THE EXPERIENCE OF INFORMATION ACQUISITION IN CHILEAN PUBLIC MARKET VIA BI IMPLEMENTATION

Claudio Loyola and Maruzella Ortíz*

ABSTRACT. Among the objectives which are possible to hold by an e-procurement public system (or e-GP), there is no doubt that there are two that stand above the rest: the first, is increasing Transparency, which means to make accessible to everyone interested the decisions and their outcomes, taken by the acquisition officials; the second is to boost Competition in the Public Market, which essentially, but not purely, means value for money. It is well known that the process of information acquisition plays an important role in determining the competitiveness of markets, especially in Public Procurement (Public Markets), where although a significant part of the GDP can be compromised, nevertheless not always the right tools to plan purchases (from the buyer point of view) neither the information nor tools to size the market and refine the bids (from the supplier point of view) are present. For instance, transaction costs related to querying information from many data sources and formats increase the likelihood of suppliers churn because it is a costly and fluctuating process, and thus directly impacting the number of bids received by each tender, which is an essential issue in order to obtain savings derived from competition.

We show the impact of public free information acquisition in a systematic way for both suppliers and buyers (acquisition officials) in the Chilean public procurement system (ChileCompra), via Business Intelligence (BI) tools, and its impact diminishing transaction costs for all players, including those not directly related to trading activities, as for instance, comptroller officials.

The main objective, however, is to show how these set of business intelligence tools allow e-Procurement systems to increase competition in public markets, strengthen specific areas of interest (for example, those n-tuples conformed by a geographical zone and a type of product, where there is a low level of competition), which derives both from the actions taken by suppliers and buyers, from the information they can process with these tools. Experiences and lessons regarding the advantage and pitfalls for public procurement of this practice are shown, as well as future research in this area.

---

* *Claudio Loyola, Head of Public Market, ChileCompra Management, Chile. Maruzella Ortiz, Former Head of Business Intelligence, ChileCompra Management, Chile*.

1. BACKGROUND

The contribution that Chilecompra has made in Government Transparency is undeniable. Nowadays anyone can keep track of public policy development done through over 350,000 yearly tenders conducted in the public market. Anyone can learn how the USD 7,700 million that are spent by the State every year in goods and services are allocated.

Having taken all this information into consideration, ChileCompra Management has developed a data analysis and information management platform. The objective is to enhance public market transparency and efficiency, and reduce market transaction cost. This is intended through better use of the available information regarding procurement, and through more effective procurement and supervision management and process control.

The transactions logged on the system database themselves are extremely valuable. With the proper technology, these records help us research the behavior of the agents within this important market. For example, a company could design precise strategies to focus its commercial effort in those public arenas, industries or business areas in which, given the right operational conditions, the probability of being awarded the bid would be higher.

In an initial stage, ChileCompra has set up a complete system of reports in the market, which details all transactions. These reports allow the calculation of behavior indexes fit for each of the agent's interests. Chilecompra has also designed and made available a basic system of OLAP cubes, which is a multidimensional analysis tool that allows the cross-reference of several metrics.

However, such tools can be deepened even more in order to reach greater transparency and lower transaction costs in the public market through more thorough information analysis processes and with better cross referenced data.

It is possible to have more successful tender processes and for vendors to have a more effective focus. This through having more and better reports, more analysis cubes, an adequate infrastructure and the incorporation of state of the art technology. For example, the

incorporation of business rules that operate on public market transaction records and allow buyers to know beforehand the probability of success for their offers or for vendors to know beforehand the probability of success of their bids, is now possible.

2. MAIN PROCESSES INVOLVED

Company Registry

Process Description: Any company that is interested in being a vendor for the State must log on to www.mercadopublico.cl and fill in the required form. They are asked to provide general information about the company and name a specific person as a main user, who will manage the password to sign up the company within the system.

Signing up is free and it then allows companies to participate in the public market by sending their bids. Additionally, companies have the option to sign up in the official registry of State vendors for Chile, www.chileproveedores.cl. Membership in this website is approximately USD 60 a year and it certifies that the company is currently able to be a contractor for the State. It also allows companies to have their information uploaded in digital format so that they don't have to resend their documents every time they want to bid or are contracted.

Process Objectives: Gather information of companies which want to do business with the State of Chile and enable them within the public procurement system in order to operate on it.

Main Process Registries and statistics: General descriptive information of the company is recorded in this process. Some information is mandatory, for example the size of the company, the address, main user and email. However, the depth and type of information will depend on who is entering the data.

