

## **BUSINESS PROCESS REENGINEERING WITH DATA MINING IN REAL ESTATE CREDIT ATTRIBUTION: A CASE STUDY**

**Catherine da Cunha<sup>1</sup>, Bruno Agard<sup>2</sup>**

*<sup>1</sup>Institut National Polytechnique de Grenoble – Laboratoire GILCO  
46 avenue Félix Viallet, 38031 Grenoble Cedex 1, France*

*[catherine.da-cunha@gilco.inpg.fr](mailto:catherine.da-cunha@gilco.inpg.fr)*

*<sup>2</sup>Département de Mathématiques et de Génie Industriel, École Polytechnique de Montréal  
C.P. 6079, succ. Centre-ville, Montréal (Québec), H3C3A7, Canada*

*[bruno.agard@polymtl.ca](mailto:bruno.agard@polymtl.ca)*

**Abstract:** The process of real estate credit attribution is complex. The risk of accepting a credit for a customer that will not pay back has to be assessed and evaluate carefully. This process requires several steps, including the search for information within different tasks and banking organisations. Nevertheless, the evolution of real estate sector induces the need to reduce the time between the reception of the client's requirements and the transmission of the acceptance or rejection decision.

This change can not be obtained by minor modifications of the existing process; a business process reengineering is needed. Using a real case study, this article proposed to investigate the interest of data mining techniques to elaborate business process reengineering prescriptions.

**Keywords:** Business Process Reengineering, data mining, estate credit attribution, risk evaluation

### **1 Introduction**

In a highly competitive environment firms have to remain innovative. It is necessary to offer products and/or services for customers with higher standards of quality. Innovation is needed in order to the products/services sold remain attractive but also to improve the internal running of the firm.

Continuous improvement is often chosen so as to gain productivity and/or to decrease costs. Nevertheless, this technique is, as its name indicates, continuous: the immediate gains can be limited and the organization needed to maintain the implication of all workers is a huge investment.

Reengineering process can be an interesting way to obtain rapidly important return on investment (ROI). While business process reengineering is not a new concept, it was formalized only recently (Hammer and Champy, 1993; Davenport, 1993). Exemplar successes of this technique, e.g. in Hallmark, popularized the idea that few investments could permit to "reinvent" a firm.

The banking industry could also benefit from a reengineering, indeed the processes are mostly the result of historical evolution and they are rarely called into question. Allen (1997) pointed out that the cost structures could in particular use managerial and infrastructure adjustments.

In order to reengineer a business process, an is-analysis enables to enlighten the points that cause slow-downs within the process, to stress the divergences between the objectives and the realization. This study can be driven using statistical tools, nevertheless techniques that include systematic identification of patterns can also be interesting.

This paper emphasizes the impact of information extraction thanks to data analysis (statistical techniques and data-mining) when reengineering business process. First the context of the study is described. Then the reengineering process is analyzed. A case study of mortgage attribution is used in order to stress the different aspects of this method; it is described in the last section.

## **2 Context**

### **2.1 Real estate sector**

Competition between banks and mortgage societies on the small sector of real estate loan is very strong. Several criteria motivate the choice of a given company: the interest rate, the reimbursement facilities, the quality of service which integrates the rapidity of response and the renegotiation possibilities. Those two last criteria can also be grouped under the label “customer service”.

The evolution of real estate sector thanks to electronic commerce also has a huge impact on clients’ expectations. The use of Web based techniques to buy or sell a house has drastically weakened the need for intermediate such as real estate agents (Crowston and Wigand, 1999). The transactions can then be done far much faster; the decisions have to be taken quasi-immediately for the buyer to be sure that its offer will be accepted. This process supposes that the house’s buyers are able to raise rapidly the needed fund. The bottleneck of the whole real estate transaction has then become the mortgage attribution phase.

This situation leads to a real urge for change in mortgage societies’ organisations; the problem is that the whole process of mortgage attribution was conceived decades ago, and this design does not take into account the new tools of information technology (IT). Minor changes will not enable to correct this situation: a total reengineering is needed.

### **2.2 Current mortgage attribution process**

It is of paramount importance that the whole process of mortgage attribution is organized transversally. Figure 1 shows the different actors that are involved in the decision process. The particularity of the operations to be performed requires the intervention of specialized workers, it is therefore important to manage this interaction between different departments or economic units.

