



PROPOSTA

**ENCOMENDA
TECNOLÓGICA DE
INSTRUÇÃO
ASSISTIDA POR
INTELIGÊNCIA
ARTIFICIAL**

Brasília, março de 2022

id: 2022PROP0212.TCU.ETEC

Sumário

| | |
|--|-----------|
| 1 Apresentação: Consórcio Modal - Intellimetri | 4 |
| 1.1 Instituto Modal de Ciência, Tecnologia e Inovação | 4 |
| 1.1.1 Missão | 4 |
| 1.1.2 Visão | 5 |
| 1.1.3 Valores | 5 |
| 1.2 Intellimetri Gestão de Dados | 5 |
| 1.2.1 Missão | 6 |
| 1.2.2 Visão | 6 |
| 1.2.3 Valores | 6 |
| 2 Introdução | 7 |
| 3 Objeto e Escopo Declarados | 8 |
| 4 Etapas de Pesquisa e de Desenvolvimento | 8 |
| 4.1 Metodologia da Pesquisa Científica | 9 |
| 4.2 Metodologia de Gestão e Desenvolvimento de Software | 10 |
| 4.2.1 Detalhamento da Trilha Desenvolvimento - Scrum | 10 |
| 4.3 Metodologia de Gerenciamento de Projetos | 12 |
| 4.3.1 Scrum e Kanban | 12 |
| 4.4 Planejamento Preliminar do Primeiro Ciclo - Aquisições Públicas | 16 |
| 4.5 Métricas | 27 |
| 5 Descrição da Solução | 30 |
| 5.1 Explicação dos Modelos de Inferência e Infraestrutura | 33 |
| 5.1.1 Marco 1 do Objeto | 33 |
| 5.1.2 Marco 2 do Objeto | 38 |
| 5.1.3 Marco 3 do Objeto | 38 |
| 5.2 Infraestrutura de Desenvolvimento | 39 |
| 5.3 Infraestrutura de Produção | 42 |
| 5.4 Infraestrutura de Segurança | 43 |
| 5.4.1 Camada de Segurança de Dados de IA Baseada em Blockchain para Mitigar “Algorithm and Data Poisoning” | 45 |

| | |
|---|-----------|
| 6 Cronograma Físico-Financeiro | 46 |
| 7 Planilha de Formação de Preços | 50 |
| 8 Considerações Finais | 50 |
| Anexo 1: Portfólio de Projetos | 52 |
| A1.1 Instituto Modal | 52 |
| A1.1.1 Experiência em Inteligência Artificial | 52 |
| A1.1.2 Portfólio: Casos e Provas de Conceito | 53 |
| A1.2 Intellimetri | 54 |
| A1.2.1 Experiência em Inteligência Artificial: ML, NLP e CV | 54 |
| A1.2.2 Portfólio: Casos e Provas de Conceito | 54 |
| Anexo 2: Qualificação dos Principais Profissionais Envolvidos no Projeto | 63 |
| Anexo 3: Atestados de Capacidade Técnica - Intellimetri | 70 |
| A3.1 Registro no Instituto Nacional de Propriedade Industrial (INPI) | 71 |
| A3.2 Diretório de parceiros Google Cloud | 72 |
| A3.3 Atestado de Capacidade Técnica - Tergos S.A | 73 |
| A3.4 Atestado de Capacidade Técnica Marco Marketing | 74 |
| A3.5 Atestado de Capacidade Técnica ISA7 | 75 |
| A3.6 Atestado de Capacidade Técnica - Paschoalotto | 76 |
| Anexo 4: Currículos dos Principais Profissionais | 77 |

1 Apresentação: Consórcio Modal - Intellimetri

O Consórcio Modal - Intellimetri foi instituído com o fim específico de participar da Encomenda Tecnológica para o desenvolvimento de uma solução que contribua com a evolução da atual ferramenta de Instrução Assistida do Tribunal de Contas da União (TCU).

Integram o consórcio o [Instituto Modal de Ciência, Tecnologia e Inovação](#), associação privada sem fins lucrativos com sede no Distrito Federal, e a [Intellimetri Gestão de Dados S.A.](#), empresa privada sediada em São Paulo. As duas organizações firmaram um termo de compromisso e, em conjunto, contam com profissionais, pesquisadores e experiência comprovada por atestados de capacidade técnica e por projetos desenvolvidos utilizando, dentre outras técnicas e ferramentas: Processamento de Linguagem Natural; *Machine Learning*; Inteligência Artificial (simbólica e subsimbólica); classificação de documentos; identificação de imagens; e redação automática de textos.

O Consórcio é *Google Cloud Partner* e conta com profissionais certificados no uso e aplicação de tecnologias baseadas em nuvem, principalmente em soluções de processamento de linguagem natural, *speech analytics* e OCR (*Document AI*).

1.1 Instituto Modal de Ciência, Tecnologia e Inovação

O Instituto Modal é uma Instituição de Ciência, Tecnologia e Inovação (ICT), sem fins lucrativos, que nasceu do encontro entre a academia e o mercado. Desenvolve pesquisa aplicada, soluções com Inteligência Artificial e projetos de novos negócios visando criar oportunidades e equidade social para o desenvolvimento do Brasil.

A base legal que fundamenta a atuação do Modal é a [Lei nº 13.243, de 11 de janeiro de 2016](#) (Marco de Ciência, Tecnologia e Inovação), regulamentada pelo [Decreto nº 9.283, de 7 de fevereiro de 2018](#).

Em 23 de junho de 2021, o Instituto Modal foi qualificado pelo Ministério da Justiça como Organização da Sociedade Civil de Interesse Público (OSCIP), com publicação no [Diário Oficial da União nº 118, de 25 de junho de 2021](#).

1.1.1 Missão

O Instituto Modal tem como foco principal é a pesquisa e a construção de modelos, estudos e aplicações baseadas em diferentes técnicas e ferramentas, especialmente

aquelas baseadas em Inteligência Artificial, através da execução de projetos voltados à promoção de novas soluções, de novos negócios, de oportunidades e de equidade social nos setores público e privado, e para a população em geral.

1.1.2 Visão

- Tornar-se referência em soluções inovadoras de Inteligência Artificial, com foco prioritário na área industrial, no prazo de 5 anos.
- A inovação é entendida como a novidade útil, com sólida fundamentação científica, que traz melhorias contínuas para o setor, beneficiando negócios, gerando emprego e transformando o mercado e a sociedade.

1.1.3 Valores

- Ética e transparência - Seguir todas as boas práticas de governança da informação.
- Inovação com foco em resultados - A inovação por si própria serve somente para alimentar o ego – o que interessa para o mercado é o resultado da ação inovadora.
- Melhoria contínua - Esforço contínuo para aprender – afinal, o Instituto Modal é uma instituição de pesquisa.

O Instituto Modal possui um forte [Programa de Integridade](#), que abrange políticas Anticorrupção, de Privacidade, de Tratamento de Dados Pessoais, Código de Ética e Ética Profissional, dentre outros documentos de *compliance* e governança corporativa.

Informações adicionais sobre o Instituto Modal estão disponíveis no [Anexo 1 - A1.1 Instituto Modal](#).

1.2 Intellimetri Gestão de Dados

Com sede em São Paulo e escritórios em Bogotá (Colômbia) e Cidade do México (México), a Intellimetri Gestão de Dados S/A é uma empresa 100% nacional, que há quatro anos desenvolve e aplica soluções baseadas em Inteligência Artificial e *Machine Learning* como forma de contribuir para a transformação digital de empresas, organizações e governos.

A Intellimetri trabalha com equipes multidisciplinares em todos os seus projetos, com o objetivo de resolver as dores de seus clientes e ajudá-los a tomar decisões baseadas em dados. Por meio da colaboração e da criação de valor em rede, a Intellimetri promove a

transferência de conhecimento e a maturidade de seus clientes no uso estratégico de dados em seus negócios.

Além de [parceira estratégica](#) para o desenvolvimento de soluções baseadas em Inteligência Artificial do Google Cloud no Brasil e na América Latina, a empresa também é parceira oficial de revenda e serviços Google Cloud.

1.2.1 Missão

Ajudar a promover a transformação digital nos setores público e privado por meio da aplicação de tecnologias cognitivas para a solução de problemas e para a geração de insights estratégicos que agreguem valor ao negócio.

1.2.2 Visão

- Ser referência no Brasil e na América Latina no desenvolvimento de soluções inovadoras baseadas em Inteligência Artificial nos próximos 5 anos;
- Desenvolver os melhores algoritmos para a construção de nuvens e bordas inteligentes, baseadas em inteligência artificial ("IA") em um mundo "AI First".
- A Inteligência Artificial é muito mais que pura tecnologia. Ela é, cada vez mais, parte de nossas vidas. Por isso, queremos contribuir para que seja onipresente.
- A Inteligência Artificial é o motor da Revolução Industrial 4.0. A tecnologia melhora a mobilidade, molda as cidades do futuro e faz as casas mais inteligentes, indústria conectada e cuidados com a saúde mais eficientes.

Em todas as esferas, a Intellimetri está trabalhando rumo a um mundo mais inteligente.

1.2.3 Valores

- Integridade, honestidade e ética.
- Paixão por tecnologia, parceiros e clientes.
- Abertura e respeito.
- Vontade de abraçar grandes desafios.
- Responsabilidade sócio-ambiental.
- Comprometimento com a melhoria e a evolução contínua da Inteligência Artificial.

Informações adicionais sobre a Intellimetri estão disponíveis no [Anexo 1 - A1.2 Intellimetri](#).

2 Introdução

As abordagens simbólica e a subsimbólica representam os dois principais ramos da Inteligência Artificial (IA). O paradigma dominante foi a IA simbólica até a década de 1980 - a partir de então, a IA subsimbólica começa a assumir a liderança e ganha atenção até os últimos anos. Há um longo e não resolvido debate entre as duas abordagens. No entanto, essa luta entre os diferentes domínios da IA está chegando ao fim, já que estamos vivendo uma terceira fase da IA, onde a onda dominante é a combinação da IA simbólica e da subsimbólica¹, que alguns autores chamam “métodos intermediários”.

De maneira resumida, as principais diferenças entre esses dois campos de IA são:

1. abordagens simbólicas produzem conclusões, enquanto as abordagens subsimbólicas fornecem resultados associativos;
2. A intervenção humana é comum nos métodos simbólicos, enquanto os subsimbólicos aprendem e se adaptam a partir dos dados fornecidos;
3. Os métodos simbólicos funcionam melhor quando lidam com objetos em domínios bem definidos e dados precisos, enquanto os subsimbólicos conseguem lidar com grandes e ruidosos conjuntos de dados.

Apesar das diferenças fundamentais entre o simbolismo e subsimbólico, nos últimos anos construíram-se ligações entre as duas abordagens com o uso de métodos intermediadores. Tais métodos consistem em esforços para preencher a lacuna entre os dois paradigmas. A ideia é criar um sistema que possa combinar as vantagens de ambos os métodos: a capacidade de aprender com o ambiente e a capacidade de raciocinar os resultados.

Nesta proposta de ETEC para o TCU, nossa abordagem consiste em usar métodos intermediários buscando resultados mais apropriados, tanto do ponto de vista de melhor capturar a semântica necessária à resolução do problema quanto à composição textual com mais sentido e conexões entre suas partes componentes. Com isso acredita-se que o uso de métodos intermediários tem grande potencial para melhores resultados, especialmente quando se considera que algumas das expectativas levantadas pelo Edital ainda se encontram em estágio de pesquisa científica.

Sob a perspectiva metodológica desta proposta, é necessário deixar claro que a palavra “metodologia” adquire um caráter polissêmico, visto que é utilizada em vários contextos. Por esse motivo, foram adotadas duas abordagens diferentes para o uso da

¹ Ilkou, E; Koutraki, M. Symbolic Vs Sub-symbolic AI Methods: Friends or Enemies? Proceedings of the CIKM 2020 Workshops, October 19-20, Galway, Ireland.

palavra "metodologia" neste documento:

1. "Metodologia da pesquisa científica", quando se refere aos aspectos deste trabalho envolvendo a pesquisa científica com risco tecnológico associado, como é a característica de uma ETEC; e
2. "Metodologia de gestão e desenvolvimento de software", para se referir ao conjunto de métodos e técnicas aplicados ao processo de desenvolvimento de software, como o *scrum*, padrão adotado nesta proposta.

3 Objeto e Escopo Declarados

Conforme o Edital e o Termo de Referência publicados pelo TCU, o objeto desta contratação é a Encomenda Tecnológica de um módulo de Instrução Assistida por Inteligência Artificial, a ser incorporado à solução de Instrução Assistida do TCU, conferindo-lhe as seguintes inovações:

1. Marco 1: Detecção de significado nas peças processuais. Inclui identificação das alegações (estruturação da petição inicial e suas alegações), exame de admissibilidade, cálculo da probabilidade de concessão de medidas cautelares.
2. Marco 2: Painel de jurimetria. Inclui priorização de processos e comparação com causas anteriores (precedentes).
3. Marco 3: Redação de peças. Inclui geração e interpretação de comunicações aos interessados e de instruções contendo sumarização de teses e predição da análise técnica de propostas de encaminhamento.

A encomenda terá como escopo os processos de Representações e Denúncias, que representam 40% do total de processos, e será executada em três ciclos, cada um relacionado a determinado conjunto temático:

1. Ciclo 1: Aquisições Públicas (compõem 20% das representações e denúncias).
2. Ciclo 2: Demais temas.
3. Ciclo 3: Futuras Aquisições Públicas sob a égide da nova Lei de Licitações (Lei 14.133/2021).

4 Etapas de Pesquisa e de Desenvolvimento

O atendimento a uma Encomenda Tecnológica que tem por objeto resultados com alto grau de inovação deve harmonizar a pesquisa acadêmica às melhores ferramentas e métodos disponíveis na indústria, como forma de atenuar o risco tecnológico e promover a inovação. Ao requerer mais do que "a mera aplicação de métodos e técnicas usuais do mercado de IA", o Edital (Item 4) incentiva a criação de soluções

inovadoras e otimizadas considerando o caso concreto.

Nesse sentido, propomos uma rota tecnológica² composta de duas trilhas metodológicas (“Trilha Pesquisa” e “Trilha Desenvolvimento”), independentes e complementares entre si, como forma de implementar a Proposta de Pesquisa & Desenvolvimento (P&D) em todo o seu percurso.

As próximas seções detalham essas trilhas, os modelos e os algoritmos que serão utilizados na rota tecnológica proposta para o projeto, mostrando a consistência e a viabilidade da pesquisa e sua aplicação na solução a ser desenvolvida.

4.1 Metodologia da Pesquisa Científica

Considerando as características cíclica e incremental da ETEC, adotaremos na “Trilha Pesquisa” a metodologia “Pesquisa-Ação”, na qual o pesquisador interage com o objeto de estudo, envolvendo-se no trabalho de pesquisa de forma participativa, buscando determinado resultado.

Uma das principais características da Pesquisa-Ação é a repetição das suas etapas em um processo cíclico que transita ora no domínio da pesquisa, ora no da ação. Existe “um ciclo de retroalimentação no qual os achados iniciais geram possibilidades para mudanças, as quais são então implementadas e avaliadas como um início para mais uma investigação”³. Além da característica do processo cíclico, a Pesquisa-Ação, segundo Thiollent⁴, possui ainda três características fundamentais:

- Prática – lida com problemas e assuntos do mundo real;
- Mudança – tanto como uma forma de lidar com os problemas práticos, como de descobrir mais sobre um fenômeno;
- Participativa – as pessoas envolvidas com o problema no mundo real são cruciais no processo de pesquisa, tendo uma participação ativa e não passiva.

Pretendemos, na “Trilha Pesquisa”, investigar, entre outras coisas, como a metodologia de modelagem conceitual fundamentada em ontologias⁵ pode contribuir para a obtenção de melhores resultados na aplicação dos métodos e técnicas da inteligência artificial. Por exemplo, sabemos o quanto a IA depende dos sistemas de classificação para o treinamento e uso dos seus modelos. As classificações são estruturas de

² Entendemos “rota tecnológica” como um recurso para suporte ao planejamento de projetos de pesquisa e desenvolvimento ao longo do tempo.

³ DENSCOMBE, M. The Good Research Guide. Buckingham: Open University Press. 2005. 310 p.

⁴ THIOLENT, M. Metodologia Pesquisa-Ação. Campinas: Cortez. 1998. 108 p.

