	RP a Pagar	Total Pago
CT-AERONÁUTICO	52.296,00	52.295,40	50.094,96	50.094,96	498.449,63	1.080.619,77	548.544,59
CT-AGRONEGÓCIO	1,00	0,00	0,00	0,00	0,00	1.820.933,86	0,00
CT-AMAZÔNIA	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CT-AQUAVIÁRIO	225.409,00	225.408,05	225.408,05	225.408,05	1.305.914,19	0,00	1.531.322,24
CT-BIOTECNOLOGIA	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CT-ENERGIA	0,00	0,00	0,00	0,00	19.939,61	3.237.996,55	19.939,61
CT-ESPACIAL	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CT-HIDRO	0,00	0,00	0,00	0,00	25.510,48	703.840,16	25.510,48
CT-INFORMÁTICA	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CT-INFRA	27.332.454,00	25.743.460,60	25.291.092,69	22.028.883,92	54.066.965,70	35.139.150,92	76.095.849,62
CT-MINERAL	0,00	0,00	0,00	0,00	0,00	9.700,00	0,00
CT-PETRO	81.294,00	81.293,75	58.339,68	58.339,68	0,00	4.385.171,22	58.339,68
CT-SAÚDE	33.598,00	13.597,74	13.597,74	13.597,74	4.526.511,89	1.810.165,08	4.540.109,63
CT-TRANSPORTE	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CT-VERDE-AMARELO	8.078.452,00	578.451,03	462.568,98	462.568,98	1.038.857,21	2.996.104,48	1.501.426,19
<b>Total</b>	<b>35.803.504,00</b>	<b>26.694.506,57</b>	<b>26.101.102,10</b>	<b>22.838.893,33</b>	<b>61.482.148,71</b>	<b>51.183.682,04</b>	<b>84.321.042,04</b>

## 2.5. Visualizar Painel Acompanhamento de Termos de Referência (TR)

ID	NOME
HST005	Visualizar Painel Acompanhamento de Termos de Referência (TR)
<p><b>COMO</b> Gestor  <b>QUERO</b> ter acesso ao painel Acompanhamento de Termos de Referência  <b>PARA</b> que eu possa visualizar informações referentes aos termos de referência por instrumentos de implementação, por agências e por ações programáticas</p>	

### 2.5.1. Critérios de Aceitação

<b>CA063.</b>	O painel Acompanhamento de Termos de Referência (TR) deverá apresentar 4 (quatro) gráficos, são eles: <ul style="list-style-type: none"> <li>▪ Quantidade de TR e Valor Aprovado por Ano;</li> <li>▪ Qdt. de TR e Valor por Instrumento de Implementação;</li> <li>▪ Qdt de TR e Valor Aprovado por Agência; e</li> <li>▪ Quantidade de TRs e Valor Pago por Ação Programática.</li> </ul>
<b>CA064.</b>	Deverá permitir filtrar os gráficos por: <ul style="list-style-type: none"> <li>▪ Ano;</li> <li>▪ Instrumento; e</li> <li>▪ Agência.</li> </ul>
<b>CA065.</b>	No totalizador <b>Total de TR</b> , deverá exibir o resultado do somatório dos TRs.
<b>CA066.</b>	No totalizador <b>Valor aprovado</b> , deverá exibir o resultado do somatório dos valores aprovados.
<b>CA067.</b>	Para o gráfico <b>Quantidade de TR e Valor Aprovado por Ano</b> , deverá apresentar um gráfico de colunas com os valores aprovados e a quantidade de TRs para cada ano base. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores aprovados e a quantidade de TR; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna da Quantidade TR: Ano e Quantidade TR; e</li> <li>- Quando for coluna Valor aprovado: Ano e Valor aprovado.</li> </ul> </li> </ul>
<b>CA068.</b>	Para o gráfico <b>Qdt de TR e Valor por Instrumento de Implementação</b> , deverá apresentar um gráfico de colunas com os valores aprovados e a quantidade de TRs para cada instrumento de implementação. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores aprovados e a quantidade de TR; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna da Quantidade TR: Instrumento e Quantidade TR; e</li> <li>- Quando for coluna Valor aprovado: Instrumento e Valor</li> </ul> </li> </ul>

	aprovado.
<b>CA069.</b>	<p>Para o gráfico <b>Qdt de TR e Valor Aprovado por Agência</b>, deverá apresentar um gráfico de colunas com os valores aprovados e a quantidade de TRs para cada agência. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores aprovados e a quantidade de TR; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna da Quantidade TR: Agência e Quantidade TR; e</li> <li>- Quando for coluna Valor aprovado: Agência e Valor aprovado.</li> </ul> </li> </ul>
<b>CA070.</b>	<p>Para o gráfico <b>Quantidade de TRs e Valor Pago por Ação Programática</b>, deverá apresentar um gráfico de colunas com os valores aprovados, quantidade de TRs e os valores empenhados para cada ação programática. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores dos valores aprovados, a quantidade de TR e os valores empenhados; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna da Quantidade TR: Fundo e Quantidade TR;</li> <li>- Quando for coluna Valor aprovado: Fundo e Valor aprovado; e</li> <li>- Quando for coluna Valor empenhado: Fundo e Valor empenhado.</li> </ul> </li> </ul>

