

Programa Nacional de Acesso Aberto a Dados de Pesquisa

17º WIRNP

Workshop RNP

30 | 31 MAIO
SALVADOR | BA

Leonardo Lazarte

IBICT/MCTI

llazarte@ibict.br
llazarte@unb.br



Foco da apresentação:

Como avançar para estruturar um Programa Nacional de Acesso Aberto a Dados de Pesquisa

- Que são dados abertos? Exemplos
- Porque são relevantes os Dados Abertos
- Porque Dados Abertos AGORA?
- Situação Internacional
- Porque um Programa Nacional
- Escopo de um Programa Nacional
- Exemplos de áreas de atuação
- Desafios na estruturação do Programa

Dados Pesquisa

Exemplo 1: Teste de um tratamento médico (processo usual)

- Hipóteses
- Escolha de população teste
- Levantamento de **dados** da população
- Aplicação do tratamento
- **Dados** do processo da aplicação
- Resultados do processo
- **Dados** dos resultados

Metodologia → Dados → Análise → Conclusões → Publicação

A publicação usualmente não inclui os dados, só um resumo dos mesmos, ou uma interpretação.

Dados Abertos de Pesquisa

Exemplo 1: Teste de um tratamento médico (dados abertos)

Suponhamos que foram disponibilizados para acesso aberto os seguintes dados:

- **Dados** da população
- **Dados** do processo da aplicação
- **Dados** dos resultados

Outro pesquisador pode avaliar se as conclusões publicadas são compatíveis com os dados publicados.

Verificabilidade

Dados Abertos de Pesquisa

Exemplo (2):

Observação num telescópio avançado de um objeto ou fenômenos astronômico específico.

Exemplo (3):

Estudo de uma partícula subatômica numa colisão produzida num acelerador de partículas.

Em ambos os casos, o instrumento (telescópio ou acelerador), produz um volume de informação enorme, do qual o pesquisador utiliza uma parte ínfima.

Outros pesquisadores podem utilizar dados que usualmente seriam descartados.

Multiplicação da **eficiência** dos equipamentos

Acesso aberto a dados de pesquisa

Relevância dos Dados Abertos

- Verificabilidade
- Reuso
- Uso de dados “excedentes”

Valor estratégico

- Acesso a pesquisa de ponta em regiões remotas ou com poucos recursos
- Incorporar a capacidade de pesquisa de cientistas fora dos grandes centros

Oportunidade de Acesso Aberto a Dados de Pesquisa

- Aumento significativo na capacidade de Armazenamento
- Conectividade com alta capacidade e capilaridade

Evolução recente

- Comunidades pioneiras, largo uso de tecnologias digitais
- Percepção do valor estratégico do acesso aberto a dados de pesquisa
- Iniciativas que vão além de comunidades de áreas específicas
- Iniciativas nacionais
- Iniciativas regionais e internacionais

Research Data Alliance

- Estados Unidos, União Europeia, Austrália
- 2013
- Plenárias semestrais
- Dezenas de Grupos de Trabalhos: Resultados a Curto Prazo
- Dezenas de Grupos de Interesse
- Ampliação para incluir outros países
- Convite ao Brasil. RDA → Itamaraty → MCTI → RNP → IBICT

Porque um Programa Nacional de Acesso Aberto a Dados de Pesquisa

- **Compartilhamento de Infraestrutura**
 - Armazenamento (RNP, Nuvem)
 - Conectividade (RNP)
 - Processamento (LNCC)
- **Políticas de incentivo ao compartilhamento de dados (MCTI, CNPq, FAPESP, CAPES, etc)**
- **Programas de capacitação (Universidades, CAPES, Coursera)**
- **Facilidade de acesso centralmente coordenado (IBICT)**
- **Ferramentas de busca (IBICT)**
- **Facilidade de cruzamento de dados de áreas diferentes**

Principais elementos que deve contemplar um Programa Nacional de Acesso Aberto a Dados de Pesquisa

- Infraestrutura
 - Armazenamento
 - Conectividade
 - Processamento
- Políticas de incentivo ao compartilhamento
 - Projetos de pesquisa devem incluir Planos de Gestão de Dados
- Armazenamento e recuperação de Dados
 - Arquiteturas de dados e metadados específicas por áreas
 - Mecanismos de autoarmazenamento
 - Ferramentas de busca
 - Mecanismos de coleta de metadados (armazenamento distribuído)
 - Preservação e arquivamento
- Capacitação
 - Do pesquisador que gera os dados
 - Do “bibliotecário de dados” ou “*data broker*”
 - Do pesquisador que utilizará os dados abertos
 - De quem implanta e cuida da infraestrutura

O que envolve o dia a dia de um Programa Nacional de Acesso Aberto a Dados de Pesquisa

- Metadados por área e catálogos de metadados, padrões (do pesquisador, lab, ou gerais)
- Identificadores persistentes
- Currículo de formação do Cientista de Dados
- Cursos rápidos de especialização dos gestores de dados
- Qualidade dos dados, curadoria
- Interoperabilidade entre sistemas já existentes
- Nuvem de dados abertos
- Privacidade versus publicidade, ética, proteção de dados
- “Datometria” (bibliometria de dados: quem, o que, quanto usou)
- Autoria
- Fluxos de publicação de dados
- Dados dinâmicos: referência e preservação

Desafios MACRO para a elaboração e implementação de um Programa Nacional de Acesso Aberto a Dados de Pesquisa

- Articular as instituições relevantes
 - Agências de fomento à pesquisa
 - Percepção da relevância estratégica do Programa
 - Desenvolvimento de políticas de incentivo aos dados abertos
 - Financiamento de aspectos do Programa (relevamento inicial, infraestrutura piloto)
 - Órgãos do governo que lidam com políticas de C&T
 - Instituições que já tem iniciativas na área
- Levantamento das comunidades existentes no Brasil
 - Ferramentas utilizadas
 - Infraestrutura utilizada
 - Envolvimento na construção do Programa
- Inserção nas comunidades internacionais
- Financiamento e sustentabilidade

- Em construção, RNP + IBICT

17º WIRNP

Workshop RNP

Obrigado!

Leonardo Lazarte
llazarte@ibict.br
llazarte@unb.br



Ministério da
Defesa

Ministério da
Cultura

Ministério da
Saúde

Ministério da
Educação

Ministério da **Ciência,**
Tecnologia, Inovações e
Comunicações

GOVERNO FEDERAL