⁵ De acordo com Gruber (1993), uma ontologia é “uma especificação de um vocabulário representacional para um domínio compartilhado de discurso – definições de classes, relações, funções e outros objetos” que pode ser utilizada para “apoiar o compartilhamento e a reutilização do conhecimento formalmente representado entre os sistemas de IA”. GRUBER, T.R. A translation approach to portable ontologies. Knowledge Acquisition, v. 5, p. 199–220, 1993.

representação do conhecimento organizadas segundo a relação taxonômica de subsunção (isA). Mesmo que as estatísticas indiquem que foi obtido um alto *F score* numa determinada tarefa de classificação, o resultado prático pode ser prejudicado devido aos problemas inerentes ao próprio sistema de classificação. Dessa forma, pretende-se, como um passo na trilha de pesquisa, revisar a adequação taxonômica dos sistemas de classificação a serem utilizados com o apoio da metodologia Ontoclean⁶ e, se necessário, da modelagem conceitual multinível⁷.

Ainda na “Trilha Pesquisa”, será investigada a possibilidade de construção de um *Knowledge Graph (KG)* específico para o domínio do caso concreto, que poderá ser utilizado, por exemplo, nas tarefas de reconhecimento de entidades nomeadas.

4.2 Metodologia de Gestão e Desenvolvimento de Software

Na “Trilha Desenvolvimento”, a metodologia ágil *Scrum*, indicada pelo Edital, é estruturada em três ciclos, subdivididos em quatro marcos, que por sua vez são subdivididos ainda em uma quantidade variável de *sprints*, conforme detalhamento especificado na próxima subseção.

4.2.1 Detalhamento da Trilha Desenvolvimento - Scrum

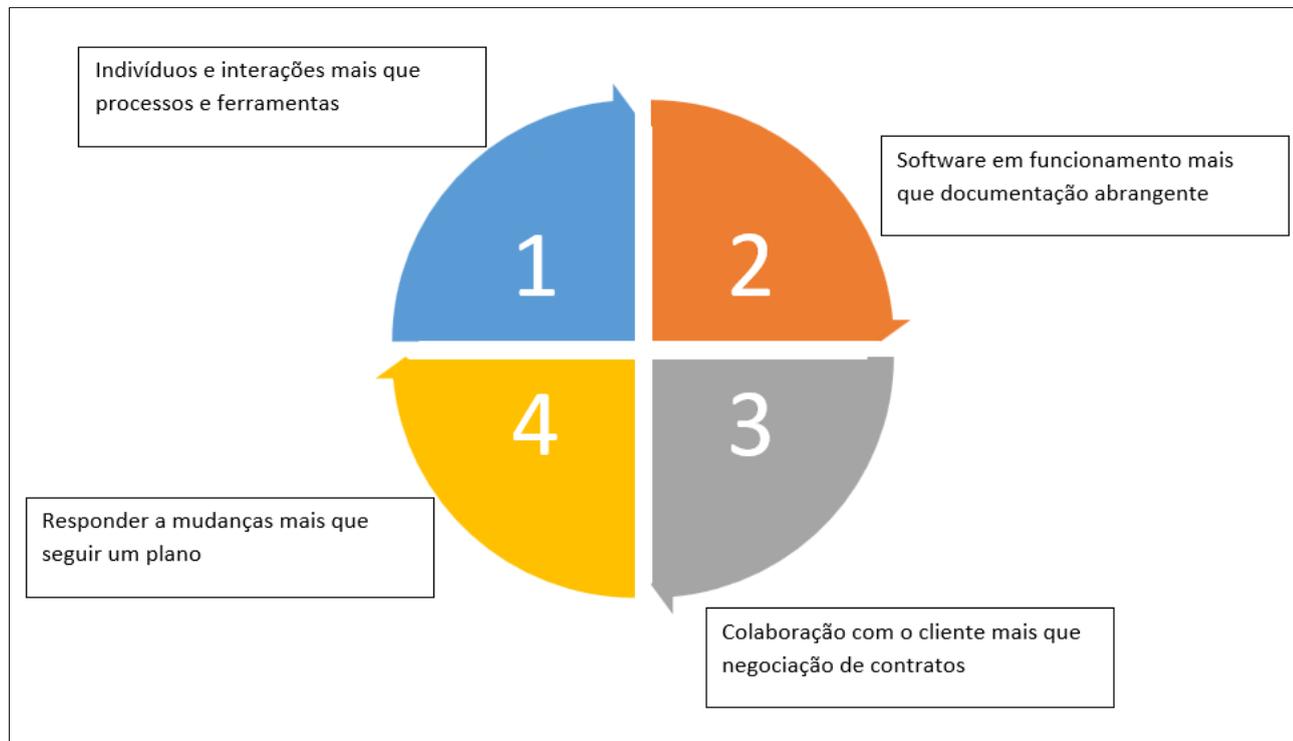
Por se tratar de uma solução marcada pelo ineditismo e pelo risco tecnológico, a ferramenta de Instrução Assistida por Inteligência Artificial deve ser desenvolvida seguindo metodologias ágeis que permitam a correção de rotas a partir do acompanhamento da execução do projeto e a transferência bilateral de conhecimento durante todo o ciclo de desenvolvimento.

De acordo com o Manifesto Ágil, os principais pontos a serem considerados são:

⁶ Guarino, N.; Welty, C. Evaluating Ontological Decisions with Ontoclean. *Communications of ACM*, v. 45, n. 2, p. 61-65, 2002.

⁷ Carvalho V.A. & Almeida J.P.A (2018). Toward a well-founded theory for multi-level conceptual modeling, *Software & Systems Modeling*, 17: 205.

Figura 1: Manifesto Ágil



Pelas características do projeto, chegamos a avaliar o uso do processo de gestão *Lean* combinado com *Google Venture Design Sprints*. No entanto, conforme sugerido pelo próprio edital, optamos por uma associação do *framework Scrum* com artefatos do *Kanban* e passagem de conhecimento seguindo o modelo *Google Venture*.

Assim, cada ciclo será dividido em *sprints*, em que são:

1. Definidos os conjuntos de objetivos (*product backlogs*);
2. Realizadas reuniões de planejamento;
3. Realizadas e desenvolvidas tarefas específicas da *sprint*; e
4. Realizadas reuniões diárias com a equipe.

Ao final de cada *sprint* será realizada uma reunião de revisão e apresentação do trabalho desenvolvido e dos resultados daquele ciclo (*sprint review meeting*).

Terminado o ciclo, será realizada uma *sprint* de uma semana para passagem de conhecimento, baseada no modelo *Google Venture*⁸, que implica a utilização de ferramentas de *design thinking* e a interação *hands on* com a equipe do TCU.

Um dos desafios deste projeto é combinar a “Trilha Pesquisa” com a “Trilha

⁸ O modelo *Google Venture* é um processo intenso de cinco dias, chamado *Design Sprint*, que tem o objetivo de viabilizar que *startups* resolvam problemas rapidamente.

Desenvolvimento”, porque ambas costumam andar em ritmos diferentes. Porém, a utilização das metodologias ágeis tem o intuito de auxiliar no gerenciamento de projetos de inovação, pesquisa e desenvolvimento (P&D) em diferentes áreas de trabalho, a fim de promover principalmente a flexibilidade frente a dificuldades e mudanças constantes.

Em **projetos de inovação** é necessário um esforço para criar um produto ou serviço para alcançar resultados eficazes. Na iniciativa privada, por exemplo, as empresas buscam inovar para melhorar desempenho, aumentar a produtividade, criar diferenciação mercadológica e diminuir custos.

Já em **P&D**, é preciso adquirir os conhecimentos necessários para um projeto de inovação efetivo. Isso faz parte do conjunto de atividades de inovação, o qual pode ser realizado em diferentes etapas do processo, como fonte de conhecimento para ideias ou para solucionar problemas que possam surgir ao longo do processo de inovação.

Projetos de **P&D** geralmente envolvem termos de cooperação com institutos de pesquisa e universidades. Esses termos eram normalmente elaborados de maneira planejada, com entregas que, dadas as incertezas da pesquisa, poderiam levar até anos. No entanto, no contexto da transformação digital, existem restrições de tempo de colocação no mercado e a necessidade de entregas rápidas para experimentar opções de solução.

A aplicação de metodologias ágeis ajudam a reduzir o tempo de entrega dos projetos de **P&D** sem comprometer a qualidade da pesquisa, na medida que:

- Apresenta um problema para pesquisa.
- Refina a história em experimentos viáveis.
- Inicia os experimentos.
- Avalia os resultados à medida que avança e em cada limite de *sprint*.
- Determina os próximos passos com base no que foi aprendido na *sprint* finalizada.

4.3 Metodologia de Gerenciamento de Projetos

4.3.1 Scrum e Kanban

O gerenciamento do desenvolvimento da ferramenta de Instrução Assistida por Inteligência Artificial do TCU seguirá um processo de metodologia ágil, utilizando o *Framework Scrum* com gestão *Kanban*. Ambos podem ser facilmente integrados, unindo

a praticidade e as questões visuais do *Kanban* aos ritos e cerimônias de acompanhamento do *Scrum*.

Scrum é um *framework* baseado no empirismo e no *lean thinking*. O empirismo afirma que o conhecimento vem da experiência e da tomada de decisões com base no que é observado. O *lean thinking* reduz o desperdício e se concentra no essencial.

Esse *framework* emprega uma abordagem iterativa e incremental para otimizar a previsibilidade e controlar o risco. Envolve grupos de pessoas que, coletivamente, possuem todas as habilidades e conhecimentos necessários para fazer o trabalho e compartilhar ou adquirir essas habilidades conforme necessário.

Para o desenvolvimento dos artefatos desta ETEC, iremos seguir os **eventos e cerimônias do *Scrum***, segmentando os trabalhos em *sprints* e realizando as reuniões pontuais previstas no *framework*: planejamento, diária, revisão e retrospectiva.

- ***Sprint*** - As sprints são o coração do *Scrum*, são eventos de duração fixa onde ideias são transformadas em valor. **Para este projeto utilizaremos uma *timebox* de 15 dias.**
- ***Sprint Planning*** - A *Sprint Planning* é uma reunião planejada realizada no início de cada *sprint* e define o trabalho a ser realizado, indicando as prioridades de todo o *Product Backlog* e formando o *Sprint Backlog*, atividades selecionadas para serem executadas na *sprint*.
- ***Daily Scrum*** - É uma reunião diária de 15 minutos para os desenvolvedores que integram a equipe *Scrum*. As *Daily Scrums* melhoram as comunicações, identificam os impedimentos, promovem a rápida tomada de decisões e conseqüentemente, eliminam a necessidade de outras reuniões.
O horário e a forma como elas acontecerão pode ser definida no início do projeto, mas como sugestão, indicamos que sejam feitas pela manhã, por meio de ferramenta de videoconferência, como o *Google Meet*, para que todos os integrantes da equipe possam participar.
- ***Sprint Review*** - Com duração máxima de 2 horas, as reuniões de revisão de *sprints* ocorrem para que as partes interessadas, ou seja, os principais *stakeholders*, validem as entregas da equipe e verifiquem se os critérios estabelecidos no planejamento foram executados.
As pautas dessa reunião serão:
 - Apresentação das entregas pela equipe;
 - *Feedback* do *Product Owner* (PO), representado pelo TCU;
 - Validação de conclusão ou não da entrega.
- ***Sprint Retrospective*** - Ao final de cada *sprint*, a reunião de retrospectiva será realizada, visando aumentar a qualidade e a eficácia. Este evento é considerado

o evento de conclusão da *sprint* e será limitado a um *timebox* de 2 horas. Nessa reunião, a equipe *Scrum* identificará as mudanças mais úteis para melhorar sua eficácia, partindo de quatro perguntas principais:

- O que fizemos bem?
- O que fizemos que pode melhorar?
- O que aprendemos na *Sprint*?
- O que faremos diferente?

As melhorias mais impactantes serão ajustadas o mais rápido possível, podendo inclusive, ser inseridas ao *Sprint Backlog* para a próxima *Sprint*, como mostra a figura do framework *scrum* abaixo.

É importante observar que a “Trilha Pesquisa” permeia toda a “Trilha Desenvolvimento” em cada Marco, com concentração maior de esforço de pesquisa no início de cada Ciclo e validação das hipóteses ao longo do desenvolvimento (Figura 2).

Figura 2. Trilha de Pesquisa no desenvolvimento

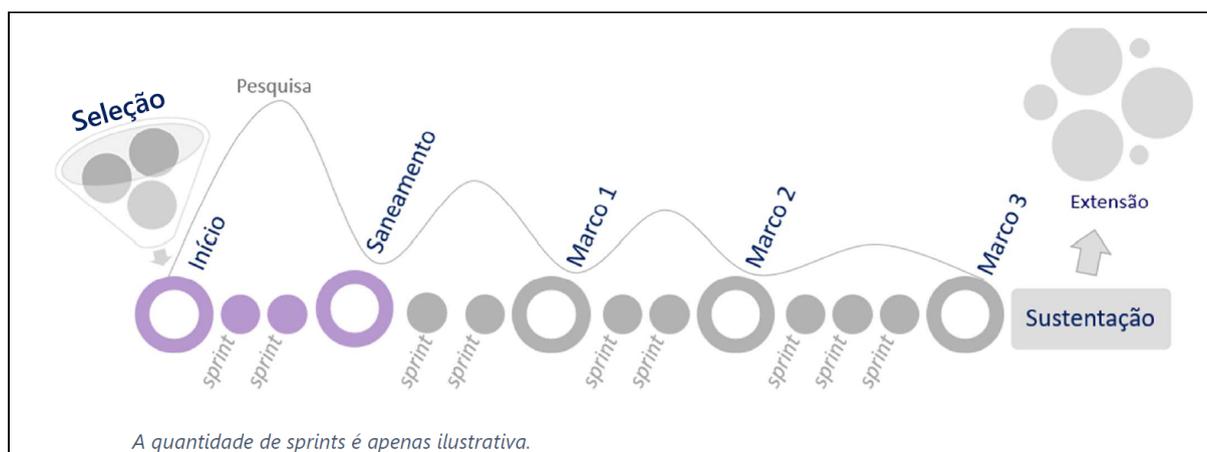
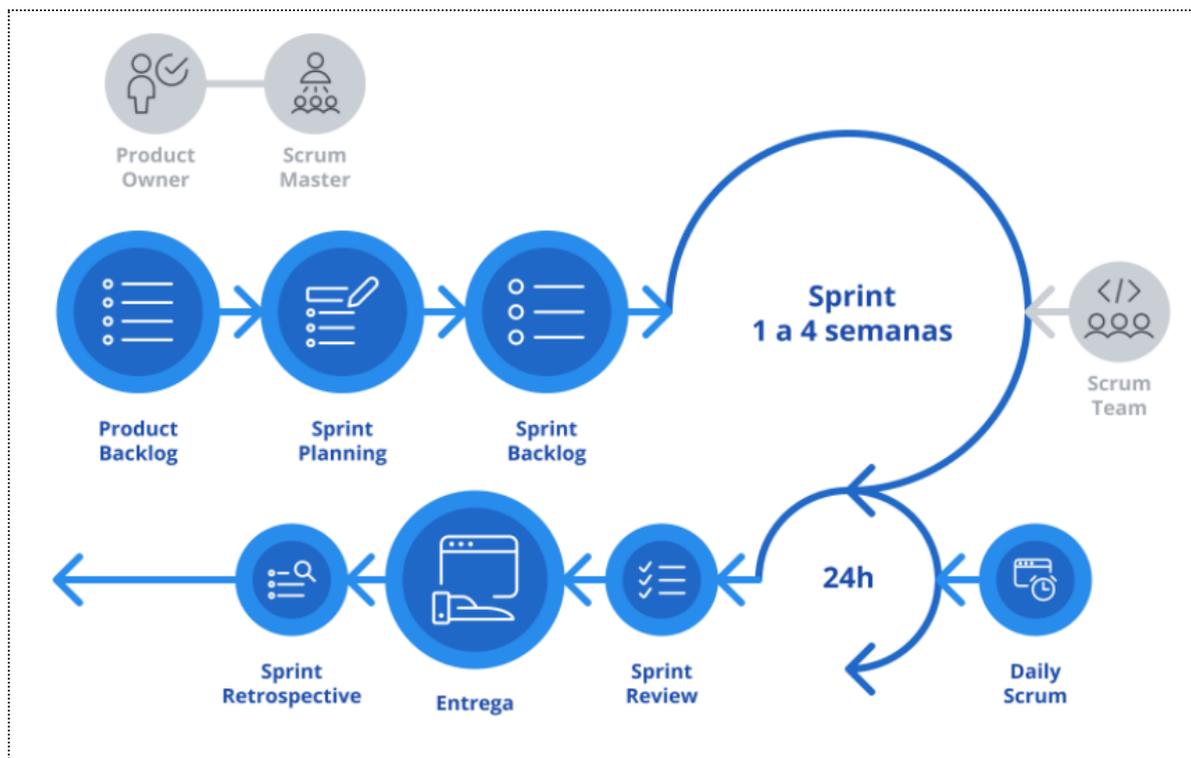


Figura 3. Framework Scrum



- **Kanban** - O *Kanban* parte do princípio de que as atividades devem ser monitoradas de maneira clara e visual, de acordo com o mapeamento de fluxos ou processos da empresa. De maneira muito simples, o *Kanban* determina que as tarefas devem ser limitadas e que uma tarefa só deve ser iniciada quando a anterior é finalizada, no que é chamado *WIP*, ou *Work in Progress*. O *Kanban* facilita a transparência, dando início ao processo, expondo gargalos, filas, desperdício e variabilidade. Essa exposição incentiva e impacta em melhorias no processo e equipe envolvida. O *WIP* será delimitado no início do projeto, em comum acordo entre as partes.

