closer

PROPOSAL

Project Tagus

JULY, 2017

Contents

Acknowledgement

We would like to thank Swiss Mobiliar for the opportunity of answering to this new challenge through the current proposal.

We are confident that the experience we have in the domains of Data Science/Advanced Analytics positions us as a partner who will effectively add a significant value to Swiss Mobiliar, strengthening its capabilities in these domains. Trust and competence are the foundations of strong and lasting relations and we see the current proposal as a first step to instigate a long and mutually beneficial relation between the two organizations.

1. Management Summary

The current proposal addresses a challenge posed by Swiss Mobiliar, through the RfP for "Project Tagus".

This proposal follows the precise structure defined by Swiss Mobiliar in the RfP, but having mind the the aim of sharing our vision and approach regarding the areas covered by the proposal. Three annexes complement the information presented in the main points.

In the service description section, we cover the (non-exhaustive) list of the tasks that, according to the RfP, Swiss Mobiliar expect to be carried out throughout the duration of the service contract. From a methodological point of view, we propose that an agile (scrum) approach - that we have adapted for analytics may be followed in development and prototyping tasks.

Security and Auditing are, understandably, due to the nature of the service, major concerns. These topics are covered in the Audit Approach section (namely, we identify and describe which main principles we consider that should be followed in an information security assessment process). The infrastructure we propose is not unrelated with these matters and Closer has been using Microsoft Azure as primary development platform. We find it to be the best solution in terms of safety, continuous monitoring and auditing, flexibility, scaling, state of the art technology, and pricing. Foremost, it guarantees complete compliance with safety and data protection requirements, such as GDPR.

Having in mind that the service will be as good as its team composition, we present some of the profiles we may involve as part of the team, as well as our recruitment process and how we see Swiss Mobiliar involvement in its later stages. The suggested training plan is also provided.

The costs presented cover both the team members and the proposed infrastructure. The team costs are detailed by each of the profiles matching the job specialization identified on the RfP. In order to provide a monthly figure for the team operation, we also present an estimation for the overall team (assuming an average of 21 working days/month).

Finally, we try to convey the company "DNA" and history. We truly believe that this service is totally aligned with our spirit, values and mission.

We are completely available to add any further insights into the proposed services as well as introducing the appropriate changes to meet any particular need we may have not anticipated.

2. Service Description

The current proposal presents a service model that will be deployed, should this proposal be accepted, with the objective to serve the business needs of Swiss Mobiliar that is currently undergoing a transformation toward an increasingly data-driven business model.

The usage of data science in the insurance sector will potentially bring many competitive advantages that can be translated into boosts of profitability, efficiency and user experience. Some examples of these benefits may include:

- Pricing optimization through the usage of unstructured data from a variety of sources, such as social media
- Fraud detection using machine learning algorithms
- Determination of the next best offer for current clients using recommendation models, coupled with tailored campaigns
- Automatization (or partial automatization) of rudimentary tasks at the different points of the process flow, from claims to underwriting
- Deployment of natural language processing, speech recognition and chatbots to respond and resolve queries and complaints fasters

In this proposal, we present our services to deploy a multidisciplinary team, with the necessary skills to respond to Swiss Mobiliar needs, build and test data science models and develop applications to be used by Swiss Mobiliar that will bring added value to their business and allow them to reap the benefits of choosing to deploy a more data-driven business model.

a) Description of services

There are a set of tasks that the nearshore team to be deployed within the scope of Project Tagus is expected to perform, as was detailed in a non-exhaustive list in the RfP provided by Swiss Mobiliar.

Below we provide some detail on these tasks.

Services

Here we describe the services to be provided by the Closer data science team for the complete model lifecycle, encompassing data-centered and model-centered steps up to deployment. The created model prototypes will be operationalized by Swiss Mobiliar, and, afterwards, monitored and maintained by the data science team.

Data

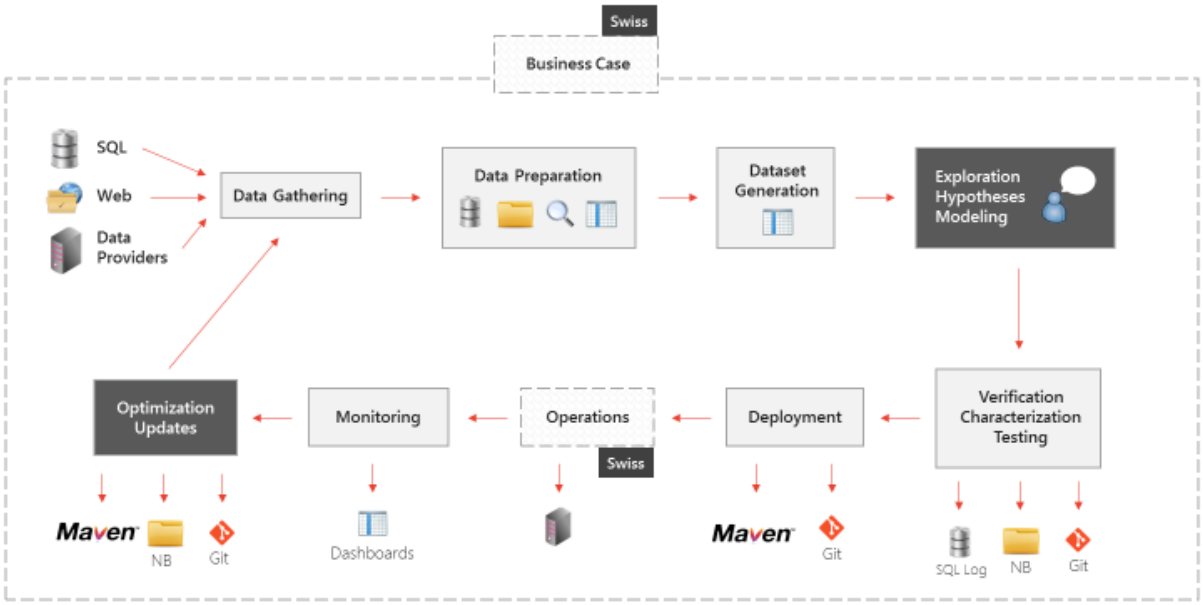


Figure 1 – Model lifecycle representation

Data is the raw material for data science and actuarial models. Naturally, all the services relevant for the preparation of data to be used in advanced analytics models are included in the scope of this proposal.

