

# ***Uma proposta de modelo de processo para publicação de dados abertos conectados governamentais***

*Programa de Pós-Graduação em Modelagem Computacional do  
Conhecimento*

*Instituto de Computação – Universidade Federal de Alagoas*

Thiago José Tavares Ávila

Mestrando

Prof. Dr. Ig Ibert Bittencourt – Orientador

Prof<sup>a</sup> Dra<sup>a</sup> Patrícia Ospina – Co-Orientadora

25/11/2015

# Agenda

- Contextualização
- Motivação
- Trabalhos Relacionados
- Problemática
- Objetivos da Pesquisa
- Proposta: Modelo de processo “Piece of Cake”
- Validação
- Limitações
- Conclusão e Trabalhos Futuros
- Referências

## Contextualizando ....

- Qual o montante destinado para a função **Educação** no **Orçamento** do Estado Brasileiro em 2014?



Download de dados do  
Orçamento Governo  
Federal

+



Download de dados dos  
Orçamentos das 27 Unidades  
Federativas

+



Download de dados dos  
Orçamentos dos 5565  
Municípios



loa2014.zip  
6,0/23,0 MB, 47 segund...



LOA\_2014-01-30.pdf  
20,2/20,5 MB, 0 segund...



## *Dados Abertos Governamentais*

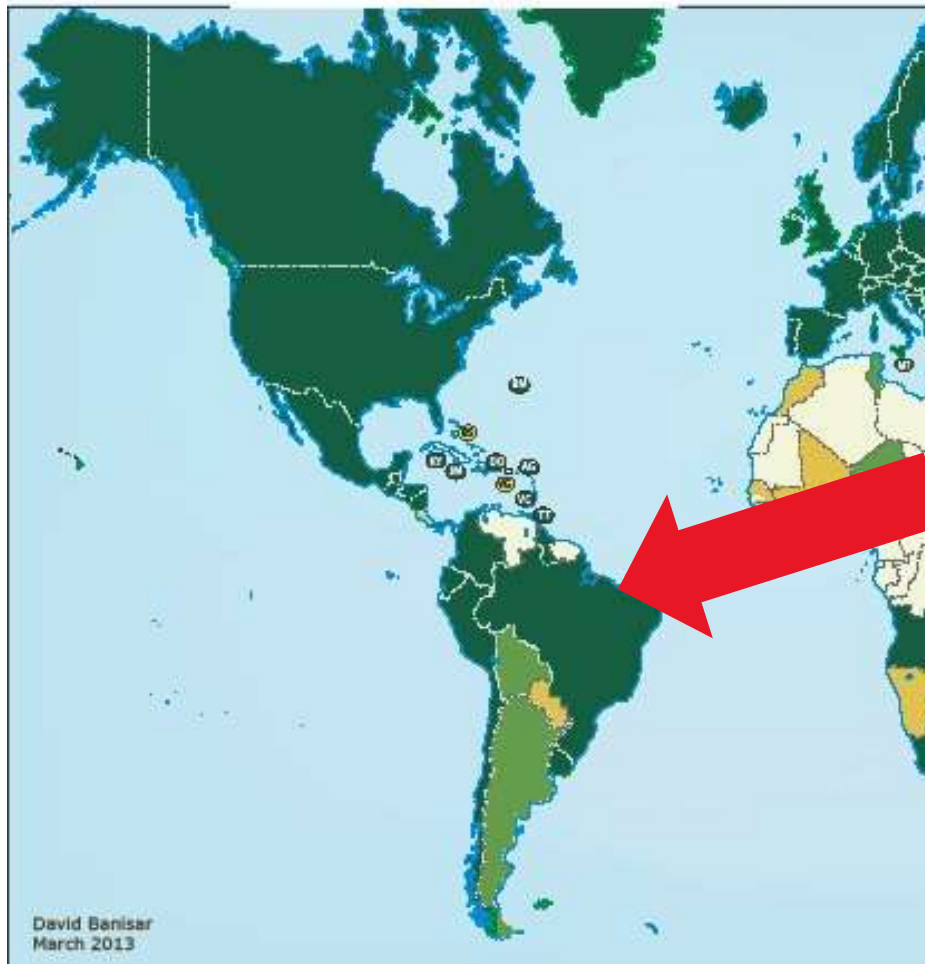
- “Dados produzidos pelo governo e colocados à disposição das pessoas de forma a:
  - tornar possível não apenas sua leitura e acompanhamento,
  - mas também **sua reutilização** em novos projetos, sítios e aplicativos;
  - seu **cruzamento com outros dados** de diferentes fontes;
  - e sua **disposição em visualizações** interessantes e esclarecedoras”.

(W3C Brasil, 2011)



# Acesso à Informação: Obrigações Legais

## National Right to Information Laws, Reg



**Lei 12.527/2011 - Art. 8 § 3º**

*II - possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações; [...]*

**Lei 12.527/2011 - Art. 8 § 3º**

*III - possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina [...]*

Figura 1 – Mapa dos países que possuem Lei de Acesso à Informação no mundo [Banisar, 2013]

## ***Mas problemas existem ...***

*“Dispor de **dados abertos governamentais** é algo ótimo, mas é muito difícil **tomar decisões sensatas sem o contexto do que significam** tais dados. Atualmente, é **necessário envolver analistas e cientistas** para gerar conhecimento a partir de tais dados.*

*Robert Schaefer – Project Scientist  
Johns Hopkins University Applied Physics Lab  
Conferência Internacional para Governo Aberto (2010)*

### **Outros problemas:**

- Difíceis de serem consumidos
- Formatos variados e/ou proprietários
- Licenciamento ineficaz (ou inexistente)





# Esquema das 5-Estrelas dos Dados Abertos

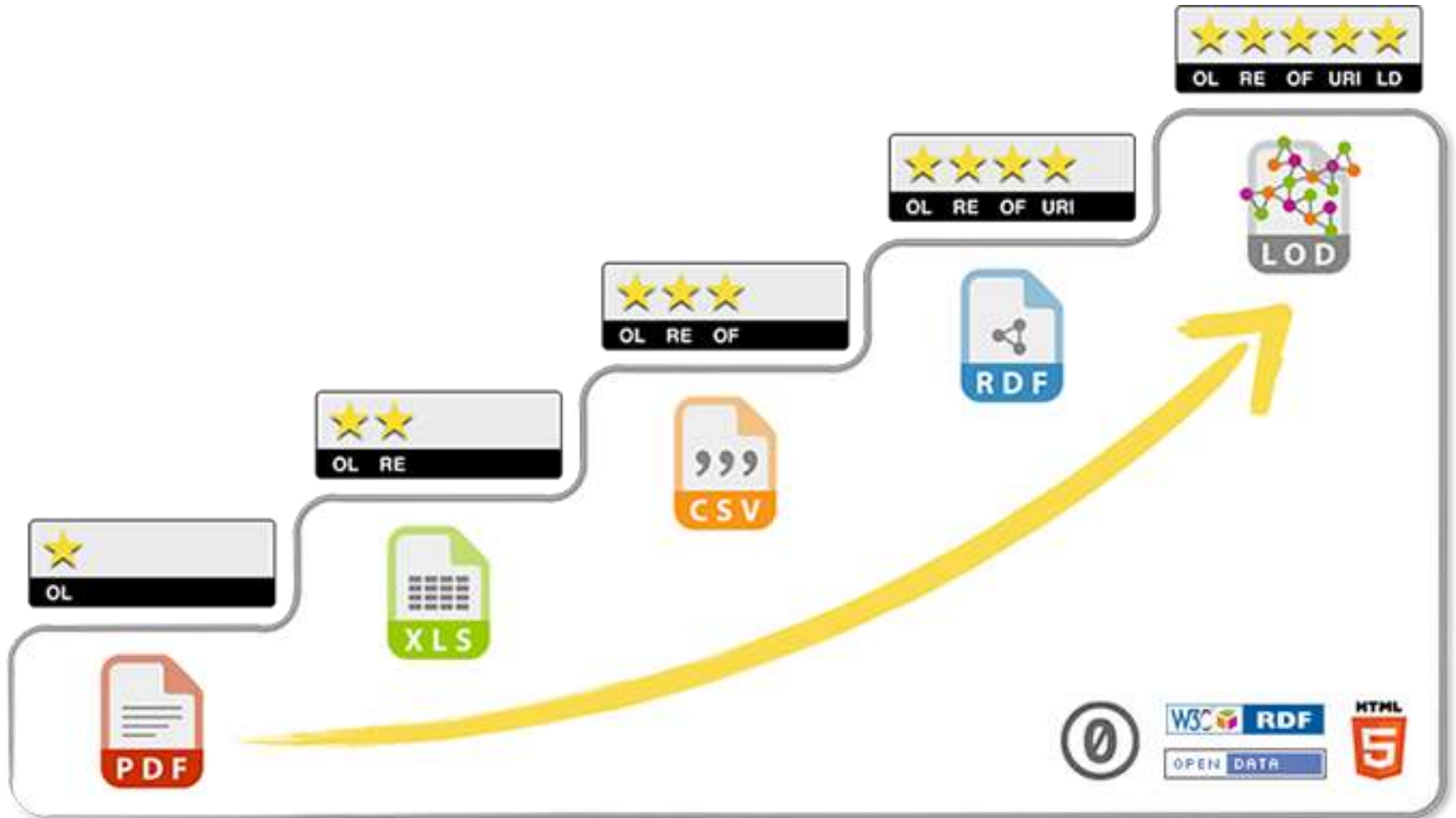
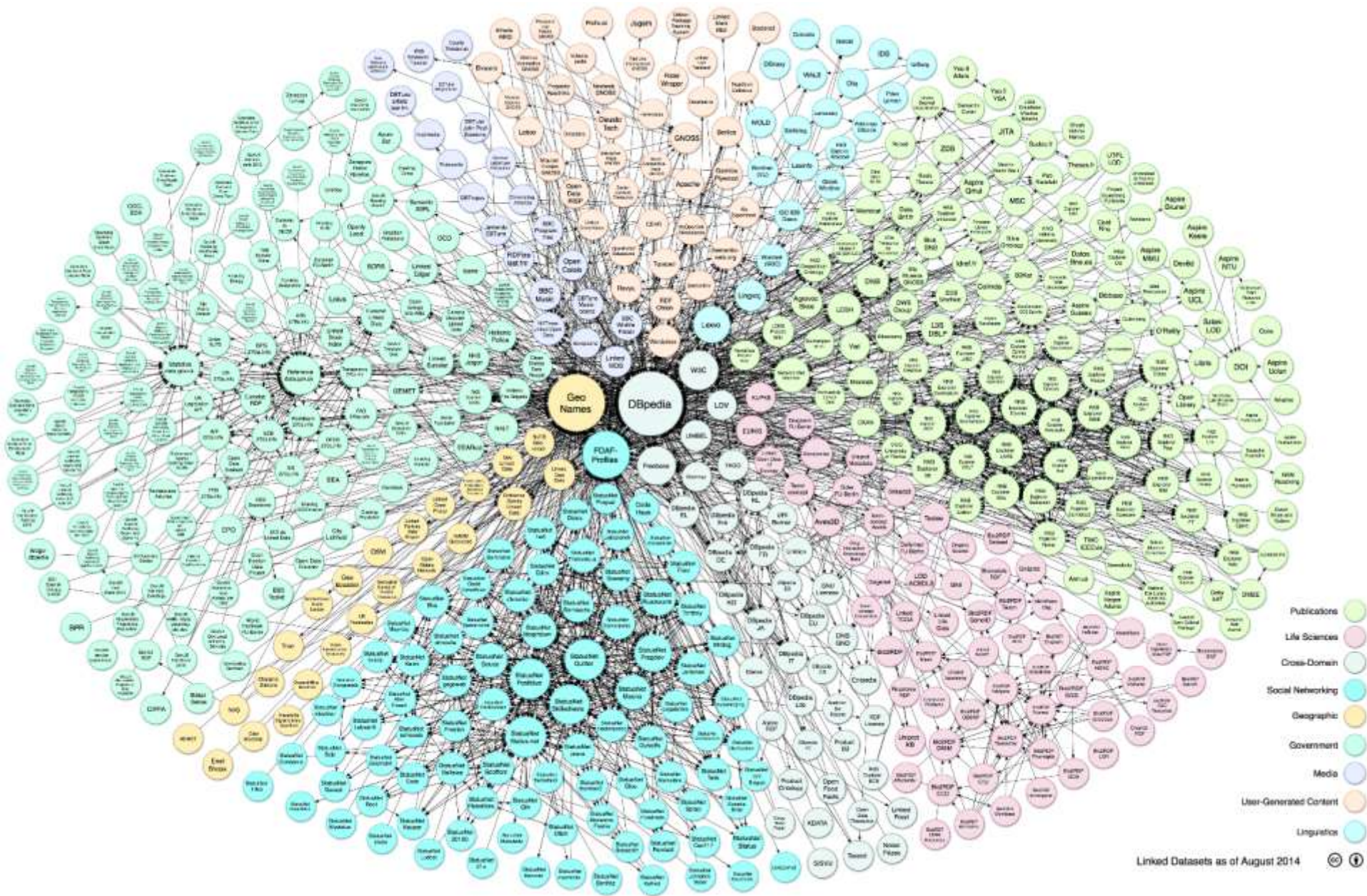


