

Talking points for the Workshop on Technology  
and Collections Management  
*Pontos de discussão para o Workshop de  
Tecnologia e Gestão de Coleções<sup>1</sup>*

Andrew Russell Green  
Instituto Mora  
Mexico City *Cidade do México*

18 November *novembro* 2011  
Museu Paulista, Universidade de São Paulo

---

<sup>1</sup>This document may be used under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs license; it may be copied and redistributed for non-commercial purposes, provided that it is not modified and the author is cited. For details, see <http://creativecommons.org/licenses/by-nc-nd/3.0/>. Also, fragments may be reproduced in accordance with standard academic practice.

# Table of Contents *Índice*

Introduction / Introdução

Free (libre) Software / Software livre

Free Software in Heritage Institutions / Software livre em instituições encarregadas do património

How to Liberate a Program in Two Minutes or Less / Como liberar um programa em dois minutos ou menos

A Rough Theoretical Framework for Archival Description /  
Aproximação de um marco teórico de descrição arquivística

Relational and Non-Relational Models / Modelos relacionais e não relacionais

Archives and Data Integration / Arquivos e integração de dados

## My (Main) Affiliation *Meu afiliação (principal)*

- ▶ **Institution:** the Instituto Mora, a Social Sciences and History research institute and graduate school in Mexico City.

***Instituição:** Instituto Mora, uma escola de pós-graduação e instituto de pesquisa em Ciências Sociais e História na Cidade do México.*

- ▶ **Workgroup:** the Audiovisual Laboratory for Social Research (acronym in Spanish: LAIS).

***Grupo de trabalho:** Laboratório de Audiovisual para a Pesquisa Social (LAIS).*

## My (Main) Affiliation *Meu afiliação (principal)*

- ▶ **Project:** I participate in a project called Image Preservation, Information Systems, Access and Research (PISIAI).  
*Projeto: Participo do projeto Preservação da Imagem, Sistemas de Informação, Acesso e Pesquisa (PISIAI),*
  - ▶ This project is based mainly at the LAIS/Instituto Mora.  
*Este projeto é baseado principalmente no LAIS/Instituto Mora.*
  - ▶ Many other institutions have also contributed to it.  
*Muitas outras instituições também têm contribuído para isso.*
- ▶ PISIAI is interdisciplinary, involving Social Science, Computer Science and Archival Science.  
*PISIAI é interdisciplinar e envolve as Ciências Sociais, Ciência da Computação e Ciência Arquivística.*

# PISIAI Project *Projeto PISIAI*

- ▶ The project's activities include:

*As atividades do projeto incluem:*

- ▶ history research that uses photographs and other visual materials as primary sources;  
*pesquisas históricas que integram fotografias e outros materiais visuais como fontes primárias;*
- ▶ the creation of methodologies for cataloguing such materials;  
*a criação de metodologias para a catalogação de tais materiais;*
- ▶ the development of software for online access to photograph collections; and  
*o desenvolvimento de software para acesso on-line para coleções fotográficas; e*
- ▶ the creation of online photograph collections using the aforementioned software and methodology.  
*a criação de fototecas on-line usando o software e metodologia acima mencionados.*

## Software: Pescador

- ▶ Over the past decade, we have (with interruptions) developed several versions of a free (*libre*) system for online access to photograph collections.

*Durante a última década, temos (com interrupções) desenvolveu diversas versões de um sistema livre para acesso on-line às coleções fotográficas.*

- ▶ The system, called Pescador, seeks to provide specialized functions for social science research.

*O sistema, chamado Pescador, visa proporcionar funções especializadas para a pesquisa em ciências sociais.*

## Software: Pescador

- ▶ Since work on Pescador began, the ecosystem of free software for archives has grown tremendously.

*Desde que o trabalho começou em Pescador, o ecossistema de software livre para arquivos tem crescido tremendamente.*

- ▶ The next version of Pescador (not yet finished) will be a set of components that will integrate with other free software for archives. (Previous versions were more monolithic.)