Data gathering

Data gathering embraces all the ETL processes, connectors and web crawlers that need to be developed and automated by the team to get the information from the different sources, both internal (belonging to Swiss Mobiliar) and external.

The team will be responsible for defining the best strategy to access the data sources based on different criteria, such as the type of database or the format of files, the data provider, and data volume. Depending on the needs expressed by Swiss Mobiliar the team will also be responsible to locate and suggest appropriate external data sources that may be useful, either from the World Wide Web or data vendors.

Data preparation

After the data is obtained, there are some specific tasks that need to be performed in order to ensure the quality and the protection of the data. The team will need to verify the quality of data regarding completeness, up to date status, relevance, consistency and accessibility. This analysis will be conducted using cleaning techniques such as NaN removal, interpolation, imputation, conversion, checks for validity and statistical significance.

In the event that our data scientists will have access to personal identifiable information (PII), such as client names or social security numbers, we will implement dynamic data masking or encryption, depending on the type of data, to ensure the protection of the confidentiality of the data. Furthermore, the team will be accountable for the actions that assure the compliance with General Data Protection Regulation (GDPR).

Dataset preparation

To convert the data into datasets that are usable to build advanced algorithms some tasks may need to be performed, such as:

- Feature selection and engineering
- Clustering
- Aggregation
- Decomposition
- Scaling
- Dimensionality reduction

This step also needs to be performed to feed data into machine learning algorithms that are already in production.

Models

Model Development/Actuarial Science

Based on best practices and recent developments in analytical areas and actuarial sciences, the Closer Data Science team will develop and deliver actuarial and customer analytics models. Examples of such models are: churn curves, churn models, propensity models, customer scoring and segmentation or clustering models.

We will perform the statistical analysis of portfolios of clients and contracts, and analyze and forecast time series of client profiles. Besides descriptive and predictive models, we may also develop prescriptive models such as recommendation/propensity systems, which can suggest custom tailored solutions/offerings to customers.

For the actuarial group, data scientists will estimate lifetimes and lifetime distributions and perform survival analysis, as well as assess goodness-of-fit via Kolmogorov-Smirnov or Darling tests. For non-life models, starting from claim-frequency and claim-severity distributions, aggregate loss models will be created, their parameters estimated and the risk will be quantified. The team will also elaborate state models by using stochastic processes and their discretization as Markov chains, and use maximum likelihood estimators to determine transition probabilities. Special importance will be given to the correct estimation of heavy-tailed function parameters due to the relevance of this type of functions in the insurance sector.

Besides the areas indicated, our core team has proven capabilities in Monte-Carlo simulation of large numbers of scenarios and ensembles of agents and learners, standard Machine Learning and deep learning techniques, numerical optimization of large-scale problems, and

the derivation and implementation of customized stochastic models for financial clients, such as stochastic processes and complex network models.

Our Agile development methodology, detailed later, allows us to rapidly compare and reliably select promising candidates from matrices of new and existing models for a given usage.

Besides that, our consultants are trained to present their results not only in a technical competent manner, but to also summarize and prioritize their findings in an appealing way, to highlight the importance of key details for decision makers, and to understand the mechanisms behind statistical findings, resorting to data storytelling techniques to *tell the story behind the data*.

Text Mining / NLP

Closer Data scientists use standard methodologies for text preparation, such as unicode conversion, tokenization, stemming and lemmatization, stopword removal, synonym replacement, spelling correction, and generation of categorized corpora and/or POS (part-of-speech) tagging.

This enables us to perform Machine Learning tasks on text corpora, such as named entity recognition, keyword extraction, email classification, or topic recognition. We also perform more advanced NLP tasks such as Intent Recognition or sentiment analysis, either via existing proprietary machine learning methods, or using third-party deep learning APIs, such as Microsoft Cortana. Finally, we are prepared to design and implement Deep Learning solutions such as recursive neural networks for sentiment recognition or Hidden Markov Models for POS tagging, for which we may use cloud VM instances for effective training.

Model verification

Our guiding principle in model development is the exhaustive critical testing of newly developed models, and the documentation of test results in the spirit of reproducible data science via notebooks.

We use established methods such as stratified cross validation, learning curves, validation curves, ROC, lift, accuracy, precision, and recall to characterize the performance of models and to select the best model with respect to confusion matrix, implementation aspects, training / retraining effort, and interpretability.

We not only enable comparison of model performance and report feature importance, but also perform statistical tests on target and feature variables, based on which we perform scenario and stress testing of our models before we consider them adequate for deployment.

While we advise our clients about limitations of models and data – in terms of quality and quantity – such that they can make an informed decision about available solutions, we are able to point out possible strategies and alternative modeling approaches, based on our exhaustive knowledge on the latest academic research and available software libraries.

Model Life Cycle Management

The deployment of a model into production is by no means the end of the line regarding the work that needs to be performed in it. Our teams implement a comprehensive policy of model governance to respond to a fundamental set of questions that our clients need to make: Which model created what results and why, how does it perform, how has its performance evolved, and how can it be improved?

For this sake, we continuously log model output to a SQL database, create automatic reports via templates, and enable model performance monitoring for accuracy measures via dashboards, which will be made available to authenticated Swiss Mobiliar users. Our policy also prescribes a versioning system for model source code and a corresponding release management, e.g. via Maven for java and Scala binaries.