Figura 2 – Esquema das 5-Estrelas dos Dados Abertos (Berners-Lee, 2006).





## Dados Abertos Conectados Governamentais (DACG)

### Key Linked Government Data initiatives in the Member States



**BE - Interconnecting Belgian National and Regional Address Data**

<http://location.testproject.eu/BEL/>

**DK - Good basic data for everyone**

<http://uk.fm.dk/publications/2012/good-basic-data-for-everyone>

**EC - Open Data Portal**

<http://open-data.europa.eu/open-data/linked-data>

**EU - Financial Transparency System**

<http://fts.publicdata.eu/page/startpage>

**GR - Legal Entities & Public Spending**

<http://linkeddata.ihu.edu.gr:8080/rdf-browser/>  
<http://publicspending.medialab.ntua.gr/>

**IT - Linked Data Guidelines**

<http://www.digitpa.gov.it/notizie/linee-guida-open-data-interoperabili>

**SE - Linked Open Data Sweden**

<http://www.ida.liu.se/~evabl45/lodsweden.sv.shtml>

**UK - Government Linked Data Working Group**

<http://data.gov.uk/linked-data>  
<http://education.data.gov.uk/>  
<http://location.data.gov.uk/>  
<http://law.data.gov.uk/>

**NL - Geonovum.nl**

<http://www.geonovum.nl/>

Figura 4 – Iniciativas de Dados Abertos Conectados Governamentais na Europa (ISA, 2013)

← → ↺ orcamento.dados.gov.br/sparql/

# OpenLink Virtuoso

This query page is designed to help you test OpenLink Virtuoso SPARQL protocol endpoint. Consult the [Virtuoso Wiki page](#) describing the service or the [Online Virtuoso Documentation](#) section [RDF Database](#).

There is also a rich Web based user interface with sample queries. In order to use it you must install the iSPARQL package.

## Query

Default Graph URI

(Security restrictions of this server do not allow you to retrieve remote data. Database administrator can change this.)

Query text

```
SELECT ?nomeFuncao (sum (?DotacaoInicial)) as ?soma
WHERE
{
  ?item rdf:type loa:ItemDespesa .
  ?item loa:temExercicio [loa:identificador 2014] .
  ?item loa:temFuncao [rdfs:label ?nomeFuncao] .
  ?item loa:valorDotacaoInicial ?DotacaoInicial.
  ?item loa:valorDotacaoInicial ?DotacaoInicial. }
GROUP BY ?nomeFuncao
```

Rigorous check of the query: ☒ Execution timeout, in milliseconds, values less than 1000 are ignored:

HTML  (The CXML output is disabled, see [details](#))

OpenLink Virtuoso version 06.01.3127, on Linux (x86\_64-unknown-linux-gnu), Single User Edition

nomeFuncao	soma
Cultura	3012548256
Defesa Nacional	38569820724
Educação	83275927095
Segurança Pública	8456349686
Judiciária	27991625424
Encargos Especiais	1345026899337
Indústria	2555003472
Administração	27206782816
Energia	2059672458
Saúde	98036362031
Desporto e Lazer	2228585737
Agricultura	24439596267
Legislativa	6950840498
Essencial à Justiça	5392088914
Gestão Ambiental	8734525210
Ciência e Tecnologia	8997879362
Assistência Social	68216343528
Relações Exteriores	2072986181
Trabalho	63838510464
Comércio e Serviços	5560101209
Previdência Social	473509193503
Direitos da Cidadania	1854107920
Organização Agrária	5156372499
Saneamento	3426319188
Transporte	24646839493
Urbanismo	8289330501
Comunicações	1637445736
Habitação	458081935
Reserva de Contingência	31577857866

Figura 6 – Consulta SPARQL sobre o Orçamento Federal Brasileiro

## *Oferta de DACG*

País	Recursos de dados abertos conectados	Total de recursos de dados	Participação
Alemanha	4	37.479	0,01%
Austrália	0	2.137	0,00%
Brasil	15	8.582	0,17%
Cingapura	0	11.977	0,00%
E.U.A.	7.051	332.568	2,12%
Espanha	1.775	9.029	19,66%
Itália	2.055	21.850	9,41%
Japão	0	18.105	0,00%
Reino Unido	362	9.497	3,81%

Tabela 1 – Participação de dados conectados nos catálogos de dados abertos governamentais (Do autor, 2015).

## *Problemas de Pesquisa*

- **Problema de Negócio:**
  - *Como guiar instituições governamentais a publicar dados abertos conectados considerando as melhores práticas e procedimentos produzidos e disponíveis na literatura?*