The main difficulty is to keep a consistent flow of information and workload while the different tasks are fragmented.

The sales representative plays the role of a “case manager”; he/she acts as a buffer between the client and the decision makers. He/she has access to information about the decision process through a data base that describes the work flows.

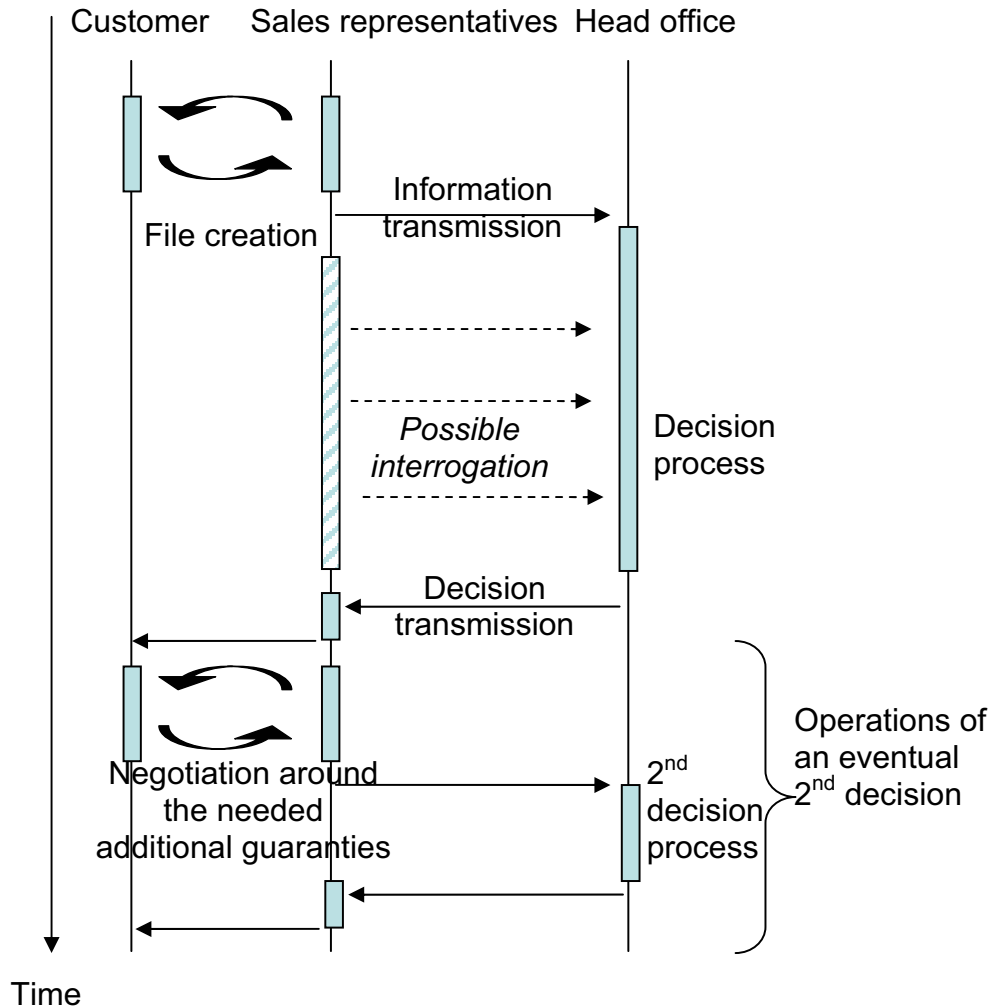


Figure 1 Decision process.

### 2.3 Process history: standards and regulation

The process of credit acceptance is mostly the result of an historical evolution. The different levels of decision, that define the scope of expertise and legitimacy of the workers, are the transcription of old structures that were in place when the only informational vector was the ordinary mail. Actual criteria are the result of an empirical evaluation of the deadbeats enabled by the deciders' expertise. For example, several decision processes exist, the attribution of a file to one of this process is based on criteria such as:

- mortgage amount: smaller amounts are adjudicated faster,
- type of operation: the buy of the main home is adjudicated faster than investments for rent,

- occupational group: files of liberal professions and handymen (supposed more risky) need more time.