We thus, not only control the current state and performance of our models, but can consistently maintain, improve, correct and substitute models, as well as perform periodical or on-demand retraining.

Report Generation

Our proposal includes the creation of dashboards for automatic and continuous monitoring of business KPIs (e.g. leads per marketing channel, claims ratio, policy sales growth), client and portfolio related indicators (such as revenue per policyholder and VaR), but also model performance (such as accuracy, precision and recall, and usage reports). This will allow Swiss Mobiliar users to obtain insights quickly and intuitively, regarding core areas of the business and also the performance of the models behind those insights.

Reports will be generated automatically, using templates, from the monitoring data with a given frequency, to be defined in conjunction with Swiss Mobiliar, or on-demand. Additionally, the monitoring system can be configured to trigger alarms to be sent to stakeholders or key users when user-defined thresholds for model usage, KPIs or security indicators are surpassed.

Just as the other components of the modeling chain, access to the dashboards will be secured and subject to monitoring.

Software development

The scope of the work developed by the Closer team will go far beyond the pure data science work. Advanced analytics algorithms and their subsequent insights need to be made available and need to be usable to serve the needs of the business. With this purpose the team will also develop custom web-based applications that the Swiss Mobiliar users will have access to, in order to, for example, trigger a new run of an algorithm, analyze its results or consult a dashboard that was prepared with relevant information for the operation.

Rapid Software prototyping

In an environment where it is necessary to produce deliverables as fast as possible. Therefore our team needs to be able to design a prototype, collect feedback and introduce that feedback in the development process as fast as possible.

For this purpose, we propose the design thinking methodology to manage rapid prototype development. Using this methodology, we create prototypes, that are often "low-fidelity" (since the aim is to learn and test a concept, rather than test a theoretically finished product), in order to present them to the key stakeholders, effectively involving them in the design process.

In this process we follow these guidelines:

1) *Start small and simple* – Employ prototyping early and often. It leaves little room for interpretation and ensures the involvement of the key users and stakeholders.

2) *Storytelling* - Visualize the concept resorting to visual aids, using as few words as possible. Complexity is added as the prototype progresses.

3) *Showing instead of telling* – The initial prototype needs to feel real through mock imagery, artifacts, and experiences.

4) *Multiple options* – The audience should have access to different choices/scenarios in order to test the prototypes and decide the best options.

5) *External validation* – Every prototype that is designed needs to be validated by its intended users or appropriate stakeholders and not be those who designed it.

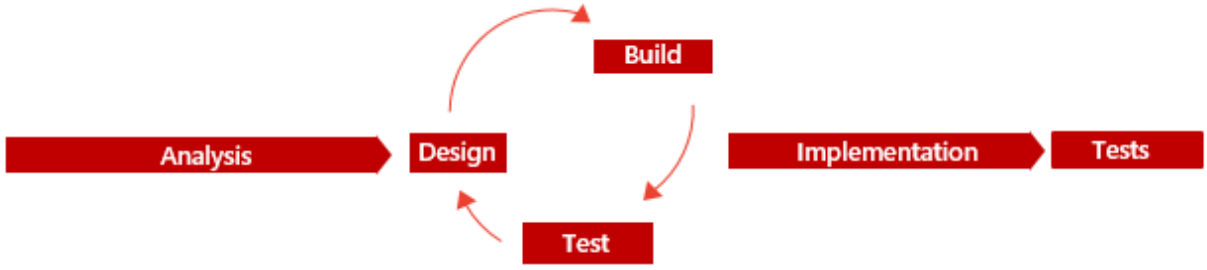


Figure 2 – Prototyping cycle

Software Testing

The Closer Data Science unit will perform standard unit and integration testing using a combination of Junit and Maven frameworks. Besides this, we will also test code performance and profile code whenever it is deemed necessary. Another aspect that will be routinely tested is code usability (UI/UX).

All test results will be documented and logged to the development database.

Server maintenance

Closer will establish a complete development and testing environment based on Microsoft Azure Cloud components. The system will encompass an SQL server for development and logging with automated services for report generation and Azure file storage for development and archival. Additionally, we will add VM instances whenever needed for development, such as Neural Network training.

Closer has been using Microsoft Azure as primary development platform and we find it to be the best solution in terms of safety, flexibility, scaling, state of the art technology, and pricing. Foremost, it guarantees complete compliance with safety and data protection requirements, such as GDPR. Closer will enable access control via VPN, log access to VMs, storage and databases, enable automated backups, and will include a disaster recovery plan, based on the availability of multiple data centers (guaranteeing that only data centers in Europe are considered), and an update management policy for virtual and local hardware.

We will also access the Microsoft Team Foundation Server site, which will be used for project management and Agile planning.