## Trabalhos Relacionados

*“Best Practices for Publishing Linked Data – BPLDs” (W3C)*

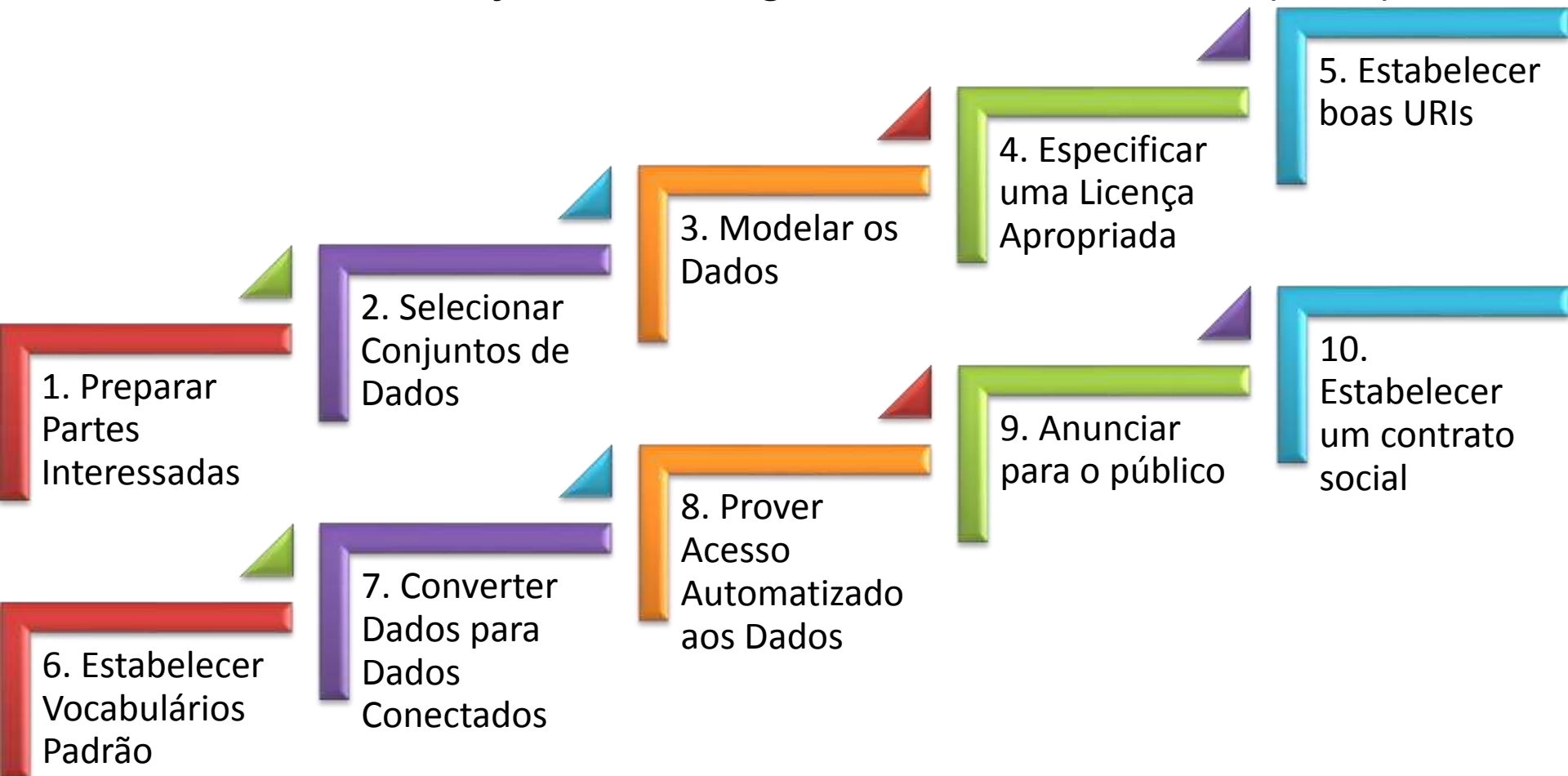


Figura 7 – 10 Melhores Práticas para Publicação de Dados Conectados - W3C (W3C, 2014)

## Trabalhos Relacionados

Identificación de  
la información

Figura 8 – Processo de public

Estructuración,  
cargue y  
publicación

ia (Colombia, 2012)

Identify

Figura 9 – Processo de publi  
and Wood, 2011)

Publish

ta on the Web (Hyland

Exploit

Specify

Model

Publish

Generate

Figura 10 – Processo de publicação de dados proposto em Methodological Guidelines for Publishing Government Linked Data (Villazón-Terrazas et. al, 2011)

# *Problemática*

- **Desafios:**
  - Existem poucos processos que suportam a publicação de **dados conectados governamentais**;
  - Ausência de **processos detalhados e softwares que apóiem o ciclo de vida de publicação de dados conectados governamentais**
    - Processos são direcionados à desenvolvedores de software, não para agentes governamentais (Villazón-Terrazas et. al, 2011);
  - Processos existentes, aparentemente, **não consideram o nível de maturidade** em gestão e publicação de dados das organizações publicadoras.

## *Problemas de Pesquisa*

- Problema de Negócio:
  - *Como guiar instituições governamentais a publicar dados abertos conectados considerando as melhores práticas e procedimentos produzidos e disponíveis na literatura?*
- Problema Técnico:
  - *“Que novos procedimentos devem ser propostos para que instituições governamentais publiquem dados abertos conectados considerando o seu nível de maturidade em gestão e publicação de dados?”*



## *Objetivos*

- Propor um **modelo de processo para publicação de dados abertos conectados governamentais** que:
  - **Considere o nível de maturidade em gestão e publicação de dados das instituições governamentais;**
  - Apresente atividades para a implementação de melhores práticas para publicação de dados conectados;
  - Seja iterativo e incremental.

# A Proposta - Modelo de Processo “Piece of Cake”

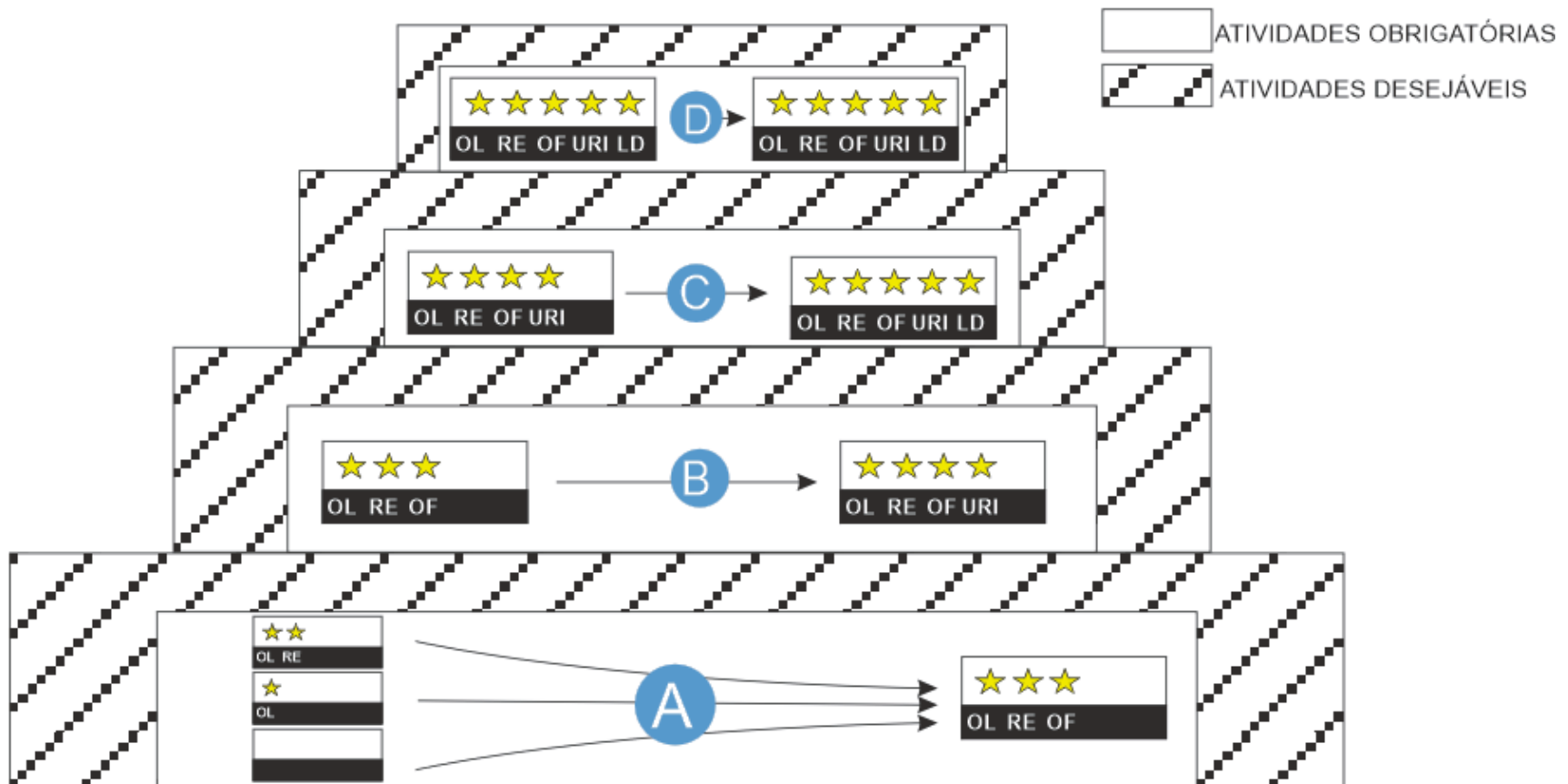
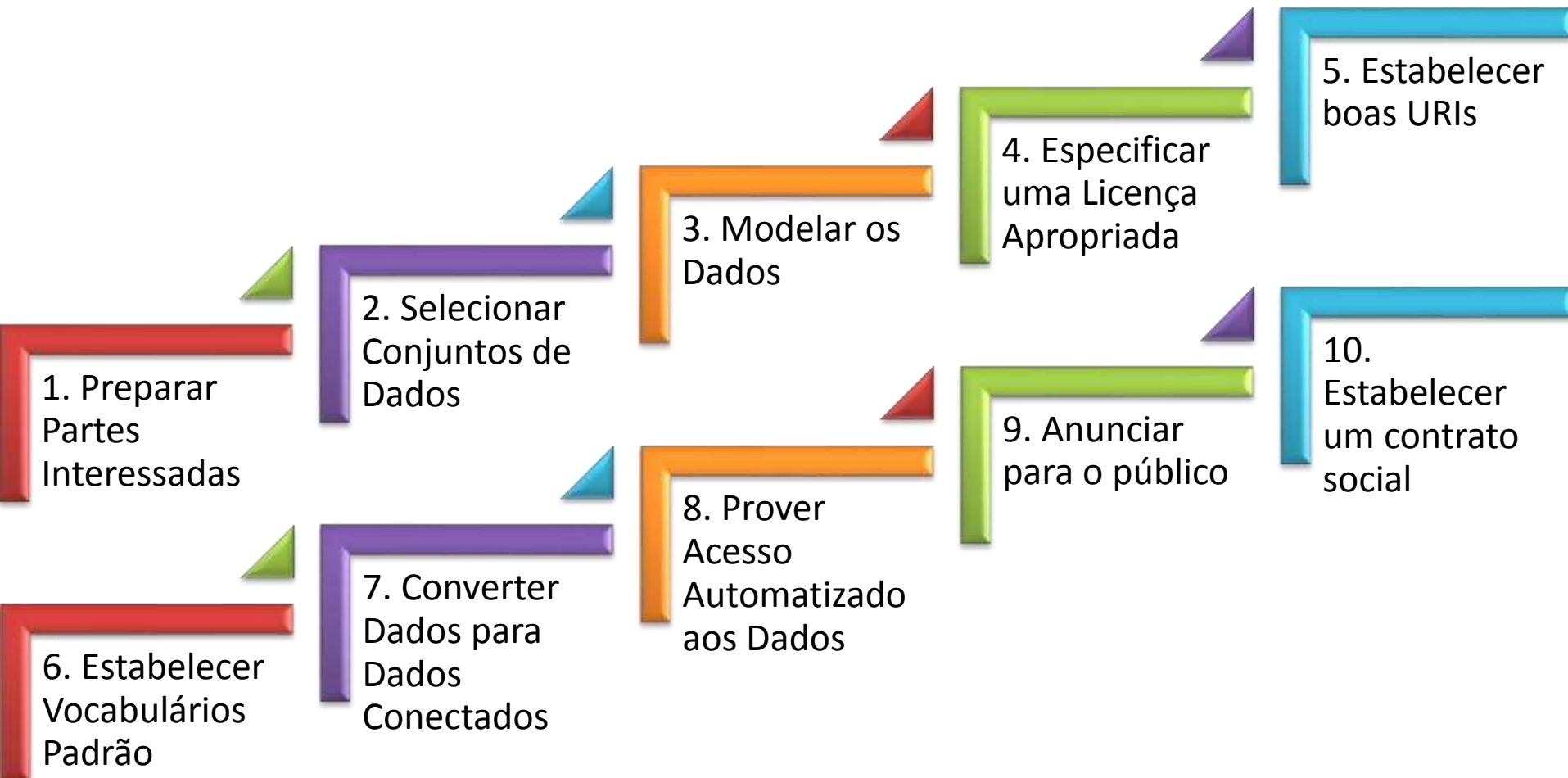