*A próxima versão do Pescador (ainda não concluído) será um conjunto de componentes que irão integrar com outros softwares livres para arquivos. (As versões anteriores eram mais monolítico.)*

## Software: Pescador

- ▶ Instead of storing information in a standard relational database, Pescador uses a Semantic Web (SW) store.  
*Em vez de armazenar informações em um banco de dados relacional padrão, Pescador usa um repositório de dados em formato Web Semântica (WS).*
- ▶ The SW is a standard for creating and sharing flexible data structures, similar to those used in some types of artificial intelligence and expert systems.  
*O SW é um padrão para criação e compartilhamento de estruturas de dados flexíveis, semelhantes aos utilizados em alguns tipos de inteligência artificial e sistemas especialistas.*



## Software: Pescador

- ▶ The SW will (hopefully) support specialized functions for research and help address the numerous data integration problems facing archive software.  
*(Esperamos que) o SW irá suportar funções especializadas para a pesquisa e irá contribuir para resolver os inúmeros problemas de integração de dados que enfrentam softwares de arquivamento.*

# Free Software in Heritage Institutions

## *Software livre em instituições encarregadas do património*

- ▶ Free software fosters interdisciplinary research and inter-institutional collaboration, is a form of knowledge-sharing, and creates re-usable public resources.  
*O software livre fomenta a pesquisa interdisciplinar e colaboração inter-institucional, é uma forma de compartilhamento de conhecimento, e cria recursos públicos reutilizáveis.*
- ▶ This matches the public mission of heritage institutions and an epistemic view of heritage.  
*Isto concorda com a missão pública de instituições encarregadas do patrimônio e uma visão epistêmica do patrimônio.*

# Free Software in Heritage Institutions

## *Software livre em instituições encarregadas do património*

- ▶ It is not a “cure-all”, and brings new challenges that heritage institutions may not be used to.  
*Não é um “remédio para todos os males”, e traz novos desafios a que as instituições encarregadas do património podem não estar acostumados.*
- ▶ There are almost no good reasons for institutions not to release in-house developments as free software.  
*Quase não existem boas razões para não liberar como software livre os desenvolvimentos próprios.*
- ▶ Vast sums of public money are spent on commercial software. In most cases, these resources would be better spent on well-managed free software projects.  
*Grandes quantidades de dinheiro público são gastos em software comercial. Na maioria dos casos, seria melhor gastar esses recursos em projetos de software livre bem gerenciado.*

# How to Liberate a Program in Two Minutes or Less

## *Como liberar um programa em dois minutos ou menos*

1

Add a copyright header to source files.

*Adicionar um cabeçalho de direitos autorais aos arquivos de código fonte.*

2

Make the source files and license available on the Web.

*Fazer os arquivos de código fonte e licença disponíveis na web.*

Note: repeat steps 1 and 2 when source files change or new ones are added.

*Nota: repita os passos 1 e 2 quando arquivos de código fonte têm alterações ou novos foram adicionados.*

# A Rough Theoretical Framework for Archival Description

## *Aproximação de um marco teórico de descrição arquivística*

- ▶ Archival description (cataloguing) is a form of communication.  
*A descrição arquivística (catalogação) é uma forma de comunicação.*
- ▶ The cataloguer is the sender of the communication, and the catalogue user (often a patron) is the recipient.  
*O catalogador é o emissor da comunicação, e o usuário do catálogo (muitas vezes um consulente) é o receptor.*

# A Rough Theoretical Framework for Archival Description

## *Aproximação de um marco teórico de descrição arquivística*

- ▶ The physical medium of the communication is often a computer system.  
*O meio físico da comunicação é freqüentemente um sistema informático.*
- ▶ The characteristics of the computer system strongly determine the characteristics of the communication that passes through it.  
*As características do sistema informático determinam as características da comunicação que passa por ela.*

# A Rough Theoretical Framework for Archival Description

## *Aproximação de um marco teórico de descrição arquivística*

- ▶ According to a cognitive view of communication, meaning is not in words per se, but in our individual and collective thought.