## 2.5.2. Protótipo

### 2.5.2.1. Tela Painel Acompanhamento de Termos de Referência (TR)

**FNDCT**  
Fundo Nacional de Desenvolvimento Científico e Tecnológico

ARRECAÇÃO
EXECUÇÃO ORÇAMENTÁRIA
ACOMPANHAMENTO DE TR
PANORAMA DE EXECUÇÃO

### ACOMPANHAMENTO DE TERMOS DE REFERÊNCIA (TR)

**Filtros**

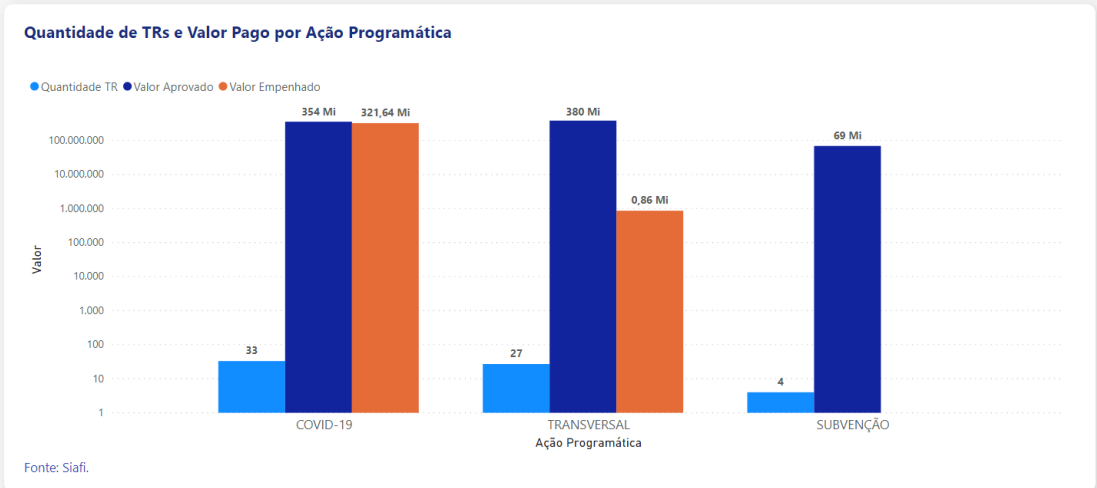
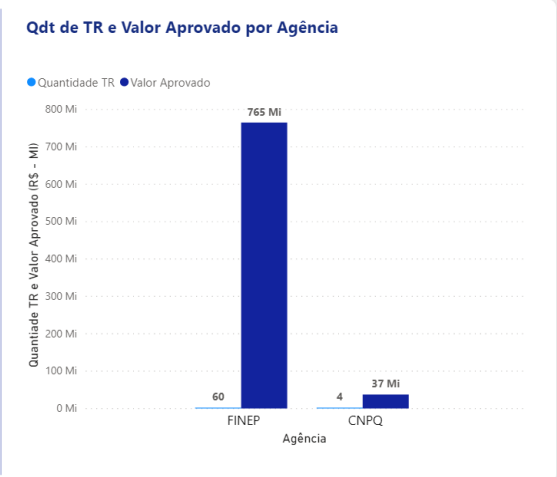
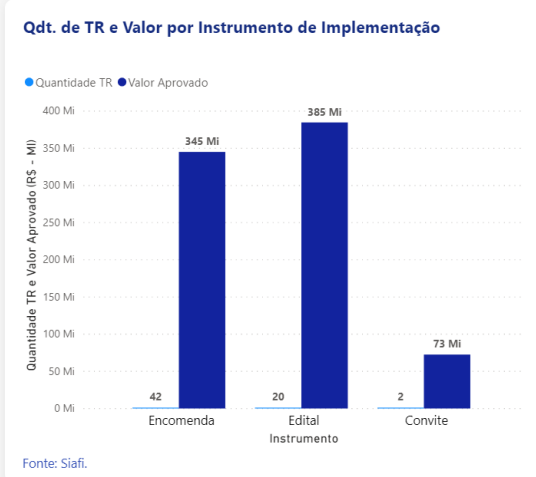
**Ano** 2020

**Instrumento** Todos

**Agência** Todos

**64**  
Total de TR

**802 Mi**  
Valor Aprovado



## 2.6. Visualizar Painel Panorama de Execução – Visão por Projeto

ID	NOME
HST006	Visualizar Painel Panorama de Execução – Visão por Projeto
<p><b>COMO</b> Gestor  <b>QUERO</b> ter acesso ao painel Panorama de Execução Visão por Projeto  <b>PARA</b> que eu possa visualizar informações referentes a quantidade de projetos por valor contratado</p>	

