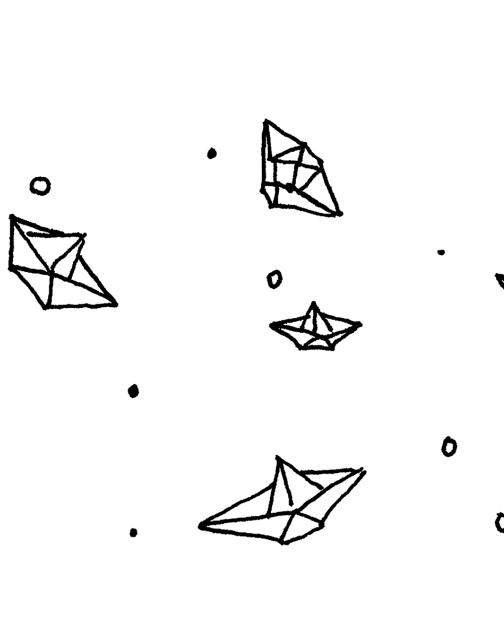




Descodificar a Legislação: o caso de uso do Estatuto de Aposentação



Aposentação

Proposta de Prestação de Serviços | Centro de Conhecimento em Data Science

Agenda

- A. O nosso entendimento 3
- B. A nossa proposta 9
- C. NA

A. O nosso entendimento





Dada a sua maturidade enquanto produto, o DRE torna-se um relevante objeto de interesse na aplicação de técnicas inovadoras

O potencial do DRE

"O acesso ao DRE permite a consulta de toda a Legislação Portuguesa, tendo ainda um grande potencial de evolução ao nível de serviço ao Cidadão, em particular, em serviços que permitam ao utilizador encontrar as normas e os procedimentos que dizem respeito à sua necessidade em particular (o motivo que o levou a consultar a legislação)."

A oportunidade e o desafio

Para a clarificação de um determinado tema por parte do utilizador, muitas vezes é necessário consultar diferentes diplomas que apresentam uma necessidade de interpretação técnica e detalhada.

"Como tal, o DRE apresenta um desafio constante de se manter em permanente monitorização, análise e evolução de melhorias que tornem a plataforma mais ágil e a solução mais assertiva no acesso e consulta do utilizar ao conteúdo pretendido."

Fonte: Memória Descritiva, Sistema de Apoio à Transformação Digital da Administração Pública (SATDAP), partilhado pela INCM com a Nova SBE.



Tirando partido do potencial e desafios do DRE, a INCM pretende apostar na investigação e experimentação focadas no serviço ao cidadão

Desta forma, os objetivos da INCM passam por:



Encontrar uma solução para as limitações comumente identificadas pelos utilizadores do DRE, a nível de articulação e agregação da informação, proporcionando uma visão única e consolidada (limitações agrupadas em 4 eixos)



1

Garantir a relevância do DRE enquanto instrumento útil e ágil para os Cidadãos, dotado das melhores técnicas ao nível da usabilidade em sistemas de informação

As limitações identificadas podem ser agrupadas em 4 eixos:

- 1. A articulação de artigos dentro de um mesmo documento legislativo;
- 2. A articulação de artigos dentro de um mesmo documento legislativo e a remissão para outros;
- 3. A forma de cálculo de um dado imposto, taxa, coima, multa, subsídio, incentivo, entre outros atributos, articulando vários documentos legislativos;
- 4. A catalogação de todos os elementos legislativos referente a um dado tema, artigo, procedimento, sejam eles de natureza nacional ou comunitária.

Fonte: Memória Descritiva, Sistema de Apoio à Transformação Digital da Administração Pública (SATDAP), partilhado pela INCM com a Nova SBE.



