




THE DEVELOPER'S CONFERENCE

**Um estudo sobre implementação de ILM
(Information Lifecycle Management) no Sicredi**
Trilha Arquitetura de Dados

Patricia Nunes Gonçalves



PHD Engenharia da Computação – Universidade de Lisboa
Arquitetura Corporativa  **Sicredi**

Agenda



- Introdução ILM
- Dicas para implementar ILM
- Entrevistas realizadas no Sicredi
- Oportunidade e Ações identificadas para ILM
- Cenários para armazenamento dados histórico



A primeira instituição financeira cooperativa do Brasil

Em um mundo cada vez mais colaborativo, queremos unir pessoas com os mesmos propósitos. Por isso, há mais de 116 anos, acreditamos que é possível lidar com o dinheiro de uma maneira diferente.

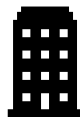
Saber mais:
www.sicredi.com.br



4 milhões de Clientes



R\$ 2,7 bilhões em resultados



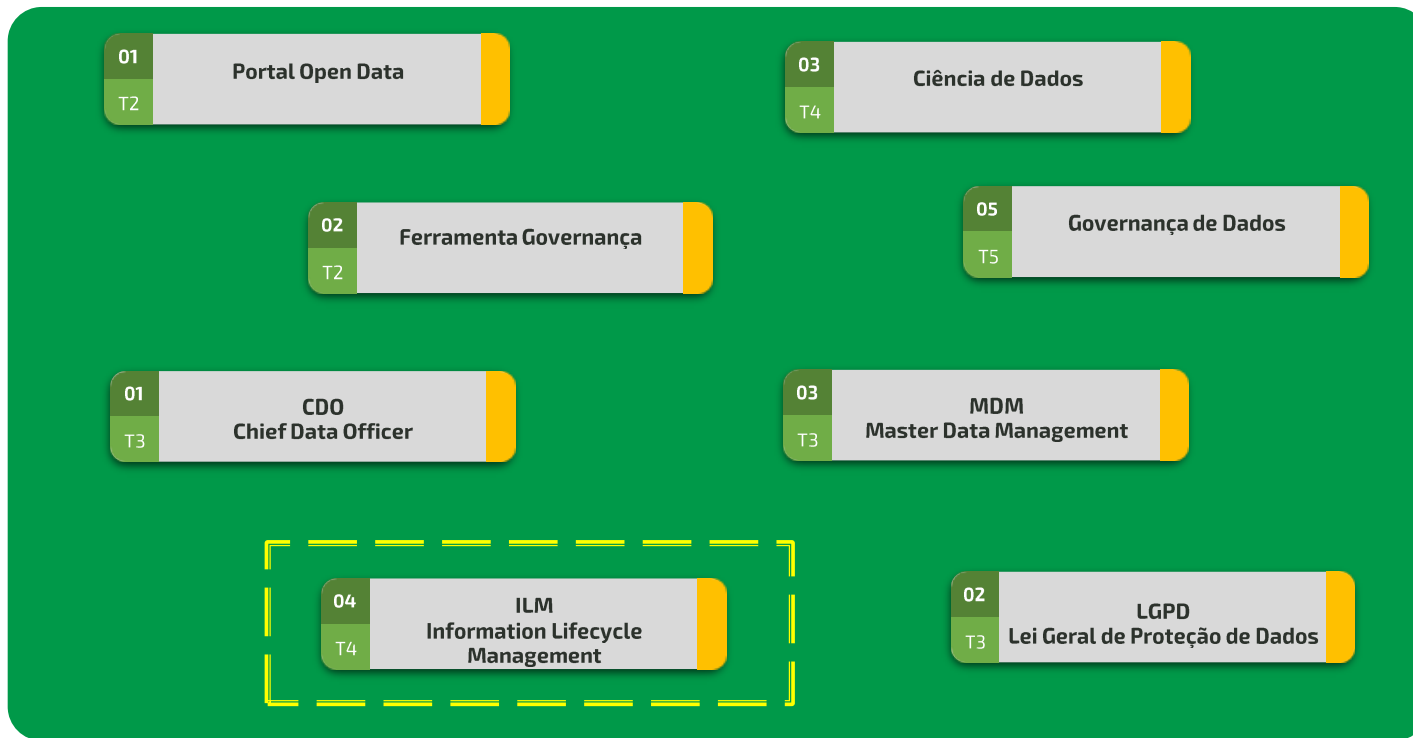
R\$ 15 bilhões patrimônio líquido

Projeto Gestão de Dados

Projeto Gestão de Dados é baseado na Arquitetura Alvo de Dados e evoluções do tema.