### 2.6.1. Critérios de Aceitação

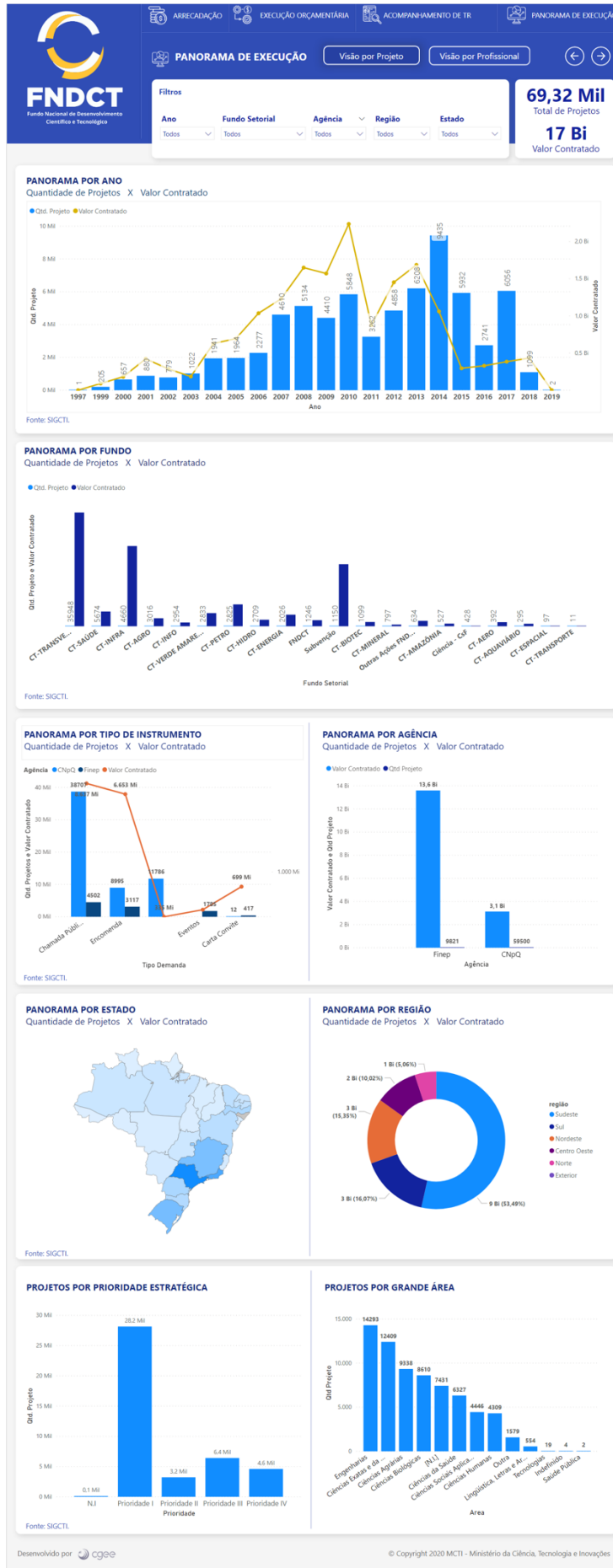
<b>CA071.</b>	O painel Panorama de Execução Visão por Projeto deverá apresentar 8 (oito) gráficos, são eles: <ul style="list-style-type: none"> <li>▪ Panorama por Ano;</li> <li>▪ Panorama por Fundo;</li> <li>▪ Panorama por Tipo de Instrumento;</li> <li>▪ Panorama por Agência;</li> <li>▪ Panorama por Estado;</li> <li>▪ Panorama por Região;</li> <li>▪ Projetos por Prioridade Estratégica; e</li> <li>▪ Projetos por Grande Área.</li> </ul>
<b>CA072.</b>	Deverá permitir filtrar os gráficos por: <ul style="list-style-type: none"> <li>▪ Ano;</li> <li>▪ Fundo Setorial;</li> <li>▪ Agência;</li> <li>▪ Região; e</li> <li>▪ Estado.</li> </ul>
<b>CA073.</b>	No totalizador <b>Total de Projetos</b> , deverá exibir o resultado do somatório de todos os projetos.
<b>CA074.</b>	No totalizador <b>Valor Contratado</b> , deverá exibir o resultado do somatório dos valores contratados.
<b>CA075.</b>	Para o gráfico <b>Panorama por Ano</b> , deverá apresentar um gráfico de colunas e linha com a quantidade de projetos e valores contratados em cada ano base. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar a cor da quantidade de projetos e a cor do valor contratado; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna: Ano e Quantidade de Projetos; e</li> <li>- Quando for linha: Ano e Valor contratado.</li> </ul> </li> </ul>
<b>CA076.</b>	Para o gráfico <b>Panorama por Fundo</b> , deverá apresentar um gráfico de colunas com a quantidade de projetos e valores contratados em cada por cada fundo setorial. Também deverá exibir: <ul style="list-style-type: none"> <li>▪ Uma legenda para representar a cor da quantidade de projetos e a cor do valor contratado; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações:</li> </ul>

