Use of multidimensional data models to increase active public transparency

Uso de modelos de dados multidimensionais para a ampliação da transparência ativa

Ricardo Ceśar Gonçalves Sant'Ana *

Fernando de Assis Rodrigues **

ABSTRACT

Environments for access to government data, via Information and Communications Technologies, may increase possibilities for citizens overseeing it, providing feedback for future demands. The aim of this study is to identify, in the available data by active transparency, the existence of elements that allow for the creation of new proposals for dimensional models, enabling anticipation of demands on data access. The theoretical-methodological framework, the text uses the concepts of Citizen Intelligence and Business Intelligence. As a result, a dimensional model was proposed, building on a dimensional model from a daily expenses query, available in the Transparency home-page of the Brazilian Federal Government.

Keywords: Public Transparency, Information and Communication Technologies, Collecting Data Collection, Citizen Intelligence, Data Warehouse.

RESUMO

Ambientes para acesso a dados governamentais, via Tecnologias de Informação e Comunicação, podem ampliar possibilidades de acompanhamento pelo cidadão, retroalimentando futuras demandas. O objetivo deste estudo é identificar nos dados disponíveis via transparência ativa, a existência de elementos que permitam a elaboração de propostas de modelos dimensionais, propiciando a antecipação de demandas de acesso a dados. Como referencial teórico-metodológico, o texto utiliza os conceitos Business Intelligence e Citizen Intelligence. Como resultado, foi elaborada a proposta de um modelo dimensional a partir da consulta de despesas diárias, disponível no Portal de Transparência do Governo Federal.

Palavras-chave: Transparência Pública, Tecnologias de Informação e Comunicação, Coleta de Dados, Citizen Intelligence, Data Warehouse.

---

* Department of Information Science, São Paulo State University (UNESP), Marília, Brazil. PhD. in Information Science. E-mail: ricardo.santana@unesp.br ORCID: https://orcid.org/0000-0003-1387-4519. Departamento de Ciência da Informação, Universidade Estadual Paulista (UNESP), Marília, Brasil. Doutor em Ciência da Informação. E-mail: ricardo.santana@unesp.br ORCID: https://orcid.org/0000-0003-1387-4519.

** Instituto de Ciências Socais Aplicadas, Universidade Federal do Pará, Belém, Brasil. Doutor em Ciência da Informação. E-mail: fernando@rodrigues.pro.br ORCID: https://orcid.org/0000-0001-9634-1202. Institute of Applied Social Sciences, Federal University of Pará (UFPA), Belém, Brazil. PhD. in Information Science. E-mail: fernando@rodrigues.pro.br ORCID: https://orcid.org/0000-0001-9634-1202.

INTRODUCTION

The transparency of government actions to society, through access to State data and information, is an integral part of discussions on trends in modernizing public administration models, mainly in representative democracies - where citizens elect their representatives, directly or indirectly, for the composition of executive and legislative powers.

This new model of public administration:

> [...] seeks to redistribute coordination skills and resources between different institutional and organizational levels, governmental and non-governmental, allowing institutional pluralism in public functions, contrary to the old model of state monopoly (MALIN, 2006, p. 1, translated by the authors).

As established by the aforementioned model, the transparency of the State's activities to society has as one of its premises to strengthen the participation of citizens in public administration. This strengthening can be guaranteed with the construction of democratic environments that, among other characteristics, create possibilities for new information flows between the State and society, thus guaranteeing greater visibility on the management of government affairs. In representative democracies, expanding these mechanisms of civil society control over public administration means going beyond voting, that is, it is about the commitment to create conditions for social monitoring in public administration beyond electoral processes (BOHMAN, 1996).

Environments for government data access, through Information and Communication Technologies (ICTs), can expand possibilities for citizen action, feedback to the user's demand for government information and also create additional motivation for transparency in the involved parties within the state sectors (SANT'ANA, 2008; SANT'ANA, 2009).

The information resources maintained by the government in digital information environments can be analyzed under two objectives for its usage: a) for internal use, which can support the internal processes of administration and management of government affairs, which are used by employees linked to the government, and b) for external use, resulting from the application of ICTs in environments for interaction between public administration and external entities, such as companies, citizens (or groups of citizens) or even other public organizations (SANT'ANA, 2009).

The advantages in the adoption of technological resources as a support tool for public management are evident and based on several studies in areas such as Administration, Information Science, and Computer Science. Nevertheless, government agencies do not explore the numerous possibilities offered in the use of ICTs for external use. Part of the official websites of the Federal Government uses technological resources only to make available information such as bulletin boards; bureau opening hours; institutional pages; to share legislation, ordinances, and guidelines; to share phone and e-mail directories (RODRIGUES, 2012).

Thus, the exploration of external use potential of the ICTs in order to intensify the relationship between State and society may become a key element for a more direct and efficient relationship – providing an increase in the transparency of the State to society, as well as the repercussion of its actions, which strengthens democracy and encourages the mobilization of groups in the defense of collective rights and interests and, also, to the citizen be able to establish a critical reflection, minimizing informational asymmetries in the relationship between society and government (SANT'ANA, 2008; SANT'ANA, 2009).