THE
DEVELOPER'S
CONFERENCE

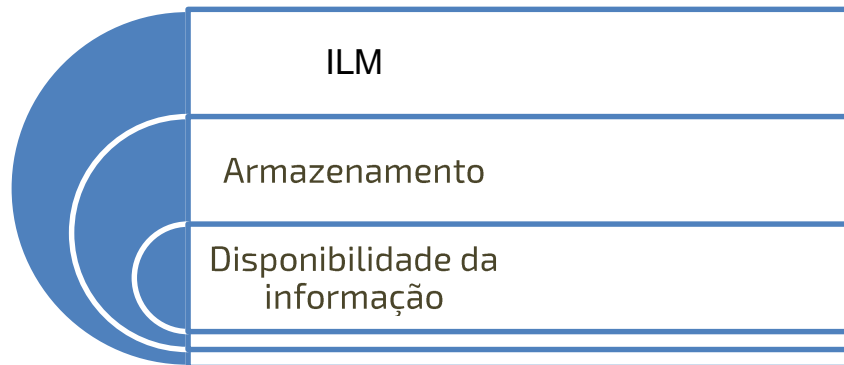


O que é ILM?



O ILM (Information Lifecycle Management) consiste na gestão racional do patrimônio de **informação** da empresa, em função do **valor** da informação e o **custo** do seu armazenamento.

ILM : onde informação deve ter em conta as exigências técnicas, **regulamentares** e **jurídicas** mais adaptadas para *armazenar* e tornar a *informação disponível*;





THE
DEVELOPER'S
CONFERENCE

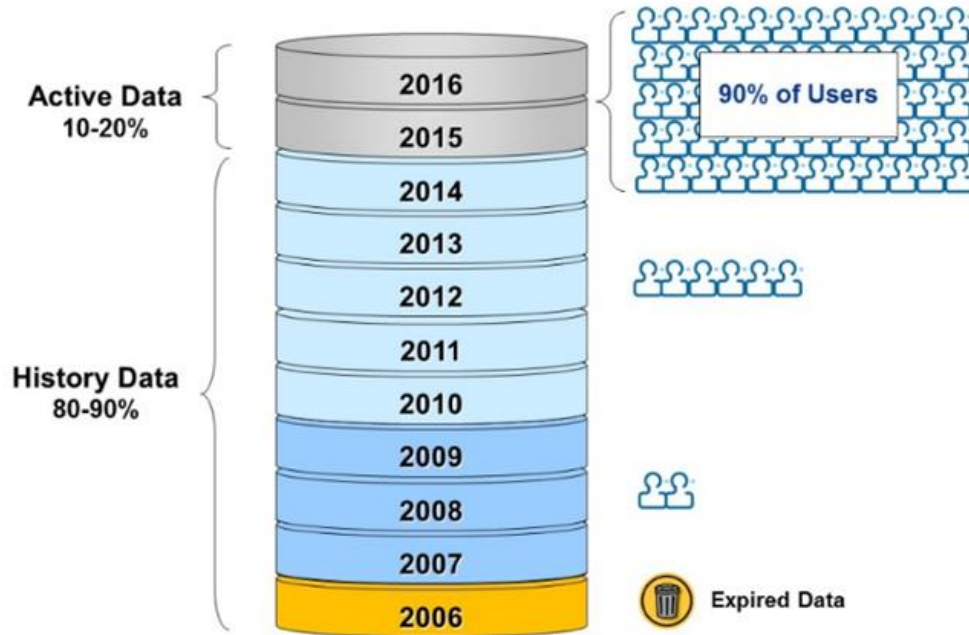
O que dizem os especialistas sobre ILM?



Qual é o valor do Dado?



Analistas do setor apontam que até 80% dos dados não são atuais e o valor dos dados diminui drasticamente ao longo do tempo. O ILM classifica os dados na criação e move os dados pelas camadas da infraestruturas com base nas regras de negócios e nas políticas de retenção.