Figura 11 – Visão Geral do modelo de processo “Piece of Cake” (Do autor, 2015)

## Objetivos

- Propor um **modelo de processo para publicação de dados conectados governamentais** que:
  - Considere o nível de maturidade em gestão e publicação de dados das instituições governamentais;
    - **Esquema 5-Estrelas dos Dados Abertos**
    - **Maturidade das instituições publicadoras**
  - **Apresente atividades para a implementação de Melhores Práticas para Publicação de Dados Conectados;**
- Seja iterativo e incremental.

## ***“Best Practices for Publishing Linked Data – BPLDs” (W3C)***





# Revisão de Literatura - Metodologia

Código	Origem	Documentos (processos)	Tipo
P1	BRASIL	Plano de Dados Abertos; Processo de Abertura de Dados; Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0 - Ministério do Planejamento, Orçamento e Gestão do Brasil	Dados Abertos Governamentais
P2	CHILE	Norma Técnica para Publicación de Datos Abiertos en Chile	Dados Abertos Governamentais
P3	COLÔMBIA	Guía para la apertura de datos en Colombia	Dados Abertos Governamentais
P4	EQUADOR	Guia de Política Pública de Datos Abiertos – Ecuador	Dados Abertos Governamentais
P5	URUGUAI	Guía rápida de publicación em datos.gub.uy	Dados Abertos Governamentais
P6	ITÁLIA	Geolinked Open Data for the Municipality of Catania	Dados Abertos Conectados Governamentais
P7	INTERNACIONAL	LOP - Capturing and Linking Open Provenance on LOD Cycle	Dados Abertos Conectados
P8	INTERNACIONAL	TWC LOGD: A portal for linked open government data ecosystems	Dados Abertos Conectados Governamentais
P9	INTERNACIONAL	Linked Open Data: The Essentials A Quick Start Guide for Decision Makers.	Dados Abertos Conectados
P10	INTERNACIONAL	W3C Linked Data Cookbook	Dados Abertos Conectados
P11	GRÉCIA	Applying Linked Data Technologies to Greek Open Government Data: A Case Study	Dados Abertos Conectados Governamentais
P12	INTERNACIONAL	Managing the Life-Cycle of Linked Data with the LOD2 Stack	Dados Abertos Conectados
P13	ESPANHA	Methodological Guidelines for Publishing Government Linked Data	Dados Abertos Conectados Governamentais
P14	UNIÃO EUROPÉIA	Methodology for publishing datasets as open data – COMSODE	Dados Abertos Governamentais
P15	INTERNACIONAL	Open Data Handbook	Dados Abertos

Tabela 2 – Processos de publicação de dados abertos analisados na pesquisa (Do autor, 2015).

## *Revisão de Literatura - Metodologia*

- **Questões de Pesquisa:**
  - **HLQ: “Processos de abertura de dados estão incentivando a produção de dados conectados ?”**

RQ 3: Que recomendações para publicação de Dados Conectados podem ser extraídas dos processos analisados?

**RQ 3.1 ... 10: O que os processos recomendam a ser feito para contemplar cada uma das 10 Melhores Práticas para Publicação de Dados Conectados do W3C?”**

# Extração de recomendações

## Melhores Práticas para Publicação de Dados Conectados - W3C

1. Preparar partes interessadas
2. Selecionar conjuntos de dados
3. Modelar os dados
4. Especificar uma licença apropriada
5. Estabelecer bons identificadores universais (URIs)
6. Utilizar vocabulários padrão
7. Converter e enriquecer dados
8. Prover acesso automatizado aos dados
9. Anunciar os conjuntos de dados para o público
10. Estabelecer um contrato social para os dados publicados

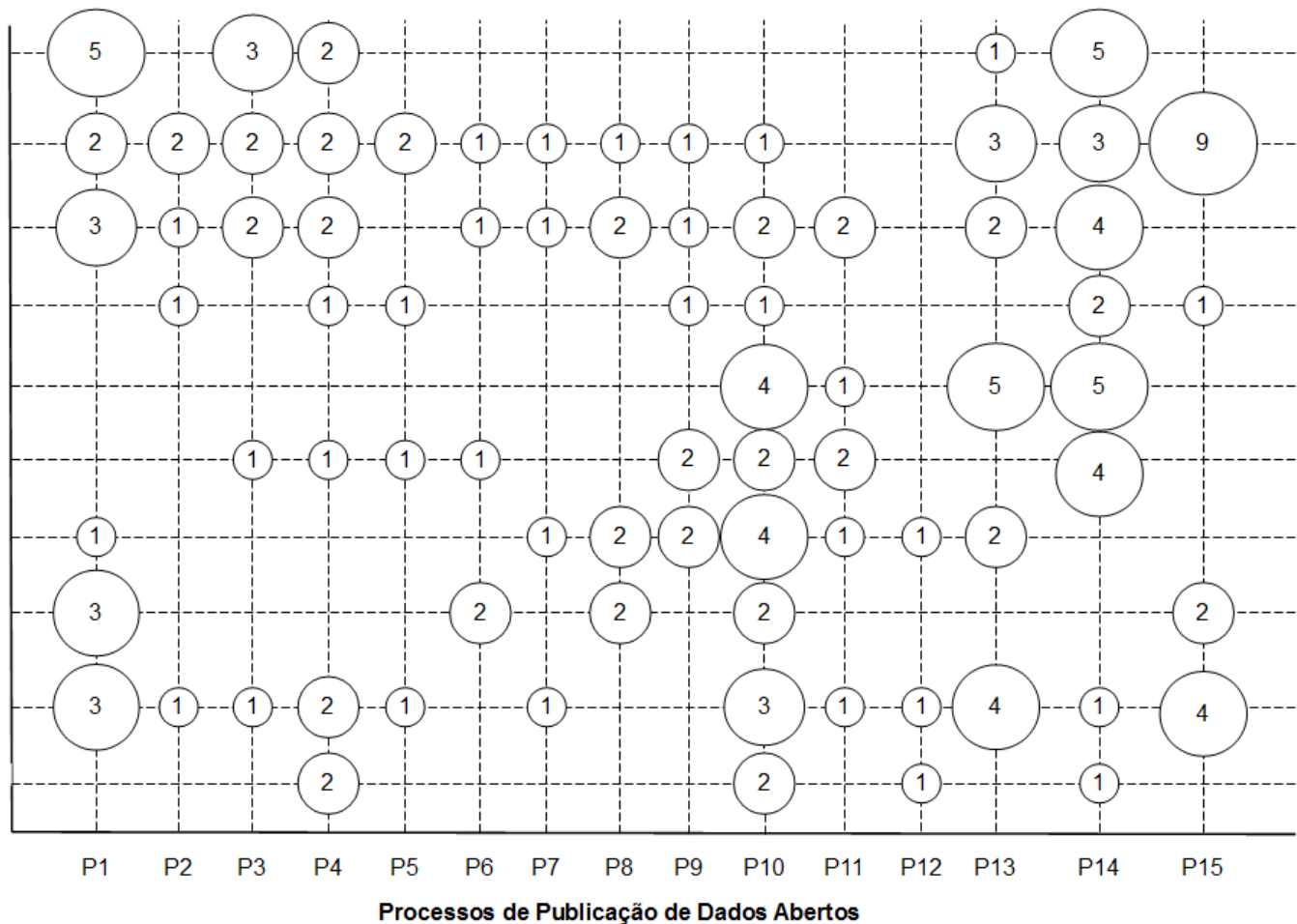


Figura 12 – Ocorrência de recomendações na comparação dos 15 processos de publicação de dados abertos com as BPLDs (Do autor, 2015)