The Brazilian Federal Constitution of 1988, fifth article, in item XXXIII highlights:

> [...] everyone has the right to receive information of their particular interest from public agencies, or of collective or general interest, which will be provided within the term of the law, under penalty of liability, except for those whose confidentiality is essential to the security of society and of the State (BRAZIL, 1988, p. 1, translated by the authors).

Also, the Brazilian Federal Constitution in the two hundred sixteenth article, second paragraph, states that the public administration must create mechanisms for consultation with those who need their documents.

> The public administration is responsible, in the form of the law, of the management of governmental documentation and the measures to open its consultation to those who need it (BRAZIL, 1988, p. 1, translated by the authors).

On November 18th, 2011 was approved the Federal Law n. 12.527/2011, applied on May 16th, 2012, turning obsolete the Federal Law n. 11.111/2005. This Law, also known as the Access to Information Law (known as the acronym *LAI*, in the Portuguese language), added new obligations to public institutions (BRAZIL, 2005; BRAZIL, 2011).

One of the important subjects that the *LAI* contextualizes is the confidentiality of governmental information: access to any governmental information is guaranteed; and secrecy becomes an exception - as opposed to previous approaches (BRAZIL, 2011).

Regardless of which sphere they are linked to, Government agencies must adopt two models for achieving the transparency of their data. The first is the active form - when government agencies must make data sets available for free access to the citizens; and the second is the passive form - in which government agencies must offer mechanisms to receive requests from citizens for access to data sets not actively available. In both models, availability is not carried out in a concentrated manner. Each Government agency must develop mechanisms to meet the legal aspects linked to active and passive transparency on its official websites, following guidelines found in Brazilian legislation.

Additionally, the legal aspects related to the promotion of access to government data are important, as well as the creation of elements that promote the intensive use of ICTs in the development of applications that assist the process of visualization of

these data - because the visualization of data is one of the components that drive public transparency (FIERRO; GIL-GARCIA, 2012).

The studies related to data visualization are multidisciplinary. They come from different areas of knowledge, such as Design, Computer Science, Information Science, among others. Data visualization is defined as "A tool or method for interpreting image data fed into a computer and for generating images from complex multi-dimensional data sets" (MCCORMICK *et al. apud* OWEN, 1993, p. 340).

One of the most important issues in the elaboration of data visualizations is the mapping of attributes from the selected data for the posterior elaboration of a graph, as simplified as possible, but effectively guaranteeing the transmission of the informational content of the data (SENAY; IGNATIUS, 1999). However, one data set presents several possibilities for mapping its attributes, leading to different models and techniques for the construction of the visualization. In this sense, the same data set can have multiple views, depending on which attributes are selected for the elaborations.

This multiplicity of data visualization studies is due to its very essence. For example, in face of a designing and generating a graph, a summary, a frame, or a table from government data sets, a number of steps are required, such as the data collection in the information resource, the data storage on a database, the treatment, and processing of data, the choice of graphic elements, colors, among other factors. In addition to aesthetic and computational factors, the target audience for the visual construction of government data is also important.

Brazilian legislation already provides for a simplification of the data visualization process:

> It is the duty of the State to guarantee the right of access to information, which will be franchised, through <u>objective and agile procedures, in a transparent, clear, and easy-to-understand language</u> (BRAZIL, 2011, p. 1, emphasis added by the authors, translated by the authors).

The *LAI* also makes it mandatory to use the internet infrastructure as an instrument for sharing and accessing government data.

> [...] public agencies and entities must use all legitimate means and instruments at their disposal, being mandatory the disclosure on official websites of the world wide web (internet) (BRAZIL, 2011, p. 1, emphasis added by the authors, translated by the authors).

In November 2004, the Brazilian Comptroller General of the Union (2004) initiated the activities of the website called the Brazilian Federal Government's Transparency Portal, in order to ensure access to data about the application of public financial resources, with the objective of increasing the transparency of public management to citizens, so that can follow the destination of public financial resources and assist in the inspection process.

However, Government websites that provide government data sets through active transparency are prepared in an attempt to meet the information needs of any interested party, in an opposed sense of elaborated applications or web sites developed to serve specific audiences, such as an application to assist the management process of an industry or an application for managing retail sales. This factor is that makes complex the construction of data visualizations from the collection of government data sets since the tables or graphs present in the Government websites will not always meet all the informational demands of society.

Summarizing, it is important to highlight the urgency for the existence of graphics, summaries, charts, and other visualization artifacts in the Government websites. Also, these websites must contemplate the possibility of collecting primary data - the most granular as possible, with the highest level of detail - to create opportunities for the construction of new perceptions and visualizations about government data, based on the elaboration of a previous structure that allows visualizations to be created outside the Government websites - mainly due to the plurality of the public interested in this type of information.