Armazenamento desnecessário de informação





A implementação de ILM passa por várias fases:

- ❑ Análise da infraestrutura de armazenamento existente;
- ❑ Classificação da informação;
- ❑ Criação de camadas de armazenamento para as diferentes classes de dados;
- ❑ Definição de políticas de armazenamento;

Podemos usar o Data Warehouse como base histórica?

Toda empresa toma decisões com base em dados históricos específicos de um período de tempo.

Determinando a taxa na qual os dados são atualizados nesses sistemas, a data de validade dos dados no Data Warehouse as estratégias de manipulação de dados expirados no data warehouse são coletivamente conhecidas como ILM (Information Life Cycle Management Strategy) para Data Warehouse.

Não! No! Nyet!



Quem deve liderar a estratégia de Dados?

À medida que as salas de reuniões em todo o mundo lidam com a definição de um futuro orientado a dados para seus negócios, há um perigo real em acreditar que a necessidade de mudança está fora da sala.

A resposta está na cultura de dados!

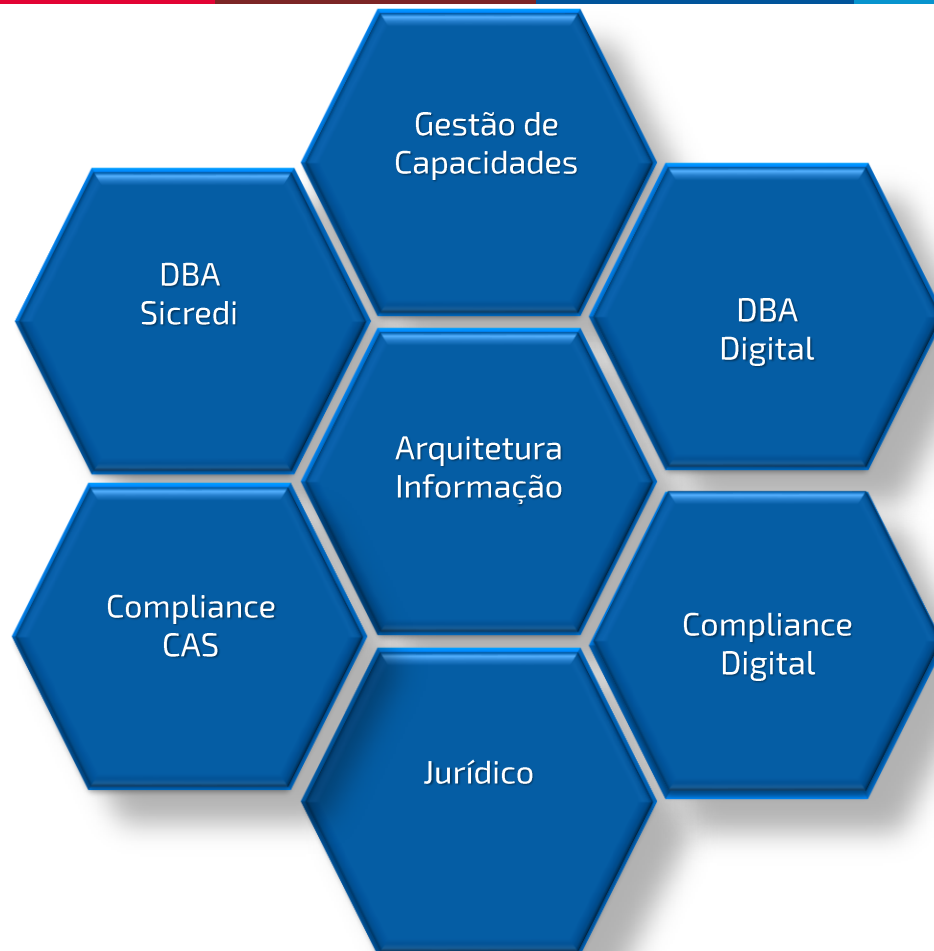
-DATA DRIVEN-

O que diz a empresa sobre ILM?





THE
DEVELOPER'S
CONFERENCE



Foram 7 Entrevistas

Num total de 65 itens entre Contexto e Dificuldades

Classificação da informação: Uso Irrestrito

Dificuldades e Limitadores

Levantamento de informações de em entrevistas realizadas com áreas do Sicredi.