# Extração de recomendações

## Recomendações

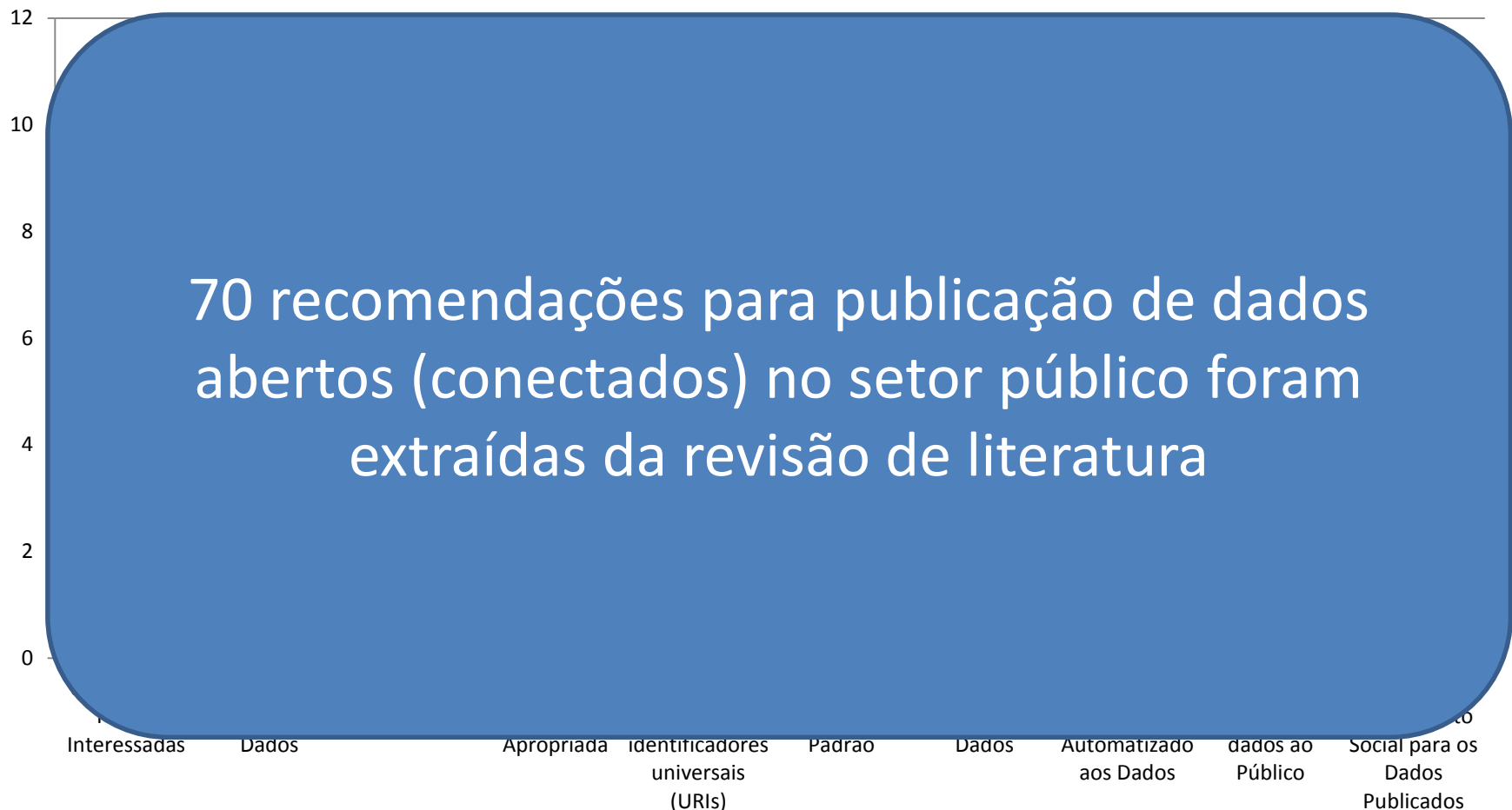


Figura 13 – Sumarização das recomendações identificadas na revisão de literatura (Do autor, 2015)



## *Exemplificando ...*

- ***BPLD 2 -Selecionar Conjuntos de Dados***
  - Estabelecer diretrizes que orientem a priorização de dados a serem abertos (2B)
  - Realizar consultas aos usuários sobre a demanda de dados (2C)
- ***BPLD 8 – Prover acesso automatizado aos dados***
  - Disponibilizar bases completas para download (*dumps*) (8A)
  - Desenvolver um *endpoint* SPARQL (8D)

## Objetivos

- Propor um **modelo de processo para publicação de dados conectados governamentais** que:
  - Considere o nível de maturidade em gestão e publicação de dados das instituições governamentais;
    - Esquema 5-Estrelas dos Dados Abertos
    - Maturidade das instituições publicadoras
  - Apresente atividades para a implementação de Melhores Práticas para Publicação de Dados Conectados;
    - **Melhores Práticas para Publicação de Dados Conectados – W3C**
    - **Recomendações da revisão de literatura**
  - **Seja iterativo e incremental.**

0	Identificar maturidade da instituição publicadora
1	Preparar partes inte
2	Selecionar conjunto
3	Modelar os dados:
4	Especificar uma lice
5	Estabelecer bons id
6	Utilizar vocabulários
7	Converter e enrique
8	Prover acesso auto
9	Anunciar os conjunt
10	Estabelecer um con
11	Retrospectiva e avalia

## 73 Atividades distribuídas entre os 4 processos

- 70 atividades oriundas das recomendações
- 3 novas atividades propostas pelo pesquisador
  - Identificar maturidade
  - Retrospectiva – Características do SCRUM
  - Decisão de Continuidade do Processo – Continuar para o próximo nível ou repetir o nível atual?

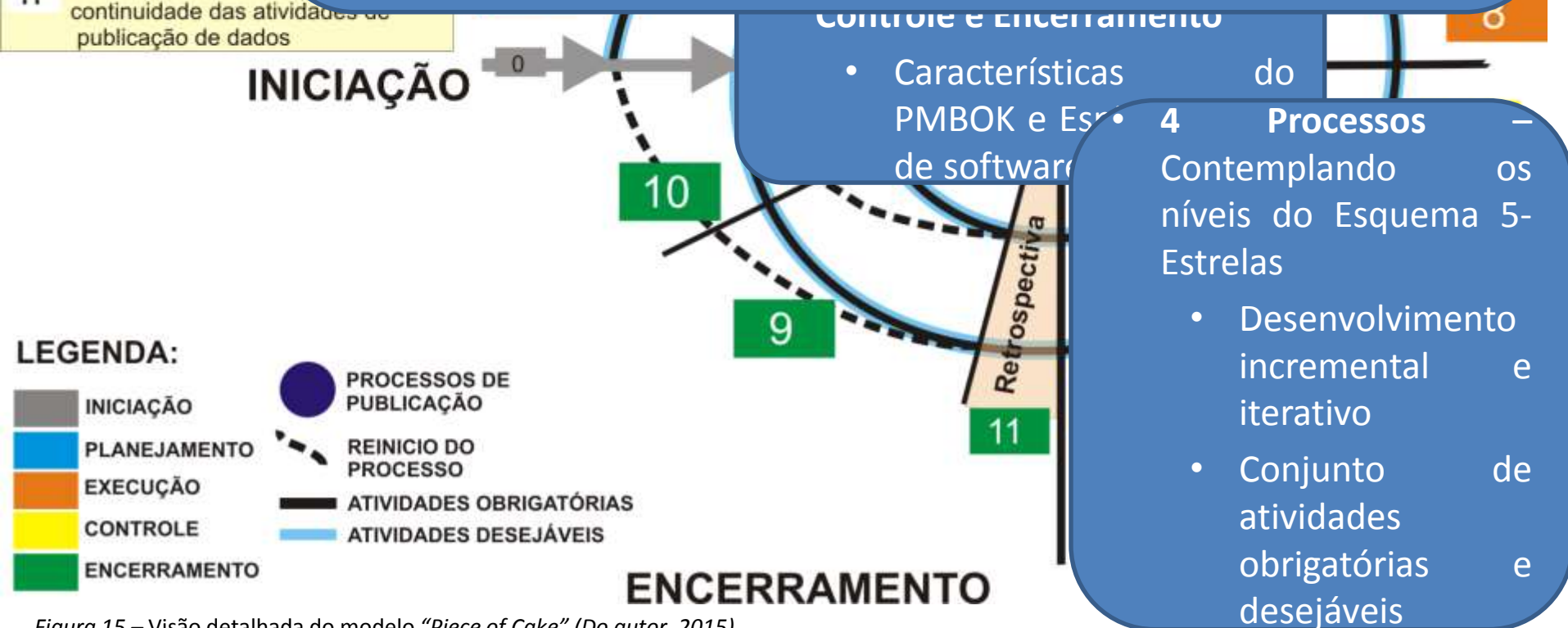


Figura 15 – Visão detalhada do modelo "Piece of Cake" (Do autor, 2015)