If the company registers in the official registry of State Vendors, the information is validated through official sources. Therefore, the information pertaining companies in the official registry is more trustworthy than that of companies that are only in the basic registry.

So far, there are approximately 410,000 companies that have signed up in the basic registry. 100,000 of these are active (i.e., made transactions this past year). There are 43,000 companies signed up in the official registry.

Making a Tender

Process Description: Public buyers create a document or tender file in which they detail their requirement, including deadlines, technical specifications for products and services and evaluation schemes which will be applied in the awarding decision making process.

Tender files associate products or services using the United Nations items classifier. It lists 20,000 generic products grouped under the 3 levels, specifically at level 3.

The tender file is created, having previously been authorized in the system by a user of the respective institution. A user with special powers publishes the file, and from that moment on, the interested vendors will be able to bid for that Tender until the set the closing date.

Process Objectives: Document all buyers requests in a structured and standardized fashion, in order to facilitate data registration and the bidding process for vendors.

Main Process Registries and Statistics: During this process, detailed information for every tender, size of the business, closing dates for bid application, estimated award dates, requirement specs, evaluation criteria and weighting are registered. Additional information -such as fines, guarantees, approximate deadlines, important landmarks within the tender - could also be contained in this document, but it will depend on the depth with which each buyer operates.

The 350,000 tenders that are published annually in the public procurement system are grouped in three categories: those under USD 7,000 which represent 85% of the total tenders. Those that range between USD 7,000 and USD 70,000 which represent 12% of

the total of tenders. And those over USD 70,000, which represent the remaining 3%.

The quantity and quality of the information gathered in the tender form, considering the documents that each user attaches to the form, is directly proportional to the amount of money of each tender notice.

Bidding

Process Description: The vendor who is interested in bidding signs in to their private section in www.mercadopublico.cl. There they can prepare their bid based on the information provided by the tender file.

The vendor must state if they are interested in bidding for all or part of the requirement outlined in the file. They must also include the bidding price for each of the required products or services. Additionally, the vendor can attach documents or appendixes to their bid. This could be mandatory in some cases, depending on what is stated in each file.

Once the bid is sent, the vendor will be sent a confirmation notice through email.

Process Objectives: Enable vendors to send online bids for the tenders in which they are interested. The information must be structured in such a way as to facilitate the evaluation process for the buyer.

Main Process Registries and Statistics: Bids are the main indicator of the real participation of companies within the public procurement system. 57,000 of the 103,000 enrolled companies send bids.

In this process, information pertaining the specific bid of each vendor is entered in the system. The information is registered in relation to the vendor's account and the respective tender. An important amount of information of the bids is contained in attached documents and only general information along with the process is the mandatory values that are entered in specific fields.

An average of 6,3 bids per tender were received in 2011. An estimated average sending time per vendor, considering all vendors that make transactions in 2 years, is approximately 18 months. There are also an estimated 43% of new companies every year who send bids. I.e., companies that in the t-1evaluation period, in which t is years, didn't send bids through the system.

Bid Awards

Process Description: once the closing date for bids is due, the system disables the option for new bids. The buyer opens the bids.

Depending on what the buyer stated on the tender document, the opening can be done in two ways: in one stage, in which all information about the bids is accessed once the opening is done. Or in two stages, in which two openings are done subsequently. First, a technical stage, in which the buyer will only have access to the technical information of each bid in order to assess and discard those that don't comply with the technical specifications. Second, a financial stage in which the buyer will see the financial information (prices) of the bids that qualified technically. Then the buyer can make a complete assessment of the bids.

Process Objectives: Choose the bid that best fits the evaluation criteria set in the tender file, and enter the result of the assessment process of bids in the system. Thus, the decision making process that backs up the awarding of the bid is more transparent.

Main Process Registries and statistics: Information regarding the bid assessment process is added in the system, such as the scores in each evaluation criteria and clarification requests of a bid.

Once the buyer awards the bid, the system generates an automatic and structured record, which can in some specific cases include complementary non-structured assessment records. These are attached as additional documents.

On average, there are 2,000 awards daily. 90% of these are done within the time limit. Over 95% of the awards are in one stage. In

95% of them, the vendor's technical appendixes are public documents.

10% of the bids are discarded during the awarding process. As of today, we cannot analyze the reasons for discarding, because they haven't been clearly structured. Regarding the awarding decision, 0,1% of the awards receives vendor complaints questioning the decision. All complaints are sent through the complaint platform of ChileCompra Management.