Nesta sequência, e em parceria com outras entidades do sistema de I&D, a INCM colocou em curso o caso de uso do Direito do Consumidor

Assim, a INCM iniciou um projeto de 18 meses, com as seguintes fases e entregáveis:

Investigação e Análise do Estado da Arte

Criação e Preparação de Conjuntos de Treino

Criação e Treino de Modelos

Desenvolvimento de Provas de Conceito

4

- Pesquisa dos melhores trabalhos desenvolvidos ou em desenvolvimento no que toca ao processamento de linguagem natural aplicado a problemas de pesquisa de informação e análise semântica.
- Criação de um dataset constituído por documentos da Legislação Portuguesa que permitam treinar os modelos de processamento de linguagem natural a desenvolver na Fase 3.
- Modelo capaz de extrair elementos do dataset da Legislação Portuguesa, e apresentar o seu resumo em Língua Natural.
- Desenvolvimento de provas de conceito que, nesta tarefa, não serão ainda integrados com o DRE.
- Desenvolvimento de uma prova de conceito, com uma interface web para perguntas sobre um documento específico da Lei dos Direitos do Consumidor.

Entregável 1 (E1): Relatório de estado da arte em processamento de linguagem natural no formato de artigo científico.

Entregável 2 (E2): Dataset processado com exemplos anotados para as tarefas de aprendizagem supervisionada ou semi-supervisionada de algoritmos de processamento de linguagem natural

Entregável 3 (E3): Interface Web perguntas sobre um documento da Lei dos Direitos do Consumidor.

Entregável 4 (E4): Código fonte dos modelos de processamento de linguagem natural criados.

Fonte: Documento "Projeto Descodificar a Legislação – Macroplaneamento", partilhado pela INCM com a Nova SBE.



A equipa encontra-se a desenvolver as tarefas previstas nas Fases 1 e 2, de anotações lógicas com base na interpretação das leis em causa

Assim, a INCM iniciou um projeto de 18 meses, com as seguintes fases e entregáveis:

Desenvolvimento e Integração de Protótipo Funcional

...

Testes e Refinamento de Protótipo Funcional

- Desenvolvimento de uma ferramenta de interface que permita receber como input texto escrito em língua portuguesa e retorne uma resposta baseada numa interpretação da Legislação Portuguesa.
- Elaboração e execução de um conjunto de testes que permita aferir a exatidão dos resultados obtidos com base nos algoritmos de processamento de linguagem natural desenvolvidos.

Entregável 5 (E5): 1o Protótipo -Modelos NLP para a Lei dos Direitos do Consumidor Entregável 6 (E6): 2o Protótipo -Modelos NLP aplicados a outras leis fora da Lei do Direito **Entregável 7 (E7):** Versão final dos algoritmos e código fonte. Relatório final do projeto.

Fonte: Documento "Projeto Descodificar a Legislação – Macroplaneamento", partilhado pela INCM com a Nova SBE.

do Consumidor



A Nova SBE, através do seu Data Science Knowledge Center, surge como o parceiro ideal de investigação e experimentação neste âmbito

Após diversas reuniões e partilhas de informação entre ambas as entidades, o entendimento da equipa relativamente aos objetivos para o projeto são os seguintes:

Gerais

- A. Apoiar o desenvolvimento do projeto em curso "Descodificar a Legislação" através de uma **rede de experts em Ciência de Dados e Inteligência Artificial**
- B. Testar um modelo ágil, eficiente e de baixo custo, onde as **equipas de alunos da Nova SBE**, nomeadamente no que concerne ao seu novo programa de Mestrado em Análise de Negócio (*Business Analytics*), **desenvolvem um projeto de Ciência de Dados com a INCM** (18 meses)

Específicos

C. Desenvolver um novo caso de uso para servir como prova de conceito no projeto "Descodificar a Legislação", **no âmbito do Estatuto de Aposentação** (de acordo com as prioridades partilhadas pela INCM a este nível)

B. A nossa proposta





A Nova SBE propõe dividir a execução do projeto em 5 macro fases, desde a análise do estado de arte, à divulgação de resultados

Principais Fases do Projeto



Análise do Estado de Arte e escolha da Framework de trabalho

Análise e Tradução da Legislação das Aposentações para linguagem máquina

Desenvolvimento de uma ferramenta de conversação

Avaliação dos Resultados e Afinação do Modelo

Investigação e Divulgação dos Resultados

- Recolha dos elementos mais relevantes da bibliografia relacionados com Processamento da Língua Natural (NLP) e Entendimento de Língua Natural (NLU) no contexto jurídico;
- Comparação e síntese das abordagens mais relevantes neste contexto;
- Escolha da abordagem a seguir;
- Workshop de transferência de conhecimento (INCM > DSKC) e preparação da Fase 2.