# A Proposta - Modelo de Processo “Piece of Cake”

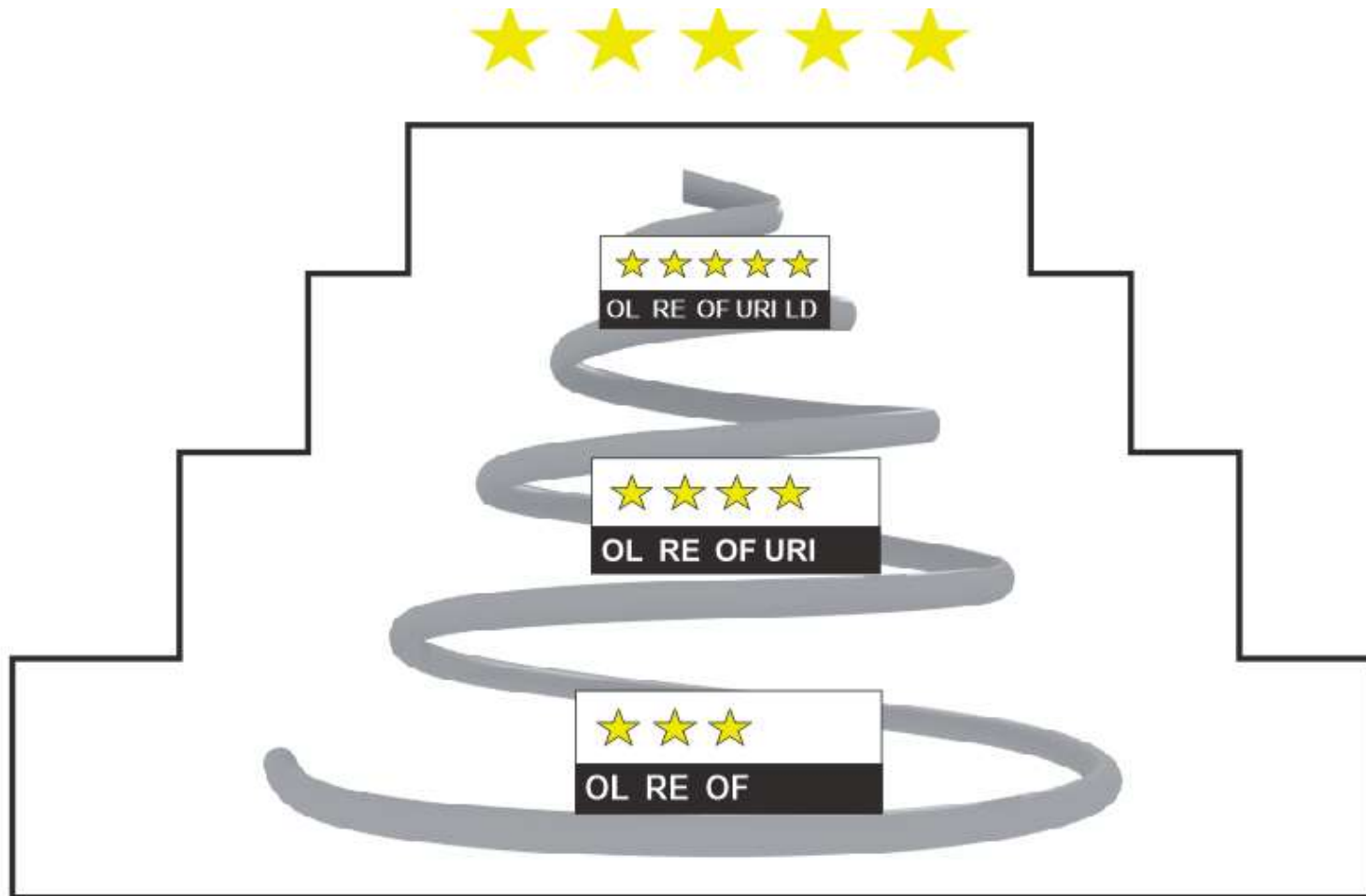


Figura 16 – Evolução do modelo de processo “Piece of Cake” dentre os níveis do esquema 5- Estrelas (Do autor, 2015)

## Objetivos

- Propor um **modelo de processo para publicação de dados conectados governamentais** que:
  - Considere o nível de maturidade das instituições governamentais;
    - Esquema 5-Estrelas dos Dados Abertos
    - Maturidade das instituições publicadoras
  - Apresente atividades para a implementação de Melhores Práticas para Publicação de Dados Conectados;
    - Melhores Práticas para Publicação de Dados Conectados – W3C
    - Recomendações
- **Seja iterativo e incremental.**
  - Espiral invertida com grupos de processos do PMBOK/Fases clássicas da espiral de software



## *Validação*

- **Questionário e Modelo de Regressão Beta para identificação de recomendações mais relevantes**
  - **Objetivo:** apresentar um subconjunto de recomendações mais relevantes para a publicação de dados (abertos e/ou conectados) no setor público, considerando a dificuldade e a relevância para implementá-las;
  - **Método de avaliação:** Modelo de Regressão Beta
  - **Ferramenta de Coleta de Dados: Questionário** com as 70 recomendações para publicação de dados (abertos e/ou conectados) no setor público coletadas na revisão de literatura, para serem avaliadas por especialistas na temática;

## Validação

- Mensuração do **nível de conhecimento dos avaliadores**;
- **Avaliação** dos subconjuntos **de recomendações obrigatórias** estabelecidas pelo pesquisador;
- **Classificação** das recomendações por **difículdade e relevância**.
- **Mensuração da avaliação geral das recomendações por avaliador**.

**HIPÓTESE:** Para a publicação de dados abertos, entendemos que as recomendações abaixo são obrigatórias:

1. Identificar os benefícios para a abertura de dados
2. Identificar as partes interessadas
3. Definir perfil profissional e serem avaliadas

**HIPÓTESE:** Para a publicação de dados abertos conectados, entendemos que as recomendações abaixo são obrigatórias:

### 4. Avaliação

## Avaliação Geral das Recomendações para Publicação de Dados Abertos Conectados

Difíc

Como você qualifica o conjunto de recomendações para dados abertos conectados apresentadas neste questionário de avaliação? \*

Rele

Atribua uma nota entre 1 e 99, onde uma nota maior representa melhor qualidade das recomendações sugeridas.

sugeridas.

ta

Figura 20 – Captura de tela de uma recomendação a ser classificação da relevância e dificuldade mediante o questionário (Do autor, 2015)

## Validação

- **Resultados – Estatística Descritiva:**
  - **Avaliação** dos subconjuntos de **recomendações obrigatórias** estabelecidas pelo pesquisador

### Dados Abertos

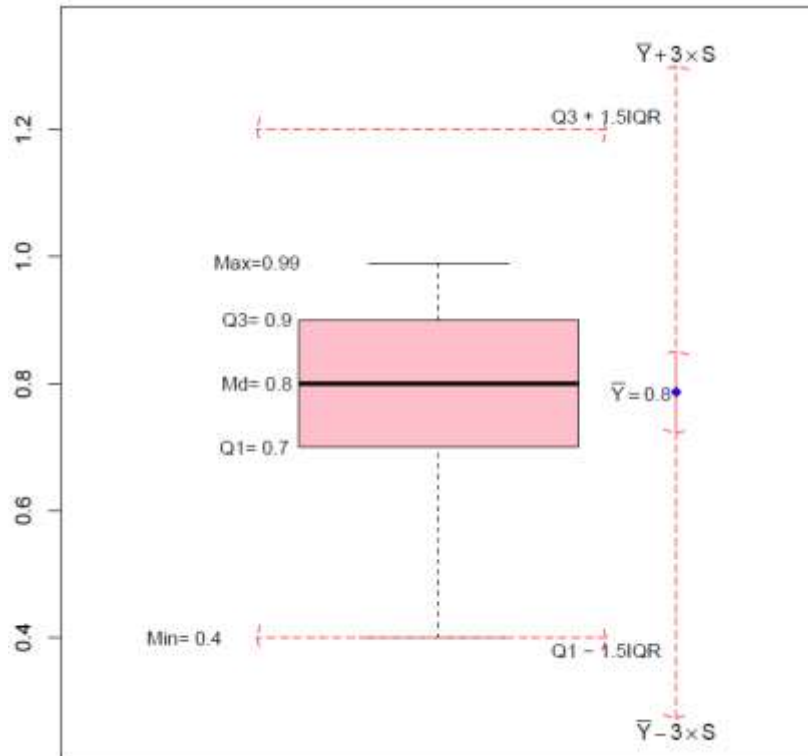


Figura 21 – Gráfico boxplot dos escores adotados para as recomendações para dados abertos pré-estabelecidas como obrigatórias (Do autor, 2015)