Furthermore, technical evolutions enabled by the information system boom were only used to speed the different phases of the process. No rethinking occurred and most important no studies were planned to analyze the pertinence of the organizational scheme with the new IT tools.

### **3 Reengineering**

#### **3.1 Objectives**

The aim of a business process reengineering is not to provide tools to improve the existing process (this problem was addressed by Agostini et al. (1993)) but to rethink and rebuild a new process.

It is important to identify prior to any action the expectations of the workers. This first contact with the people that will have to implement the reengineering has 2 main objectives: first the presentation of the method and second the identification of improvement sources.

The primal objective was to improve the service for the customers: reduction of the decision delays and better knowledge of acceptance/rejection criteria. In the present case, the major expectation for the workers (sales representatives and deciders) was a better use of the resource (i.e. a smoothing of activity's repartition trough time and a better management of the delegations). This improvement of the work environment can not be achieved without a better understanding of the activity and of the real processes.

Indeed in many firms, the processes used often differ from the description (this is true whatever the sector): the reengineering should therefore begin by an is-analysis.

Some « bad habits » are trans-sectorial: useless backward moves, multiple data capture, tasks remake due to lack of information, etc. Those characteristics should be looked for and be corrected.

For example, the contact points of the process with the customer should be identified and limited; indeed, the reduction of entry points will reduce the risk of having divergent information and also permits to focus on the verification of the information that enter the process.

#### **3.2 Techniques**

The existing process should be modelled, in order to obtain a flow sheet that represents the different steps or stages. When considering a production process, the adding-value operations have to be identified, but this technique has to be adapted when the aim of the process is to reach a decision. Other criteria have to be found for example one could focus on the information retrieval for each activity since information is the input for the decision operation (the output being acceptance /rejection of a mortgage demand).

Redundant activities should be eliminated. This statement stands for any Business Process Reengineering.

### 3.3 Data mining interest

Anand and Büchner (1998) defined data mining as “the discovery of non-trivial, implicit, previously unknown, and potentially useful and understandable patterns from large data sets”.

Data mining uses statistical tools and dedicated techniques on large data bases; the algorithms enable to find behaviour rules whose pertinence (i.e. representativeness) can be measured. Furthermore, cluster analysis enables to determine collections of objects (clusters) that are similar to one another within the same collection and dissimilar to objects in other collections. This operation may stress similarities between individuals (regarding the characteristics or behaviours described by the considered data). Many algorithms for data mining are available by Fayyad (1996).

## 4 Case study

The problem presented here derives from a real case provided by a French mortgage company. The data are real data, which were submitted to a pre-treatment aiming to respect bank secrecy: masking of all names, address (zip codes were conserved), sources of funds, etc.

### 4.1 Activity description

The first question that has to be addressed is: What is the purpose of the whole organization? This question may seem trivial; nevertheless in most sectors the goal of the organization is to improve the performance indicators in use. This could be pertinent if the indicators were controlled and revised from time to time, to keep in touch with the real objectives.

For example, a performance indicator of a mortgage society is the volume of accepted mortgages, without consideration at the operational level of the major risk factor which is non payment.

Treatment delays are also considered as a good indicator of the performances. Credit attribution faces a major problem for the evaluation of their services: if everything went well the client will not come back, once the contract signed no more contacts are required. It is therefore difficult to have a real picture of the situation as only bad feedbacks are given!

To proceed a reengineering the processes that *a priori* need rework have to be identified. In our case, we focused on several criteria to evaluate the reengineering needs:

- Processes that are qualified as “problematic” by the workers: for example the acquisition of the data (data accuracy).
- Processes that have the most impacts upon the customers: processes that endanger the duration guarantee.
- Processes that have the most probability to be successfully reconfigured: the decision process.

## 4.2 Data treatment

The study of the past activity can permit to enlighten that some of the criteria in the decision process, are not justified. For example, let's focus on the following criterion: "the smaller the mortgage amount, the smaller are the risks for the bank/credit company".

### 4.2.1 Available data

The data consist of information about the activity in 2004 and the problematic files before 2004. Some pieces of information are at disposal: occupational group, date of the file opening, mortgage's amount, type of operation, guaranties.