Sending of purchase orders

Process Description: Having been awarded a bid, buyers send purchasing orders to the vendors through a specific online form. It is generated from the data of the awarding process and it can be complemented by each buyer.

Each purchase order must be sent by a user with applicable powers, and it must then be accepted by the respective vendor. All of this is done through the system. Additionally, once the purchase order has been accepted by the vendor, the buyers report in the same document that they have received the products or services.

Process Objectives: Perfect the agreement between the parties regarding the products or services that will be delivered/received and register the commercial interaction between buyer and vendor.

Main Process Registries and statistics: 2,000,000 purchase orders are generated each year. 50% come from tender notices.

The purchase order contains the essential information of the commercial transaction: Price, product or service, quantity and other additional data that is incorporated depending on each of the parties.

Products or services registered in the purchase order are consistent to those of the tender notice file, and are organized in the same way in that document, i.e. through the item classifier of United Nations.

Therefore the level of detail reaches a generic product, that is part of a classifier items of 3 levels.

The United Nations Classifier has over 20,000 generic products associated to the items of level 3. The analysis product by product is particularly complex, given that the detailed description is the buyer's responsibility and it isn't completely structured.

3. PRODUCTS

Datawarehouse

Objective: The institutional datawarehouse has been implemented not only to ensure pertinence, quality and availability of information for procurement, but also to facilitate the exploitation of this information. This for both external users of ChileCompra Management, be them public buyers that optimize their management decisions, State Vendors that perfect their offer of products and services, and citizens that are looking for information.

Creation: The institutional datawarehouse extracts information from the different original databases. Some of them are transactional databases, user information registration databases or incident management databases. Having previously manually obtained the information of exchange rates, the datawarehouse builds a temporary database through ETLs, called DPA. This is the source for the construction of the different datamarts that make up the institutional datawarehouse.

Before executing the ETL that loads the datamarts, an error detection process is executed in order to ensure the quality of the data. It is made up by ETL of erroneous transaction detection and a manual inspection of the more borderline cases.

Currently there are 4 operational datamarts: DM_Transactions; DM_Processes; DM_Store; DM_User; DM_Formation. There are 5

more that are under construction: DM_Invoice and payment; DM_Contracts; DM_Companies; DM_Complaints; and DM_Indicators. A SQL script makes improvements, like: add sector to the institutions; location to the buyers; source for the purchase orders and update of the depending ministries; evaluate outliers in the amounts for the purchase orders.

The tender, Purchase orders, Framework Agreement and Bid OLAP cubes are created from the aforementioned datamarts. The dimensions are processed first and then the cubes. The cubes are what allow an in-depth and transversal analysis of the public market data. Additionally, these cubes enable coverage to the need of user information, both internal and external to ChileCompra Management.

Technologies: The institutional datawarehouse operates under Microsoft infrastructure, specifically Microsoft SQL Server 2008 Management Studio y Microsoft Analysis Services Client Tools. The system has been programmed in Microsoft Visual Studio 2008 (Version 9.0.30729.1 SP), Microsoft .NET Framework (Version 3.5 SP1), SQL Server Analysis Services, Microsoft SQL Server Analysis Services Designer (Version 10.0.1600.22)

Report System.

Objective: The report system covers a particular need for users that work with periodic and structured reports for generating diverse public procurement information, be it for management reasons or just for research.

Creation: Reports are generated from 2 data sources: the institutional datawarehouse and the DCCPProcurement transactional database. The reason is because on the one hand, certain reports cover the needs of users who are seeking consolidated information- for example, for interpreting some series, or making in depth research. On the other, some reports are generated for those seeking spot information to focus on specific business.

The reports that are generated from the transactional bases are called “daily” and “highlights”. In both cases the information is

offered in Excel format from SQL queries that are executed in batch overnight.

The information provided on daily reports is of the tenders that at the moment are in "published" status. This means that vendors are still able to send bids.

The Highlights reports also state the tenders that are under "published" status, but they are additionally filtered through an algorithm that has been previously defined. This algorithm is used to define a tender as "highlighted". It also includes parameters such as the amount of the tender, area of business, deadline and other.

Technologies: Reports are generated on Excel and are visualized through a web interface developed on HTML. The code is built in such a way that once a report is generated, it will be available in the corresponding page through SQL queries.

Online analysis system.

