

# Repositórios de Dados Científicos

## Estado da Prática no Estado de São Paulo

Fabio Kon

Coordenação Adjunta de Pesquisa para Inovação  
FAPESP

(Ciência da Computação - IME/USP)



2017

Data Management  
Policy

Compulsory Data  
Management Plans

WG - 7 public  
universities

Establish network of  
**OPEN**  
Research data  
repositories

# OPEN SCIENCE BY DESIGN

Realizing a Vision for 21st Century Research



National Academies of Sciences,  
Engineering, Medicine

July 2018

Open science =

Open access = papers

**Open data**

Open methods = open source

# OPEN DATA = OPEN METADATA

Metadata on papers - and papers...

Metadata on data  
Metadata on software  
Metadata on everything associated with experiment

OPEN METADATA REPOSITORIES

DATA MAY BE OPEN OR NOT - DEPENDS on SCIENTIST

FILES FILES FILES FILES FILES

...

FILES FILES FILES FILES FILES

# Política de gestão de dados

- Planos de gestão de dados obrigatórios na submissão
- Inicialmente para grandes projetos (TEM, JP)
  - Até hoje, cerca de 600 projetos já aprovados com avaliação de planos de gestão de dados
- Estendidos progressivamente para outras modalidades

# Uso do dmptool

- Estado SP = Segundo maior usuário do dmptool norteamericano
- Instituições estão criando seus templates - USP, UNICAMP, UNESP, UFABC
- FAPESP não impõe template obrigatório



**Welcome to the DMPTool**

Create data management plans that meet institutional and funder requirements.



A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) convida para o



LANÇAMENTO DA  
**REDE DE REPOSITÓRIOS**  
DE DADOS CIENTÍFICOS  
DO ESTADO DE SÃO PAULO

---

16 de dezembro de 2019  
das 13h às 16h - FAPESP



# GT - Rede de repositórios de SP



7 IES públicas, aprox. 48 campi

11,5 mil docentes

170 mil alunos

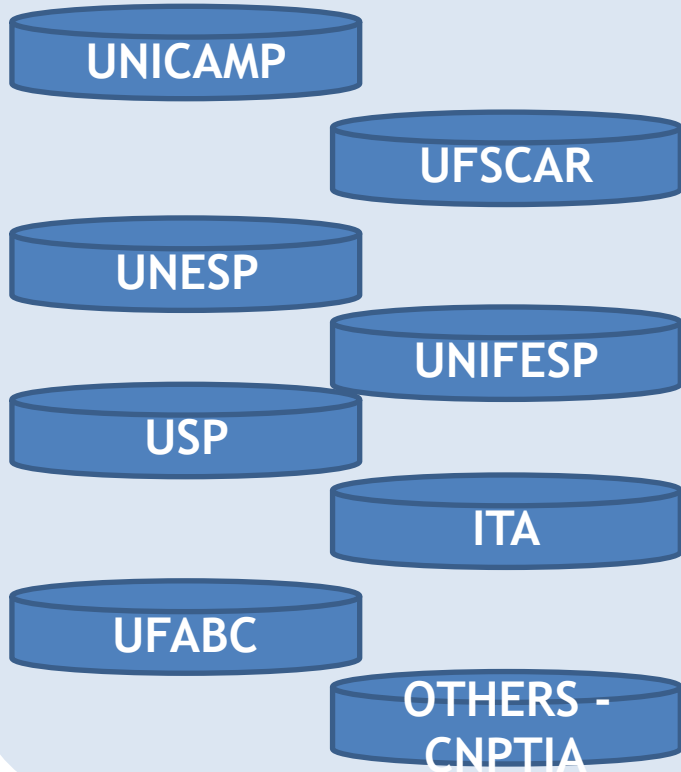
+ pesquisadores em informática em  
agricultura (CNPTIA)



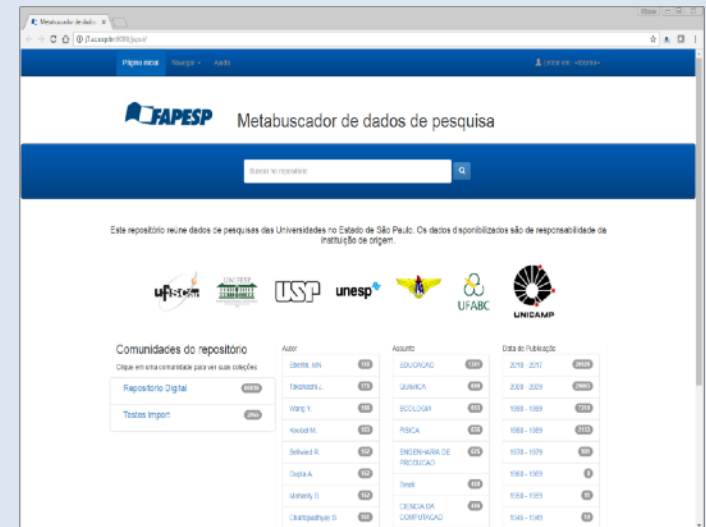
# GT - Rede de repositórios

- Cada instituição tem seu próprio sistema
- Interface única de busca de metadados (*harvester*)
- Reuniões mensais desde agosto de 2017

## Storage



## RETRIEVAL



# Federação de repositórios de dados

- Harvester de metadados implementado em DSpace (STI da USP)
- Coleta metadados diariamente - aceita comunicação com DSpace, Dataverse e ckan
- Usa OAI - PMH
- Dublin Core qualificado
- Conjunto de cerca de 200 campos de metadados

# Gestão em nível institucional

- Cada instituição criou um grupo de políticas de gestão de dados
  - Composição e responsabilidades definidas por portarias
  - Bibliotecários, analistas dos centros de computação e usuários-cientistas de vários domínios
- Repositórios considerados ação prioritária das reitorias
- Decisões de implementação e de gestão são locais

# Decisões de implementação

- Dataverse = Unicamp, Unifesp, UFABC, ITA, CNPTIA
- DSpace = UNESP, UFSCAR
- DSpace com casca própria - USP
  
- Sistema de identificação dos dados
  - DOI (Datacite) - UNICAMP e CNPTIA
  - Demais - por enquanto, Handle.Net

# Pessoal envolvido na implementação

- Todos em tempo parcial, com fases de tempo integral
- USP - 7 pessoas
- UNIFESP, ITA, UFABC, UFSCAR - 2 a 3 pessoas
- UNESP, UNICAMP - 3 a 4 pessoas

# Estágio atual

- Todas instituições estão treinando usuários selecionados e inserindo dados
- A rede irá desenvolver um conjunto de políticas e portarias para garantir sua sustentabilidade (cada universidade já está definindo suas portarias)

# Contato

Profa. Claudia B. Medeiros  
cmedeiros@fapesp.br