|  |
| --- |
|  |
|  |
| Cronograma |
| Projeto BigData |

# Sumário

[Sumário 2](#_Toc523325812)

[Controle de Versão 3](#_Toc523325813)

[Necessidades 4](#_Toc523325814)

[Cronograma 5](#_Toc523325815)

[Sugestões Próximas Atividades 8](#_Toc523325816)

# Controle de Versão

|  |  |  |  |
| --- | --- | --- | --- |
| Data | Responsável | Alteração | Versão |
| 29.08.2018 | Nome completo | Versão inicial (não contem alterações). | 1.0 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

As informações contidas neste documento são CONFIDENCIAIS, protegidas pelo sigilo legal e por direitos autorais. A divulgação, distribuição, reprodução ou qualquer forma de utilização do teor deste documento depende de autorização do emissor, sujeitando-se o infrator às sanções legais. Caso este documento tenha sido recebida por engano, favor avisar imediatamente o emissor.

# Necessidades

|  |  |
| --- | --- |
| 1 | Definir tecnologia do ETL |
| 2 | Definir ferramenta(s) e/ou tecnologia de carga e transformação da informação.  Em analise para discussão:  Data Cleaner – utilizar os processos existentes de data cleansing e enriquecimento.  Python – replicar os processos existentes do DC para Autor, Afiliação (Niveis 1, 2 e 3) e País de Afiliação. |

# Cronograma

|  |  |
| --- | --- |
| 1 | DBA e Bireme  Reunião de análise dos campos existentes nos arquivos CSVs, identificação da estrutura de origens dos dados do FI-Admin e analise do XML gerado para o Lucene.  Reunião realizada em 23.08.2018.  Atividade concluída. |
| 2 | DBA  Analise da estrutura de dados da origem.  16 horas. |
| 3 | DBA  Modelagem de dados do ETL.  40 horas. |
| 4 | Bireme  Aprovação item 3. |
| 5 | Bireme  Liberação da maquinas para implementação do ETL.  Montar 3 máquinas ETL1, ETL2 e ETL3 (1GB RAM e 32GB de HD) |
| 6 | DBA  Implementação do software de banco de dados.  16 horas. |
| 7 | DBA  Criação dos bancos de dados.  8 horas. |
| 8 | DBA e Bireme  Definição dos processos de carga.  4 horas. |
| 9 | DBA e Bireme  Definição dos processos de transformação.  4 horas. |
| 10 | DBA  Confecção dos processos de carga.  Estimar horas depois de definido item 8. |
| 11 | DBA  Confecção dos processos de transformação.  Estimar horas depois de definido item 9 |
| 12 | DBA  Testes de desenvolvimento.  Estimar horas depois de definido os itens 10 e 11. |
| 13 | DBA  Implementação dos processos de carga e transformação.  16 horas. |
| 14 | DBA e Bireme  Rodar a primeira carga. |
| 15 | DBA e Bireme  Analise dos dados da primeira carga |
| 16 | DBA e Bireme  Reunião de atividades realizadas referente aos itens 1 a 15, fazer aqui um Status Report e/ou discussão de historio e lições aprendidas. |
| 17 | DBA e Bireme  Reunião de planejamento das próximas atividades. |

# Sugestões Próximas Atividades

|  |  |
| --- | --- |
| 1 | Definir geração do XML para o Lucene. |
| 2 | Definir ferramenta de análise e visualização de informações. |
| 3 | Definir processos de análise e visualização. |
| 4 | Definir governança de sustentação do ambiente. |