O projeto será desenvolvido de acordo com o *Product Backlog* abaixo:

- Rotulagem de dados
- Classificação por ontologia
- Criação do servidor sFTP
- Criação da infraestrutura de segurança
- Preparo de ambiente em *Python*
- Preparo e limpeza dos dados
- Desenvolvimento da infraestrutura do modelo
- Desenvolvimento do modelo de NLP
- Curadoria humana de resultados do modelo

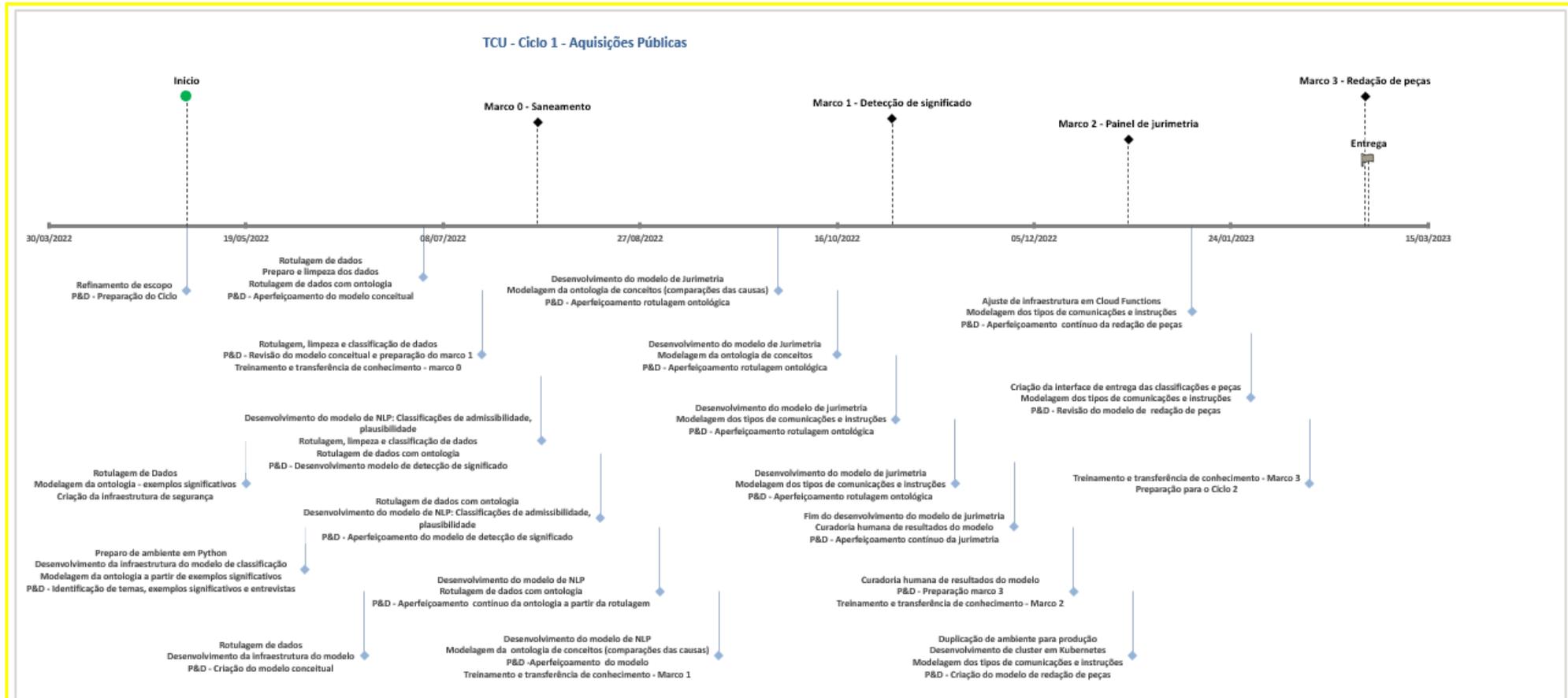
- Duplicação de ambiente para produção
- Desenvolvimento de cluster em *Kubernetes*
- Ajuste de infraestrutura em *Cloud Functions*
- Criação da interface de entrega das classificações e peças
- Treinamento e transferência de conhecimento

4.4 Planejamento Preliminar do Primeiro Ciclo - Aquisições Públicas

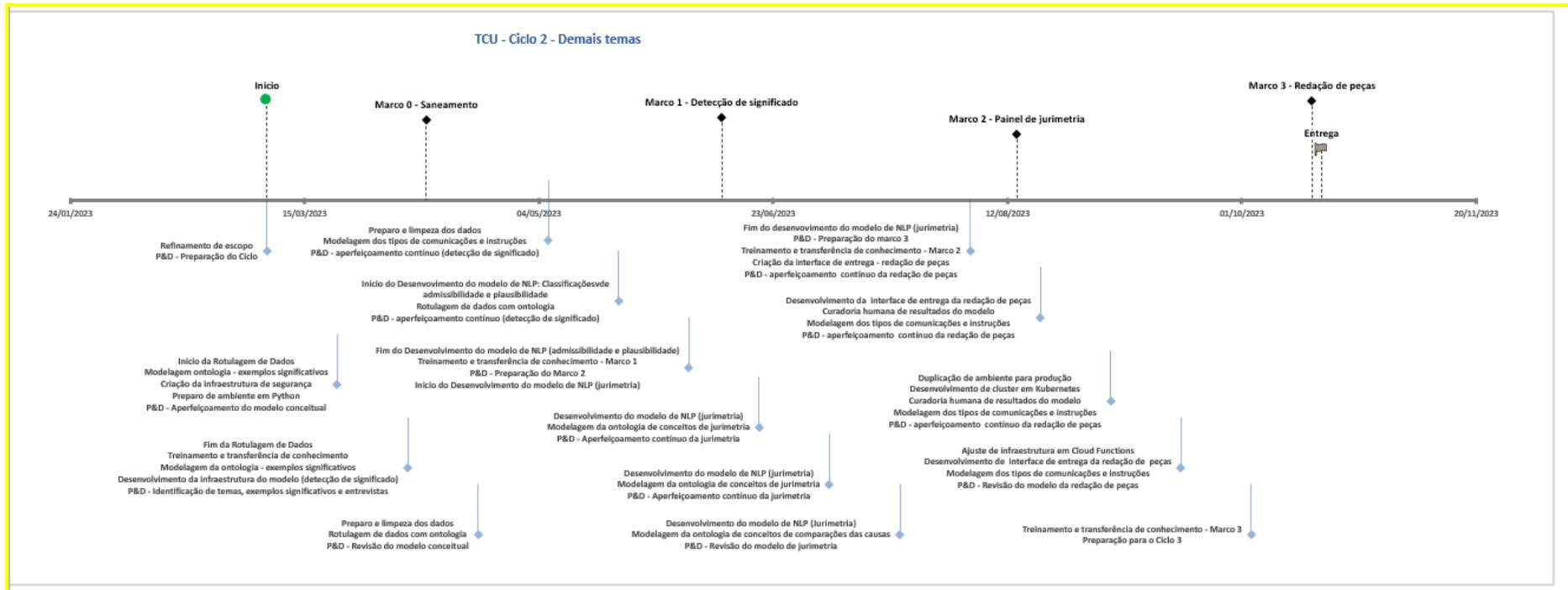
Será utilizado um *timebox* de 15 dias para o projeto, prevendo 20 *sprints* para o primeiro ciclo e 14 *sprints* para os segundo e terceiro ciclos. As duas *sprints* iniciais dos ciclos 2 e 3 são usadas para saneamento dos dados, que aqui chamamos “Marco 0”.

Nesta proposta, detalhamos as *sprints* de desenvolvimento dos Ciclos 1, 2 e 3, como indicado nas Figuras [4](#), [5](#) e [6](#). Mas é importante salientar que esse pré-planejamento é apenas uma sugestão para dar a ideia do planejamento do desenvolvimento, uma vez que não é característica das metodologias ágeis fazer planejamentos longos com antecedência. Afinal, o *backlog* da *sprint* é definido no início de cada *sprint* e depende do aprendizados e objetivos alcançados na *sprint* anterior.

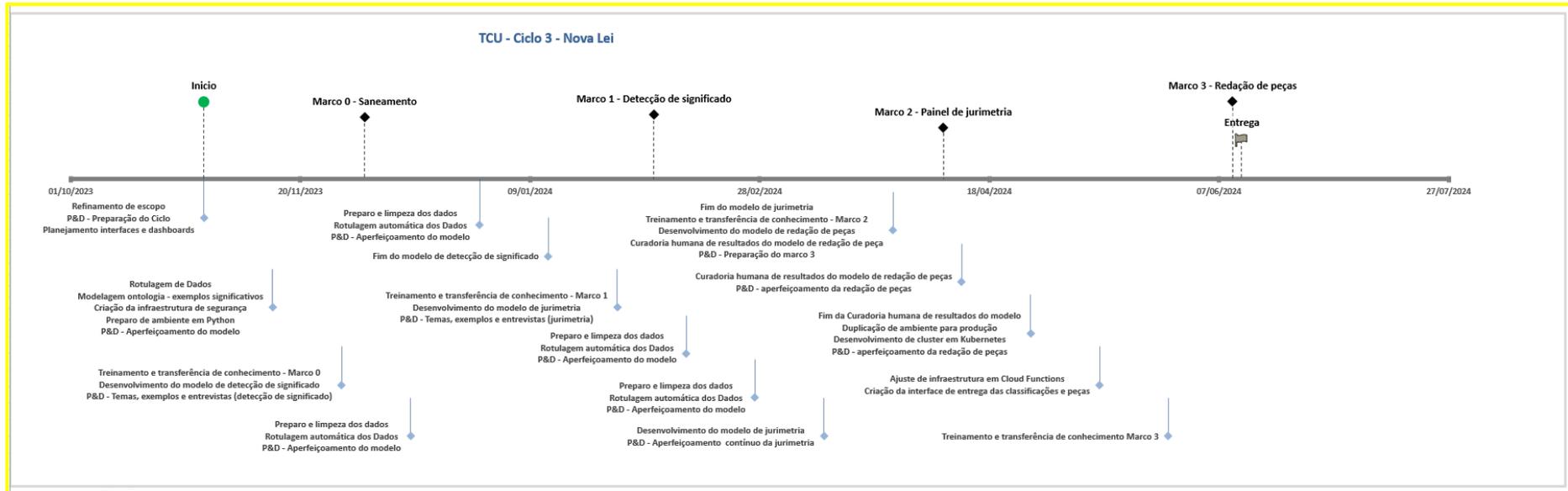
Figura 4: Pré-planejamento das *sprints* do Ciclo 1 (duração: 10 meses)



Figuras 5: Pré-planejamento das sprints do Ciclo 2 (duração: 7 meses cada)



Figuras 6: Pré-planejamento das *sprints* do Ciclo 3 (duração: 7 meses cada)



Nesta proposta detalhamos as sprints do Ciclo 1, de 10 meses, mas trata-se apenas de uma sugestão, uma vez que o planejamento de longo prazo não é característico das metodologias ágeis.

Tabela 1: *Backlog* das sprints do Ciclo 1 (datas ilustrativas)

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---|--------|-----------|-----------|---------|--|
| Marco 0 Saneamento | 1 | 4/5/2022 | 18/5/2022 | 15 | Refinamento de escopo P&D - Preparação do Ciclo |
| | 2 | 19/5/2022 | 2/6/2022 | 15 | Rotulagem de Dados Modelagem da ontologia - exemplos significativos Criação da infraestrutura de segurança |
| | 3 | 3/6/2022 | 17/6/2022 | 15 | Preparo de ambiente em Python Desenvolvimento da infraestrutura do modelo de classificação Modelagem da ontologia a partir de exemplos significativos P&D - Identificação de temas, exemplos significativos e entrevistas |
| | 4 | 18/6/2022 | 2/7/2022 | 15 | Rotulagem de dados Desenvolvimento da infraestrutura do modelo P&D - Criação do modelo conceitual |
| | 5 | 3/7/2022 | 17/7/2022 | 15 | Rotulagem de dados Preparo e limpeza dos dados Rotulagem de dados com ontologia P&D - Aperfeiçoamento do modelo conceitual |
| | 6 | 18/7/2022 | 1/8/2022 | 15 | Rotulagem, limpeza e classificação de dados P&D - Revisão do modelo conceitual e preparação do marco 1 Treinamento e transferência de conhecimento - marco 0 |
| Marco 1 – Detecção de Significado | 7 | 2/8/2022 | 16/8/2022 | 15 | Desenvolvimento do modelo de NLP: Classificações de admissibilidade, plausibilidade Rotulagem, limpeza e classificação de dados Rotulagem de dados com ontologia P&D - Desenvolvimento modelo de detecção de significado |
| | 8 | 17/8/2022 | 31/8/2022 | 15 | Rotulagem de dados com ontologia Desenvolvimento do modelo de NLP: Classificações de admissibilidade, plausibilidade |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---------------------------------|--------------|------------|------------|---------|---|
| | | | | | P&D - Aperfeiçoamento do modelo de detecção de significado |
| | 9 | 1/9/2022 | 15/9/2022 | 15 | Desenvolvimento do modelo de NLP Rotulagem de dados com ontologia P&D - Aperfeiçoamento contínuo da ontologia a partir da rotulagem |
| | 10 - Marco 1 | 16/9/2022 | 30/9/2022 | 15 | Desenvolvimento do modelo de NLP Modelagem da ontologia de conceitos (comparações das causas) P&D -Aperfeiçoamento do modelo Treinamento e transferência de conhecimento - Marco 1 |
| Marco 2 Painel de Jurimetria | 11 | 1/10/2022 | 15/10/2022 | 15 | Desenvolvimento do modelo de Jurimetria Modelagem da ontologia de conceitos (comparações das causas) P&D - Aperfeiçoamento rotulagem ontológica |
| | 12 | 16/10/2022 | 30/10/2022 | 15 | Desenvolvimento do modelo de Jurimetria Modelagem da ontologia de conceitos P&D - Aperfeiçoamento rotulagem ontológica |
| | 13 | 31/10/2022 | 14/11/2022 | 15 | Desenvolvimento do modelo de jurimetria Modelagem dos tipos de comunicações e instruções P&D - Aperfeiçoamento rotulagem ontológica |
| | 14 | 15/11/2022 | 29/11/2022 | 15 | Desenvolvimento do modelo de jurimetria Modelagem dos tipos de comunicações e instruções P&D - Aperfeiçoamento rotulagem ontológica |
| | 15 | 30/11/2022 | 14/12/2022 | 15 | Fim do desenvolvimento do modelo de jurimetria Curadoria humana de resultados do modelo P&D - Aperfeiçoamento contínuo da jurimetria |
| | 16 | 15/12/2022 | 29/12/2022 | 15 | Curadoria humana de resultados do modelo P&D - Preparação marco 3 Treinamento e transferência de |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|--------------------------------|--------|------------|-----------|---------|---|
| | | | | | conhecimento - Marco 2 |
| Marco 3 Redação de Peças | 17 | 30/12/2022 | 13/1/2023 | 15 | Duplicação de ambiente para produção Desenvolvimento de cluster em Kubernetes Modelagem dos tipos de comunicações e instruções P&D - Criação do modelo de redação de peças |
| | 18 | 14/1/2023 | 28/1/2023 | 15 | Ajuste de infraestrutura em Cloud Functions Modelagem dos tipos de comunicações e instruções P&D - Aperfeiçoamento contínuo da redação de peças |
| | 19 | 29/1/2023 | 12/2/2023 | 15 | Criação da interface de entrega das classificações e peças Modelagem dos tipos de comunicações e instruções P&D - Revisão do modelo de redação de peças |
| | 20 | 13/2/2023 | 27/2/2023 | 15 | Treinamento e transferência de conhecimento - Marco 3 Preparação para o Ciclo 2 |