*De acordo com uma perspectiva cognitiva da comunicação, o significado não é em palavras per se, mas sim em nosso pensamento individual e coletivo.*

- ▶ “In its communicative function, language is a set of tools with which we attempt to guide another mind to create within itself a mental representation that approximates one we have.”

*“Na sua função comunicativa, a linguagem é um conjunto de ferramentas com as quais tentamos orientar outra mente para criar dentro de si uma representação mental que se aproxima a que temos.”*

—Scott Delancy

# A Rough Theoretical Framework for Archival Description

## *Aproximação de um marco teórico de descrição arquivística*

- ▶ Formal and controlled languages—including formal logic, computer languages and languages for archival description—are all means for human expression.

*Linguagens formais e controladas—including a lógica formal, linguagens de computador e linguagens de descrição arquivística—são meios de expressão humana.*

- ▶ We process such languages using many of the same (individual and collective) mechanisms that we use with natural languages. (I try to demonstrate this in a thesis I'm working on.)

*Nós processamos tais linguagens usando muitos dos mesmos mecanismos (individuais e coletivas) que usamos com as línguas naturais. (Tento demonstrar isso em uma tese que estou elaborando.)*



# A Rough Theoretical Framework for Archival Description

## *Aproximação de um marco teórico de descrição arquivística*

- ▶ Many formal and controlled languages differ from natural languages in that, in addition to conveying meaning, they are designed to function as a part of mechanistic systems (like computer systems).

*Muitas linguagens formais e controladas diferem de linguagens naturais em que, além de transmitir significado, eles são projetados para funcionar como parte de sistemas mecanicistas (como sistemas informáticos).*

# Relational and Non-Relational Models

## *Modelos relacionais e não relacionais*

- ▶ Relational and non-relational database systems are not as different as they may appear.  
*Os sistemas de banco de dados relacional e não relacional não são tão diferentes quanto parecem.*
- ▶ With systems of either type, a given idea can almost always be modeled in a variety of ways.  
*Com os sistemas de qualquer tipo, uma determinada idéia quase sempre pode ser modelado em uma variedade de maneiras.*

# Relational and Non-Relational Models

## *Modelos relacionais e não relacionais*

- ▶ One may consider the advantages and disadvantages of different models of an idea in a given system (ease of programming and maintenance, functions to be offered, etc.).  
*Pode-se considerar as vantagens e desvantagens dos diferentes modelos de uma idéia em um determinado sistema (facilidade de programação e manutenção, as funções a serem oferecidos, etcetera).*
- ▶ Different types of system may be seen as different types of formal languages that provide different ways of expressing a given idea; experimentation is a good idea.  
*Diferentes tipos de sistema podem ser vistos como diferentes tipos de linguagens formais que fornecem maneiras diferentes de expressar uma idéia determinada; experimentação é uma boa idéia.*

# Archives and Data Integration

## *Arquivos e integração de dados*

- ▶ The definition of whole catalogue formats may be replaced by the definition of reusable conceptual record components.  
*A definição de formatos de catálogo completo pode ser substituída pela definição de reutilizáveis componentes conceituais de registros.*
- ▶ This is a potential solution for problems of data integration.  
*Esta é uma solução potencial para problemas de integração de dados.*
- ▶ Data integration within and among archives, libraries and museums is one of the greatest of all enterprise software engineering problems.  
*A integração de dados dentro e entre arquivos, bibliotecas e museus é uma das maiores de todos os problemas de engenharia de software da empresa.*

## Links and Address / Links e endereços

`http://lais.mora.edu.mx`

`http://ceur-ws.org/Vol-348/worm08_contribution_10.pdf`

`http://durito.nongnu.org`

`andrew.green.df@gmail.com`

Thanks! *Obrigado!* ¡Gracias!