> Primary Data Access is considered to be one of the success key factors in the construction of alternatives for accessing data, avoiding eventual misinterpretation and still allowing varied paths of analysis due to the high degree of freedom of combinations and possible processing that a primary data offers (SANT'ANA; RODRIGUES, 2013, p. 50, translated by the authors).

Therefore, the objective of this research is to identify in the government data available via active transparency, the existence of elements that allow the elaboration of proposals for dimensional models based on the concept of Data Warehouse, enabling the anticipation of demand for data access.

As a theoretical-methodological framework, this text makes use of concepts such as Business Intelligence, Data Warehouse, Data Mart e Citizen Intelligence. As a result, the proposal for a dimensional model was elaborated based on the query 'Daily Expense Details', available on the Federal Government's Transparency Portal.

BUSINESS INTELLIGENCE

In business environments, managers can count on the assistance of processes to transform data sets into information to support their decisions. The set of theories, methodologies, and technologies to assist the decision-making process is the concept known as Business Intelligence (BI), that is directly related to the decision making itself; to the use of Database Management Systems (DBMS) that allow for the creation of a database that is compatible with a Data Warehouse (DW) for generating reports and data analysis; and pattern-finding processes in data sets stored by computer programs, such as data mining (BARBIERI, 2001).

Data Warehouse is a subject-oriented database that is integrated, variable in time, non-volatile in helping the management decision-making process (INMON, 1992).

Unlike the transactional database, used and structured for use and support to the daily activities of a company or public agency, the DW is structured with a specific focus to carry out consultations and analyses oriented to a specific demand (KIMBALL, 1996).

Furthermore, private institutions can make use of the DW to improve the decision-making process: agencies, divisions, government entities, and development agencies, among other areas, can adopt the DW to support the decision. For example, Mussi (2005) concludes that, by adopting a DW, the Brazilian National Health Surveillance Agency (ANVISA) creates the possibility to seek and to interpret

> [...] stored information, ensuring greater accuracy in decision making; it allows establishing an integrated and collaborative approach to capture, to create, to organize and to use all the information assets of a company; can anticipate sudden changes in the market; to understand the business; have a clear view on new businesses; to assist in the implementation of new management tools; to have knowledge about new technologies, products or services that may, directly and indirectly, affect your business (MUSSI, 2005, p. 6, translated by the authors).

For the elaboration of a database in a dimensional model, proposed by the Data Warehouse theories, it is important to determine and identify two terms: fact and dimension. DW is formed by multiple tables, each representing facts and dimensions. Each interconnected set of facts and dimensions is called Data Mart.

The Data Mart (DM) is a flexible data set based on data in its most granular state, extracted from a transactional database, presented in a dimensional model more adaptable to user queries than in a transactional database. The DM represents data from a single business process for an institution, such as sales (KIMBALL; ROSS, 2002).

For example, a transactional database for a company in the retail industry has a table for storing sales data (Figure 1). This table contains several attributes, and in each of these attributes, a data type is allocated, such as (i) the unique numeric identifier for each sale (as the primary key), (ii) the date it was made, (iii) the total amount, (iv) the information about the employee who made the sale and (v) the information about the department of the store.

Figure 1 – Example of sale table in a transactional database

	Sale
PK	**Sale ID**
	Date
	Total
	Employee
	Department

Source: Authors.

In the moment of the beginning of the elaboration of a dimensional model from a transactional database of the aforementioned hypothetical company, it is required to identify the attributes that are facts. A fact table is a primary table, the central table, and the starting point, prerequisite to be elaborated in a dimensional model, where the quantifiable attributes must be stored. For example, the elaboration of the sale fact table contains the following attributes: the total sale value; and the unique numeric identifiers (foreign keys) that will represent each employee, each department, and each sale date (Figure 2). These unique identifiers are necessary to link each data record contained in the sale fact table with the tables containing information about the departments, the sales dates, and the employees (KIMBALL; ROSS, 2002).

Figure 2 – Example of a sale fact table in a dimensional model

	Sale (fact table)
FK 1	**Date ID**
FK 2	**Employee ID**
FK 3	**Department ID**
	Total

Source: Authors.

Just like the fact table, dimension tables are essential components for the functioning of a DM. The dimension tables contain the textual descriptions for each unique identifier (foreign keys) found in the fact table. Each of these foreign keys must relate to a dimension table.

For example, the employee's unique identifier must be directly linked to a dimension table called employee, and that dimension table must contain not only the descriptive element that represents his name but also a much broader context: descriptions such as his gender, his age, the shift in which exercise his function, the personal IDs, among others (Figure 3).

Figure 3 – Example of the employee dimension table in a dimensional model

	Employee (dimension table)
PK	**Employee ID**
	Name
	Gender
	Age
	Shift
	Personal ID
	...

Source: Authors.

The dimension tables are the input interfaces for data queries in a DM. The attributes of these tables will serve as a primary resource in the queries to be performed (KIMBALL; ROSS, 2002).

Finally, the union of the relationships between all tables that represent the dimensions with the table that represents the facts generates the dimensional model (Figure 4). This representation model, with the fact table in the center and the dimension tables around it, is known as a star scheme, adopted in this study.