Tabela 2: *Backlog* das sprints do Ciclo 2 (datas ilustrativas)

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---------------------------------------|--------|------------|------------|---------|--|
| Marco 0 Saneamento | 21 | 31/10/2023 | 14/11/2023 | 15 | Refinamento de escopo P&D - Preparação do Ciclo |
| | 22 | 15/11/2023 | 29/11/2023 | 15 | Início da Rotulagem de Dados Modelagem ontologia - exemplos significativos Criação da infraestrutura de segurança Preparo de ambiente em Python P&D - Aperfeiçoamento do modelo conceitual |
| | 23 | 30/11/2023 | 14/12/2023 | 15 | Fim da Rotulagem de Dados Treinamento e transferência de conhecimento - marco 0 |
| Marco 1 Detecção de Significado | 23 | 30/11/2023 | 14/12/2023 | 15 | Modelagem da ontologia - exemplos significativos Desenvolvimento da infraestrutura do modelo |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---------------------------------|--------|------------|------------|---------|--|
| | | | | | (detecção de significado) P&D - Identificação de temas, exemplos significativos e entrevistas |
| | 24 | 15/12/2023 | 29/12/2023 | 15 | Preparo e limpeza dos dados Rotulagem de dados com ontologia P&D - Revisão do modelo conceitual |
| | 25 | 30/12/2023 | 13/1/2024 | 15 | Preparo e limpeza dos dados Modelagem dos tipos de comunicações e instruções P&D - aperfeiçoamento contínuo (detecção de significado) |
| | 26 | 14/1/2024 | 28/1/2024 | 15 | Início do Desenvolvimento do modelo de NLP: Classificações de admissibilidade e plausibilidade Rotulagem de dados com ontologia P&D - aperfeiçoamento contínuo (detecção de significado) |
| | 27 | 29/1/2024 | 12/2/2024 | 15 | Fim do Desenvolvimento do modelo de NLP (admissibilidade e plausibilidade) Treinamento e transferência de conhecimento - Marco 1 P&D - Preparação do Marco 2 |
| Marco 2 Painel de Jurimetria | 27 | 29/1/2024 | 12/2/2024 | 15 | Início do Desenvolvimento do modelo de NLP (jurimetria) |
| | 28 | 13/2/2024 | 27/2/2024 | 15 | Desenvolvimento do modelo de NLP (jurimetria) Modelagem da ontologia de conceitos de jurimetria P&D - Aperfeiçoamento contínuo da jurimetria |
| | 29 | 28/2/2024 | 13/3/2024 | 15 | Desenvolvimento do modelo de NLP (jurimetria) Modelagem da ontologia de conceitos de jurimetria P&D - Aperfeiçoamento contínuo da jurimetria |
| | 30 | 14/3/2024 | 28/3/2024 | 15 | Desenvolvimento do modelo de NLP (Jurimetria) Modelagem da ontologia de |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|-----------------------------|--------|-----------|-----------|---------|--|
| | | | | | conceitos de comparações das causas P&D - Revisão do modelo de jurimetria |
| | 31 | 29/3/2024 | 12/4/2024 | 15 | Fim do Desenvolvimento do modelo de NLP (jurimetria) P&D - Preparação do marco 3 Treinamento e transferência de conhecimento - Marco 2 |
| Marco 3 Redação de Peças | 31 | 29/3/2024 | 12/4/2024 | 15 | Criação da interface de entrega - redação de peças P&D - aperfeiçoamento contínuo da redação de peças |
| | 32 | 13/4/2024 | 27/4/2024 | 15 | Desenvolvimento da interface de entrega da redação de peças Curadoria humana de resultados do modelo Modelagem dos tipos de comunicações e instruções P&D - aperfeiçoamento contínuo da redação de peças |
| | 33 | 28/4/2024 | 12/5/2024 | 15 | Duplicação de ambiente para produção Desenvolvimento de cluster em Kubernetes Curadoria humana de resultados do modelo Modelagem dos tipos de comunicações e instruções P&D - aperfeiçoamento contínuo da redação de peças |
| | 34 | 13/5/2024 | 27/5/2024 | 15 | Ajuste de infraestrutura em Cloud Functions Desenvolvimento de interface de entrega da redação de peças Modelagem dos tipos de comunicações e instruções P&D - Revisão do modelo da redação de peças |
| | 35 | 28/5/2024 | 11/6/2024 | 15 | Treinamento e transferência de conhecimento - Marco 3 Preparação para o Ciclo 3 |

Tabela 3: *Backlog das sprints do Ciclo 3 (datas ilustrativas)*

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---------------------------------------|--------|-----------|-----------|---------|---|
| Marco 0 Saneamento | 35 | 7/3/2023 | 21/3/2023 | 15 | Refinamento de escopo P&D - Preparação do Ciclo Planejamento interfaces e dashboards |
| | 36 | 22/3/2023 | 5/4/2023 | 15 | Rotulagem de Dados Modelagem ontologia - exemplos significativos Criação da infraestrutura de segurança Preparo de ambiente em Python P&D - Aperfeiçoamento do modelo |
| | 37 | 6/4/2023 | 20/4/2023 | 15 | Treinamento e transferência de conhecimento - Marco 0 |
| Marco 1 Detecção de Significado | 37 | 6/4/2023 | 20/4/2023 | 15 | Desenvolvimento do modelo de detecção de significado P&D - Temas, exemplos e entrevistas (detecção de significado) |
| | 38 | 21/4/2023 | 5/5/2023 | 15 | Preparo e limpeza dos dados Rotulagem automática dos Dados P&D - Aperfeiçoamento do modelo |
| | 39 | 6/5/2023 | 20/5/2023 | 15 | Preparo e limpeza dos dados Rotulagem automática dos Dados P&D - Aperfeiçoamento do modelo |
| | 40 | 21/5/2023 | 4/6/2023 | 15 | Fim do modelo de detecção de significado |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|---------------------------------|--------|-----------|-----------|---------|---|
| | 41 | 5/6/2023 | 19/6/2023 | 15 | Treinamento e transferência de conhecimento - Marco 1 |
| Marco 2 Painel de Jurimetria | 41 | 5/6/2023 | 19/6/2023 | 15 | Desenvolvimento do modelo de jurimetria P&D - Temas, exemplos e entrevistas (jurimetria) |
| | 42 | 20/6/2023 | 4/7/2023 | 15 | Preparo e limpeza dos dados Rotulagem automática dos Dados P&D - Aperfeiçoamento do modelo |
| | 43 | 5/7/2023 | 19/7/2023 | 15 | Preparo e limpeza dos dados Rotulagem automática dos Dados P&D - Aperfeiçoamento do modelo |
| | 44 | 20/7/2023 | 3/8/2023 | 15 | Desenvolvimento do modelo de jurimetria P&D - Aperfeiçoamento contínuo da jurimetria |
| | 45 | 4/8/2023 | 18/8/2023 | 15 | Fim do modelo de jurimetria Treinamento e transferência de conhecimento - Marco 2 |
| Marco 3 Redação de Peças | 45 | 4/8/2023 | 18/8/2023 | 15 | Desenvolvimento do modelo de redação de peças Curadoria humana de resultados do modelo de redação de peça P&D - Preparação do marco 3 |
| | 46 | 19/8/2023 | 2/9/2023 | 15 | Curadoria humana de resultados do modelo de redação de peças P&D - aperfeiçoamento da redação de peças |

| Marco | Sprint | Início | Fim | Duração | Backlog |
|-------|--------|-----------|-----------|---------|--|
| | 47 | 3/9/2023 | 17/9/2023 | 15 | Fim da Curadoria humana de resultados do modelo Duplicação de ambiente para produção Desenvolvimento de cluster em Kubernetes P&D - aperfeiçoamento da redação de peças |
| | 48 | 18/9/2023 | 2/10/2023 | 15 | Ajuste de infraestrutura em Cloud Functions Criação da interface de entrega das classificações e peças Treinamento e transferência de conhecimento Marco 3 |

4.5 Métricas

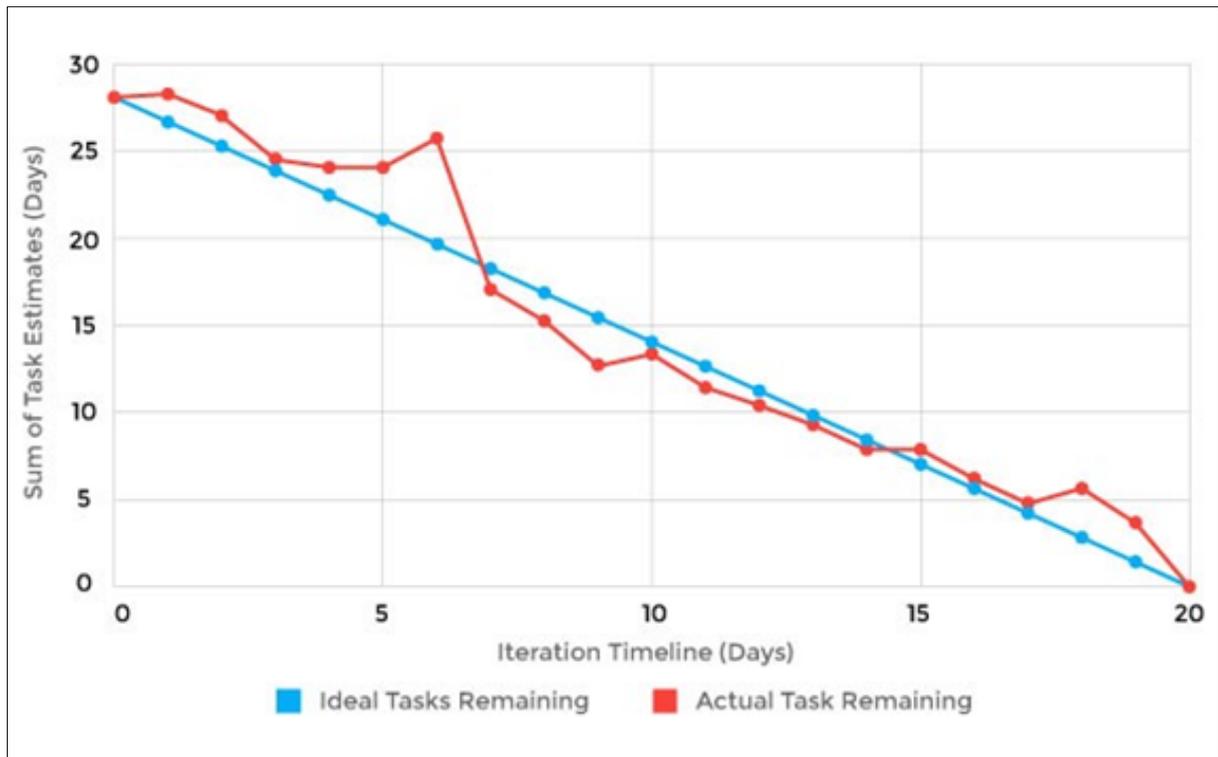
Para acompanhamento da evolução das *sprints* e conclusão das tarefas, serão utilizadas as métricas abaixo:

Sprint Burndown

O gráfico de *burndown* é uma representação visual que mostra a quantidade de trabalho restante em um projeto *versus* a quantidade de tempo em que ele deve ser executado. Permite a visualização e entendimento do progresso para facilitar o alcance aos objetivos. Existem dois tipos diferentes de gráficos de *burndown*:

- Gráfico de *burndown* de *Sprint*: mede o trabalho restante em uma tarefa específica ou *sprint*
- Gráfico de *burndown* de trabalho: avalia o trabalho restante em um projeto inteiro

Figura 7: Gráfico de *Burndown*

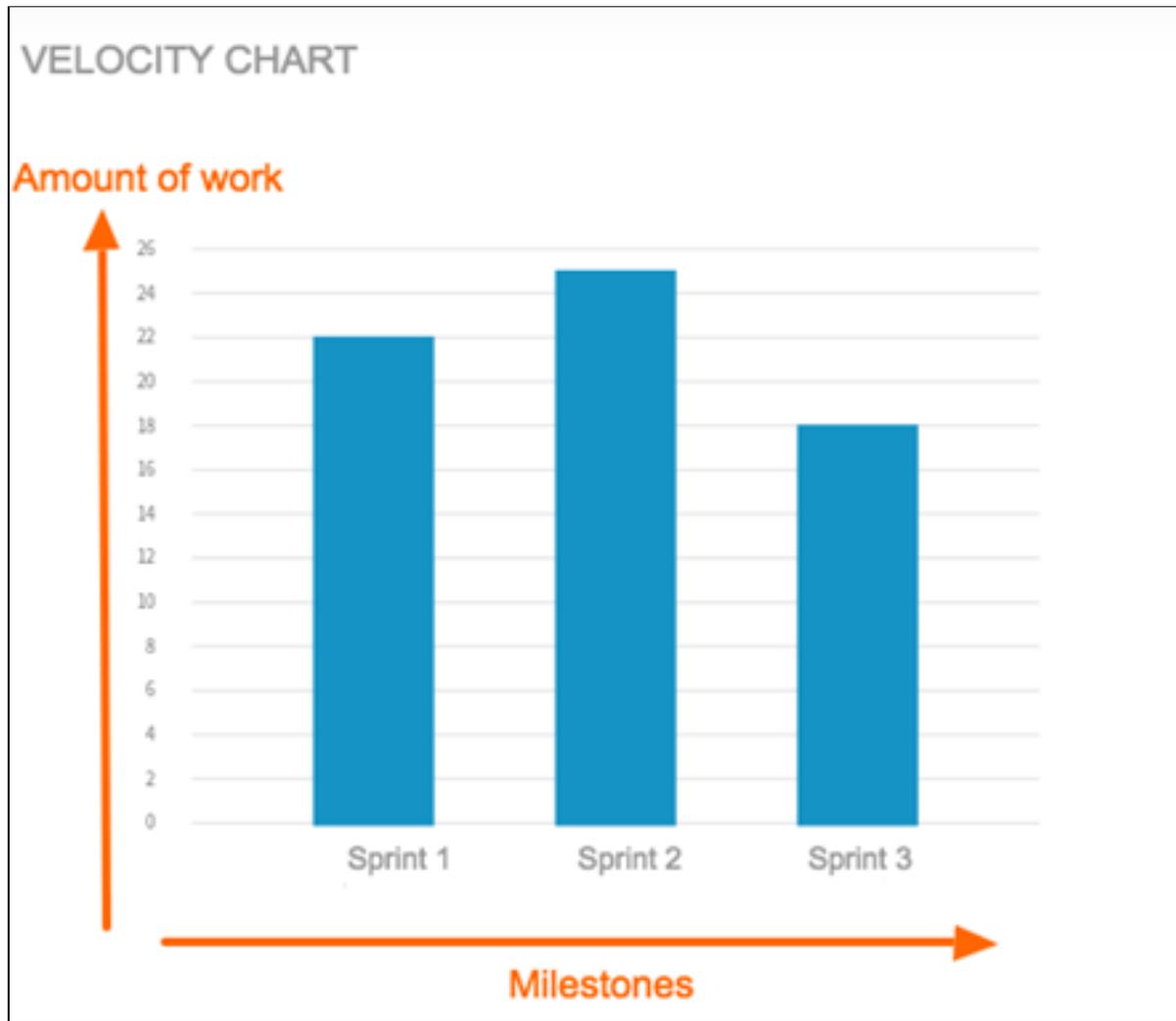


A unidade de medida pode ser horas ou pontos da história e a equipe prevê a carga de trabalho no início de um *Sprint*. O objetivo é concluir a carga de trabalho até o final do *Sprint*.

Velocity Chart

Velocidade é a quantidade média de trabalho que uma equipe de *scrum* conclui durante uma *sprint*, medida em horas ou pontos de história, considerando as estimativas de tempo ou tamanho (planejado) *versus* o executado (realizado) ao final da *sprint*.