	<ul style="list-style-type: none"> <li>- Quando for coluna Qtd Projeto: Fundo setorial e Quantidade projetos; e</li> <li>- Quando for coluna Valor Contratado: Fundo setorial e Valor contratado.</li> </ul>
<b>CA077.</b>	<p>Para o gráfico <b>Panorama por Tipo de Instrumento</b>, deverá apresentar um gráfico de colunas e linha com a quantidade de projetos e valores contratados em cada tipo de demanda. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar as cores de cada agência e a cor do valor contratado; e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: <ul style="list-style-type: none"> <li>- Quando for coluna: Tipo demanda, Agência, Quantidade projeto e Valor contratado; e</li> <li>- Quando for linha: Ano e Valor contratado.</li> </ul> </li> </ul>
<b>CA078.</b>	<p>Para o gráfico <b>Panorama por Agência</b>, deverá apresentar um gráfico de colunas com a quantidade de projetos e valores contratados em cada agência. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar a cor da quantidade de projetos e a cor do valor contratado; e</li> <li>▪ Um hint em cada coluna do gráfico com as seguintes informações: Agência e Valor contratado.</li> </ul>
<b>CA079.</b>	<p>Para a <b>Panorama por Estado</b>, deverá apresentar um mapa do Brasil com a representação da quantidade de projetos e valor contratado em cada estado do país. Também deverá exibir um hint em cada estado com as seguintes informações: Cidade, Quantidade Projeto e Valor contratado.</p>
<b>CA080.</b>	<p>Para o gráfico <b>Panorama por Região</b>, deverá apresentar um gráfico de rosca com a quantidade de projetos e valor contratado em cada região. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar a cor de cada região (Sudeste, Sul, Nordeste, Centro Oeste, Norte e Exterior); e</li> </ul> <p>Um hint em cada ponto do gráfico com as seguintes informações: Região, Valor Contratado e Qtd Projeto.</p>
<b>CA081.</b>	<p>Para o gráfico <b>Projetos por Prioridade Estratégica</b>, deverá apresentar um gráfico de colunas com a quantidade de projetos por prioridade. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Prioridade, Quantidade projeto.</p>
<b>CA082.</b>	<p>Para o gráfico <b>Projetos por Grande Área</b>, deverá apresentar um gráfico de colunas com a quantidade de projetos por área. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Área e Quantidade projeto.</p>

## 2.6.2. Protótipo

### 2.6.2.1. Tela Painel Panorama de Execução – Visão por Projeto







## 2.7. Visualizar Painel Panorama de Execução – Visão por Profissional

ID	NOME
HST007	Visualizar Painel Panorama de Execução – Visão Profissional
<p><b>COMO</b> Gestor  <b>QUERO</b> ter acesso ao painel Panorama de Execução visão Profissional  <b>PARA</b> que eu possa visualizar informações referentes aos profissionais</p>	

### 2.7.1. Critérios de Aceitação

<b>CA083.</b>	<p>O painel Panorama de Execução Visão por Profissional deverá trazer os dados sobre a equipe envolvida nos projetos. Os profissionais estão classificados em estudantes, Apoio Técnico, Apoio Administrativo, Bolsistas (em nível médio), Pesquisadores e coordenadores.</p> <p>O painel deverá apresentar 5 (cinco) gráficos, são eles:</p> <ul style="list-style-type: none"> <li>▪ Panorama por Ano;</li> <li>▪ Panorama por Titulação;</li> <li>▪ Panorama por Grande Área;</li> <li>▪ Quantidade de Profissionais por Estado; e</li> <li>▪ Quantidade de Profissionais por Região.</li> </ul>
<b>CA084.</b>	<p>Deverá permitir filtrar os gráficos por:</p> <ul style="list-style-type: none"> <li>▪ Ano;</li> <li>▪ UF;</li> <li>▪ Titulação;</li> <li>▪ Fundo setorial;</li> <li>▪ Região; e</li> <li>▪ Área.</li> </ul>
<b>CA085.</b>	<p>No totalizador <b>Total de Profissionais</b>, deverá exibir o resultado do somatório de todos os profissionais.</p>
<b>CA086.</b>	<p>Para o gráfico <b>Panorama por Ano</b>, deverá apresentar um gráfico de colunas com a quantidade de profissionais envolvidos nos projetos em cada ano base. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Ano e Quantidade de profissionais.</p>
<b>CA087.</b>	<p>Para o gráfico <b>Panorama por Titulação</b>, deverá apresentar um gráfico de colunas com a quantidade de profissionais envolvidos nos projetos em cada titulação. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Titulação e Quantidade de profissionais.</p>
<b>CA088.</b>	<p>Para o gráfico <b>Panorama por Grande Área</b>, deverá apresentar um gráfico de colunas com a quantidade de profissionais envolvidos nos projetos em cada área. Também deverá exibir um hint em cada coluna do gráfico com as seguintes informações: Área e Quantidade de profissionais.</p>
<b>CA089.</b>	<p>Para a <b>Quantidade de Profissionais por Estado</b>, deverá apresentar</p>

	um mapa do Brasil com a representação da quantidade de profissionais em cada estado do país. Também deverá exibir um hint em cada estado com as seguintes informações: Estado e Quantidade Profissionais.
<b>CA090.</b>	<p>Para o gráfico <b>Quantidade de Profissionais por Região</b>, deverá apresentar um gráfico de rosca com a quantidade de profissionais em cada região. Também deverá exibir:</p> <ul style="list-style-type: none"> <li>▪ Uma legenda para representar a cor de cada região (Sudeste, Sul, Nordeste, Centro Oeste e Norte); e</li> <li>▪ Um hint em cada ponto do gráfico com as seguintes informações: Região e Quantidade de profissionais.</li> </ul>