Figure 4 – Example of a dimensional model in a star scheme
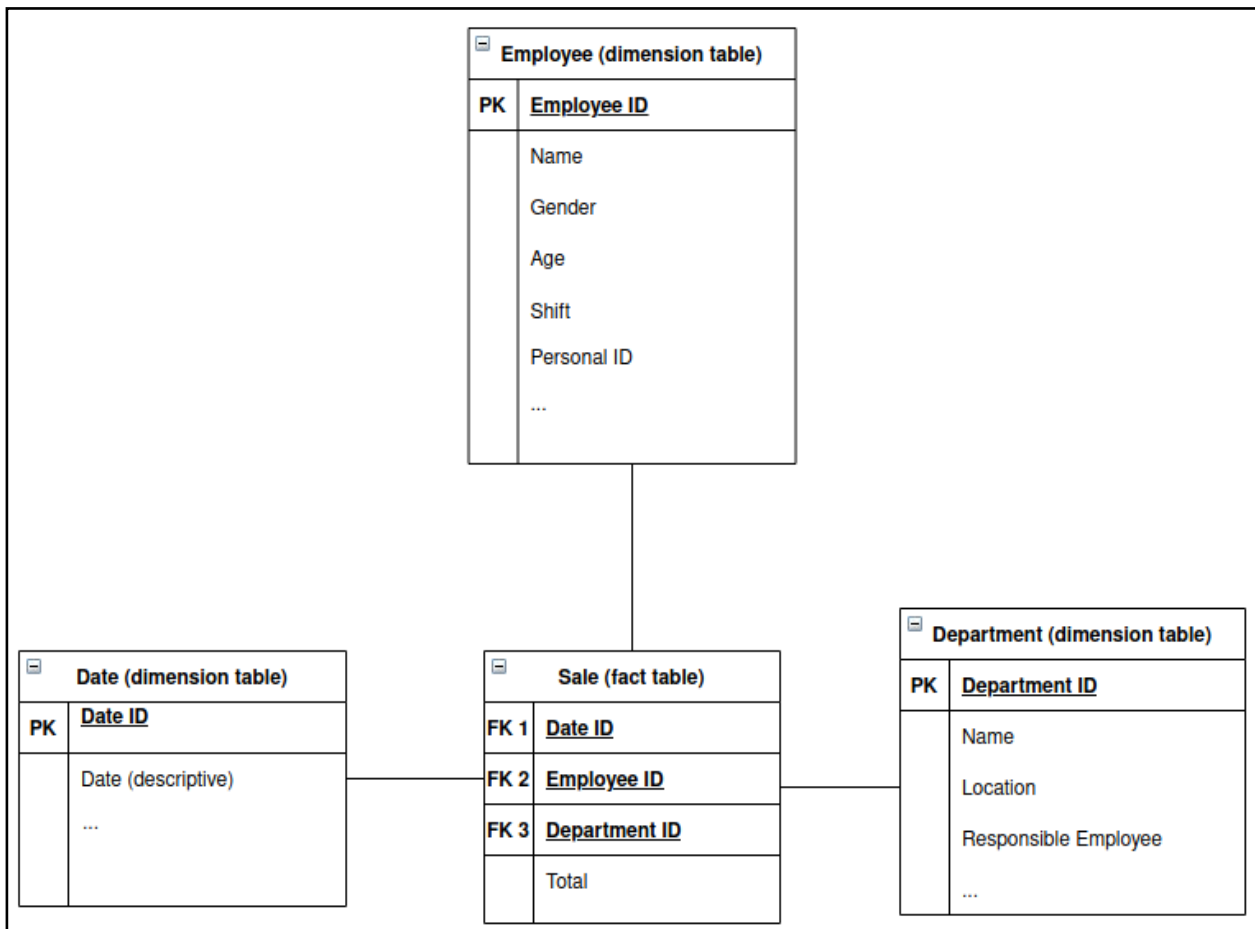
Source: Authors.

Making queries to meet specific objectives for decision making is simplified with the developed dimensional model because the decision-making process starts from the attributes found in each dimension. For example, to retrieve data on the sales amount sold only by employees who work in a given shift, it is necessary to start querying the sale fact table from the employee dimension table - avoiding the use of unnecessary data from the date and the department dimensions to answer. In another perspective, to answer the query of the total sales on a given day, the only requirement is to use the results from the sale fact table from the date dimension table. It is important to emphasize that the success of the results of the queries in a dimensional model is linked directly to the formulation of the right questions.

One of the applications of the dimensional model in organizations occurs in an attempt to anticipate demand from information stored in companies' databases. With the dimensional model, it is possible to apply techniques to recognize patterns or needs of a specific action in the short or long term. With this, managers can predict events and prepare strategies to make the best use of future opportunities or avoid crises.

The process of developing a dimensional model may be applied in contexts such as consultations on government data designed to meet the demand for active transparency.