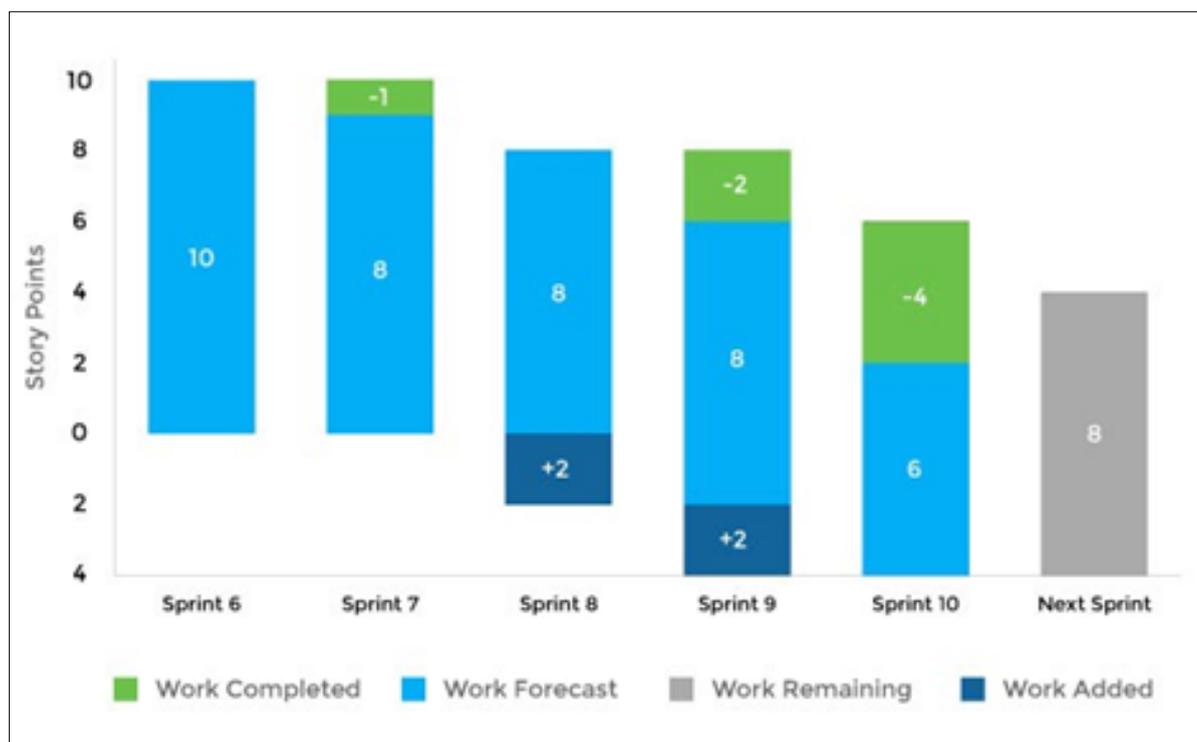
Figura 8: *Velocity Chart*



Release Burndown

Ao contrário de um *burndown* de *sprint*, o *Release Burndown* concentra-se no quadro geral, acompanhando o progresso em um grande corpo de trabalho. Existem muitos épicos e versões de trabalho em uma *sprint*; portanto, é importante acompanhar tais progressos, bem como o de cada *sprint*.

Figura 9: Release Burndown



6 Cronograma Físico-Financeiro

As tabelas [4](#), [5](#) e [6](#) apresentam o cronograma físico-financeiro para cada ciclo do projeto, alinhado às *sprints* associadas. Dessa forma, é possível ter uma visão do custo do projeto durante toda a sua execução, observados os seguintes aspectos:

- Como se trata de solução tecnológica com P&D, os valores apresentados são estimados, podendo sofrer ajustes no decorrer do projeto em função da natureza inerente à ETEC.
- O planejamento do *backlog* das *sprints* deverá ser ajustado conforme o andamento do projeto, como é típico do *Scrum*. Isso pode trazer a necessidade de redimensionamento de custos.
- Conforme previsto no Edital, o Ciclo 4 - Sustentação e Evolução será negociado entre as partes. Essa definição é acertada, uma vez que seria irreal prever valores antes do projeto alcançar um certo nível de maturidade, já que a natureza do projeto trata diversas variáveis cujo comportamento e impacto no desenvolvimento e na pesquisa ainda é incerto.

Tabela 4: Cronograma Físico-Financeiro - Ciclo 1: Aquisições Públicas

| CICLO | SPRINT | DIAS DO PROJETO | | | CRONOGRAMA DE DESEMBOLSO (R\$) | | | | | | | | | |
|-------|--------|-----------------|-----|-------------------|--------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| | | INÍCIO | FIM | DURAÇÃO DA SPRINT | MÊS 1 | MÊS 2 | MÊS 3 | MÊS 4 | MÊS 5 | MÊS 6 | MÊS 7 | MÊS 8 | MÊS 9 | MÊS 10 |
| 1 | 1 | 0 | 15 | 15 | 196.899,07 | | | | | | | | | |
| | 2 | 15 | 30 | 15 | | | | | | | | | | |
| | 3 | 30 | 45 | 15 | 154.624,93 | | | | | | | | | |
| | 4 | 45 | 60 | 15 | | | | | | | | | | |
| | 5 | 60 | 75 | 15 | 148.306,77 | | | | | | | | | |
| | 6 | 75 | 90 | 15 | | | | | | | | | | |

| CICLO | SPRINT | DIAS DO PROJETO | | | CRONOGRAMA DE DESEMBOLSO (R\$) | | | | | | | | | |
|------------------------|--------|-----------------|-----|-------------------|--------------------------------|-------|-------|------------|------------|------------|------------|------------|------------|------------|
| | | INÍCIO | FIM | DURAÇÃO DA SPRINT | MÊS 1 | MÊS 2 | MÊS 3 | MÊS 4 | MÊS 5 | MÊS 6 | MÊS 7 | MÊS 8 | MÊS 9 | MÊS 10 |
| | 7 | 90 | 105 | 15 | | | | 227.542,62 | | | | | | |
| | 8 | 105 | 120 | 15 | | | | | | | | | | |
| | 9 | 120 | 135 | 15 | | | | | | | | | | |
| | 10 | 135 | 150 | 15 | | | | | 227.542,62 | | | | | |
| | 11 | 150 | 165 | 15 | | | | | | | | | | |
| | 12 | 165 | 180 | 15 | | | | | | 194.460,62 | | | | |
| | 13 | 180 | 195 | 15 | | | | | | | | | | |
| | 14 | 195 | 210 | 15 | | | | | | | 377.325,55 | | | |
| | 15 | 210 | 225 | 15 | | | | | | | | | | |
| | 16 | 225 | 240 | 15 | | | | | | | | 293.276,91 | | |
| | 17 | 240 | 255 | 15 | | | | | | | | | | |
| | 18 | 255 | 270 | 15 | | | | | | | | | 326.596,01 | |
| | 19 | 270 | 285 | 15 | | | | | | | | | | |
| | 20 | 285 | 300 | 15 | | | | | | | | | | 360.971,69 |
| TOTAL DO CICLO: | | | | | R\$ 2.507.546,80 | | | | | | | | | |

Tabela 5: Cronograma Físico-Financeiro - Ciclo 2: Outras Aquisições

| CICLO | SPRINT | DIAS DO PROJETO | | | CRONOGRAMA DE DESEMBOLSO (R\$) | | | | | | | |
|------------------------|--------|-----------------|-----|-------------------|--------------------------------|------------|------------|------------|------------|------------|--------|------------|
| | | INÍCIO | FIM | DURAÇÃO DA SPRINT | MÊS 11 | MÊS 12 | MÊS 13 | MÊS 14 | MÊS 15 | MÊS 16 | MÊS 17 | |
| 2 | 21 | 300 | 315 | 15 | 207.943,88 | | | | | | | |
| | 22 | 315 | 330 | 15 | | | | | | | | |
| | 23 | 330 | 345 | 15 | | 173.110,77 | | | | | | |
| | 24 | 345 | 360 | 15 | | | | | | | | |
| | 25 | 360 | 375 | 15 | | | 196.946,46 | | | | | |
| | 26 | 375 | 390 | 15 | | | | | | | | |
| | 27 | 390 | 405 | 15 | | | | 230.028,47 | | | | |
| | 28 | 405 | 420 | 15 | | | | | | | | |
| | 29 | 420 | 435 | 15 | | | | | 230.028,47 | | | |
| | 30 | 435 | 450 | 15 | | | | | | | | |
| | 31 | 450 | 465 | 15 | | | | | | 265.697,82 | | |
| | 32 | 465 | 480 | 15 | | | | | | | | |
| | 33 | 480 | 495 | 15 | | | | | | | | 361.067,08 |
| | 34 | 495 | 510 | 15 | | | | | | | | |
| TOTAL DO CICLO: | | | | | R\$ 1.664.822,95 | | | | | | | |

Tabela 6: Cronograma Físico-Financeiro - Ciclo 3: Futuras Aquisições Públicas

| CICLO | SPRINT | DIAS DO PROJETO | | | CRONOGRAMA DE DESEMBOLSO (R\$) | | | | | | | | |
|------------------------|--------|-----------------|-----|-------------------|--------------------------------|------------|------------|------------|------------|------------|--------|--|------------|
| | | INÍCIO | FIM | DURAÇÃO DA SPRINT | MÊS 18 | MÊS 19 | MÊS 20 | MÊS 21 | MÊS 22 | MÊS 23 | MÊS 24 | | |
| 3 | 35 | 510 | 525 | 15 | 207.943,88 | | | | | | | | |
| | 36 | 525 | 540 | 15 | | | | | | | | | |
| | 37 | 540 | 555 | 15 | | 173.110,77 | | | | | | | |
| | 38 | 555 | 570 | 15 | | | | | | | | | |
| | 39 | 570 | 585 | 15 | | | 196.946,46 | | | | | | |
| | 40 | 585 | 600 | 15 | | | | | | | | | |
| | 41 | 600 | 615 | 15 | | | | 230.028,46 | | | | | |
| | 42 | 615 | 630 | 15 | | | | | | | | | |
| | 43 | 630 | 645 | 15 | | | | | 230.028,46 | | | | |
| | 44 | 645 | 660 | 15 | | | | | | | | | |
| | 45 | 660 | 675 | 15 | | | | | | 265.697,82 | | | |
| | 46 | 675 | 690 | 15 | | | | | | | | | |
| | 47 | 690 | 705 | 15 | | | | | | | | | 361.067,07 |
| | 48 | 705 | 720 | 15 | | | | | | | | | |
| TOTAL DO CICLO: | | | | | R\$ 1.664.822,93 | | | | | | | | |

7 Planilha de Formação de Preços

A [tabela 7](#) apresenta a planilha estimativa de formação de preços do projeto. Convém fazer as seguintes observações:

- Como se trata de solução tecnológica com P&D, os valores apresentados são estimados, podendo sofrer ajustes no decorrer do projeto em função da natureza inerente à ETEC.
- Não estão inclusos valores para a operação contínua da solução após a finalização do projeto - esses valores incluem custos de nuvem e serviços de software de terceiros.

Tabela 7: Planilha de Formação de Preços

| CICLO | CATEGORIAS DE DESPESAS | | |
|--|-------------------------|-------------------------|-------------------------|
| | SETUP INFRAESTRUTURA | SERVIÇOS DE SOFTWARE | PESSOAS |
| Ciclo 1 - Aquisições Públicas | R\$ 171.336,11 | R\$ 254.400,65 | R\$ 2.081.810,04 |
| Ciclo 2 - Demais Aquisições | R\$ 171.336,11 | R\$ 254.400,70 | R\$ 1.239.086,14 |
| Ciclo 3 - Futuras Aquisições Públicas | R\$ 171.336,09 | R\$ 254.400,70 | R\$ 1.239.086,14 |
| SUBTOTAL: | R\$ 514.008,31 | R\$ 763.202,04 | R\$ 4.559.982,32 |
| DISTRIBUIÇÃO DE DESPESAS (%): | 9% | 13% | 78% |
| TOTAL DO PROJETO: | R\$ 5.837.192,68 | | |

8 Considerações Finais

Por se tratar de um instrumento de contratação pública novo, a Encomenda Tecnológica traz um conjunto de riscos associados, tanto por sua própria natureza - o objeto intrínseco relacionado à pesquisa, desenvolvimento e inovação - quanto pela relativa fragilidade normativa brasileira e falta de experiências relevantes nesta modalidade. Nesse sentido, é válido louvar a iniciativa do TCU em abrir processo de contratação utilizando esse novo instrumento, que certamente se tornará referência para outros órgãos.

A construção da proposta para atendimento à presente ETEC baseou-se na ampla

experiência dos colaboradores das empresas consorciadas, que contam com pesquisadores com formações ao nível de pós-doutorado, doutorado, mestrado e especializações de diferentes áreas, além de técnicos com experiência no desenvolvimento de soluções tecnológicas de ponta com a utilização de métodos, técnicas, modelos e ferramentas de inteligência artificial.

A rota tecnológica adotada para atender o objeto do Edital abrange tanto um esforço na construção de fundamentos e artefatos científicos, utilizados como pilares para a aplicação de modelos e ferramentas de IA, quanto no desenvolvimento de solução baseada em tecnologias de referência no setor, reforçadas por algoritmos e modelos a serem criados especificamente para o caso em pauta.

As abordagens, modelos, ferramentas e técnicas aplicadas têm plenas condições de cumprir todos os requisitos do Edital, garantindo resultados adequados, conforme as expectativas do TCU, e integrando aspectos de inovação que poderão fundamentar outras aplicações, agregando valor, escalabilidade, segurança, *compliance* e rastreabilidade a toda a solução.

Brasília, 25 de março de 2022.

Mamede Lima-Marques
Representante do Consórcio Modal-Intellimetri
assinado eletronicamente

Anexo 1: Portfólio de Projetos

As empresas que constituem o consórcio - Instituto Modal e Intellimetri - têm ampla experiência em pesquisa e desenvolvimento de soluções baseadas em Inteligência Artificial, Machine Learning e Processamento de Linguagem Natural para os setores público e privado, conforme comprovação a seguir.

A1.1 Instituto Modal

A1.1.1 Experiência em Inteligência Artificial

O Instituto Modal é uma Instituição de Ciência, Tecnologia e Inovação (ICT), sem fins lucrativos, que tem como foco principal a pesquisa e a construção aplicações baseadas em diferentes técnicas e ferramentas, especialmente Inteligência Artificial, voltadas à promoção de novas soluções, de novos negócios, de oportunidades e de equidade social nos setores público e privado.

Para garantir o sucesso de um projeto de IA, o Modal adota um método de experimentação com evolução incremental, até alcançar o estágio de produção.

1) Prova de Conceito

A partir das informações e dados disponibilizados, é feita uma análise exploratória para identificar padrões preliminares, avaliar abordagens possíveis e determinar a viabilidade de cada abordagem de IA. A PoC é apresentada através de um painel ou protótipo para discussão com o cliente.

2) Desenvolvimento

Uma vez validada a PoC, a solução começa o ciclo de desenvolvimento e preparação para produção, que consiste na automatização de rotinas de extração, carga e tratamento de dados, refinamento dos algoritmos de IA, incorporação de regras de negócio e desenvolvimento da interface de usuário.

3) Evolução

Finalizado o ciclo de desenvolvimento, o Modal oferece suporte técnico corretivo e/ou evolutivo, bem como integrações adicionais, expansão de casos de uso, transferência de tecnologia e capacitação, conforme a necessidade da empresa.

A1.1.2 Portfólio: Casos e Provas de Conceito

O Instituto Modal disponibiliza algumas Provas de Conceito (PoCs) e casos desenvolvidos para clientes. Todas as informações publicadas foram devidamente autorizadas pelos clientes.

A relação a seguir apresenta os casos e PoCs autorizados ou que não violam informações confidenciais:

- [Inteligência Artificial para o Setor Jurídico](#)
Prova de Conceito que "varre" decisões do Superior Tribunal de Justiça, identifica padrões utilizando técnicas de análise textual e permite criar relações, identificar gargalos e organizar as informações.
- [Inteligência Artificial e Similaridade da Legislação Federal](#)
Prova de Conceito de análise textual não supervisionada para verificar a similaridade semântica entre diferentes leis, destacando pontos de convergência e identificando atos que tratam do mesmo objeto.
- [Inteligência Artificial e a PEC da Previdência](#)
Prova de Conceito para analisar a Proposta de Emenda Constitucional da Previdência enquanto seu texto ainda estava em discussão, utilizando uma combinação de análise textual e *datamining*.
- [Inteligência Artificial para Automação Industrial](#)
Roteamento automático de cabos e configuração da linha de produção
- [Inteligência Artificial para a Eficiência Energética na Indústria](#)
Prova de Conceito para o sistema de refrigeração industrial da Brasal Refrigerantes, com simulador de eficiência energética
- [Inteligência Artificial na Prevenção do Abuso Sexual Infantil na Internet](#)
Identificação de padrões de abordagem textual de mensagens entre assediadores e crianças para prevenção do abuso sexual infantil *on-line*.
- [Inteligência Artificial e a Patologia do Concreto](#)
Prova de Conceito capaz de analisar e identificar a ocorrência de patologias do concreto, viabilizando melhores condições de manutenção das estruturas.
- [Painel COVID-19 no Brasil e no Mundo](#)
Com a crise de saúde sem precedentes, o Instituto Modal construiu um painel com o mapeamento diário dos casos confirmados, mortos e recuperados no

mundo. A abordagem é exploratória e traz dados complementares.