## 2.7.2. Protótipo

### 2.7.2.1. Tela Painel Panorama de Execução – Visão por Profissional

**FNDCT**  
Fundo Nacional de Desenvolvimento Científico e Tecnológico

ARRECAÇÃO
EXECUÇÃO ORÇAMENTÁRIA
ACOMPANHAMENTO DE TR
PANORAMA DE EXECUÇÃO

Visão por Projeto
Visão por Profissional

**Filtros**

Ano	UF	Titulação
Todos	Todos	Todos
Fundo	Região	Área
Todos	Todos	Todos

109,07 Mil

Total de Profissionais

**PANORAMA POR ANO**  
Pesquisadores Envolvidos em Projetos

Ano	Qtd. Profissional
1999	136
2000	685
2001	1704
2002	5415
2003	5739
2004	13019
2005	13530
2006	19042
2007	11950
2008	16343
2009	15254
2010	20519
2011	7416
2012	10454
2013	8789
2014	3442
2015	19

Fonte: SIGCTL

**PANORAMA POR TITULAÇÃO**  
Pesquisadores Envolvidos em Projetos

Titulação	Qtd. Profissional
Doutorado	27 Mil
Graduação	23 Mil
Mestrado	17 Mil
Pós-Doutorado	15 Mil
(Em branco)	13 Mil
Especialização	6 Mil
N.I.	3 Mil
Livre Docência	2 Mil
Ensino Médio	1 Mil
Curso Técnico Professi...	1 Mil
Mestrado Profissional...	1 Mil
Aperfeiçoamento	0 Mil
Residência Médica	0 Mil
Ensino Fundamental	0 Mil

Fonte: SIGCTL

**PANORAMA POR GRANDE ÁREA**  
Pesquisadores Envolvidos em Projetos

Área	Qtd. Profissional
Ciências exatas e da terra	18,2 Mil
Engenharias	17,8 Mil
N.I.	17,4 Mil
(Em branco)	13,3 Mil
Ciências biológicas	11,7 Mil
Ciências agrárias	10,1 Mil
Ciências da saúde	6,4 Mil
Ciências sociais aplicadas	6,2 Mil
Ciências Humanas	4,5 Mil
Outros	2,5 Mil
Linguística letras e artes	1,0 Mil

Fonte: SIGCTL

**QUANTIDADE DE PROFISSIONAIS POR ESTADO**

Fonte: SIGCTL

**QUANTIDADE DE PROFISSIONAIS POR REGIÃO**

região	Quantidade	Porcentagem
Sudeste	32,33 Mil	43,97%
Nordeste	15,04 Mil	20,45%
Sul	6,75 Mil	9,18%
Centro Oeste	4,63 Mil	6,29%
Norte	14,00 Mil	19,11%

Desenvolvido por cgée

© Copyright 2020 MCTI - Ministério da Ciência, Tecnologia e Inovações

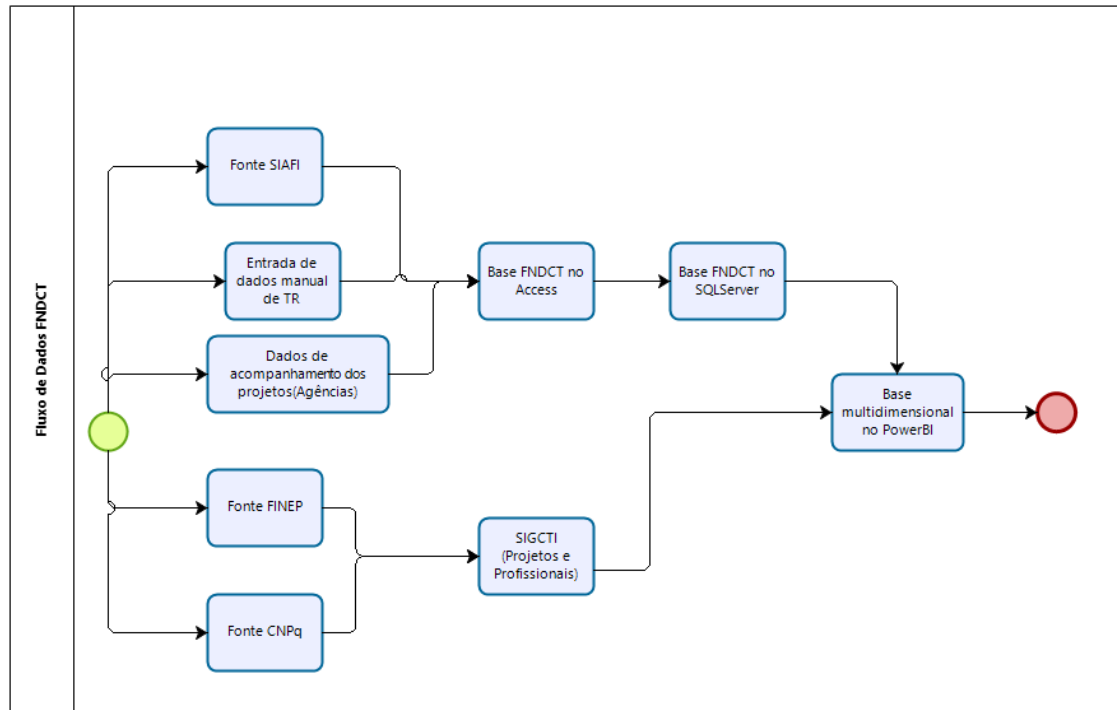
Observações:

- Diferença de “Profissional” e “Pesquisador”:

- Profissional: Membro da equipe de um projeto
- Pesquisador: Membro especialista da equipe de um projeto

### 3. DOCUMENTAÇÃO DATA PIPELINE

Visão Geral Fluxo de dados FNDCT



Powered by  
**bizagi**  
Modeler

### 4. DICIONÁRIO DE VARIÁVEIS

Variável	Fato	Dimensão	Métricas	Perspectivas	Fonte de dados
<b>Valor arrecadado</b>	Execução orçamentária FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi
<b>Valor autorizado LOA</b>	Execução orçamentária FNDCT	Execução orçamentária	Valor total por fundo setorial Total por grupo de ação programática Total por instrumento	Tempo (Ano) Fundo setorial Ação programática Grupo de ação programática	Siafi
<b>Valor empenhado</b>	Execução orçamentária FNDCT	Execução orçamentária	Valor total por fundo setorial Total por grupo de ação programática Total por instrumento	Tempo (Ano) Fundo setorial Ação programática Grupo de ação programática	Siafi
<b>Despesa paga</b>	Execução financeira FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi
<b>Despesa executada</b>	Execução financeira FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi

<b>Valor pago</b>	Execução financeira FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi
<b>Restos a pagar pago</b>	Execução financeira FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi
<b>Restos a pagar a pagar</b>	Execução financeira FNDCT	Execução orçamentária	Valor total por fundo setorial	Tempo (Ano) Fundo setorial	Siafi

## 5. REFERÊNCIAS

- Documento: Manual-Operativo-FNDCT.pdf
- Documento: FNDCT\_Identidade\_Visual.pdf
- Documento: Tema FNDCT.xmind
- Documento: Modelo\_De\_Dados\_Painel\_Gerencial.pdf

## Anexo V – Tema Estratégico Indicadores da COICT

### Detalhamento da demanda:

ArqDigMCTI - tema estratégico INDICADORES COIND

Escopo de trabalho: trazer tabelas dos indicadores COIND para HDFS em conformidade com os padrões de curadoria.

Objetivo geral: percorrer todos os conjuntos de indicadores CT&I da página web

[http://www.mctic.gov.br/mctic/opencms/indicadores/indicadores\\_cti.html](http://www.mctic.gov.br/mctic/opencms/indicadores/indicadores_cti.html)  
e para cada tabela extrair metadados e conteúdo conforme exemplo abaixo.

Exemplo:

Recursos Aplicados

Indicadores Consolidados

2.1.1 Brasil: Dispêndio nacional em ciência e tecnologia (C&T), em valores correntes, por atividade, 2000-2017

Exportação CSV (pode ser utilizado a exportação XLSX também):

Tabela 2.1.1;;;;;;;;;

;;;;;;;;;

Brasil: Dispêndio nacional em ciência e tecnologia (C&T)(1) por atividade, 2000-2017;;;;;;;;;;

;;;;;;;;;[unintelligible] (em[milh]es de R\$ correntes);;

Ano ;Ciência e Tecnologia (C&T)(1);;;;;;;;;

;Total;Pesquisa e Desenvolvimento (P&D);;"Atividades Científicas e  
Técnicas Correlatas (ACTC)";;

;;Total;Orçamento executado;Ensino superior(2);Total;Orçamento executado;Ensino  
superior(2);;

2000;15.839,1;12.560,7;9.349,3;3.211,4;3.278,4;3.278,4;-;;

2001;17.655,6;13.973,0;10.444,4;3.528,6;3.682,6;3.682,6;-;;

2002;19.756,7;15.031,9;10.957,4;4.074,6;4.724,8;4.724,8;-;;

2003;22.278,8;17.169,0;12.590,3;4.578,7;5.109,8;5.109,8;-;;

2004;25.437,7;18.861,6;14.109,4;4.752,2;6.576,1;6.576,1;-;;

2005;28.179,8;21.759,3;16.764,3;4.995,0;6.420,5;6.420,5;-;;

2006;30.540,9;23.807,0;18.018,3;5.788,7;6.733,9;6.733,9;-;;

2007;37.468,2;29.416,4;21.331,0;8.085,4;8.051,8;8.051,8;-;;

2008;45.420,6;35.110,8;25.730,8;9.380,0;10.309,8;10.309,8;-;;

2009;51.398,4;37.285,3;27.713,1;9.572,2;14.113,1;14.113,1;-;;

2010;62.223,4;45.072,9;33.662,6;11.410,2;17.150,5;17.150,5;-;;

2011;68.155,0;49.875,9;35.981,5;13.894,3;18.279,2;18.279,2;-;;

2012;76.432,7;54.254,6;38.547,6;15.707,0;22.178,1;22.178,1;-;;

2013;85.646,4;63.748,6;45.149,0;18.599,6;21.897,8;21.897,8;-;;