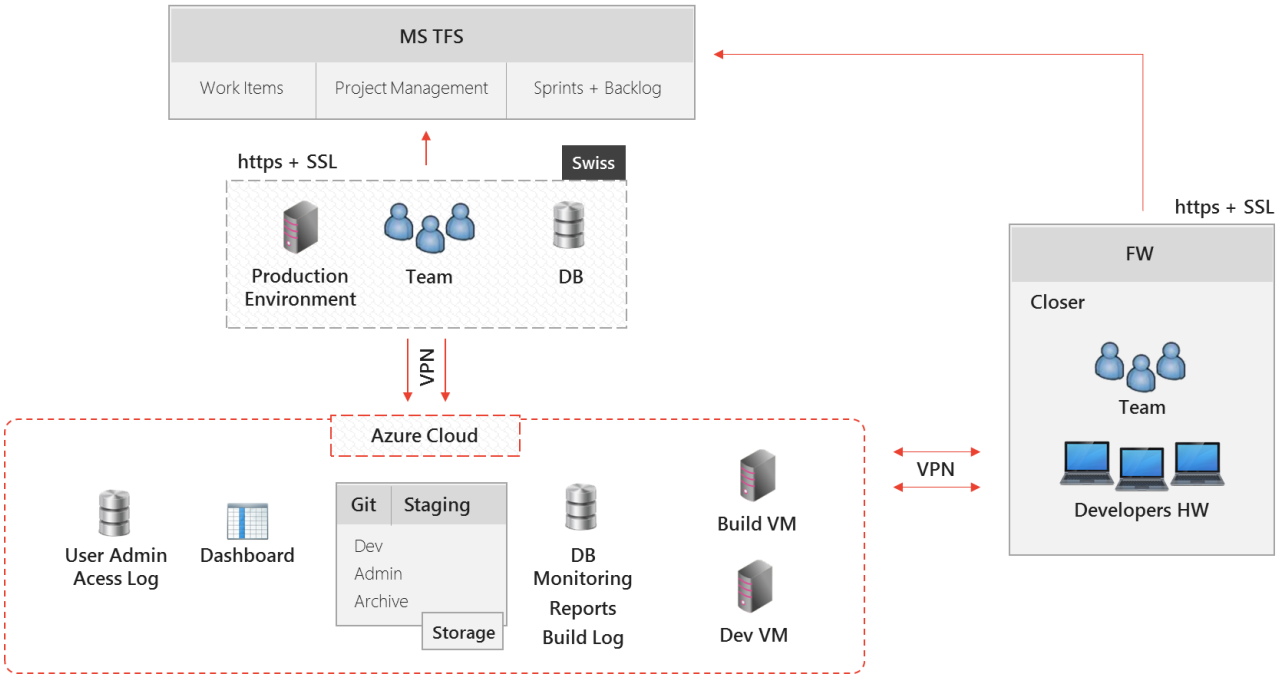


Figure 3 – Proposed Architecture

User support

Following our software versioning and documentation policy, Closer will ensure complete documentation of all models and provide user manuals via central storage. Both source code and the executables in the staging area will be accessible to authorized Swiss Mobiliar users in order to facilitate handover, debugging and maintenance of executables.

Swiss Mobiliar users can formulate feature or bug requests either through the SPOC, or selected users can obtain authorization to directly submit work requests via Visual Studio Team Services. Upon request, we are available to organize training events and workshops for the relevant users.

Methodology

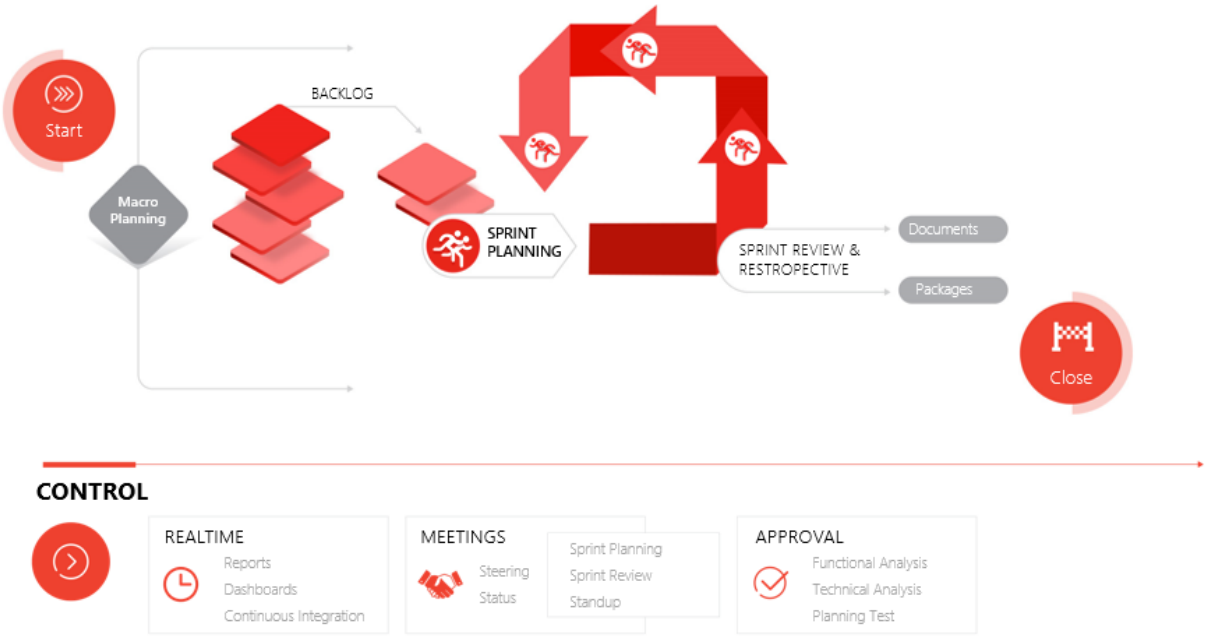


Figure 4 – Agile methodology

Agile Methodology[1] - Tailor made for analytics (e.g. analytics)

Throughout its 11 years of existence, Closer has evolved its development methodology that is based on renowned and increasingly accepted approaches.

Our methodology has Agile as its base, to which we have added a strong relational and behavioral component, centered on the well-being of the individual members of the teams to obtain high performance teams, that can effectively respond to market demands, and can produce the necessary deliverables with a high-quality standard and within the established deadlines.

For Project Tagus, we envision a service that can be viewed as a succession of projects, all with a defined scope, objectives and plan, and so, for each of these iterations, we will go through four macro stages:

[1] Agile methodology should be seen as a suggestion we are proposing. We may tune the team operation with Swiss Mobiliar according to what may serve better its needs and expectations.