– **Relacionamentos entre Entidades por Grafos**

Aplicação de grafos para explorar as relações entre os objetos de um determinado conjunto em situações complexas.

A1.2 Intellimetri

A1.2.1 Experiência em Inteligência Artificial: ML, NLP e CV

A Intellimetri é uma empresa focada em soluções e serviços baseados em três pilares da Inteligência Artificial:

1. *Machine Learning* e Algoritmos Preditivos (predição de receita, manutenção preditiva e modulação e posologia de medicamento digital e predição de falhas de equipamentos de CTIs, redação automatizada de artigos)
2. Processamento de Linguagem Natural - NLP (*chatbots* e *voicebots*, monitoramento e auditoria de *call centers*, *growth hacking*, automação de redação de artigos noticiosos)
3. Visão Computacional (prevenção de fraude, análise de comportamento, monitoramento de processo produtivo)

A empresa tem patente registrada no Instituto Nacional da Propriedade Industrial de algoritmo de processamento de linguagem natural - análise de voz em chamadas telefônicas (INPI: BR512020002649-4) e é [parceira estratégica do Google Cloud para soluções de Inteligência Artificial, NLP e ML na América Latina](#), tanto para a iniciativa privada quanto para o setor público.

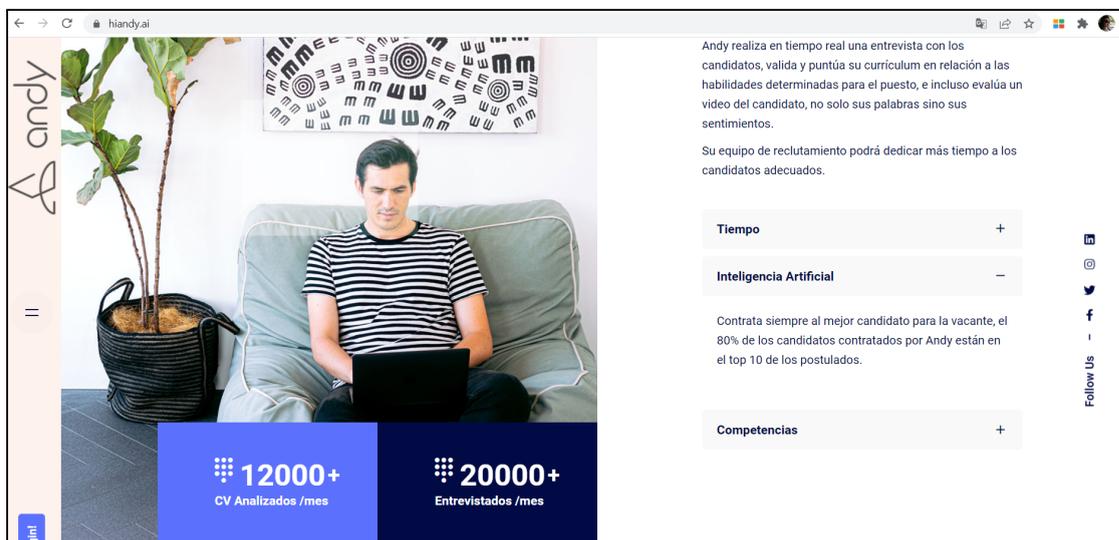
Tanto o certificado de registro no INPI quanto os atestados de capacidade técnica estão no **Anexo 3** desta proposta.

A1.2.2 Portfólio: Casos e Provas de Conceito

A Intellimetri disponibiliza algumas Provas de Conceito (PoCs) e casos de Processamento de Linguagem Natural (NLP) desenvolvidos para clientes da iniciativa privada e do setor público. Todas as informações publicadas foram devidamente autorizadas e não violam informações confidenciais:

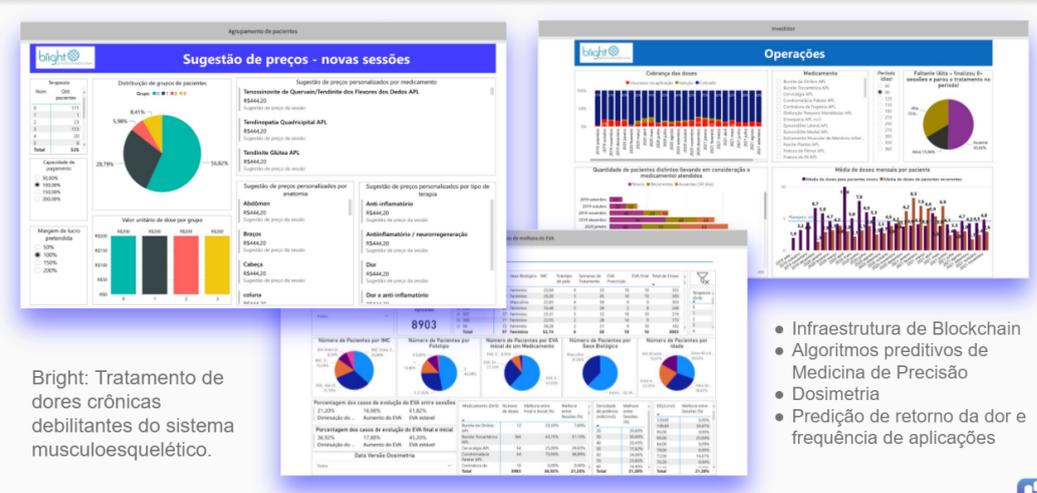
- **[Caso 1] Andy**: Extração de Texto, Entidades e NLP (processamento de linguagem natural) de vídeo currículos e currículos em PDF e Microsoft Word não estruturados e também algoritmos de classificação e previsão em NLP para

um *chatbot* de recrutamento e seleção de pessoal – Marco Marketing. O Atestado de Capacidade Técnica referente a este projeto encontra-se no item 11.3 do Anexo 3.



- [Caso 2] **Bright Medicine**: Extração de Texto, Entidades Nomeadas (NER) e NLP (processamento de linguagem natural) de prontuários médicos não estruturados e armazenados em *Microsoft Word*. Também extração de entidades nomeadas e NLP para um *chatbot* de acompanhamento terapêutico de pacientes (todos os algoritmos treinados em português brasileiro). Por se tratar de uma empresa privada, o acesso a esta ferramenta não está aberto para público geral. O Atestado de Capacidade Técnica referente a este projeto encontra-se no item 11.2 do Anexo 3.

Caso de sucesso: Bright Biomedicine - Remédio Digital



Bright: Tratamento de dores crônicas debilitantes do sistema musculoesquelético.

- Infraestrutura de Blockchain
- Algoritmos preditivos de Medicina de Precisão
- Dosimetria
- Predição de retorno da dor e frequência de aplicações

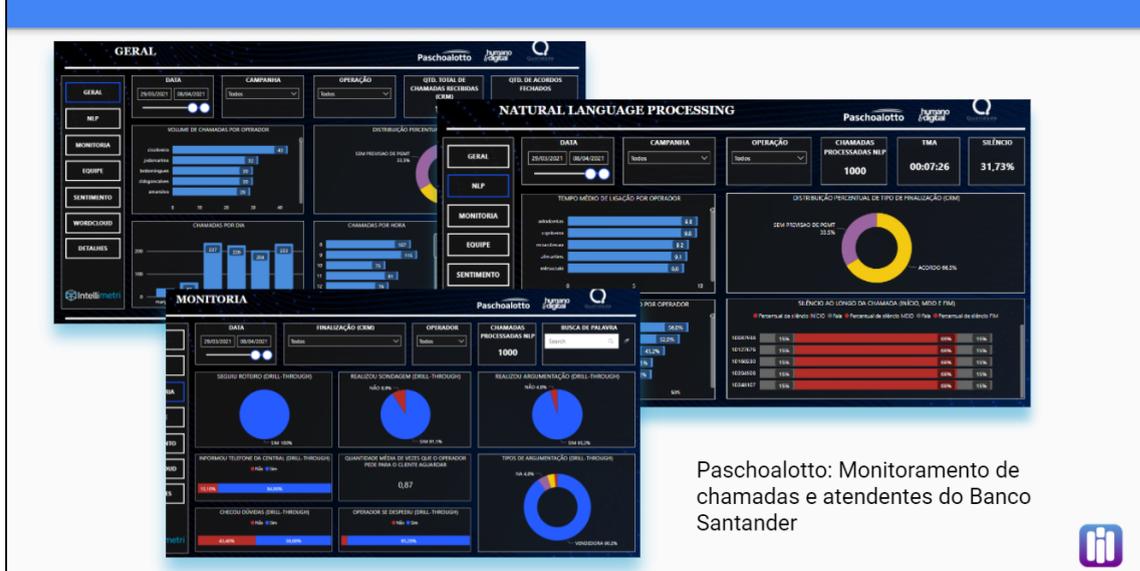
- **[Caso 3] Agência Broadcast - NLP + Extração de Texto + Predição de Texto – PtBr** no desenvolvimento de robô para geração automática de artigos e notícias jornalísticas na área de economia. Projeto desenvolvido em três fases para a agência de notícias financeiras do Grupo Estado:
 - Geração automática de textos simples baseados em modelos pré-definidos a partir de dados capturados de bases de dados não estruturadas;
 - Produção automática de textos mais elaborados baseados em modelos de árvore de decisão e com alteração de templates via Machine Learning para criação de histórias únicas para os vários públicos-alvo;
 - Redação de textos originais por ML, a partir do input de ideias, contextos, frases e dados estatísticos. **Projeto atualmente na fase 1 de desenvolvimento. Previsão de conclusão 11/2022.**
- **[Caso 4] CDF - NLP + Extração de Texto + Classificação + Reconhecimento de Entidades Nomeadas (NER)** em solução para monitoramento da central de atendimento ao cliente com base na análise dos áudios das chamadas, da transcrição de conversas do chatbot, dos dados de CRM e dos sentimentos dos atendentes e dos clientes. Desenvolvimento de dashboards analíticos a partir do cruzamento dos dados analisados. O acesso a esta ferramenta não está aberto para público geral.

Caso de sucesso: CDF - Monitoramento de call center - Speech Analytics



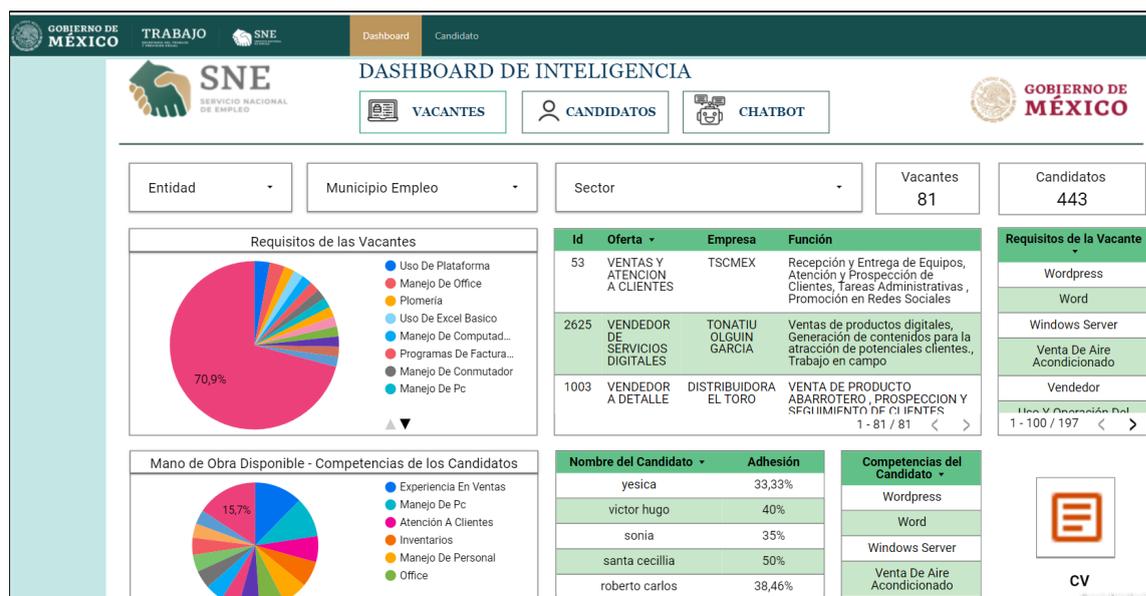
- [Caso 5] [Paschoalotto](#): NLP + Extração de Texto + Classificação + Predição de texto - (PtBr) solução para monitoramento dos atendimentos da central de cobrança na análise dos áudios das chamadas, dos dados de CRM e dos sentimentos dos atendentes e dos clientes. Desenvolvimento de *dashboards* analíticos a partir do cruzamento dos dados analisados. O acesso a esta ferramenta não está aberto para público geral. O Atestado de Capacidade Técnica referente a este projeto encontra-se no item 11.5 do Anexo 3.

Caso de sucesso: Paschoalotto - Monitoramento de chamadas



- [PoC 1] [Virtual Career Center - Servicio Nacional del Empleo \(México\)](#): Extração

de Texto, Entidades e NLP (Processamento de Linguagem Natural) de currículos em PDF e Microsoft Word não estruturados e também algoritmos de classificação e previsão em NLP para *chatbot* de recrutamento e seleção de pessoal. Combinação de qualificações do candidato com requisitos da vaga e indicação de vagas por geolocalização, e indicação de cursos de capacitação para melhorar a empregabilidade do candidato.



- [PoC 2] ISA7 / Via: NLP + Extração de Texto + Classificação em solução para auditoria da equipe de monitoramento da central de atendimento ao cliente da rede de varejo *online* Via, com base na análise dos áudios das chamadas, dos dados de CRM e dos sentimentos dos atendentes e dos clientes. Desenvolvimento de *dashboards* analíticos a partir do cruzamento dos dados analisados. O acesso a esta ferramenta não está aberto para público geral.

Caso de sucesso: Via- Monitoramento de chamadas



Isa7: Monitoramento das chamadas do centro de contato da gigante de ecommerce Via. Desenvolvido como White Label

- [\[PoC 3\] Secretaria da Fazenda \(SEFAZ\) de Salvador](#): Extração de Texto, Entidades e NLP em *chatbot* tributário de atendimento ao cidadão para tirar dúvidas sobre impostos e postos de atendimento.



- [\[PoC 4\] Secretaria da Fazenda \(SEFAZ\) de Salvador](#): Extração de Texto + Classificação no desenvolvimento de ferramenta para identificar não conformidades em documentos fiscais não estruturados (imagens de notas fiscais). Os algoritmos extraem e higienizam o texto da descrição do produto ou serviço, comparando-os com o CNAE da empresa informado na NF. Foram processados 50 milhões de documentos em cinco horas e encontrados 5% de discrepância.

POC Tribunal Superior Eleitoral (TSE) - Prestação de contas eleitorais

TSE: Extração de informações de notas fiscais, recibos e fotos de Google Street View para identificar discrepâncias nas prestações de contas de campanhas eleitorais.

E-mail: `financeiro@bahianoticias.com.br`

DISCRIMINAÇÃO DOS SERVIÇOS
Pesquisa de opinião no Estado da Bahia.
9

Conta corrente para depósito:
Banco Santander (033)
Ag. 4307
C/C 13000324-3
VALOR TOTAL DA NOTA = **R\$25.000,00**

NOTA FISCAL DE SERVIÇOS ELETRÔNICA - Nota Salvador

PREFEITURA MUNICIPAL DO SALVADOR
SECRETARIA MUNICIPAL DA FAZENDA
NOTA FISCAL DE SERVIÇOS ELETRÔNICA - Nota Salvador

PRESTADOR DE SERVIÇOS
CINIS DE
NOME: NOME
P.E. ATIVIDADE: P.E. ATIVIDADE
RUA: RUA
CNPJ: CNPJ

TOMADOR DE SERVIÇOS
CINIS DE
NOME: NOME
P.E. ATIVIDADE: P.E. ATIVIDADE
RUA: RUA
CNPJ: CNPJ

VALOR TOTAL DA NOTA = **R\$25.000,00**

TERMINAL: (240, 179) (240, 181) (225, 181)
[grafica: 10 31]

GRÁFICA Win

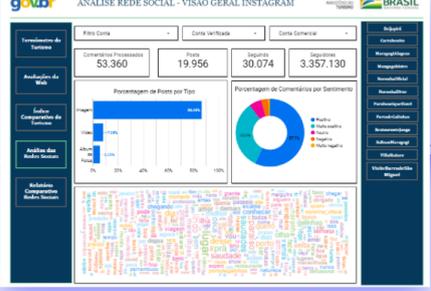
KEROX IMPRESSÃO

KEROX SCANNER

- [POC 7] **Virtual Career Center - Prefeitura de Bogotá (Colômbia):** Extração de Texto, Entidades e NLP (processamento de linguagem natural) em ferramenta para entrevista de candidatos, análise de currículos em PDF e Microsoft Word não estruturados, geração de currículos em PDF e também algoritmos de classificação e previsão em NLP para *chatbot* de recrutamento e seleção de pessoal. Combinação de qualificações do candidato com requisitos da vaga e indicação de vagas por geolocalização, e indicação de cursos de capacitação para melhorar a empregabilidade do candidato. *Dashboards* analíticos de vagas e candidatos. **Etapas do projeto: refinamento de escopo. Entrega estimada: julho/2022.**
- [POC 8] **Ministério do Turismo:** Extração de Texto + Reconhecimento de Entidades Nomeadas (NER) + NLP: Desenvolvimento do Termômetro do Turismo Nacional a partir dos cruzamentos de dados de *Call Detail Record (CDR)*, *Google Trends*, *Google Maps*, *Google Places*, redes sociais e outras bases de dados estruturadas, com *dashboards* analíticos que fornecem, entre outros *insights*, informações sobre infraestrutura e serviços dos destinos turísticos.