CITIZEN INTELLIGENCE

With new possibilities offered in collaborative environments endeavored at strengthening the Society-State relationship, the concept of Citizen Intelligence (CI) emerges. Citizen Intelligence is the set of theories, information sources, and methods that determine necessary elements in the construction of mechanisms to support citizens' decision-making based on government data. However, unlike BI which is used as a support to the internal decision-making of an institution, the Citizen Intelligence is completely open and transparent not only for the involved parties but also to society as a whole (acting as the final users); and its results are focused in the public and collective interest and not in the interests of a specific government agency or an industry or company management group. Because it is transparent, the process and result of Citizen Intelligence can be shared without restrictions, as there is no secrecy in the methods for the elaboration of the process and the results obtained (P2P FOUNDATION, 2011).

When a particular group or citizen has an interest in a set of government data, that interest comes from reflection - in one issue. For example, a group may be interested in quantifying financial resources spent in a municipality in public health, or the daily amount is spent on projects daily by the Federal Government. This type of problem is concomitant with the questions that a DW, DM, and dimensional model may try to answer, but in another scenario.

Hence, government data collecting is not performed by itself (collect only to collect, without purpose): the collection of government data presupposes answering a question originating from citizens or groups.

The elaboration of a star scheme model to carry out queries on government data sets can assist in this data collection process, allowing to answer future questions from citizens or groups. The elaboration of this model inherited to the idea of active transparency is an anticipation demand: make government data sets available on Government websites before citizens request access via passive transparency.

In this scenario, the use of BI and DW significant differential is the need to anticipate future demands from managers, and, in a related way, Citizen Intelligence also requires having a mapping of the dimensions so that a structure capable of meeting these future data access demands.

DIMENSIONAL MODEL FOR TRANSPARENCY

The Brazilian Federal Government's Transparency Portal has an information resource called Expenses, where it is possible to monitor the daily information regarding the

financial transfer of federal funds to states and municipalities, and payment of services and materials.

Within this information resource, the Daily Details of Expenses presents "[…] detailed and daily updated data on the acts practiced by the managing units of the Federal Executive departments during the execution of their expenses" (BRAZIL, 2004, p. 1, translated by the authors).

In the current query form (Figure 5), it is possible to refine the results from seven filters, namely: Period, Expense Phase, Related Agency, Agency / Linked Entity, Management Unit, Expense, and Recipient.

Figure 5 – Forms of refinement in the Daily Details of Expenses query

Source: Brazil (2006).

An important feature in this refinement is the restrictions of date periods to be consulted: the results of the query only return data for periods of a maximum of 31 days if a higher organ is selected, or return data of only 1 day if a superior organ is not selected. This restriction is necessary for the current adopted query model because, as it is an access point to the data contained in a transactional database, the amount of Federal Government expense records daily generated is large.

Additionally, it is only possible to query expenditure data from Recipient if the citizen knows in advance the number of the National Register of Legal Entities (*CNPJ - Cadastro Nacional da Pessoa Jurídica*, in the Portuguese language) or the Register Identification of Individuals (*CPF - Cadastro de Pessoas Físicas*, in the Portuguese language), or the Register of Management Unit (*UG-Gestão*, in the Portuguese language). The Register of Management Unit is necessary when the destination of the expenditure is a government agency, such as a city hall or a city council.

Therefore, if a citizen wishes to identify expenses for a specific agency in an annual period, he will have to carry out 12 consultations to attend his goal. If the goal of the citizen is to find the total amount spent in the annual period by this specific agency, he will have to calculate the total amount outside the official website, as the result of the query is a list increasingly and ordered by date, divided into pages with 15 records each, containing the attributes: Date, Phase, Document, Species, Related Agency, Agency / Linked Entity, Management Unit, Expense, Recipient, and Value (Figure 6).