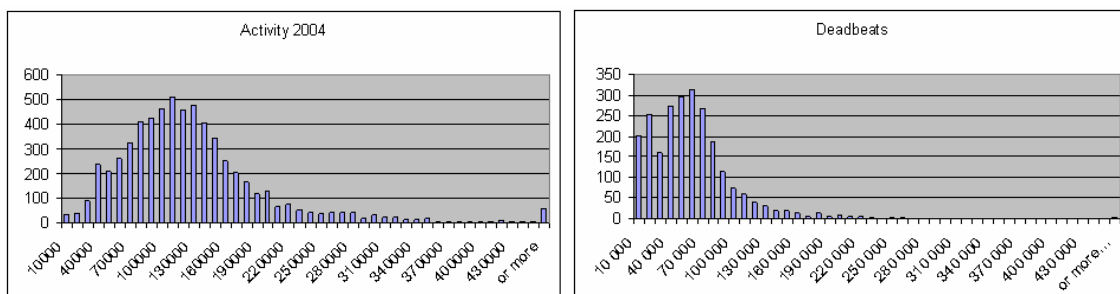
Let's approach here the question of the data's bias. An important point that can not be put aside is the incompleteness of the deadbeat data. At a given time  $t$ , the data available are:

- the characteristics of the files (accepted or rejected) before  $t$  (Called activity description), those data are static.
- the state of the deadbeats.

This latter actually groups the mortgages that were accepted between  $t - 30$  and  $t$  (30 years being the greater accepted mortgage duration) and that have had repayment problems. These data are of course dynamic, repayment problems occurring every day at the mortgage company level. At year  $t$ , one can only have accurate deadbeat's data for the mortgage accepted in year  $t - 30$ . Nevertheless, it is not acceptable for the company to wait so long to have a feed-back on their acceptance process's pertinence. Therefore, we will use here "partial" data on the deadbeats.

### 4.2.2 Classical statistical analysis

The study of the past activity enables to draw the repartition of contracts in function of the mortgage amount. The same can be done when studying the deadbeats for the same period. By comparing those two graphs, Figure2 shows that the risk of non payment is higher for small amount than for high amount.



**Figure 2** Number of accepted contracts (resp. deadbeats) inventoried in 2004, in function of the mortgage amount.

The explanation of this phenomenon can not be done without the help of bank experts; the sources can indeed be multiple. Two main reasons can justified this fact:

1 - The process of attribution of high mortgage is so strict that only applications that have very strong guarantees are accepted. The potential deadbeats are filtered through the process.

2 - The applicants for low mortgage correspond to a population that is more subject to “life accidents<sup>1</sup>”. The process is not responsible of this phenomenon, the securities required for this kind of mortgage are under protective and supplementary requests of guaranties will slow down the decision process but could drastically reduce the deadbeats.

The study of the rejection rates for this particular kind of mortgage can also permit to point the right explanation. In this particular case, the choice of the mortgage value as a criterion for the duration of the decision process seem only based on prejudices, the reengineering of this task will enable to only use criterion that are validate with the data history.