1) Start: The Start phase comprises the activities leading to the alignment of expectations among all parties involved in the project. It is done at the beginning of the project, with the Kick-Off meeting and the production of its respective Kick-Off document.
2) Macro planning: Phase that includes the analysis and planning of the functionalities to be made available in the solution that will be developed. The Backlog produced in this phase will have priorities defined based on business needs and does not have to be complete at the end of the Macro Planning phase, since it can grow and be alterable during the subsequent phases as business or client priorities evolve.
3) Development: Set of activities developed by the participating resources to materialize the deliverables defined to achieve the objectives of the business, according to the Backlog. This phase occurs in sprints of one to four weeks, leading to Releases. At the beginning of each Sprint, the necessary tasks for the stories to be developed that are recorded in the Backlog with the highest priority are identified and planned.
   On a daily basis, the development team, led by the Scrum Master, will hold a Standup Meeting where they discuss the evolution of the tasks performed the previous day, set concrete goals for the current day, and discuss impediments to development progress. Possible improvement measures with potential to increase the team's performance are also identified.

   At the end of each Sprint a presentation is made to the Product Owner of the work developed and a retrospective analysis is performed to highlight what went well and identify possible improvement points. During this phase, the Product Owner presents the developments to the users in order to obtain early feedback. They should also update the Backlog and its priorities to take effect in the following Sprints.

   We have successfully employed Sprints for data science modeling, specifically for concurrently testing different modeling approaches, by formulating them as Sprints and having different teams or team members execute them. By carefully defining requirements and a common metric, we thus can quickly and reliably identify the most suitable model.
4) Close: In the closing phase, a set of activities developed at the end of the project are carried out for final validation of the project and evaluation of results. It is during this phase that the activities of previous phases are reviewed. It is also at this stage that final acceptance tests are included.

As an additional macro stage, we also consider a Control component that is transversal to the other four stages. This component can be viewed as the set of activities developed by the management team and direction of the project to monitor its evolution and find the necessary corrective measures to solve potential hurdles.

At the Beginning of the Project, an alignment of the Governance Model to be adopted in the Project is done in conjunction with Swiss Mobiliar. Profiles for the team and their respective responsibilities are defined, the meetings to be held, periodicity, attendance and their respective objectives are determined. The following are the main meetings to be held:

- Steering Meeting: These meetings should be attended by the steering committee and any other relevant people at the moment depending on the subjects to be discussed. They should occur on a monthly basis.

- Status Meeting: In these meetings, the work carried out in the previous week is analyzed and the main Milestones, Risks and Dependencies are reviewed. The necessary attendees to these meetings will depend on the subjects to be discussed and may include the Product Owner, Scrum Master, Architect, Team Leader and Tester. They should occur on a weekly basis.

- Standup Meeting: Daily meeting where each Team member presents what has been done, what will be done and the impediments they have found to fulfill their objectives. These meetings should be no longer than 15 minutes, led by the Scrum Master that will later facilitate the resolution of the indicated impediments. The various elements of the development team and Scrum Master should be present.

The proposed meetings and their respective frequency correspond to our methodology but may be susceptible to adjustments in the beginning of our collaboration to better fit Swiss Mobiliar needs.

The various tasks to be performed within the scope of this proposal will be recorded in a central project management tool where all team members update the progress of their work. Using this information, Reports and Dashboards are made available to the various stakeholder involved in the Project which allows the monitoring, in real-time, of the evolution of the Project.

The Dashboards will allow the visualization of KPI's regarding the project's evolution. They are created and configured to serve the needs of the different team roles (e.g. Management, Team Leader, Scrum Master, Development team, Analysts, Testers). The Reports will help the team to increase its performance. We consider the following reports: Burndown Chart, Release Burndown, Velocity Chart, Cumulative Flow Diagram and other custom reports to visualize the progress and results of work developed in a given moment.

As tools, Visual Studio Team Services and JIRA are two options that effectively meet the requirements and they have been our top choices. However, this choice will be made at the beginning of the Project together with Swiss Mobiliar.