Figure 6 – Results of the Daily Details of Expenses query in list form

Resultado da Consulta

Data	Fase	Documento	Espécie	Órgão Superior	Órgão / Entidade Vinculada	Unidade Gestora	Elemento de Despesa	Favorecido	Valor (R$)
11/07/2013	Pagamento	2013NL000001	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	COORD-GERAL PROG. DESENV. SUBMAR. NUCLEAR	MATERIAL DE CONSUMO	DCN CHERBOURG	345,662,590.15(*)
11/07/2013	Pagamento	2013OB800237	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	SECRETARIA DE ECON. E FINANCAS DA AERONAUTICA	PRINCIPAL CORRIGIDO DIVIDA CONTRATUAL RESG.	BBVA-BANCO BILBAO VIZCAYA ARGENTARIA S.	7,405,426.25
11/07/2013	Pagamento	2013NL000003	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	COORD-GERAL PROG. DESENV. SUBMAR. NUCLEAR	MATERIAL DE CONSUMO	DCN CHERBOURG	5,487,058.29(*)
11/07/2013	Pagamento	2013OB800063	Original	MINISTERIO DA DEFESA	COMANDO DO EXERCITO	D CONT - SETORIAL FINANCEIRA		BANCO SANTANDER (BRASIL) S.A.	4,096,512.00
11/07/2013	Pagamento	2013OB800238	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	SECRETARIA DE ECON. E FINANCAS DA AERONAUTICA	JUROS SOBRE A DIVIDA POR CONTRATO – LC141/12	BBVA-BANCO BILBAO VIZCAYA ARGENTARIA S.	1,963,941.92
11/07/2013	Pagamento	2013OB800312	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	ARMJ OMPS - I APOIO DE BASE	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	BANCO DO BRASIL SA (*)	1,445,138.57
11/07/2013	Pagamento	2013OB802333	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	GRUPAMENTO DE APOIO LOGÍSTICO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	DIGEX AIRCRAFT MAINTENANCE S.A.	798,805.04
11/07/2013	Pagamento	2013OB800405	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO DE IMPLANT.DO SIST.DE CONTR.DO E AER	Múltiplo	OMNISYS ENGENHARIA LTDA	692,563.60
11/07/2013	Pagamento	2013OB801510	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	MATERIAL DE CONSUMO	VILLARES METALS SA	690,590.25
11/07/2013	Pagamento	2013OB801496	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	MOODY INTERNATIONAL BRASIL LTDA.	639,867.05
11/07/2013	Pagamento	2013OB803559	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	GRUPAMENTO DE INFRA-EST E APOIO DE S J CAMPOS	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	BANCO DO BRASIL SA (*)	588,114.70
11/07/2013	Pagamento	2013OB801497	Original	MINISTERIO DA DEFESA	COMANDO DA MARINHA	CENTRO TECNOLOGICO DA MARINHA EM SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	MOODY INTERNATIONAL BRASIL LTDA.	565,022.95
11/07/2013	Pagamento	2013NS000635	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO AERONAUTICA BRASILEIRA NA EUROPA	MATERIAL DE CONSUMO	THALES RAYTHEON SYSTEMS	537,727.57(*)
11/07/2013	Pagamento	2013OB800403	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	COMISSAO DE IMPLANT.DO SIST.DE CONTR.DO E AER	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	ATECH - NEGOCIOS EM TECNOLOGIAS S.A.	529,595.26
11/07/2013	Pagamento	2013OB800917	Original	MINISTERIO DA DEFESA	COMANDO DA AERONAUTICA	PARQUE DE MATERIAL AERONAUTICO DE SAO PAULO	OUTROS SERVICOS DE TERCEIROS-PESSOA JURIDICA	IAS INCREASE AVIATION SERVICE LTDA	420,950.61

(*) Este favorecido é um intermediário que recebe o recurso e repassa-o para os favorecidos finais.

(*) Valor em moeda estrangeira convertido para real (R$).

Source: Brazil (2006).

If there is a requirement to query other information regarding each expense, the values of the Document attribute have a hyperlink to the expense detailed. Figure 7 presents how the expense data is found in the query result.

Figure 7 – Details of documents linked to each expense

Source: Brazil (2006).

In the current form of query daily expenses data, the result displays the attributes organized and distributed in three groups: Basic Data, Detailed Data, and Related Documents.

The citizens will be able to see descriptive attributes that identify the sphere linked to the expenditure only at this moment, including data about the source of the financial resources that will pay the expense and the identifying number of the recipient (such as the *CPF*, the *CNPJ* or the *UG-Gestão*).

Therefore, if the citizen is interested in drawing a graph showing values of the total amount of expenses transferred annually to a private company, but only those that come from a certain government branch, the data collection process becomes infeasible. For example, it would be necessary to carry out 12 queries, one for each year, one for each month of the chosen year, with all results linked to the selected private institution. Second, the citizen will need to enter all documents linked to the contracts to identify the branch; transform the data results collected on query's pages into records in an electronic spreadsheet, filter only the branch of his interest, and generate an annual total sum.

The elaboration of a DM may assist in this kind of limitation in the direct queries to data stored in a transactional database since the expense fact can be linked to dimensions to work as entry points.

The attributes found in the listing resulting from the query, associated with the attributes found in the Basic Data and Detailed Data groupings can be reorganized in a dimensional model as proposed by Kimball and Ross (2002), as it is possible to identify fact table attributes (such as expenses values) and dimensional table attributes, such as the branch, the source of funds, the date, the related agency, agency or linked entity, the management unit, and the recipient.

Based on identifying the attributes themselves in the detailing of expenses and reorganizing the attributes in fact and dimension tables, it is possible to prepare a proposal for a dimensional model in a star scheme (Figure 8) for the Daily Details of Expenses query. The fact table is at the center (expense) with each of the dimensions around it. The unique attributes are identified with a PK (primary key) and FK (foreign key), the descriptive attributes (containing texts or numbers) are identified with non-bold labels.