- Análise e entendimento compreensivo da legislação relacionada com o caso de uso selecionado;
- Desenvolvimento de uma ontologia que se enquadre e compreenda toda a legislação das aposentações;
- Criação de uma metodologia consistente para a tradução de linguagem natural de foro jurídico em linguagem lógica representada por modelos matemáticos e assente na ontologia previamente definida.

- Desenvolvimento dos fluxos de conversa mais relevantes no contexto selecionado;
- Implementação dos fluxos de conversa segundo as regras derivadas anteriormente;
- Desenvolvimento de mecanismo de deteção de alterações na legislação¹.

- Análise de resultados contra painel de avaliadores da qualidade da ferramenta desenvolvida;
- Cíclo de melhoria considerando o feedback gerado anteriormente.
- Desenvolvimento de investigação com base no trabalho realizado e experimentado;
- Realização de evento para partilha e discussão de resultados.

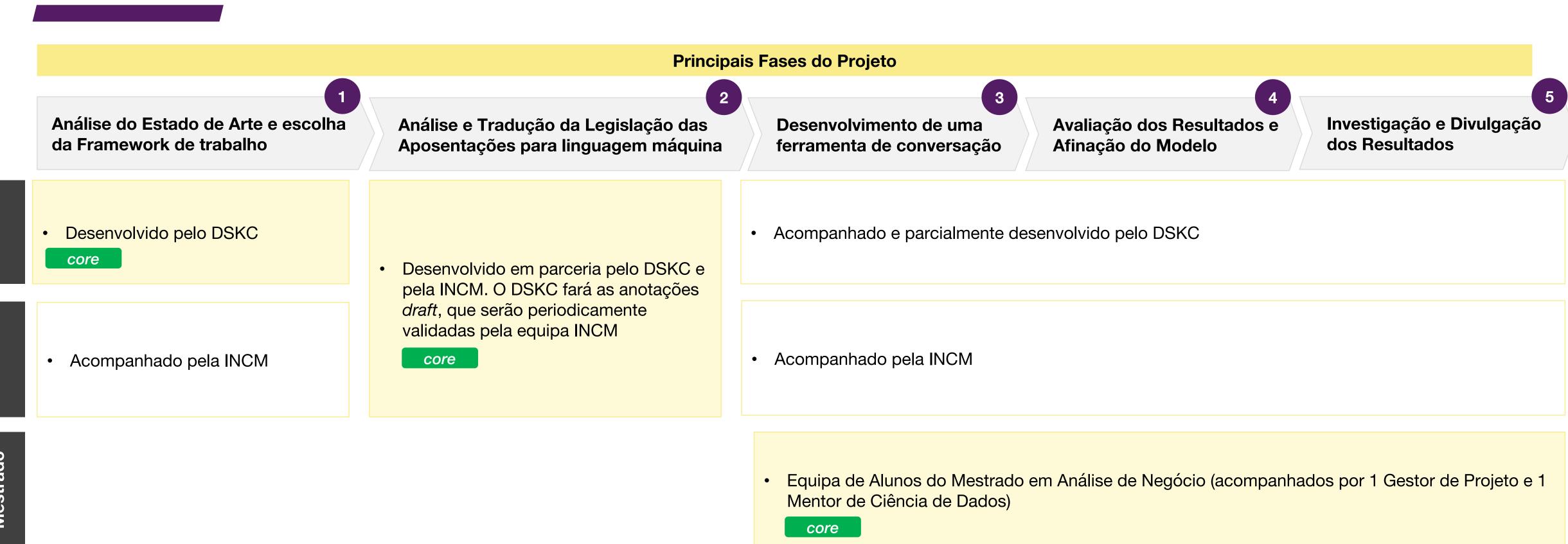
- Documento sumário da pesquisa bibliográfica realizada, tal como o comparativo de diferentes abordagens ao problema;
- Recomendação de metodologia a adotar.
- Documento com a representação lógica das Leis da Aposentação;
- Anotações e observações relativas à análise das Leis da Aposentação.
- Protótipo de uma ferramenta conversacional baseada na lei das aposentações.
- Avaliação por painel da ferramenta desenvolvida.
- Ferramenta final de conversação.