### Dados Abertos Conectados

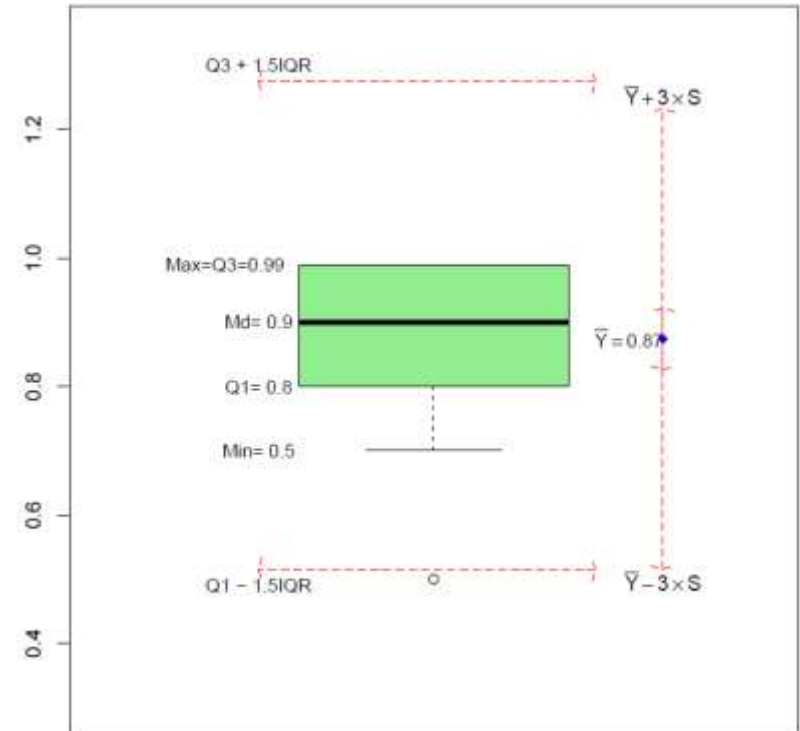
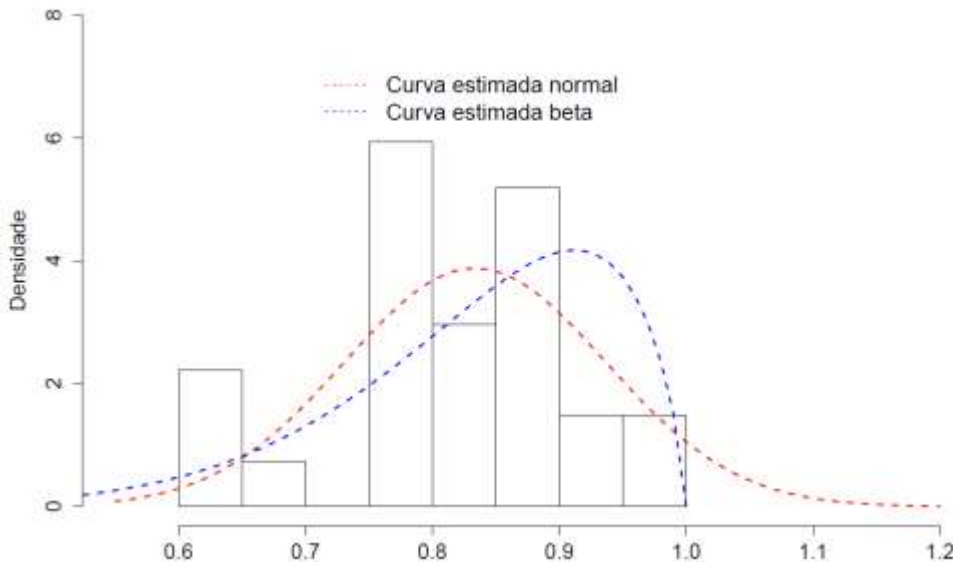


Figura 22 – Gráfico boxplot dos escores adotados para as recomendações para dados abertos conectados pré-estabelecidas como obrigatórias (Do autor, 2015)

## Validação

- Resultados – Estatística Descritiva:
  - Avaliação global das recomendações apresentadas no questionário

Dados Abertos



Dados Abertos Conectados

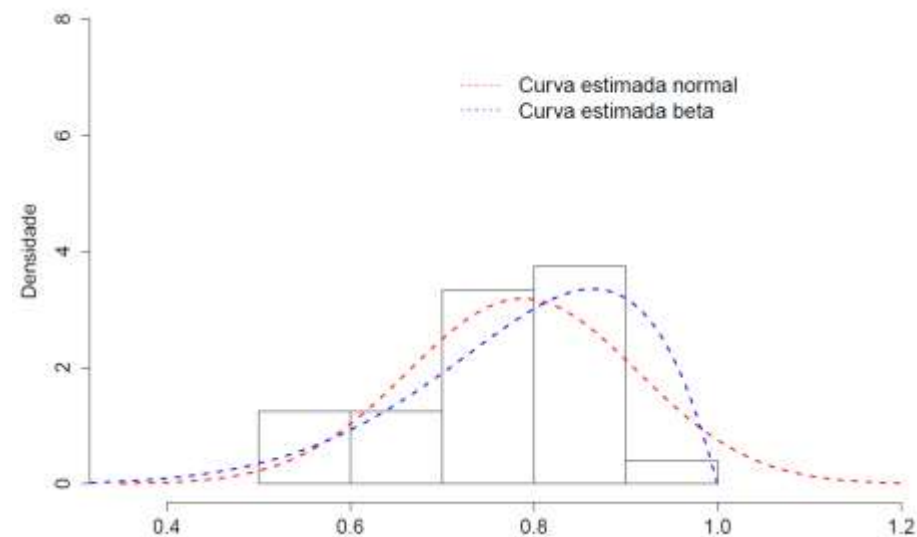


Figura 23 – Gráfico histograma dos escores adotados para a avaliação global das recomendações para dados abertos (Do autor, 2015)

Figura 24 – Gráfico histograma dos escores adotados para a avaliação global das recomendações para dados abertos conectados (Do autor, 2015)

## Validação

- Resultados – Modelo de Regressão:
  - Recomendações sugeridas para publicação de Dados Abertos

Variáveis	Estimativas	p_Valor
ScoreConhecimento	-0.10	0.06
Recomendações		
1A, 1B, 1C, 2B, 2D, 2E, 2H, 3A, 3B, 4A, 6A, 7A, 8A, 9A, 9B, 9H	1.7	0.03
1D	0.10	0.06
2A	0.25	0.00
2C	0.21	0.01
8B	0.12	0.10
9C	0.02	0.11
10A	0.15	0.03
10B	0.14	0.00

Analisar a estrutura organizacional (2A)

Tabela 3 – Recomendações para publicação de dados abertos sugeridas pelo modelo de regressão (Do autor, 2015).



## Validação

### Resultados – Modelo de Regressão:

- Recomendações sugeridas para publicação de Dados Abertos Conectados

Variáveis	Estimativas	p_Valor
ScoreConhecimento	-0.80	0.03
Recomendações		
3F, 5A, 5C, 5D, 7C	0.53	0.00
2K	1.17	0.00
5B	0.70	0.06
6B	0.02	0.08
6D	0.72	0.04
6F	0.15	0.07
9I	1.20	0.00

Disponibilizar dados conectados em servidores de triplas (9I)

Tabela 4 – Recomendações para publicação de dados abertos conectados sugeridas pelo modelo de regressão (Do autor, 2015).

## Validação

- **Classificação das recomendações entre os níveis do Esquema 5-Estrelas**
- **Objetivo:** Orientar a distribuição das recomendações entre os níveis do Esquema 5-Estrelas

**Distribuição das atividades do modelo de processo conforme o Esquema 5-Estrelas**

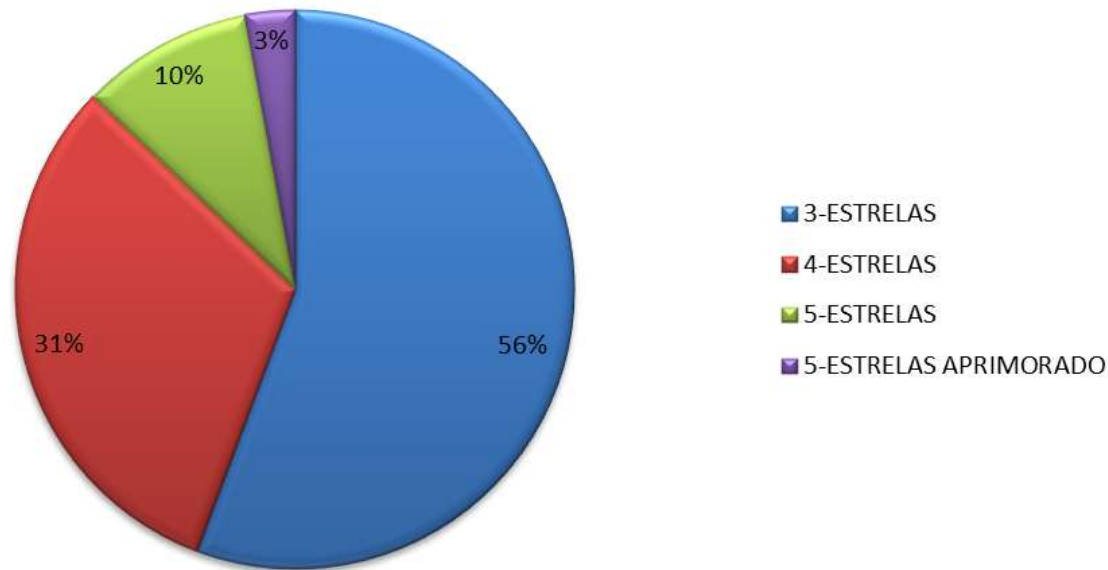


Figura 25 – Gráfico com o resultado da classificação das recomendações entre os níveis do Esquema 5-Estrelas (Do autor, 2015)

## Validação

- **Estudo Empírico**
  - **Objetivo:** Verificar a eficácia do modelo de processo “Piece of Cake” para publicação de dados abertos governamentais (DAG) e dados abertos conectados governamentais (DACG)
    - OBS.: Foram fornecidos dois *datasets* no nível “sem estrelas” para que fossem convertidos para 5-Estrelas
  - **Método:** Estudo empírico em laboratório estruturado mediante a metodologia GQM, contendo 3 objetivos (Goals), 9 questões (Questions) e 15 métricas (Métrics).
  - **Planejamento do Estudo:**
    - **Foco da qualidade:** Verificar a eficácia na publicação de DACG;
    - **Ferramenta de Coleta de Dados:** Planilha contendo a relação de atividades do Modelo “Piece of Cake” e artefatos gerados pelas equipes;
    - **Contexto:** 3 equipes, com níveis de experiência distintos em publicação de dados governamentais, compostas por um profissional e alunos de Pós-Graduação e Graduação.