Figure 8 – Dimensional model proposal in the star scheme
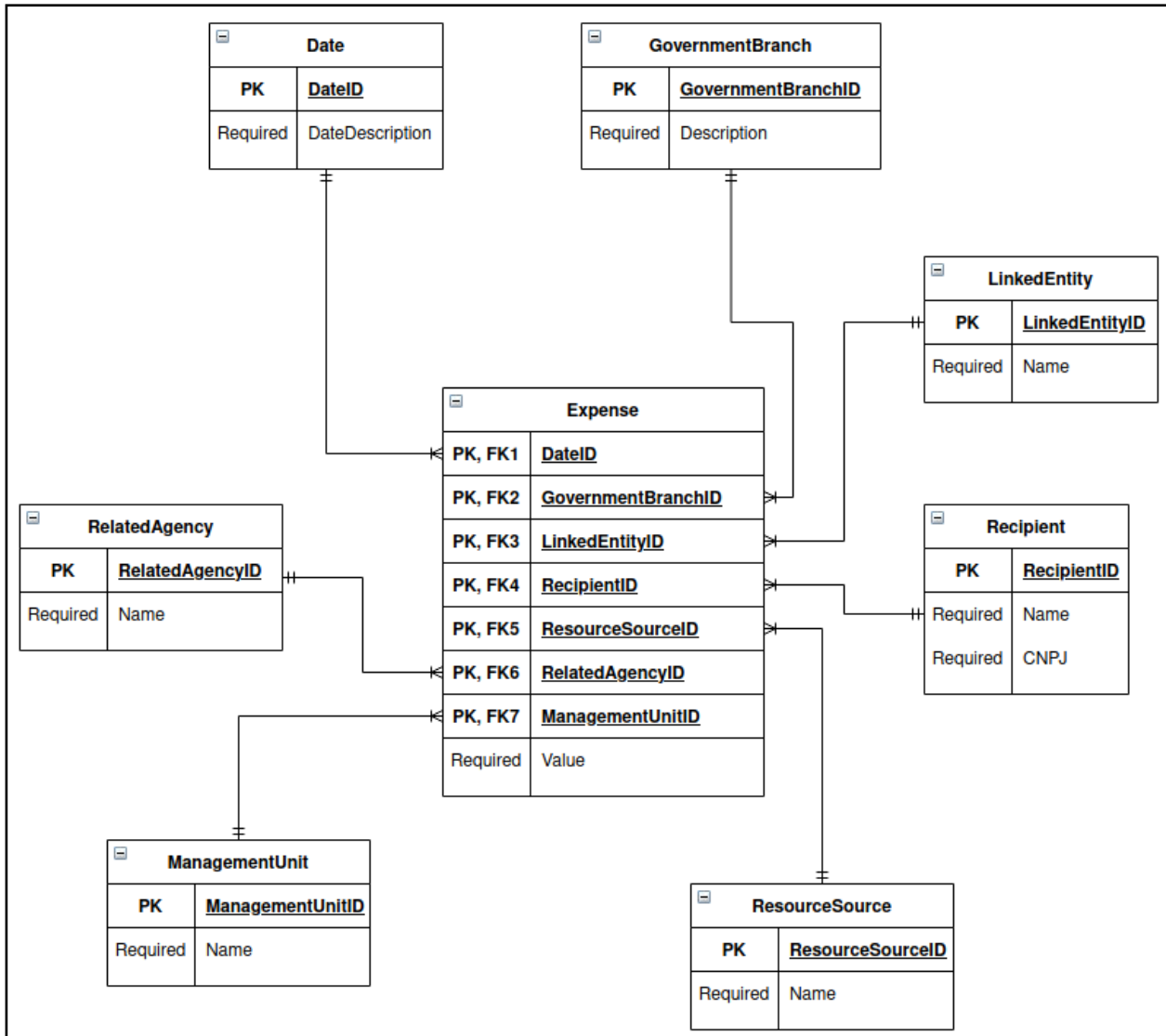
Source: Authors.

As a result, the attributes of each of the dimension tables can represent a starting point for different types of queries for the same fact (expenses), creating a thematic perception in the elements of origin and destination of expenses, for a specific audience (in this case, citizens interested in such kind of information).

CONCLUSIONS

The creation of a dimensional model by a DW may contribute to the construction of new queries to the information resources found on the Brazilian Federal Government's Transparency Portal, which is directly linked to the active transparency process. The elaboration of DM sets is a key element for public agents to anticipate demands for government data, as they can assist in the data collection process to satisfy questions from society.

For example, through the proposed DM, the process of collecting government data that deals directly with the values of expenses and the path taken between the source and the destination of an expense could be simplified. Besides, dimensions may contain more descriptive attributes, which are not present in the retrieval of information in the transactional model proposed in the current query. The dimensional model could bring more information about the branches, the agencies, recipients, among others.

Also, there are no restrictions for a DW of government data to be elaborated with sets of DM taking into account other scenarios, such as DM's centered on other facts (such as agreements, transfers of resources, among others) or with dimensions serving other interests (such as different dimensions for actual expenses). In both scenarios are important in the development of a DW and, also, in the process of accessing public management data with more quality and efficiency - and in favor of the concept of Citizen Intelligence.

It is also important to note that, despite the proposal to develop a DW of government data that is transparent to society, government web sites should continue to provide access to primary data and access to the existing set of queries. The DW has a complementary role to such queries, mainly helping to interpret the existing queries themselves.