Project HR Management

Recruitment

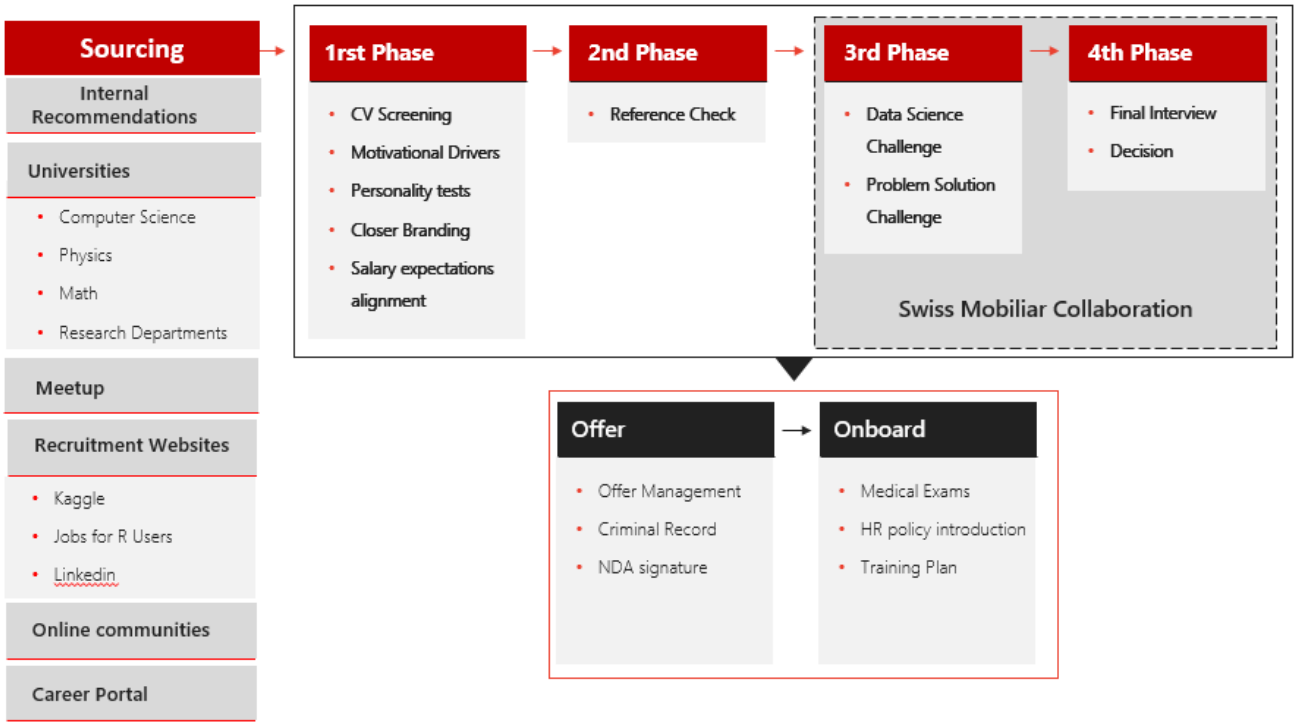


Figure 5 – Recruitment process

Swiss Mobiliar will be included in the recruitment process, regarding the later-stage selection of team elements. We therefore include in this proposal an explanation about how the recruitment process will be implemented to feed this project.

For this type of collaboration, we suggest the active involvement of Swiss Mobiliar on the 3rd and 4th phase of our recruitment process, as is presented in the following diagram:

We would like to highlight some key features for each of the steps displayed on the Figure 5.

Sourcing – we have strong relationships with universities, namely with the research and doctoral students departments, that were built based on the synergy between our PhD´s team and universities. This is our main strategy to attract data scientists to our team. We complement this stage with other activities, such as participation in meetups, online communities and specific recruitment online platforms.

Recruitment phase

- 1st Phase: In this phase, we execute a screening of the CV of applicants and the first interview with the candidates is held. In it we follow a set of procedures, such as assessing motivational drivers, applicant experience, align expectations regarding salary, provide a brief introduction on project and carry out personality tests.

- 2nd Phase: After the initial interview, the provided references are verified and applicants are selected to move on to the following stages of the recruitment process.

- 3rd Phase: Data science challenges are introduced in this phase, usually for profiles with previous experience in similar projects. For other profiles, we choose to apply problem solving challenges. In this Phase we suggest the involvement of Swiss Mobiliar, both for the selection or suggestion of challenges and the evaluation of the applicants' performance in conjunction with Closer.

- 4th Phase: Final interview of selected applicants, held with Swiss Mobiliar team leader, followed by the final decision.

Offer

This stage includes the job offer management and the contract signature. Before signing the contract, the candidates must fulfill the following three requirements:

- Deliver all the necessary documents to the contract signature, such as habilitations and criminal record.

- Sign a Non-Disclosure Agreement

Onboard

In this stage, we will start our admission plan, regarding the onboarding of the new team elements. The most important steps in the admission plan for this project are:

- Project briefing and team introduction

- Training covering roles, responsibilities, processes and controls for handling and protecting customer's data and confidential information.

- Introduce the training plan for the project

- Other general obligations, regarding Closer policy (e.g. medical exams, internal procedures, time report system introduction)

We include in appendix I our general HR screening policy

Training plan

Below we present our suggestion for the training plan for the first year of the service. This suggestion is based on the specific tools required for the project combined with our standard training plan (appendix II).

		Data Science		Business		Wilis Tower Watson			Text Mining/NLP			Development	
	Quarter	IST Machine Learning School	Data Science Specialization (Courses)	Actuarial Science	Swiss Moobiliar Trainnig	Radar	Emblem	Classifier	OpenNLP	Stanford CoreNLP	LingPipe	D3.JS	SCALA
2017	1st Quarter												
	2nd Quarter		X										
	3rd Quarter	X		X					X	X	X	X	X
	4th Quarter		X		X	X	X	X					
2018	1st Quarter		X										
	2nd Quarter	X											

Figure 6 – Training Macro Plan

This plan intends to be a base to work on, so all the necessary training will be agreed and planned between the Closer and Swiss Mobiliar before the project Kickoff. Also, Closer will ensure that every new team element will have access to the necessary training before their integration in the project.

b) Team composition

Organogram

The project team is represented by the following organogram:

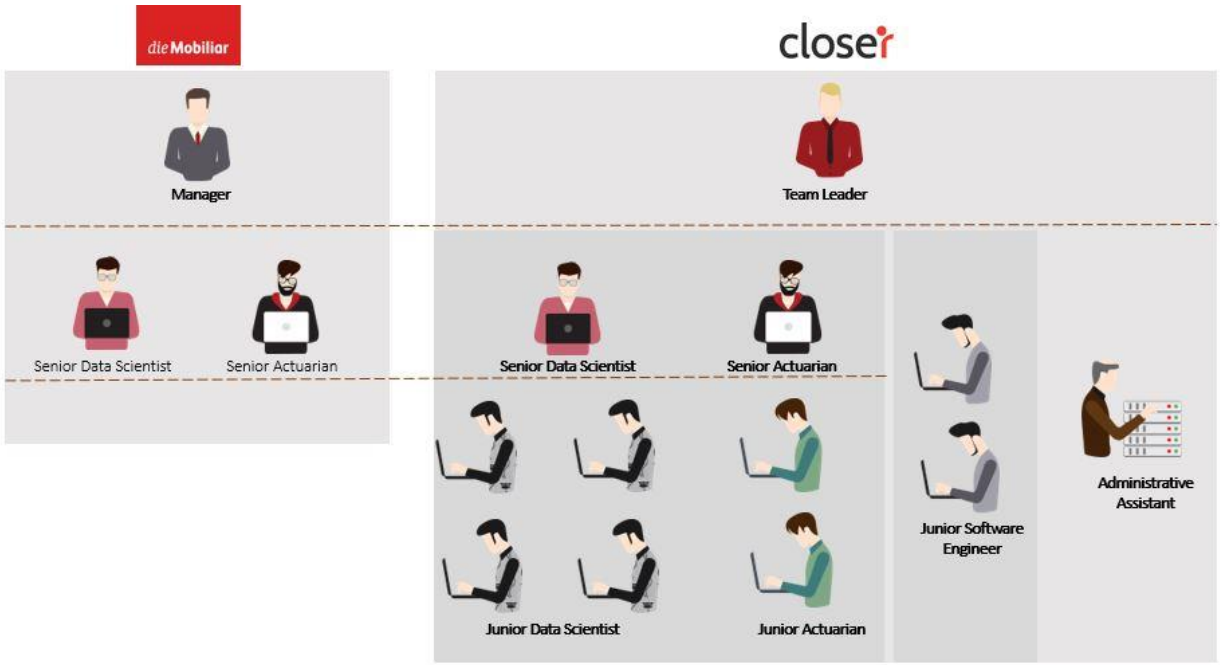


Figure 7 – Team composition

Team Leader – As the most experienced element on the team, he will accompany the team and manage the formal meetings with Swiss Mobiliar team. The Team Leader will also be the primary point of contact. Based on our methodology, the team leader will accumulate the product owner role. He will also be responsible for managing the project's operations/financial reports together with the administrative assistant.

Senior Data Scientist – Highly experienced data scientist who, besides his own tasks, will be responsible for coordinating the data science team and determining/evaluation the best modeling solution to be investigated by the data science team. We envision this role to have an open communication channel with Swiss Mobiliar teams.

Senior Actuarian – Highly experienced in actuarial sciences, he will accumulate the responsibility for accompanying the remaining actuarial team.

Junior Data Scientist – Experienced in data science projects, he will be responsible for the execution of the data science projects, as well as for the completion of routine reporting tasks.

Junior Actuarian – Experienced as actuarial projects, he will be responsible for the execution of the actuarial projects, as well as for the completion of routine reporting tasks.

Junior Software Engineer – Responsible for development and testing tasks, namely in prototype development.

Administrative Assistant – this element will be responsible for the administrative work that supports the project (e.g. travel booking and expenses, meetings schedule, invoices, effort reports, contracts)

Professional experience

This project demands a carefully balanced team, so we will deploy a team that combines academic background (Physics, Mathematics, Computer Engineering and Economics) with relevant experience in data science/actuarial projects.

In terms of language skills, the entire team will have proficiency in English. Furthermore, Closer intends to step up the team with German, French and Italian speakers.

c) Audit approach

Auditing

The information security risk must be permanently addressed and monitored. This can only be performed effectively if done continuously. Microsoft Azure delivers an impressively robust and comprehensive set of audit and logging[2] features which ensure the ongoing monitoring of the infrastructure. Furthermore, they are being constantly updated to face new threats as they are identified.

This is, however, not enough as the complete architecture may include systems not held in the cloud. The overall information security may, therefore, be periodically assessed through security audits. In our perspective, there are some fundamental principles that the auditing process should consider.

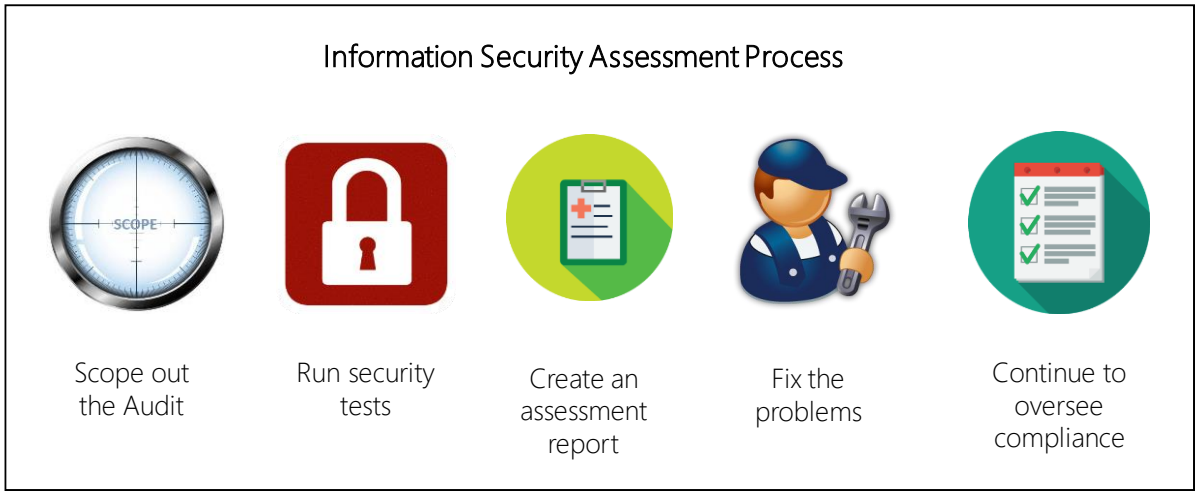


Figure 8 – Steps in the Information Security Assessment Process. Each of them is briefly described in the paragraphs below.

[2] https://www.microsoft.com/en-us/trustcenter/security/auditingandlogging

Definition of Scope

This is arguably the most important phase of a solid information security assessment. The defined scope must ensure that all of the critical systems are under observation since all it takes is one seemingly benign system, network segment or security process to put everything in jeopardy. The whole architecture – whose precise details will be defined with Swiss Mobiliar and may evolve over time – must be taken into account, considering external systems, internal systems and systems hosted by third parties in the cloud.