#### 4.2.3 Data mining analysis

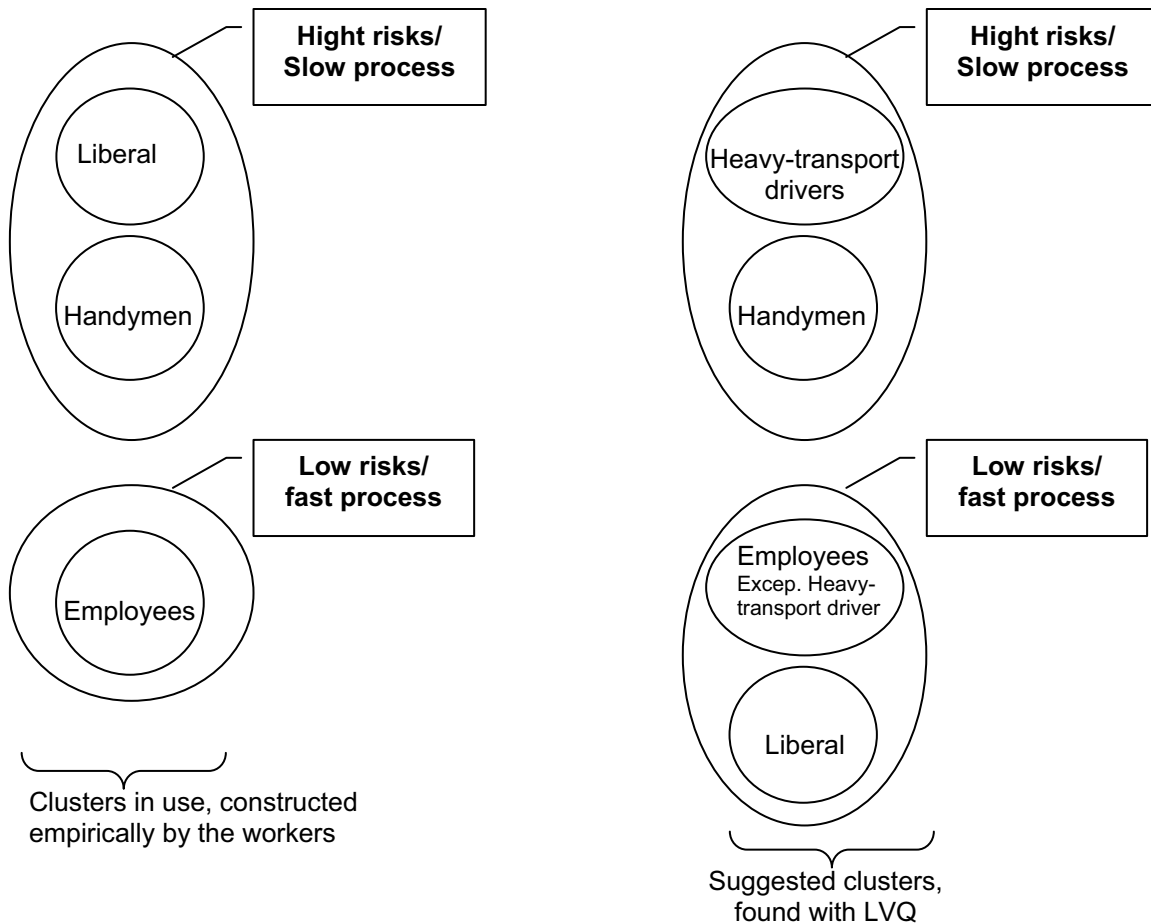
A data mining analysis enables to study the relationship that exists between occupational group and risks. The actual decision process is based on the fact that employees are less subject to “live accidents”: for example they have less bankruptcy risk than the handymen.

A clustering approach stresses the fact that the actual sections are not made of homogenous members (in terms of deadbeats’ risk). Here the LVQ algorithm (Kohonen, 1988) was used. The initialization (manual clustering) was done using the past activity, the label data being high risks (for files that had repayment problems) and low risks.

One of the clusters in use at the moment groups liberal professions and handymen, the common denominator being the risk of bankruptcy linked to their status, nevertheless the result of the clustering show that some employees (the heavy-transport drivers) are more risky than the liberal professions. New sections could therefore be used to better match the risk characteristic to the file analysis process.

---

<sup>1</sup> This term groups events such as divorce, death, unemployment, bankruptcy



### 4.3 Reengineering prescription

Information extracted from raw data enables to manage the reengineering. The rules obtained stressed out some of the characteristic of the bad files (i.e. mortgage demands that were accepted and not paid back). This knowledge is important to rethink the risk policy of the mortgage firm but also to redefine the decision process.

#### 4.3.1 Prescription on the process

It is important to dissociate two phases in the reengineering process: the prescription on the process and the prescriptions on the process' organization.

Those two stages have to be done sequentially: indeed the organization of a stage can not be pertinently done when having a myopic vision of the process, the knowledge of the whole business process has to be at disposal before any organizational change occurs.

Choice criterion for the decision process and therefore the duration of this process, is then unfounded (Cf. previous section).



#### 4.3.2 Prescription on future use (data to be collected)

Non crucial missing information could be constructed with consideration to the past experiences. This technique enables to give the decision maker a better view of the file without loss of time (the time needed for the look for the real information).