THE
DEVELOPER'S
CONFERENCE

Dificuldades e Limitadores

Papel e
Responsabilidades

Ferramenta

Desenvolver
Cultura

Estratégia

LGPD

Expurgo
Dados

Armazenamento
Dados Histórico

Segurança

Disponibilização
Informações

Qualidade
Dados

Dificuldades e Limitadores

Levantamento de informações de em entrevistas realizadas com áreas do Sicredi.



Direcionamentos		Entrevistas
1	Definir papéis e responsabilidades sobre dados engajando áreas de Negócio e Sistemas	Papel e Responsabilidades
2	Definir processos claros para expurgo de dados	Papel e Responsabilidades
3	Levantar custos de armazenamento de dados	Armazenamento Dados Histórico
4	Definir e divulgar as regras do ciclo de vida do dado apoiando as áreas de negócio e apoiando a frente LGPD	Expurgo Dados LGPD
5	Identificar os casos críticos e montar um plano de ação unindo Infraestrutura, Gestão de Capacidades, Área de Negócio e Sistemas.	Expurgo Dados
6	Fomentar a cultura de dados no Sicredi	Desenvolver Cultura

Dificuldades e Limitadores

Levantamento de informações de em entrevistas realizadas com áreas do Sicredi.

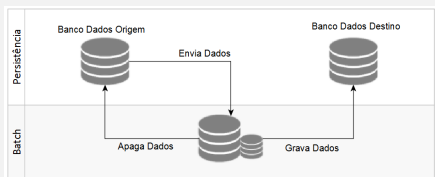


THE
DEVELOPER'S
CONFERENCE

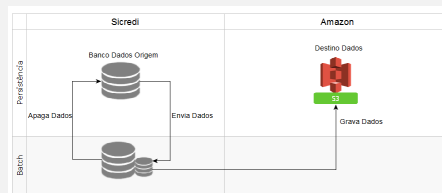
Direcionamentos		Entrevistas	
6	Implementar a Ferramenta de Governança de dados <ul style="list-style-type: none">• Inventário de dados• Regras de Ciclo de Vida• Linhagem de dados	Qualidade Dados Ferramenta Qualidade	
7	Realizar um Benchmark em outras empresas que implementam ILM nos apoiando dos parceiros.	Desenvolver Cultura	
8	Colaborar com a área de Gestão de Capacidades para divulgar indicadores referente ao armazenamento dos dados	Disponibilização Informações	
9	Desenvolver a maturidade sobre dados nos times de Desenvolvimento	Desenvolver Cultura	
10	Promover, juntamente com a frente da LGPD, o engajamento dos executivos e áreas de negócio	Estratégia LGPD	
11	Apoiar times de sistemas e áreas de negócio para gestão eficiente do armazenamento de dados históricos	Armazenamento Histórico	Dados