REFERENCES

BARBIERI, C. **BI - Business intelligence modelagem & tecnologia.** Rio de Janeiro: Excel Books, 2001. p. 424.

BOHMAN, J. BOHMAN, J; **Public Deliberation, Pluralism, Complexity and Democracy.** London: MIT Press, 1996.London: MIT Press, 1996.

BRAZIL. **Constituição da República Federativa do Brasil de 1988.** Portal do Planalto. Brasília, DF. Disponível em: <http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm>. Acesso em: 20 jan. 2012.

BRAZIL. **Lei número 11.111, de 05 de maio de 2005.** Regulamenta a parte final do disposto no inciso XXXIII do caput do art. 5o da Constituição Federal e dá outras providências. Portal do Planalto. Brasília, DF, 5 maio 2005. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Lei/L11111.htm>. Acesso em: 31 jan. 2012.

BRAZIL. **Lei número 12.527, de 18 de novembro de 2011.** Regula o acesso a informações previsto no inciso XXXIII do art. 5o, no inciso II do § 3o do art. 37 e no § 2o do art. 216 da Constituição Federal; altera a Lei no 8.112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. Portal do Planalto. Brasília, DF, 18 nov. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acesso em: 31 jan. 2012.

BRAZIL. **Portal da transparência do Governo Federal.** Brasília, 2006. Disponível em: <http://www.portaltransparencia.gov.br>. Acesso em: 20 jun. 2013.

BRAZIL. **Portal da transparência do Governo Federal - Detalhamento das despesas diárias.** Brasília, 2004. Disponível em: <http://www.portaltransparencia.gov.br/despesasdiarias/>. Acesso em: 20 jun. 2013.

FIERRO, A. E.; GIL-GARCIA, J. R. Transparency websites as tools for decision making in a democratic government. In: **TRANSATLANTIC CONFERENCE ON TRANSPARENCY RESEARCH**, Utrecht, Netherlands: Utrecht University, 2012. Disponível em <http://www.transparencyconference.nl/wp-content/uploads/2012/05/Fierro.docx>. Acesso em: 6 jul. 2013.

INMON, W.H. **Building the data warehouse.** John Wiley & Sons, 1992. 543 p.

KIMBALL, R. **The data warehouse toolkit.** John Wiley & Sons, 1996. 388 p. ISBN 978-0-471-15337-5.

KIMBALL, R.; ROSS, M. **The data warehouse toolkit**: the complete guide to Dimensional Modeling. John Wiley & Sons, 2002. 393 p. ISBN 0-471-20024-7.

MALIN, A. M. B. Gestão da informação governamental: em direção a uma metodologia de avaliação. **DataGramaZero**, v. 7, n. 5, out., 2006.

MUSSI, C. Data warehouse - a experiência da Agência Nacional de Vigilância Sanitária do Brasil. In: **CONGRESO INTERNACIONAL ONLINE SOBRE TECNOLOGÍAS DE LA INFORMACIÓN EN CIENCIAS DE LA VIDA**,1. (TICvida'2005), 2005, Madrid.

OWEN, G. S. Visualization education in the USA. **Education and Computing**, v. 8, n. 4, p. 339–345, 1993. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/0167928793904376>. Acesso em: 4 maio 2017.

P2P FOUNDATION. **Introduction to Citizen Intelligence Sources and Methods.** P2P Foundation, 11 fev. 2011. Disponível em: <http://p2pfoundation.net/Introduction_to_Citizen_Intelligence_Sources_and_Meth ods>. Acesso em: 1 jul. 2013.

RODRIGUES F. A. Restrições tecnológicas e de acesso a dados disponíveis sobre destinos de repasses financeiros federais para a saúde pública em ambientes informacionais digitais. In: **ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO**,13. Rio de Janeiro, 2012. ISSN: 978-85-62454-02-8. Disponível em: <http://www.eventosecongressos.com.br/metodo/enancib2012/arearestrita/pdfs/194 35.pdf>. Acesso em: 1 jul. 2013.

SANT'ANA, R. C. G. Mensuração da disponibilização de informações e do nível de interação dos ambientes informacionais digitais da administração municipal com a sociedade. Tese de Doutorado em Ciência da Informação – Universidade Estadual Paulista, Marília, São Paulo, 2008. 153 p.

SANT'ANA, R. C. G. **Tecnologia e gestão pública municipal**: mensuração da interação com a sociedade. São Paulo: Cultura Acadêmica, 2009. 181 p.

SANT'ANA, R. C. G.; RODRIGUES, F. A. Visualização de afinidades entre parlamentares mediante dados de votações no Senado Brasileiro. **Informação & Sociedade: Estudos**, v. 23, n. 1, jan./ abr. 2013. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/12903>. Acesso em: 1 jul. 2013.

SENAY, H.; IGNATIUS, E. Rules and principles of scientific data visualization. ACM SIGGRAPH Education Committee. **Association for Computing Machinery**, 1999. Disponível em: <http://www.siggraph.org/education/materials/HyperVis/percept/visrules.htm>. Acesso em: 20 jun. 2013.