Caso de sucesso: Ministério do Turismo - Termômetro do Turismo





Ministério do Turismo: Análise de sentimentos do turista em relação aos destinos

Anexo 2: Qualificação dos Principais Profissionais Envolvidos no Projeto

Os profissionais estão relacionados em ordem alfabética e não correspondem a toda a equipe que participará do projeto. Os currículos estão disponíveis no [Anexo 4](#).

– **Bruno Carvalho Castro Souza - Diretor Técnico**

Bruno Souza é Mestre em Engenharia de Produção pela Universidade Federal de Santa Catarina (2001), Especialista em Inovação e Empreendedorismo pela Universidade para o Desenvolvimento do Estado e da Região do Pantanal (2022), Especialista em Gestão de Negócios pela Fundação Dom Cabral (2010) e Bacharel em Comunicação Social pelo Centro Universitário de Brasília (1993). Possui experiência em gestão de projetos de ciência, tecnologia e inovação (PD&I) em órgãos públicos e privados, de âmbito nacional e internacional; em gestão de instituições de ensino de grande e pequeno porte (vivências como Diretor Executivo, Diretor Acadêmico e Coordenador de Cursos de graduação e pós-graduação); como professor universitário para cursos de graduação e pós-graduação. Bruno é empreendedor, auto-motivado e trabalha em equipe. Atualmente atua como Diretor Técnico do Instituto Modal de Ciência, Tecnologia e Inovação, do qual também é associado fundador.

Currículo Lattes: <http://lattes.cnpq.br/9218729989844596>

– **Emiliana Maria da Costa Coelho Veras - Scrum Master**

Tem 24 anos de experiência na área de Tecnologia da Informação, dos quais 14 são dedicados à gestão de projetos, aplicando os conhecimentos adquiridos e as melhores práticas do PMBOK, Metodologias Ágeis e gestão de serviços. Profissional certificada em PMP, CSM Scrum Master, auditoria CMMi / MPS-BR / ISO 9001 e prática em Pontos de Função. Atuou em funções de desenvolvedora, DBA e gerente de projetos em contratos complexos, técnica e estrategicamente. Dentre as atividades executadas, destacam-se: análise e estimativa de *softwares* em Pontos de Função, Homem/Hora e Unidade Serviço Técnico (UST), identificação e monitoramento de riscos, foco em alcançar os resultados, controle de custo e faturamento, acompanhamento físico/financeiro dos projetos, prospecção de novos projetos, negociação de conflitos, *mentoring* e suporte para melhoria pessoal e profissional do time, perfil de liderança e engajamento das pessoas ao projeto, valorizando o relacionamento interpessoal com todos os envolvidos.

LinkedIn: <https://www.linkedin.com/in/emiliana-veras-gp/>

– **Gerson Alberge Rolim - Diretor Geral**

Sócio e Diretor Geral da Intellimetri, Sócio e diretor da Observatore, diretor da camara-e.net e membro do Conselho Consultivo do PCI-DSS Brasil. Gerson é especialista em Inovação Tecnológica, Economia Digital e Internet das Coisas. Com graduação em processamento de dados e pós-graduação em rede de computadores, atua, desde 1993, na área de Inovação Tecnológica Latino-Americana, É *Top 100 Global Influencer* de transformação digital (<https://bit.ly/3glkTh5>), mentor da Aceleradora Sueco-Noruega de *HealthTechs - Digitalwell Ventures* (<https://bit.ly/3JrT7BD>) - e pioneiro na Internet brasileira, tendo conectado as primeiras redes públicas e privadas do Brasil e participado do desenvolvimento dos primeiros projetos de *eCommerce* no país, em meados da década de 90. É sócio fundador e diretor da 1ª MVNO com foco total em Internet das Coisas (IoT) da América Latina, a Vecto Mobile. Ampla experiência incluindo governo, área acadêmica, multinacionais e organizações não-governamentais e multilaterais como o Mercosul. Carreira nas áreas técnica, comercial e gerencial, sempre com foco em inovação, principalmente na oferta de soluções em Tecnologia da Informação, Comunicação e *eBusiness*.
LinkedIn: <https://www.linkedin.com/in/gersonrolim/>

– **Gisele Ribeiro, Gerente de Desenvolvimento de Negócios**

Jornalista com sólida experiência em mídia impressa e online, Gisele é Bacharel em Comunicação Social - Jornalismo, com especialização em jornalismo científico pela UC Berkeley Extension e pós-graduada em divulgação científica pela USP. Em 35 anos de carreira, trabalhou na cobertura do setor de Tecnologia da Informação em jornais e revistas de circulação nacional, como O Globo, Veja e Folha de S.Paulo. No início da Internet no país, participou da construção do principal portal de conteúdo brasileiro, o UOL, onde foi editora executiva de Tecnologia. Durante três anos trabalhou como gerente de projetos da IMS - Internet Media Services, na Flórida (EUA). Nesta função, coordenou equipes multifuncionais e multilíngues na entrega de projetos de conteúdo para países como Ilhas Bahamas, Curaçao, EUA, Brasil e Argentina. De volta ao Brasil, ela passou a atuar como assessora de imprensa, atendendo contas como Câmara Brasileira de Economia Digital (camara-e.net), Associação Brasileira de Bancos (ABBC) e Fundação Getulio Vargas (FGV). Foi editora das revistas Case Studies e Nós (SCG), até receber convite para integrar a equipe da Intellimetri, onde coordena a parceria da empresa com o Google Cloud e os projetos de inteligência artificial, Machine Learning e Processamento de Linguagem Natural para os setores público e privado, no Brasil e na América Latina. Tem as

seguintes certificações: Google Cloud Sales Credential, Google Education Sales Credential, Maps Sales Fundamentals Credential, Chrome Enterprise and SMB Sales Credential e Google Workspace Sales Credential.

LinkedIn: <https://www.linkedin.com/in/gisele-ribeiro-59bba8/>

– **Gustavo Henrique Gouvea - Cientista de Dados Pleno**

Participou do Projeto de Iniciação Científica Jr. da OBMEP em Matemática durante 2013 presencialmente na Unicamp em Campinas e durante 2016 virtualmente no primeiro semestre. Graduou-se do Ensino Médio na International School of the Americas - TX/US - no primeiro semestre de 2017 e no Colégio Técnico de Limeira no segundo, obtendo diploma de conclusão do Ensino Médio Brasileiro e concluindo o ensino médio Americano. Obteve o título de técnico em Informática pela escola técnica da Unicamp COTIL em 2017. Iniciou sua graduação em 2018 na Universidade de São Paulo - USP - em Matemática Aplicada e Computação Científica. É certificado profissionalmente como engenheiro de dados e líder digital em nuvem pelo Google Cloud e atualmente trabalha como Cientista de Dados Pleno na Intellimetri. Possui experiência nas seguintes áreas: Excel avançado, Power BI avançado, programação Python, PHP, HTML, CSS, JavaScript, MySQL, Banco de Dados, Linux, Programação Orientada a Objetos, VisualBasic, Delphi, C, C++, C#, Manutenção de Redes, Hardware computacional, computação na nuvem, aprendizado de máquina, visão computacional, processamento de linguagem natural, modelos preditivos e classificadores, redes neurais e metodologias ágeis. Currículo Lattes: <http://lattes.cnpq.br/3642058427730049>

– **João Alberto de Oliveira Lima - Analista de Negócio e Cientista de Dados**

Possui graduação em Ciência da Computação pela Universidade Federal da Paraíba (1990), mestrado em Ciência da Computação pela Universidade de Brasília (2000), doutorado em Ciência da Informação pela Universidade de Brasília (2008) com estágio na Universidade de Bolonha (2007) e doutorado em Direito pela Universidade de Brasília (2019). Desde 1995, trabalha como Analista de Informática Legislativa no PRODASEN onde lidera a implementação do Projeto LexML (www.lexml.gov.br).

Currículo Lattes: <http://lattes.cnpq.br/5218258152780642>

– **Lauro César Araujo - Cientista de Dados**

Doutor em Ciência da Informação, especialista em Arquitetura da Informação e Ontologia, com graduação em Ciência da Computação. Possui mestrado e doutorado em Arquitetura da Informação pela Faculdade de Ciência da

Informação da Universidade de Brasília; pós-graduação em Engenharia de Software e graduação em Ciência da Computação. Tem experiência nas áreas de Informação e Computação, atuando especialmente com modelagem conceitual, inovação e modernização de informação legislativa e jurídica, desenvolvimento de arquiteturas de informação e de software, programação em lógica, teoria da informação, processos de software, administração de modelos de dados, gerenciamento de configuração, gerenciamento de conteúdo corporativo e gerenciamento de projetos. Pesquisador de tecnologia da informação em projetos de pesquisa da área de organização da informação jurídica e consolidação normativa em sinergia com pesquisas de direito e saúde na idealização teórica e na modelagem estrutural normativa.

Currículo Lattes: <http://lattes.cnpq.br/6937462274429029>

– **Luis Felipe Borges de Mesis - Cientista de Dados Junior**

Bacharel em Matemática no Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (ICMC - USP), Luís tem interesse na área de Análise Matemática e Ciência de Dados e é membro do grupo de pesquisa de Integração Não Absoluta e Equações Diferenciais Funcionais do ICMC-USP. Atualmente exerce a função de cientista de dados júnior, atuando em projetos ligados à inteligência artificial, processamento de linguagem natural, aprendizado de máquina e business intelligence.

Currículo Lattes: <http://lattes.cnpq.br/3395989727316495>

– **Mamede Lima-Marques - Arquiteto de P&D**

Foi Professor Titular da Universidade de Brasília, Faculdade de Ciência da Informação e foi Professor Titular em Ciência da Computação da Universidade Federal de Uberlândia, Faculdade de Computação. Possui Pós-doutorado em Lógica Matemática Aplicada à Computação, no Centro de Lógica, Epistemologia e História da Ciência, UNICAMP; Ph.D. em Ciência da Computação pela *Université Toulouse III Paul Sabatier*, França; Mestrado em Ciência da Computação pela *Université Toulouse III Paul Sabatier*, França; Mestrado em Ciência da Computação - Instituto de Computação, UNICAMP; Bacharelado e Licenciatura em Filosofia pela Pontifícia Universidade Católica de Campinas e cursou Eletrônica no Centro Federal de Educação Tecnológica de Minas Gerais. Fundador e diretor do Centro de Pesquisa em Arquitetura da Informação - CPAI (FCI/UnB). Atuou no Programa de Pós-graduação em Ciência da Informação (mestrado e doutorado) onde foi fundador da linha de Arquitetura da Informação. Foi professor na *Ecole Nationale de l'Aviation Civile*, Toulouse, França e no *Conservatoire National des Arts et Métiers*, Brive-la-Gaillarde, França. É autor de vários artigos nacionais e internacionais. Atualmente é presidente e

cofundador do Instituto Modal de Ciência, Tecnologia e Inovação.

Currículo Lattes: <http://lattes.cnpq.br/7550414412276509>

– **Mariana Meger - Cientista de Dados Júnior**

Bacharel em Matemática Aplicada e Computação Científica pela Universidade de São Paulo. Mariana domina as linguagens de programação Python, Java e C++. Desenvolve projetos de Inteligência Artificial e Analytics e Aprendizado de Máquina. Na Intellimetri, é a responsável pela criação de *chatbots* e *voicebots* utilizando algoritmos de Processamento de Linguagem Natural. **(Ver currículo no Anexo 4)**

– **Michelle Gandor Diaz - Gerente de Projetos**

Pós-graduanda em gerenciamento de projetos na Fundação Getulio Vargas (FGV) com cursos em metodologias ágeis, Michelle atua há 10 anos como gerente de projetos. Na Intellimetri, é responsável pelo planejamento e pela gestão de projetos de diversos segmentos, fazendo gerenciamento de equipe multidisciplinar com foco em Machine Learning e AI, acompanhamento, suporte e manutenção evolutiva dos projetos, análise de viabilidade técnica de recursos e sistemas e estudo e implantação das soluções de melhorias. **(Ver currículo no Anexo 4)**

– **Paulo Apoloni - Diretor de Tecnologia**

Com graduação em processamento de dados pelo Centro Universitário de Rio Preto (UNIRP), Paulo tem mais de 28 anos de experiência em desenvolvimento, arquitetura e gerenciamento de projetos e é um profissional versátil. Ele participou de projetos em algumas das maiores empresas privadas e instituições governamentais do país e é especialista em metodologias *Scrum* e *RUP*, linguagens C#, Java, Javascript, Python, VB, C++, Apex, X++, PHP; back-end .Net Core, Asp.Net, Node.js, Asp.Net; front-end jQuery, Bootstrap e React; plataformas: Salesforce, MS SharePoint, ERP Dynamics AX; mobile Xamarin, Android e iOS; e SGBDs: SQL Server, Oracle, MySQL, PostgreSQL e MongoDB. **(Ver currículo no Anexo 4)**

– **Romualdo Alves Pereira Junior - Cientista de Dados**

Atualmente é Coordenador de Modernização Institucional da Defensoria Pública da União - DPU, desde de 2020. Coordenador de Análise de Dados e Inteligência da Informação da Presidência da República, de 2017 a 2020. Analista em C&T Sênior do Instituto Nacional de Pesquisas Espaciais - INPE, de 1997 a 2005 e 2017. Chefe da Divisão de Informática e outros cargos na Agência Espacial Brasileira, de 2005 a 2016. Foi Coordenador de

Desenvolvimento de Sistemas e outros cargos no Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, de 1984 a 1997. Pós-Doutor em Ciência da Computação pela University of Ottawa, Canada (2016-2017), Linha de Pesquisa em Data Science (Machine Learning e Textual Analysis). Doutor em Ciência da Informação pela Universidade de Brasília (2013), Linha de Pesquisa em Arquitetura da Informação. Mestrado em Informática pela Universidade Federal da Paraíba - UFPb (Sistemas de Segmento Solo para Satélite Científico), Especialização em Engenharia de Software pela Universidade Católica de Brasília - UCB (idealização do projeto do banco de currículos do CNPq); e Graduação em Tecnologia de Processamento de Dados pela Universidade de Brasília - UnB. Professor Associado do Curso de Pós-Graduação (lato-sensu) em Gestão de Tecnologia da Informação da UnB (de 2007 a 2016), ministrando as disciplinas de "Plataformas e Sistemas de Informação", "Cenários em Tecnologia da Informação", "Ética em Tecnologia da Informação" e "Tópicos Especiais em Gestão de TI". Orientador e membro de bancas examinadoras de dezenas de trabalhos de pós-graduação em Gestão de TI na UnB e UCB. Líder do Grupo de Pesquisa "Arquitetura da Informação para o Programa Espacial Brasileiro", certificado em 2000 pelo Instituto Nacional de Pesquisas Espaciais - INPE. Membro do Grupo de Pesquisa em Crimes Cibernéticos da Polícia Federal, desde 2018. Foi Coordenador de Desenvolvimento de Sistemas do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - 1986 a 1995. No Instituto Nacional de Pesquisas Espaciais (INPE) exerceu atividades de Pesquisa & Desenvolvimento, como Analista em C&T Sênior. Cedido para a Agência Espacial Brasileira em 2005, assumiu a chefia da Divisão de Política Espacial e Investimentos Estratégicos. Em 2009, passou a responder pela Chefia da Divisão de Informática da AEB. Atua nas áreas de Ciência de Dados, Inteligência Artificial, Filosofia da Informação, Fenomenologia, Tecnologia da Informação, Arquitetura da Informação, Inteligência Organizacional, Sistemas de Segmento Solo para Satélites. Consultor de TI, Palestrante em Eventos de TIC e Ciência de Dados. Currículo Lattes: <http://lattes.cnpq.br/6348527995871693>