Objective: the online analysis system is thought to enable each user to build directly those reports that are not configured on a standardized way. The benefit is that users don't have to make special requests to ChileCompra Management.

Creation: The online analysis system is built from the OLAP cubes contained in the institutional datawarehouse. These cubes are made available online through DUNDAS control.

The online analysis system doesn't need any authentication, and therefore it is open to all current and potential public procurement system users. Due to infrastructure restrictions, the cube version available online is bounded in its dimensions and doesn't allow reaching the atomic data of each transaction. As of today a project is in process in order to reach this capacity for the online analysis system.

The adjusted version of tenders, purchase orders, Framework Agreement and bids is available online by the ChileCompra

Management Team. It has its own infrastructure and is visualized by users through an interface that allows working online on their own reports. These can be saved to be worked on later, so there is no need to create them again every time the user logs on to the system.

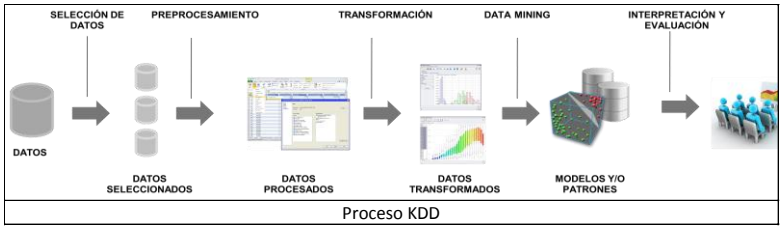
Technologies: The online analysis system operates under Microsoft Infrastructure, specifically Microsoft SQL Server 2008 Management Studio y Microsoft Analysis Services Client Tools. The system has been programmed using Microsoft Visual Studio 2008 (Version 9.0.30729.1 SP), Microsoft .NET Framework (Version 3.5 SP1), SQL Server Analysis Services, Microsoft SQL Server Analysis Services Designer (Version 10.0.1600.22)

The web interface has been developed in HTML. The DUNDAS control has been incorporated to allow visualization and online management of data.

4. MODELS OF TENDER SUGGESTIONS

Preparing the Information

The methodology used to build the model was Knowledge Discovery in Databases business intelligence. The objective is the integration and treatment of data in order to convert it into information that supports the decision making process.



Proceso KDD

The data went through its respective cleansing treatment, selection and transformation in order to reach an optimal data mining stage.

The selected data came from the datawarehouse and the process and transaction datamarts. This is a significant part of the data contained in the Tender files.

Then, the data was cleansed. The points that were out of line and infrequent data were taken out. Techniques of replacement, imputation by the mean or mode and elimination were used.

Finally the data was transformed due to the fact that the scales were very different among numeric attributes. We therefore decided to normalize all the numeric attributes and transform the descriptive attributes into dummy attributes.

Once all the registries were clean and each attribute was transformed, data mining was applied.

Data mining

Initially we applied the decision tree in order to find the business opportunity prediction pattern for each vendor. However, due to the great amount of products, the tree was over adjusted, and it only predicted the business areas with the most bids. Consequently, the other business areas were null.

But what the tree did teach us were the attributes that cause discrimination in the sample in order to predict the bid. These were: tender type, bid on product or service and Country Region in which the bid is made. This information helped us determine the probability of success for each vendor

As a second option, we decided to divide the problem in knowledge of the vendor and the business areas in which he is successful. The chosen solution was 2 statistics tools (Cluster and probability Algebra) and an association rule (Marketbasket analysis). The objective was to determine business opportunities that were probably going to be successful for each vendor.

Define Vendor Cluster: Its objective is to determine the number of vendor clusters in order to learn the different vendor profiles. They must also represent them all according to the use in time, location and business success.

First, the type of tenders in the public market is defined.

L1: Tenders under 100 UTM

LE: Tenders ranging between 100 and 1000 UTM

LP: Tenders over 1000 UTM

LS: Service Tenders.

Next, each cluster and its main characteristics are described:

Cluster1: Large vendor, successful in type L1 Tenders. Highly involved each month.

Cluster2: Mid-size vendor, successful in type L1 Tenders. Highly involved each month.

Cluster3: Micro and small vendor, successful in type L1 Tenders. Low involvement during the month.

Cluster4: Micro and small vendor, successful in type LE Tenders. Low involvement each month.

Cluster5: Micro and small vendor, successful in type LP Tenders. Highly involved each month.