- Teses de Mestrado;
- Outras peças de investigação a definir (ex.: artigos, papers, etc.);
- Evento de apresentação de resultados.





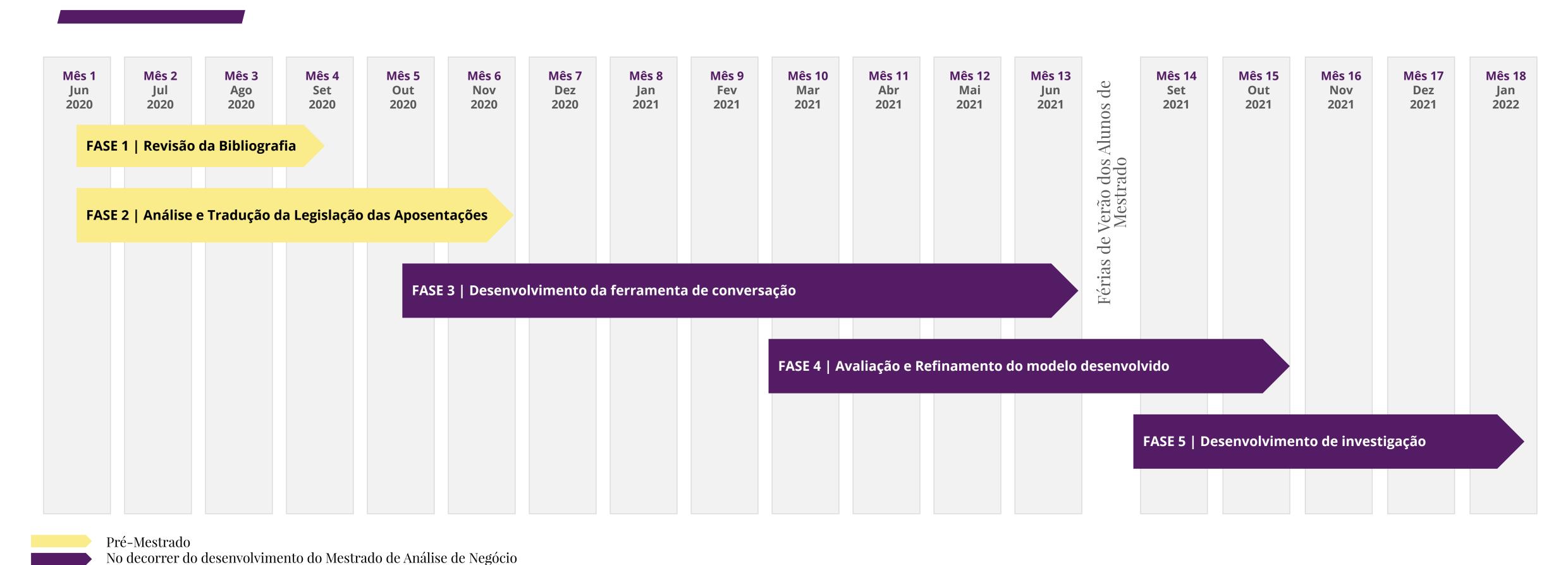
Tendo em conta as diversas competências necessárias, sugere-se uma equipa multidisciplinar entre a Nova SBE, a INCM e atuais parceiros



NCM



O calendário estimado prevê 18 meses de projeto, incluindo uma fase de preparação antes do início do trabalho dos alunos