– **Rubens Zimbres - Cientista de Dados Sênior, PhD**

Cientista de dados e engenheiro de dados profissional certificado do *Google Cloud* (CN-128764) e *Google Developer Expert* (GDE) em *Google Cloud*. Integra o grupo de Cientistas de Dados e desenvolvedores da Intellimetri, onde é responsável pela orientação e instrução da equipe. Rubens trabalha com algoritmos de *Machine Learning*, Pesquisa e Desenvolvimento de soluções de NLP, além de apresentar algoritmos e detalhes de infraestrutura para diretores e executivos *C-level*. Trabalha com *Machine Learning* Supervisionado, Não

Supervisionado, *Deep Learning*, *Transfer Learning*, Processamento de Linguagem Natural nas áreas de telecomunicações, finanças e atendimento ao cliente. Possui experiência em *SPSS*, *Wolfram Mathematica*, R, Python, PySpark ML, SQL. Tem mestrado e doutorado duplos em Administração de Empresas e Engenharia Elétrica pela Universidade Presbiteriana Mackenzie e pós-doutorado pela *Brown University* (Wolfram Research). Sua tese de doutorado gerou patente de um programa de computador no INPI (RS 10950-1) em modelagem de autômatos celulares. Mais tarde, em 2020, outra patente foi gerada na Intellimetri (BR512020002649-4), referente ao algoritmo de Processamento de Linguagem Natural em Python para análise de chamadas de voz em *contact centers*. Possui experiência prática em AWS (*IoT Core*, *EC2*, *Lambda*, *Kinesis*, *Firehose*, *SageMaker*, *Comprehend*, *Forecast*) e Google Cloud (*PubSub*, *Compute Engine*, *Kubernetes*, *Functions*, *Big Query*, *Dataflow*, *Dataproc*, *AI Platform*, *Speech API*). Criou dois *Pull Requests* no GitHub referentes a uma incompatibilidade de *pycocotools* entre Python 2 e 3 e no *hugging face*, devido a uma variável ausente no ambiente Anaconda. É colaborador do *StackOverflow* e *AskUbuntu: Python*, *Keras*, *OpenCV*, *Tensorflow*, *Deep Learning*, *AWS*, *Google Cloud* e *Ubuntu*. Repositório do GitHub em <https://github.com/RubensZimbres>. LinkedIn: <https://www.linkedin.com/in/rubens-zimbres/>

– **Wellington de Souza Evangelista - Diretor Financeiro**

Profissional com sólida carreira desenvolvida na prestação de serviços de tecnologia da informação para seguimento público e privado com predominância nos setores: Elétrico, Transportes, Automação Bancária e Fundos de Pensão. Conhecimentos avançados dos procedimentos de compra de governo, com experiência na elaboração de PDTI e Projetos básicos. Gestão de equipes multifuncionais, compostas por técnicos de diversos níveis e formações, destacando as habilidades de negociação e foco em resultados. Ampla experiência na especificação, implementação e implantação de sistemas corporativos. Especialista em Logística, Transporte e Mobilidade pela Câmara Interamericana de Transportes (2017) e graduado em Ciência da Computação pela Universidade Católica de Brasília (1998).´

Currículo Lattes: <http://lattes.cnpq.br/6264183167526090>

Anexo 3: Atestados de Capacidade Técnica - Intellimetri

Neste anexo estão os documentos que atestam a capacidade técnica da Intellimetri no desenvolvimento de soluções baseadas em Inteligência Artificial e Processamento de Linguagem Natural.

A3.1 Registro no Instituto Nacional de Propriedade Industrial (INPI)



REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DA ECONOMIA
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS INTEGRADOS

Certificado de Registro de Programa de Computador

Processo Nº: **BR512020002649-4**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 01/01/2019, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

Título: Algoritmo de processamento de linguagem natural para análise de chamadas de centrais de atendimento

Data de publicação: 01/01/2019

Data de criação: 01/06/2018

Titular(es): INTELLIMETRI GESTAO DE DADOS LTDA

Autor(es): RUBENS DE ALMEIDA ZIMBRES; GERSON ALBERGE ROLIM

Linguagem: PYTHON; C#

Campo de aplicação: AD-01; AD-05; CO-04; MT-02

Tipo de programa: AT-06; AV-01; CD-01; DS-04; DS-05; IA-01; SM-01; TC-01; UT-01

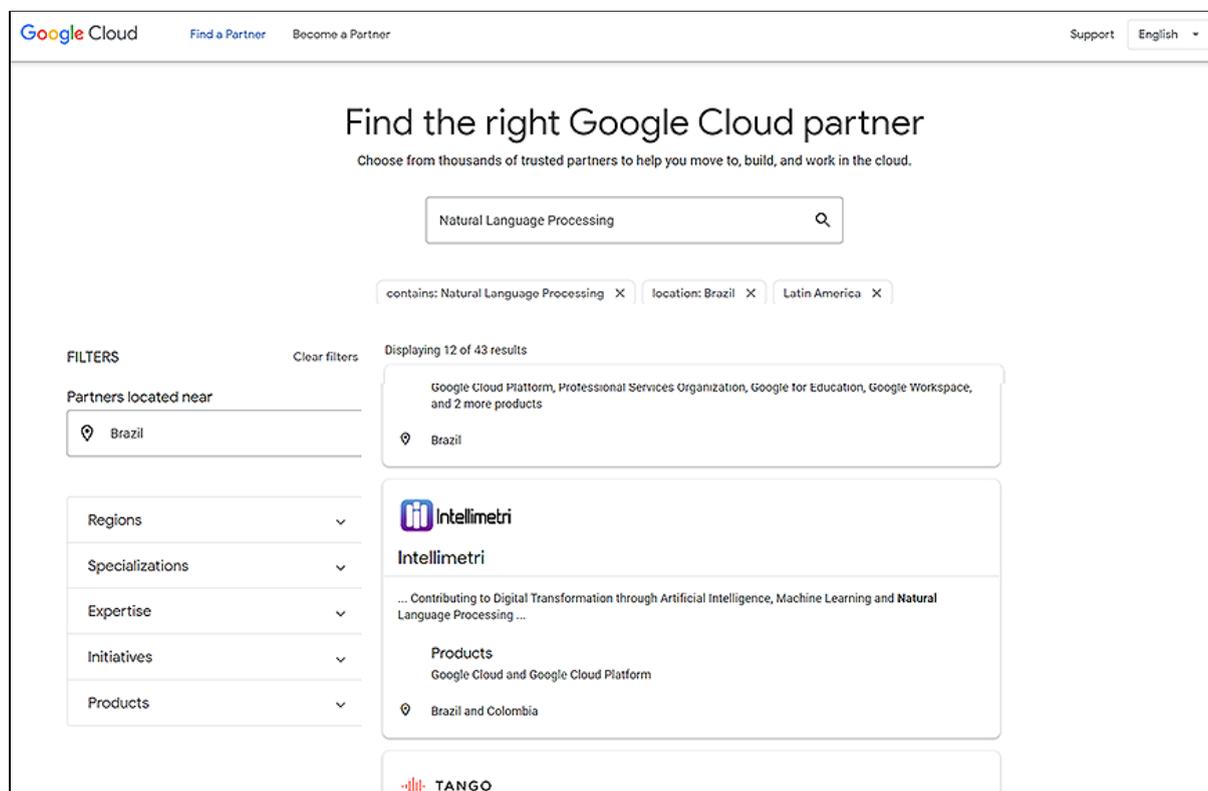
Algoritmo hash: SHA-512

Resumo digital hash:
c172e8d5e10b839f66bb750087879fcd470dacfb440d4c6c0969895ffba02895d24f3fac8a0da9b9dc2b8934fa426498434af4ac5daf0cb27f21386358fa4ad0

Expedido em: 01/12/2020

Aprovado por:
Helmar Alvares
Chefe da DIPTO - Portaria/INPI/DIRPA Nº 09, de 01 de julho de 2019

A3.2 Diretório de parceiros Google Cloud



Google Cloud Find a Partner Become a Partner Support English

Find the right Google Cloud partner

Choose from thousands of trusted partners to help you move to, build, and work in the cloud.

Natural Language Processing

contains: Natural Language Processing × location: Brazil × Latin America ×

FILTERS Clear filters

Displaying 12 of 43 results

Partners located near

📍 Brazil

Regions ▾

Specializations ▾

Expertise ▾

Initiatives ▾

Products ▾

Google Cloud Platform, Professional Services Organization, Google for Education, Google Workspace, and 2 more products

📍 Brazil

Intellimetri

Intellimetri

... Contributing to Digital Transformation through Artificial Intelligence, Machine Learning and Natural Language Processing ...

Products

Google Cloud and Google Cloud Platform

📍 Brazil and Colombia

TANGO

https://cloud.google.com/find-a-partner/?search=Dialogflow®ions=LATAM_REGION

A3.3 Atestado de Capacidade Técnica - Tergos S.A



ATESTADO DE CAPACIDADE TÉCNICA

Atestamos para os devidos fins que a **INTELLIMETRI GESTÃO DE DADOS S.A.**, pessoa jurídica de direito privado com CNPJ 30.046.988/0001-90, prestadora de serviços estabelecida à Av. Doutor Cardoso de Melo, nº 1666, 9º andar, conjunto 91, Vila Olímpia, São Paulo Capital, na categoria de prestadora de serviços de um projeto envolvendo as seguintes áreas de conhecimento: Extração de Texto, Entidades Nomeadas (NER) e NLP (processamento de linguagem natural) de prontuários médicos não estruturados e armazenados em Microsoft Word e também extração de entidades nomeadas e NLP para um Chatbot de acompanhamento terapêutico de pacientes – todos algoritmos treinados em português brasileiro, objeto da Encomenda Tecnológica de um módulo de Instrução Assistida por Inteligência Artificial, a ser incorporado solução de Instrução Assistida do TCU, prestou serviços para esta empresa **TERGOS S.A.** (Bright Photomedicine - www.brightmed.com.br), pessoa jurídica de direito privado com CNPJ 21.389.427/0001-01, na condição de cliente usuária dos serviços acima especificados.

Atestamos ainda, que tais serviços foram executados de acordo com os parâmetros técnicos de qualidade exigidos e no prazo pactuado, não existindo, em nossos registros, até a presente data, fatos que desabonem sua conduta e responsabilidade com as obrigações contratuais.

São Paulo 11 de fevereiro de 2022



TERGOS S.A.

Por: Marcelo Victor Pires de Sousa

Diretor Presidente

A3.4 Atestado de Capacidade Técnica Marco Marketing

DocuSign Envelope ID: 005438EA-6E9D-4800-8237-83C0A2BC2E96



ATESTADO DE CAPACIDADE TÉCNICA

Atestamos para os devidos fins que a INTELLIMETRI GESTÃO DE DADOS S.A., pessoa jurídica de direito privado com CNPJ 30.046.988/0001-90, prestadora de serviços estabelecida à Av. Doutor Cardoso de Melo, nº 1666, 9º andar, conjunto 91, Vila Olímpia, São Paulo Capital, na categoria de prestadora de serviços de um projeto envolvendo as seguintes áreas de conhecimento: Extração de Texto, Entidades e NLP (processamento de linguagem natural) de vídeo currículos e currículos em PDF e Microsoft Word não estruturados e também algoritmos de classificação e previsão em NLP para um Chatbot de recrutamento e seleção de pessoal propriedade de Marco Marketing e Comercialização LTDA (<https://www.hiandy.ai/>), persona jurídica de derecho privado com CNPJ 44.998.961/0001-02 – todos os algoritmos treinados em português brasileiro e espanhol latino, objeto da Encomenda Tecnológica de um módulo de Instrução Assistida por Inteligência Artificial, a ser incorporado solução de Instrução Assistida do TCU.

Atestamos ainda, que tais serviços foram executados de acordo com os parâmetros técnicos de qualidade exigidos e no prazo pactuado, não existindo, em nossos registros, até a presente data, fatos que desabonem sua conduta e responsabilidade com as obrigações contratuais.

São Paulo 11 de fevereiro de 2022

Marco Marketing e Comercialização LTDA.

Por: Maximiliano Gallo
Diretor de Tecnologia

DocuSigned by:
Maximiliano Gallo
988A7A576F29429...

Marco Marketing e Comercialização LTDA.

A3.5 Atestado de Capacidade Técnica ISA7



ATESTADO DE CAPACIDADE TÉCNICA

Atestamos para os devidos fins que a **INTELLIMETRI GESTÃO DE DADOS S.A.**, pessoa jurídica de direito privado com CNPJ 30.046.988/0001-90, prestadora de serviços estabelecida à Av. Doutor Cardoso de Melo, nº 1666, 9º andar, conjunto 91, Vila Olímpia, São Paulo Capital, na categoria de prestadora de serviços de um projeto envolvendo as seguintes áreas de conhecimento: execução bem sucedida de Prova de Conceito (POC) baseada em Extração de Texto de arquivos de áudio de chamadas de atendimento de Call Center, NLP (processamento de linguagem natural), Classificação, Predição de Texto e Desenvolvimento de Dashboards Dinâmicos e Interativos em Google Data Studio e Microsoft Power BI – todos algoritmos treinados em português brasileiro, objeto da Encomenda Tecnológica de um módulo de Instrução Assistida por Inteligência Artificial, a ser incorporado solução de Instrução Assistida do TCU, prestou serviços conjuntamente com a empresa **International Sales Associates Apoio Administrativo LTDA. (ISA7)**, pessoa jurídica de direito privado com CNPJ 33.400.602/0001-95, para a empresa **VIA VAREJO S.A.**, pessoa jurídica de direito privado com CNPJ 33.041.260/0001-64, na condição de cliente usuária dos serviços acima especificados.

Atestamos ainda, que tais serviços foram executados de acordo com os parâmetros técnicos de qualidade exigidos e no prazo pactuado, não existindo, em nossos registros, até a presente data, fatos que desabonem sua conduta e responsabilidade com as obrigações contratuais.

São Paulo 14 de fevereiro de 2022,

INTERNATIONAL SALES
ASSOCIATES APOIO
ADMINISTRATI:33400602000195

Assinado de forma digital por
INTERNATIONAL SALES ASSOCIATES
APOIO ADMINISTRATI:33400602000195
Dados: 2022.02.14 14:18:52 -03'00'

International Sales Associates Apoio Administrativo Ltda (ISA7)

Por: Valter Amorim dos Santos Junior

Diretor de Operações

Av. Eng. Armando de Arruda Pereira, 2937
Bloco B - Sala 103 - São Paulo - SP - CEP 04309-011
Phone: +55 11 99196-2512

www.isa7.net
contact@isa7.net

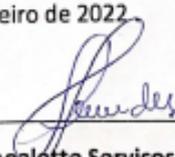
A3.6 Atestado de Capacidade Técnica - Paschoalotto

ATESTADO DE CAPACIDADE TÉCNICA

Atestamos para os devidos fins que a **INTELLIMETRI GESTÃO DE DADOS S.A.**, pessoa jurídica de direito privado com CNPJ 30.046.988/0001-90, prestadora de serviços estabelecida à Av. Doutor Cardoso de Melo, nº 1666, 9º andar, conjunto 91, Vila Olímpia, São Paulo Capital, na categoria de prestadora de serviços de um projeto envolvendo as seguintes áreas de conhecimento: execução bem sucedida de Prova de Conceito (POC) baseada em Extração de Texto de arquivos de áudio de chamadas de atendimento de Call Center, NLP (processamento de linguagem natural), Classificação, Entidades Nomeadas (NER), Predição de Texto e Desenvolvimento de Dashboards Dinâmicos e Interativos em Google Data Studio e Microsoft Power BI – todos algoritmos treinados em português brasileiro, objeto da Encomenda Tecnológica de um módulo de Instrução Assistida por Inteligência Artificial, a ser incorporado solução de Instrução Assistida do TCU, prestou serviços para esta empresa **Paschoalotto Serviços de Call Center Ltda.**, pessoa jurídica de direito privado com CNPJ 22.355.415/0001-10, na condição de cliente usuária dos serviços acima especificados.

Atestamos ainda, que tais serviços foram executados de acordo com os parâmetros técnicos de qualidade exigidos e no prazo pactuado, não existindo, em nossos registros, até a presente data, fatos que desabonem sua conduta e responsabilidade com as obrigações contratuais.

São Paulo 14 de fevereiro de 2022.



Paschoalotto Serviços de Call Center Ltda.
Por: Isabel Fernandes
Gerente de Qualidade

Anexo 4: Currículos dos Principais Profissionais

Neste anexo estão os currículos dos principais profissionais que atuarão no projeto.