#### 4.3.3 Prescription on indicators to implement

Indicators should not only permit to control the existing situation, they should also give information about future improvement directions. It is indeed necessary that the reengineering process emulation (cf. 4.4) doesn't disappear once the new processes are implemented. Some specific indicators should be used such as the duration of the whole decision process, or the percentage of % of aborted files (with the reasons: another proposal cheaper or sooner, etc.).

#### **4.4 Managing the reengineering**

When reengineering business process, special care should be given to the implementation of the prescriptions. Indeed, the new tools and methods should be understood and approved by the employees in order for the change to be successful (Mossholder et al., 2000).

Folger and Skarlicki (1999) claim that "organizational change can generate skepticism and resistance in employees, making it sometimes difficult or impossible to implement organizational improvements".

To minimize the risks of rejection several stages should be respected: communication of the policy to all the workers (even those not concerned by the process reengineered), creation of an implementation schedule, and wide diffusion of the first indicators' values.

Furthermore the profile of the different users (sales representative, secretaries, etc. ) should be considered when building the communication plan because the organisation's ability to deal with change provides a competitive advantage (Skinner et al., 2002).

## **5 Conclusion**

Having processes that are fast, not consuming in resources and that produce quality output (products or services) is the dream of every firms. Reengineering plans aim at improving drastically an existing process in order to retarget the objectives and better use the means to achieve it.

If this technique is not new, it is mostly applied on process that aimed at designing/producing physical products, the application to services production process are rare.

This article considered a real case of a mortgage company. The reengineering process aimed at rethinking the decision taking operation. First of all the real objectives of this operation were stressed and indicators that reflect these goals were developed. Then data mining techniques were used in order to extract information to conduct this reengineering in an efficient way.

Further researches should consider the extraction of customers' behaviour in order to identify the risky customers (i.e. people who are likely not to pay back their mortgage)

## 6 References

- Agostini, A., de Michelis, G., Grasso, M.A. and Patriarca, S. (1993) Reengineering a business process with an innovative workflow management system: a case study, Proceedings of the *conference on Organizational computing systems*, Milpitas, California, United States , pp. 154–165.
- Allen, P. (1997) Reengineering the bank: a blueprint for survival and success, Probus Pub. Co, Chicago, Ill.
- Anand, S. and Büchner, A. (1998) *Decision Support Using Data Mining*, Financial Times Pitman Publishers, London, UK.
- Crowston, K. and Wigand, R. (1999) Real Estate War in Cyberspace: An Emerging Electronic Market? *International Journal of Electronic Markets*, 9 (1-2), pp.1-8.
- Davenport, T. (1993) *Process innovation: reengineering work through information technology*. Harvard Business School Press, Boston, Mass
- Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (1996) *Advances in Knowledge Discovery and Data Mining*, AAAI Press/The MIT Press, Cambridge, MA.
- Folger, R. and Skarlicki, D. (1999) Unfairness and resistance to change: hardship as mistreatment, *Journal of Organizational Change Management*, pp. 35-50.
- Hammer, M. and Champy, J. (1993) *Reengineering the corporation: a manifesto for business revolution*. Harper Business, New York, NY.
- Kohonen, T. (1988) *Self-organization and associative memory*, Springer-Verlag, Berlin, Germany.
- Mossholder, K., Settoon, R., Armenakis, A. and Harris, S. (2000) Emotion during organizational transformation: an interactive model of survivor reactions. *Group & Organization management*, **September**, pp. 220-243.
- Skinner, D., Saunders, M. and Thornhill, A. (2002) Human resource management in a changing world, *Strategic Change*, 11 (7), 341-345.

## 7 Biography

CATHERINE DA CUNHA received her Ph.D from the INP Grenoble in 2004. She is now a teaching assistant in the ENSGI (industrial engineering school of the INP Grenoble) and a research assistant at GILCO laboratory. Her main interests are in the area of information use within the firm, in the production as well as in the organizational field.

BRUNO AGARD received his Ph.D from the INP Grenoble in 2002. He is now assistant professor in the department of Mathematics and Industrial Engineering at École Polytechnique de Montréal. His main interests are the design of products families (including manufacturing and logistics) and applications of data mining in engineering.