2014;96.755,0;73.468,2;51.697,5;21.770,7;23.286,8;23.286,8;-;;  
 2015;102.042,9;80.501,8;58.108,3;22.393,5;21.541,1;21.541,1;-;;  
 2016;95.602,1;79.228,3;53.937,6;25.290,6;16.373,8;16.373,8;-;;  
 2017(3);99.849,9;82.792,2;56.619,4;26.172,8;17.057,7;17.057,7;-;;

Fonte: Coordenação de Indicadores e Informação (COIND) - CGGI/DGE/SEXEC - Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC);

Elaboração: Coordenação de Indicadores e Informação (COIND) - CGGI/DGE/SEXEC - Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC);

"Notas: (1) ciência e tecnologia (C&T) = pesquisa e desenvolvimento (P&D) + atividades científicas e técnicas correlatas (ACTC);"

2) considerados os gastos da pós-graduação como proxy dos dispêndios em P&D das instituições de ensino superior (IES).;

3) dados preliminares;

Objetivo imediato para cada tabela:

A partir do CSV ou XLSX da tabela no site web, gerar:

- 1) "tab\_2\_1\_1\_E" => arquivo baixado
  - 2) "Tabela 2.1.1-meta.txt" => metadados => nome da tabela, título, cabeçalho das colunas, notas de rodapé
  - 3) "Tabela 2.1.1-extracao.csv" => conteúdo => cabeçalho (), conteúdo (utf-8, separador ";", QUOTE\_MINIMAL)
- ( ) Cabeçalho => Ano; C&T\_Total ; P&D\_Total ; C&T\_P&D\_Orçamento\_executado ; C&T\_P&D\_Ensino\_superior ; C&T\_ACTC\_Total ; C&T\_ACTC\_Orçamento\_executado ; C&T\_ACTC\_Ensino\_supeior

Objetivo final:

Estrutura de pastas mantendo relação com o site, e na pasta folha, os três arquivos acima.

Opções de formato final:

- CSV (por arquivo),
- CSV (completo),
- R Studio RDS,
- SQL Server (uma tabela por arquivo),
- SQL Server (uma tabela para todos os arquivos)

Quantidade tabelas na página Web	Quantidade de arquivos extraídos (CSV)	Quantidade de tabelas resultantes (RData)
146 (*)	143	143

(\*) Algumas tabelas na página Web apontavam para mais de um arquivo fonte. Exemplo: "patentes\_tab\_6\_1\_7\_e" e "patentes\_tab\_6\_1\_7\_e\_nr".

### Documentação de variáveis - Formato de saída (Json):

{

```

"nome": "Nome da tabela (conjunto de dados).",
"titulo": "Título da tabela.",
"subtitulo": "Subtítulo da tabela, se houver.",
"descricao": "Detalhe complementar contido na tabela.\n",
"fontededados": "Fontes de dados.\n",
"notasrodape": "(1) Uma nota de rodapé \n
                (2) Segunda nota \n
                (3) Notas de rodapé podem ser das variáveis ou dos metadados\n"
],
"variaveis": [
  [
    { "nome variável 1": "descrição variável (nome da coluna tratado)" },
    { "nome variável 2": "descrição variável (nome da coluna tratado)" },
    { "nome variável N": "descrição variável (nome da coluna tratado)" }
  ]
]
}

```

## Extração e carga:

Nome	Tipo	Tamanho
▼ bolsas_formacao	Pasta	--
tabela-41-extracao.csv	Documento CSV	637 bytes
tabela-42-extracao.csv	Documento CSV	846 bytes
tabela-43-extracao.csv	Documento CSV	584 bytes
tabela-44-extracao.csv	Documento CSV	729 bytes
tabela-45-extracao.csv	Documento CSV	5 KB
tabela-46-extracao.csv	Documento CSV	4 KB
tabela-47-extracao.csv	Documento CSV	5 KB
tabela-48-extracao.csv	Documento CSV	802 bytes
tabela-49-extracao.csv	Documento CSV	474 bytes
tabela-410-extracao.csv	Documento CSV	951 bytes
tabela-411-extracao.csv	Documento CSV	968 bytes
tabela-412-extracao.csv	Documento CSV	6 KB
tabela-413-extracao.csv	Documento CSV	7 KB
tabela-414-extracao.csv	Documento CSV	756 bytes
bolsas_formacao.json	JSON Document	43 KB
> comparacoes_internacionais	Pasta	--
comparacoes_internacionais.json	JSON Document	106 KB
> dados_socio_economicos	Pasta	--
dados_socio_economicos.json	JSON Document	23 KB
> inovacao	Pasta	--
inovacao.json	JSON Document	22 KB
> patentes	Pasta	--
patentes.json	JSON Document	109 KB
> producao_cientifica	Pasta	--
producao_cientifica.json	JSON Document	30 KB
> recursos_aplicados	Pasta	--
recursos_aplicados.json	JSON Document	174 KB
> recursos_humanos	Pasta	--
recursos_humanos.json	JSON Document	181 KB