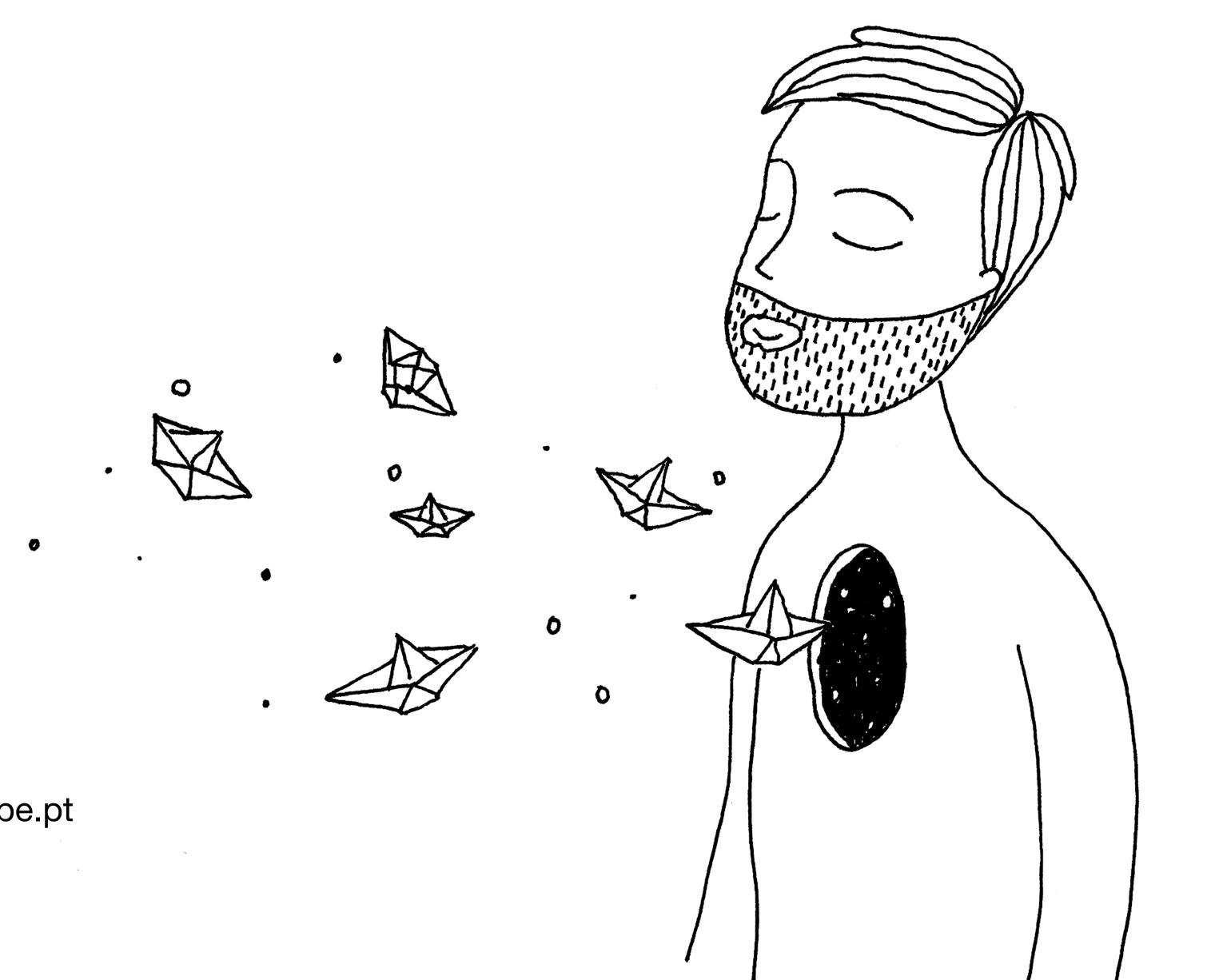






De forma a garantir o sucesso do projeto, é relevante identificar os riscos associados e discutir a sua possível mitigação

- O Desenvolvimento atempado das Fases 1 e 2, de forma a garantir os *input*s essenciais ao início do trabalho da equipa de alunos
- O Limitações da aplicação dos modelos conversacionais existentes na Língua Portuguesa
- Limitações da aplicação dos modelos conversacionais existentes no domínio jurídico
- O Ambiguidade inerente à legislação e a sua subjetividade
- O Subjetividade da avaliação da qualidade dos modelos linguísticos
- O Definição de métricas de avaliação da qualidade das respostas (certo/ errado)



Lénia Mestrinho lenia.mestrinho@novasbe.pt +351 96 210 42 28

