

### Descrição da solução nos termos no subitem 5.3.3

#### 1. Fase 1 – Estruturação e carga da base de dados

a. Será estruturada uma nova base de dados única, com estrutura normalizada para receber os dados extraídos através de ferramenta de ETL (Extração, transformação e carregamento).

b. A ferramenta de ETL irá possibilitar conexão com diversos tipos de base de dados para extrair os textos, links, arquivos binários e demais informações necessárias para análise na fase 2.

c. Durante a extração dos dados, será executado processo OCR (reconhecimento de caractere óptico) para extrair o texto dos arquivos identificados.

d. Após carga dos dados, será realizado trabalho de indexação na nova base de dados para extrair palavras que constam nos textos.

#### 2. Fase 2 – Desenvolvimento do sistema web

a. Será desenvolvida aplicação web responsiva.

b. A aplicação deverá conter formulário de pesquisa com filtros que possibilite identificar dados específicos ou simplesmente uma pesquisa genérica por palavras chaves.

i. O sistema irá retornar uma lista com bloco de informações que contenham os dados solicitados na pesquisa, sendo possível visualizar todas as informações da linha e realizar o download do arquivo original