## Exemplo de Arquivo JSON



```

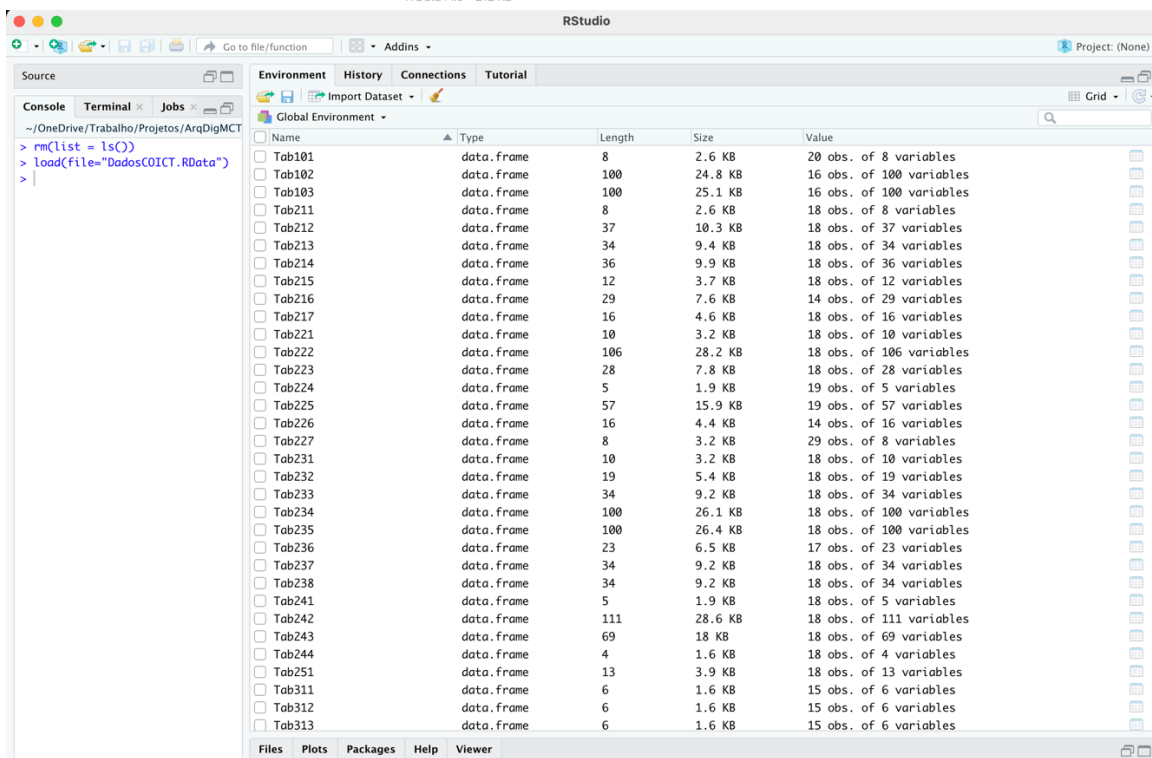
1025     {
1026       "nome": "tabela-4.8-extracao.csv",
1027       "titulo": "Brasil: CNPq - Bolsas(1) de Iniciação à Pesquisa concedidas no país por modalidades, 1990-2017",
1028       "subtitulo": "",
1029       "descricao": "Elaboração: Coordenação de Indicadores e Informação (COIND) - CGGI/DGE/SEXEC - Ministério da Ciência, Tecnologia, Inovações e Comunica
1030       "fontededados": "Fonte: http://fomentonacional.cnpq.br/dmfomento/home/fmthome.jsp? - acessado em 24/08/2018 às 10:32 h.\n",
1031       "notasrodape": "Nota: 1) bolsas-ano: representa a média aritmética do número de mensalidades pagas de janeiro a dezembro (12 meses).\n",
1032       "variaveis": [
1033         {
1034           "ano_4_8": "Ano"
1035         },
1036         {
1037           "iniciacao_cientifica_ic": "Iniciação Científica - IC"
1038         },
1039         {
1040           "icjr_junior": "Iniciação Científica Júnior - ICJR"
1041         },
1042         {
1043           "iniciacao_tecnologica": "Iniciação Tecnológica"
1044         },
1045         {
1046           "total_4_8": "Total"
1047         }
1048       ]
1049     },
1050     {
1051       "nome": "tabela-4.9-extracao.csv",
1052       "titulo": "Brasil: Capes - Total de bolsas de pós-graduação concedidas no país e total de bolsistas de pós-graduação no exterior, 1995-2019",
1053       "subtitulo": "",
1054       "descricao": "Elaboração: Coordenação-Geral de Gestão Institucional (CGAI) - DGI/SEXEC - Ministério da Ciência, Tecnologia e Inovações (MCTI)\n",
1055       "fontededados": "Fonte: Capes: http://geocapes.capes.gov.br/geocapesds ; extraído em 15/07/2020\n",
1056       "notasrodape": "",
1057       "variaveis": [
1058         {
1059           "ano_4_9": "Ano"
1060         },
1061         {
1062           "bolsas_pais": "Bolsas no país"
1063         },
1064         {
1065           "bolsista_exterior": "Bolsista no exterior"
1066         }
1067       ]
1068     }
1069 ]

```

**Formatos finais de entrega:**



DadosCOICT.RData  
R Data File - 212 KB



The screenshot shows the RStudio environment with the following data frames loaded:

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
Tab211	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

**Total de variáveis trabalhadas: 5083**