## Validação

<b>Objetivo:</b>		<b>G3: Publicar Dados Abertos Governamentais seguindo o modelo de processo “Piece of Cake”</b>		
<b>Questão:</b>	<b>Métricas:</b>	<b>EQUIPE A</b>	<b>EQUIPE B</b>	<b>EQUIPE C</b>
<b>Q5: Quantas atividades do modelo de processo foram utilizadas para publicação de DAG?</b>	<b>M7:</b> Percentual de Atividades do modelo utilizadas para publicação de DAG com 3 estrelas	46%	38%	40%
	<b>M8:</b> Percentual de Atividades do modelo utilizadas para publicação de DAG com 4 estrelas	54%	71%	60%
	<b>M9:</b> Percentual de Atividades do modelo utilizadas para publicação de DAG com 5 estrelas	53%	63%	40%

### Balanço Geral:

Equipe A utilizou 28 (40%). Equipe B utilizou 38 (54%). Equipe C, menos experiente, utilizou 42 (60%).

# Dados Abertos Governamentais Conectados

terça-feira, 20 de outubro de 2015

## Lei de Acesso à Informação (LAI)

A Lei de Acesso à Informação (LAI) estabelece que qualquer cidadão pode solicitar informações a órgãos e entidades do Poder Executivo Federal.

Mais Informações podem ser consultadas em:

[www.acessoainformacao.gov.br](http://www.acessoainformacao.gov.br)

Postado por **Josmário Albuquerque** às 12:22 Nenhum comentário:



Recomende isto no Google

## Dados Abertos do Estado de Alagoas

Cadastro de Empresas Inidôneas ou Suspensas no Estado de Alagoas – CEIS/AL

[Baixar Arquivo CSV](#)

[Baixar Arquivo Torrent](#)

[Baixar Arquivo RDF/XML](#)

Fonte: Controladoria Geral do Estado de Alagoas

SALÁRIOS DOS SERVIDORES DO ESTADO DE ALAGOAS - JANEIRO / 2015

[Baixar Arquivo CSV](#)

[Baixar Arquivo Torrent](#)

Fonte: Diretoria de Teatros do Estado de Alagoas

Quem sou eu



**Josmário Albuquerque**

[Seguir](#)

124

[Visualizar meu perfil completo](#)

Arquivo do blog

▼ 2015 (2)

▼ Outubro (2)

[Lei de Acesso à Informação \(LAI\)](#)

[Dados Abertos do Estado de Alagoas](#)



## Validação

Objetivo:	G3: Publicar Dados Abertos Governamentais seguindo o modelo de processo “Piece of Cake”	
Q6: Quantas atividades do modelo de processo foram utilizadas por todas as equipes?	<b>M10:</b> Atividades do modelo utilizadas pelas 3 equipes	<b>12</b>
Q7: Quantas atividades do modelo de processo não foram utilizadas por todas as equipes?	<b>M11:</b> Atividades do modelo utilizadas por 2 equipes	30
	<b>M12:</b> Atividades do modelo utilizadas por 1 equipe	13
	<b>M13:</b> Atividades do modelo não utilizadas	15

Tabela 6 – Resultados para as métricas M10, M11, M12 e M13 do estudo empírico (Do autor, 2015).

## *Atividades do Modelo de Processo*

Definição das atividades obrigatórias e desejáveis a partir da validação empírica:

**A**

**Modelo de Regressão**

- Obrigatórias: 33 (de 70)

**B**

**Estudo Empírico**

- Obrigatórias: 12 (de 70)

$$A \cup B = A + B - A \cap B$$

$$= 33 + 12 - 8$$

**Total de atividades obrigatórias: 37 (de 70)**

# Atividades do Modelo de Processo

PROCESSOS:	A	B	C	D
ETAPA 0:	0A			
ETAPA 1:	1A, 1B, 1C, 1D, 1E, 1F			
		Obrigatórias	Desejáveis	Total
Processo A (Para 3-Estrelas)		21	18	39
Processo B (Para 4-Estrelas)		12	10	22
Processo C (Para 5-Estrelas)		3	4	7
Processo D (Para 5-Estrelas aprimorado)		1	1	2
Subtotal		<b>37</b>	<b>33</b>	<b>70</b>
Atividade Inicial		1	0	1
Comuns para todos os processos:		2	0	2
Totais		<b>40</b>	<b>33</b>	<b>73</b>
ETAPA 9:	9A,9B,9C, 9D,9E,9F,9G		9I, 9H	
ETAPA 10:	10A,10B, 10C,10D,10E			
ETAPA 11:	11A,11B			

## *Ameaças à Validade*

- **Questionário e Modelo de Regressão:**
  - **Cansaço dos avaliadores ao responder o questionário**, que por ser extenso, pode ter levado ao cansaço de alguns avaliadores ao longo de sua resposta, o que pode ter prejudicado a qualidade de algumas respostas;
  - **Respostas emitidas por avaliadores inexperientes**, pelo questionário estar online. Entretanto, tal ameaça consegue ser tratada pelo escore que avalia o nível de conhecimento dos avaliadores (ScoreConhecimento), que considera a auto-avaliação do conhecimento dos avaliadores para sugerir as atividades mais relevantes ao modelo.

## *Ameaças à Validade*

- **Estudo Empírico:**
  - **Utilização de apenas dois conjuntos de dados;**
  - **Avaliadores escolhidos subjetivamente, onde não houve** medição para classificar corretamente os participantes, o que pode comprometer os resultados;
  - **Cansaço dos avaliadores durante o estudo, pois houve** preparação inicial, análise de um documento extenso com as recomendações/atividades para publicação de DACG, e ainda, o desenvolvimento das atividades para a publicação dos dados. Assim, o volume de trabalho foi significativo o que pode ter influenciado no desempenho de algumas equipes;
  - **Todos os participantes do estudo tem formação em computação, o que pode ter reduzido a complexidade para a execução do estudo.** Uma maior variação na formação dos participantes seria desejável para uma maior aproximação de um cenário real;
  - **Ambiente de desenvolvimento do estudo em laboratório.**

# *Conclusões e Trabalhos Futuros*

- **Contribuições da pesquisa**

- Instrumento para implementação e disseminação das Melhores Práticas para Publicação de Dados Conectados do W3C;
- Modelo de Processo que permite que instituições em estágios variados de maturidade em produção de dados possam utilizá-lo para guiar suas tarefas de publicação de dados governamentais;
- Identificação de limitações nas abordagens existentes para publicação de Dados Abertos e Dados Abertos Conectados no setor público.



# *Conclusões e Trabalhos Futuros*

## • **Trabalhos Futuros**

- Aprimorar a proposição de novas atividades para os processos do modelo;
  - Especialmente para os processos B, C e D
  - Desenvolver uma revisão de literatura mais ampla;
- Simplificar e popularizar processos e atividades de publicação de DACG;
- Desenvolver ferramentas e softwares que simplifiquem a publicação de dados abertos conectados governamentais;
- Validar o modelo de processo atual com estudos de caso reais, preferencialmente em cenários e países distintos para verificar a validade do modelo em contextos técnicos, organizacionais e políticos distintos;
- Melhorar o detalhamento das atividades dos processos do modelo
- Discutir sobre o desenvolvimento de processos específicos às necessidades particulares do setor público.
  - Processos de publicação de DACG aplicáveis a contextos particulares como a saúde, educação, energia e recursos hídricos, geociências, dentre outros.

## Referências

- [1] W3C BRASIL. Manual dos dados abertos: governo (2011). Acessado em: out. 2015. Disponível em: <[http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual\\_Dados\\_Abertos\\_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf)>.
- [2] Banisar, David (2013). List of Countries with Access to Information. Acesso em: nov. 2015. Disponível em: <http://home.broadpark.no/~wkeim/foi-list.htm>
- [3] Berners-Lee, Tim. (2006). Linked Data. Acesso em: jul. 2015. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>
- [4]. W3C (2014) Best Practices for Publishing Linked Data. W3C. Acesso em: nov. 2015. Disponível em: <http://www.w3.org/TR/ld-bp>
- [5] Wood, David. et al (2013). Linked Data: Structured data on the Web. [S.l.]: Manning Publications, 336 p. ISBN 9781617290398.
- [6] Heath, Tom; Bizer, Christian. (2011) *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool.
- [7] COLOMBIA. Guía para la apertura de datos en Colombia. (2012). 67 p. Disponível em: [http://programa.gobiernoonlinea.gov.co/apc-aa-files/da4567033d075590cd3050598756222c/Datos\\_Abiertos\\_Guia\\_v2\\_0.pdf](http://programa.gobiernoonlinea.gov.co/apc-aa-files/da4567033d075590cd3050598756222c/Datos_Abiertos_Guia_v2_0.pdf).
- [8] Hyland, Bernadette; Wood, David. (2011) The Joy of Data - Cookbook for Publishing Linked Government Data on the Web. Acesso em: mar. 2015. Disponível em: [http://www.w3.org/2011/gld/wiki/Linked\\_Data\\_Cookbook](http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook).
- [9] Villazón-Terrazas, Boris; Vilches-Blázquez, Luis. M.; Corcho, Oscar; Gómez-Pérez, Asunción (2011). Methodological Guidelines for Publishing Government Linked Data. Linking Government Data. pp 27-49

**Muito obrigado !!!**



**Thiago José Tavares Ávila**

Mestrando em Modelagem Computacional do Conhecimento – IC/Ufal  
Núcleo de Excelência em Tecnologias Sociais - NEES