Cenários de Solução

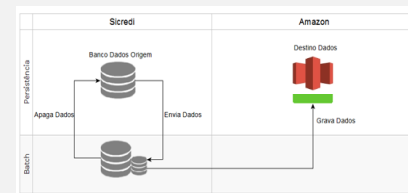
A Base Histórica On-promise



B S3 - Amazon



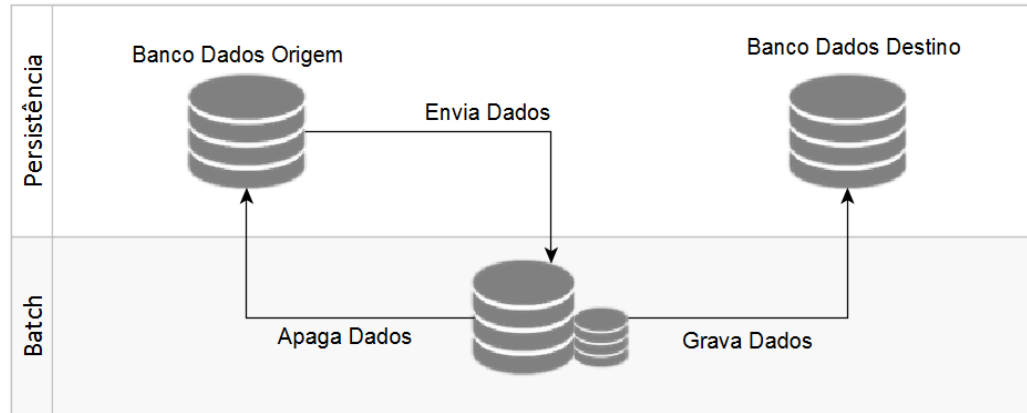
C Glacier Amazon





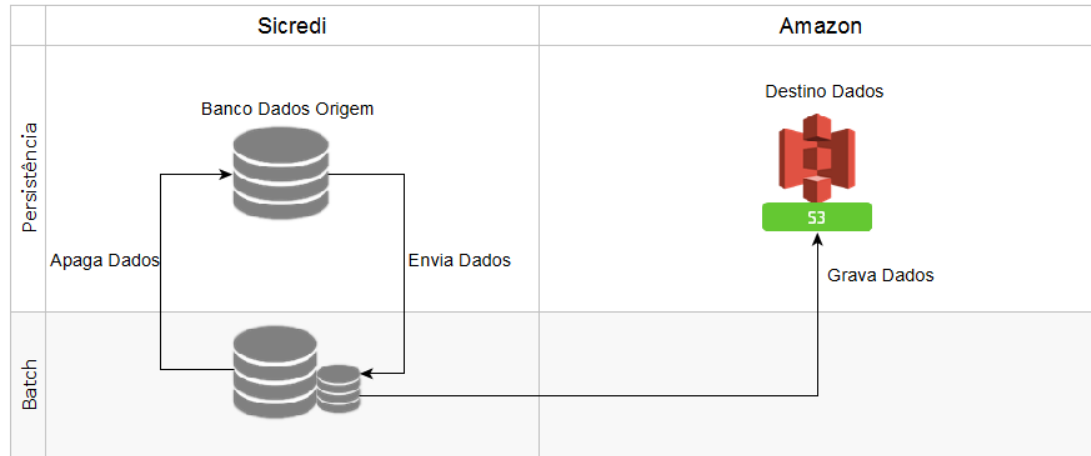
A

Base Histórica
On-premise



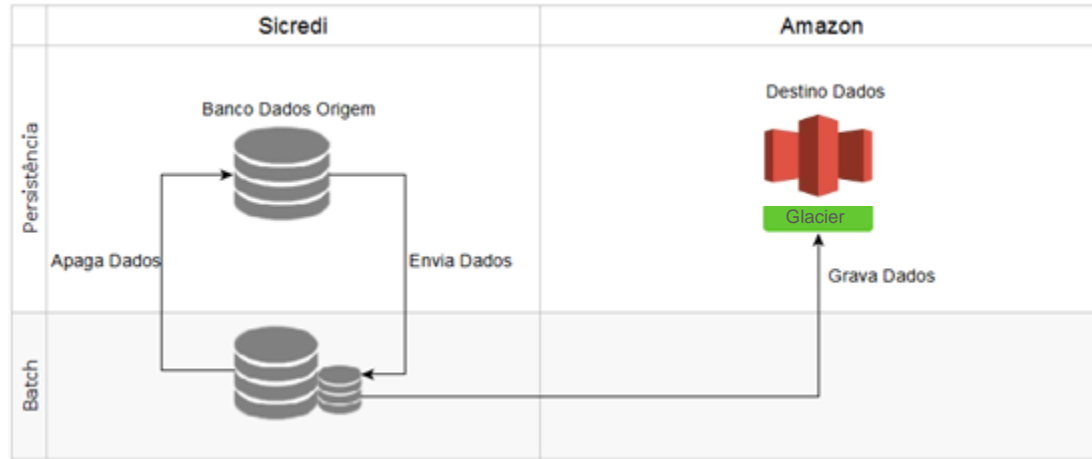


B S3 - Amazon

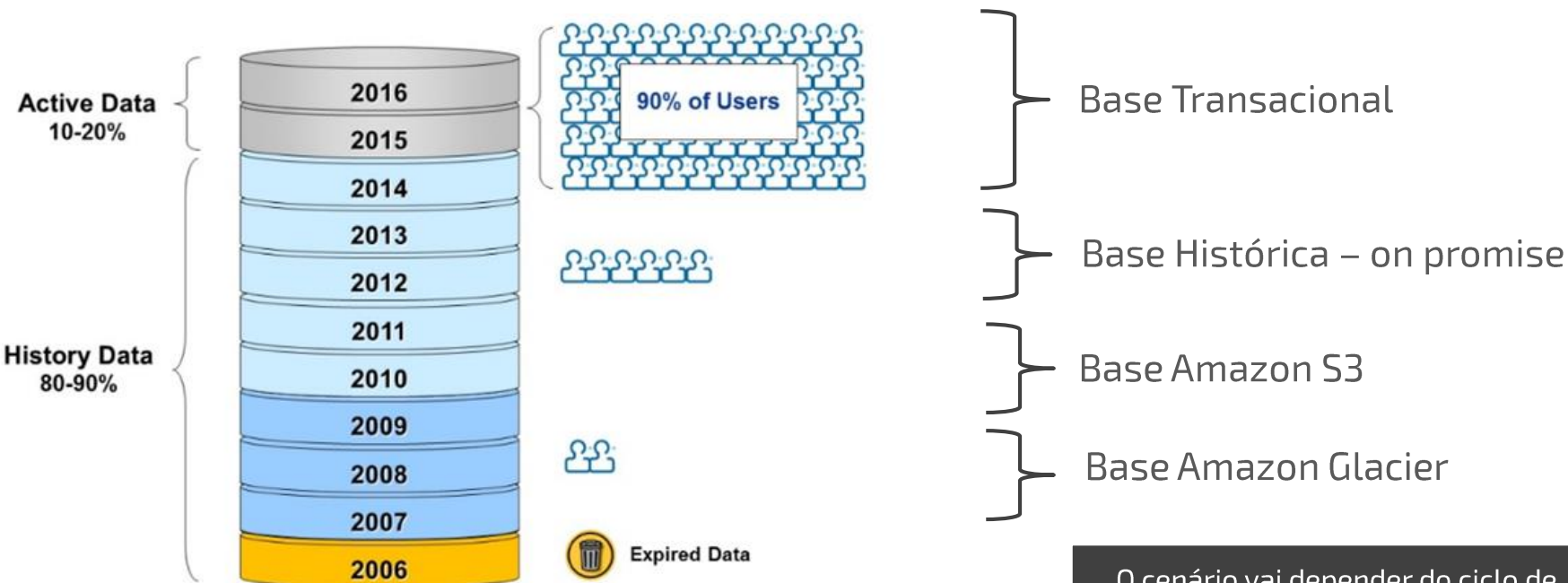




C *Glacier Amazon*



Não há cenário ideal – depende do valor da informação....



O cenário vai depender do ciclo de vida do dado!



Patricia Nunes Gonçalves

Sicredi

QUARTA | 27/11

Trilha Inteligência Artificial

O mundo está mudando e a inteligência artificial e a grande responsável por essa revolução, que vai mudar praticamente tudo como conhecemos.

QUINTA | 28/11

Trilha Machine Learning

Machine Learning e Deep Learning além do buzzword

Muito se fala de Machine Learning, Deep Learning e seu poder supostamente revolucionário, mas poucas vezes se mostra o que há por trás desses



THE DEVELOPER'S CONFERENCE

QUINTA | 28/11

Trilha Arquitetura de Dados

Discutir e compartilhar experiências sobre arquitetura de dados RDBMS e NoSQL

Nos dias de hoje com a infinidade de tecnologias de



12:10
às
13:00

Experimento de Extração de Relações com algoritmo Conditional Random Field ▾

Patricia Nunes Gonçalves

Machine Learning Canvas: Da coleta de dados à geração de valor ▾

Alexandre Ray Silva



11:10
às
12:00

Usando LDA para seleção de tópicos para Recuperação de Informação aplicado a Inteligência de Mercado ▾

Patricia Nunes Gonçalves

OCR: o que é e como usar? ▾

Guilherme Malta



10:10
às
11:00

Um estudo sobre implementação de ILM (Information Lifecycle Management) no Sicredi ▾

Patricia Nunes Gonçalves

Kafka e Dataflow: uma história de amor ▾

Leonardo Alves Miguel

Perguntas?



 patricia_nginx@sicredi.com.br

 <https://dblp.org/pid/14/6534>

 <https://www.linkedin.com/in/patricianunesgonc/>