Probability Algebra: Its objective is to determine the success rate and expected probability of success. We applied the Favorable case /Possible Case rule of probability, and the Independence of events rule. The calculation of the success probability is as follows:

$$P\left(E/TLC_h \cap P_j \cap R_k\right)=P(E \cap TLC_h \cap P_j \cap R_k)$$

In Which:

TLC_h = Bids in the h-th type of tender.

P_j = Bids in the j-th product or service

R_k = Bids in the k-th region

E = Successful Bids

Therefore, the calculated probability represents

$$P = \frac{\#\text{ of successes in the h-type of tender i-product k-region}}{\#\text{ of successes in h-type of tender i-product k-region}}$$

Define Association Rules to determine Tender Baskets: The objective is to find association rules to determine products and/or services baskets related to the commercial activity in accordance to the vendor's transaction history and movement. The Market Basket Analysis tool is applied and the baskets are selected.

The baskets were built linking together the product, tender type and region. This was taken as product within the analysis. Then, the Market Basket Analysis methodology was applied and 3 rules were determined:

Relative Importance (Support): Indicates the percentage of transactions that carry the antecedent and consequent with respect to the total of analyzed transactions.

Reliability (Confidence): Indicates the percentage of transactions that carry the antecedent and consequent together, with respect to the total of the transactions that carry the antecedent

Benefit (Gain or Improvement): Score that represents the rise in the probability of the consequent selection when compared in conjunction with the antecedent.

Database Construction

Algorithms and databases for the implementation of the model over the public procurement system are built from the tool used in the data preparation stage. The objective is to determine the adequate business opportunities in accordance to the probability of each. This is done through interrelated databases that generate a list of suggested tenders.

Vendor Users: Along with a previously calculated Cluster attribute, it carries general information obtained from the datawarehouse. It has a 3 month update and success probability built by the first algorithm (number of awarded bids in 18 months/ total of offered bids in 18 months)

Product Baskets: Carries product information along with region and tender type. It must be concatenated and the basket to which they belong is executed by the second algorithm.

Tenders: Carries type L1- LE- LP and LS published tenders, along with the product information, region of the purchase unit and tender type. Must be concatenated and is executed by the third algorithm.

Once the databases are built, the fourth algorithm is executed. It allocates to the vendor users the bids that entail products in which the probability of success is over 0.7

Each of these databases is linked through an algorithm which takes the published tender and allocates a basket. Then for each vendor that is associated to that basket and has a probability of success over 0.7, the bid is selected to be shown in the vendor's desktop.

Model Results

As a result, the 161,064 vendor users that have traded in the last 18 months will have an average of 3 offers with a success probability over 0.7. This means that the model will generate business opportunities for the vendors in line with their trading profile on www.mercadopublico.cl with a 78% rate of reliability.

Plus, the 60,317 product and/or services baskets will be on the suggested tenders showcase on www.mercadopublico.cl. Therefore, all the published tenders will be made public on the website mentioned.

At last, the specific objectives of this Business Intelligence Project will be fulfilled: encouraging vendors to bid, stimulating competition and increasing the percentage of savings.

5. CONCLUSIONS

The implementation of a business intelligence system for public procurement perfects the match between the offer and demand. It makes it easier for vendors to prepare bids in the business areas they are most competitive. They have the possibility to generate a more comprehensive market analysis from the system.

Plus, vendors can access specific information that is relevant to them from the implemented models. Consequently, there is a cost reduction of the availability of information, for vendor analysis of adequate business opportunities is narrowed down. Besides, in those cases in which the tender suggestions model is used, the success probability for vendors increases. This is a positive externality on perception of satisfaction.

Given that the tender evaluation process is one of the most expensive stages of a bidding process, the number of offers per process requires a fair balance in order to ensure competition a price savings. This balance is given by the tender suggestion model, which focuses companies towards processes in which their participation will be a benefit for the system as a whole.

Moreover, another relevant variable that the system provides is the value in terms of transparency. Considering the amount of tenders that are published annually, the bids that are made and the purchase orders that are generated, the yearly amount of transactions that can

be monitored simultaneously through the system is over 3 million. This does not include additional data such as user information and other variables of interest. It would be impossible to follow up such volume of transactions without this type of tool.

Finally, the relationship between multiple sources of information allows transversal analysis for both the companies as promoters of public procurement and other stakeholders. It also helps to detect abnormal and even in some cases unlawful behavior which cannot be detected on a level of specific analysis.