Security Testing

These are the actual vulnerability scans which will be part of the security assessment. One should look at the whole environment from the perspective of a malicious user, see what can be exploited and then demonstrate what can happen so that the issue can be analyzed and, if necessary, resolved.

Assessment Report

To be effective, a security assessment report must be concise, outlining prioritized, common sense findings and recommendations. The final report should not be unnecessarily long: it just needs to outline specific areas of weakness that require attention from the perspective of a security professional – naturally, taking the context of the systems and the business into account. It can (and should) also incorporate recommendations for fixing/mitigation strategies.

Problem Resolution

In the event of problem detection, prioritization of the problem fixing can be defined and this process should be done in a totally transparent way and in coordination with Swiss Mobiliar. The next information security assessment, will determine whether or not the issues have been resolved. Alternatively, and also considering the severity, a remediation validation of critical and high-priority findings as a follow-up to your security assessment, can be performed soon (30 to 60 days) after the report has been delivered and the findings have been assigned.

Compliance Oversight

Ensuring ongoing security between the security assessments may require tweaking of existing systems and software, possible implementation of new technical controls and an outright overhaul of the policies and processes. This process must be continuously monitored in order to assure that compliance is not only achieved but maintained.

Other Security Issues

Building

All work pertinent to the outsourcing contract will be carried out at Closer Consultoria offices at Amoreiras Tower 1 in Lisbon. This office space provides 24h security.

Information and Data - GDPR

Microsoft Azure public cloud services are used daily by millions of developers and IT professionals. Their infrastructure is designed for hosting millions of customers simultaneously,

and is recognized for its high security standards, upon which businesses can meet their security requirements. In order to earn customers' trust, Azure is basing its security approach on four pillars: a secure and state-of-the-art Platform, Privacy & Controls- which enable customers to rely on standards security and privacy settings and gives them complete control over data access, Compliance, and Transparency.

Due to its commitment to security, Microsoft has received extensive certification relevant for data privacy -such as EU-U.S. Privacy Shield-, government security and financial industry Service Organization Controls standards for operational security (SOC1, SOC2, SOC3). As the provider of the Azure Cloud Services, Microsoft is also committing to be GDPR compliant across all cloud services when enforcement begins on May 25th 2018.

From Closer's perspective, data is an asset whose property is defined by default. This means that any manipulation of data that is not our property needs specific authorization. According to the Laws of Portugal, EU and USA and of most of countries of our customers, manipulation of data is regulated and, in some sectors like banking, non-authorized manipulation of data can be considered as crime. Thus, all projects that need to deal with our customer data, including those that involve their costumers' data are considered Information Critical Projects and must be developed under partner level security policies with separated credentials from the general company credentials.

In practical terms, we will provide hardware exclusively dedicated to this project, which will be linked via VPN to the Azure Cloud. The access to the hardware will be logged internally.

We will then assure compliance of the cloud services by enabling:

- Role based access to Azure storage components
- Write SQL Server Audit Events to the Security Log
- VM internal access log (as for the Closer hardware)
- The architecture presented will be located in a regional European data center (North Europe, which is in Netherlands, or West Europe, which is in Dublin), and it is thus guaranteed by Microsoft that user data will not exit Europe (http://azuredatacentermap.azurewebsites.net/).

Backups and Disaster recovery plan

Our architecture is running all storage on Azure Blobs and managed disks for the SQL VM.

The SQL server managed disks, the VM disks and the storage will be backed up to an additional European blob storage in order to provide protection against user errors, for archival purposes, and for regulatory reasons.

The blob storage itself employs Geo-redundant storage (GRS), which replicates data to a secondary region (hundreds of miles away from the primary location of the source data) within Europe.

As part of the Business continuity and disaster recovery (BDCR) plan, we use Azure Site Recovery, which manages replication for Azure VMs replicating between Azure regions. As such, in case of disaster, all data will be available in a secondary region within Europe.

d) Specific skills

Below we present some of the profiles we may involve as part of the team.

Name	Specialty	Education	Main Competencies	Technologies	Languages Skills
João Jarego	Project Management	MSc Physics SCRUM Master Data Platform (MSFT) Azure Machine Learning (MSFT)	Risk evaluation Human & Economic Dynamics Banking Industry and Financial Markets Complex Systems Big Data Data Science Machine Learning Data Visualization	R SQL Server Power BI ETL Processes (SSIS) Azure Machine Learning	Portuguese English
Cátia Silva	Senior Data Scientist Team Leader	PhD Physics	Data Science Machine Learning Data Visualization Banking Industry and Financial Markets Complex Systems Big Data	Python R SQL Server Power BI ETL Processes (SSIS) Azure Machine Learning	Portuguese English French
Felipe Trenk	Junior Data Scientist	MSc Physics Engineering Machine Learning (Stanford U. via Coursera)	Data Analysis Medical Imaging Machine Learning Data Visualization	Matlab Python Java SQL	Portuguese English German
João Gonçalves	Team Leader Architect	M.Sc. in Informatics Engineering	Banking industry Big Data Data Science Data Visualization ETL processes Data Quality Data warehousing	Teradata SAS Miner SAS Guide PL/SQL SAS DIS SQL Microstrategy	Portuguese English
Pedro Neves	Business Analyst	Post-Doc in Physics	Business analysis Human & Economic Dynamics Complex Systems Data Visualization	Python C++ SAS Enterprise Guide SAS DIS Teradata	Portuguese English
Marta Rolho	Junior Data Scientist	MSc. in Physics Engineering	Data visualization Data analysis and reporting	SQL Phyton R	Portuguese English

Name	Specialty	Education	Main Competencies	Technologies	Languages Skills
			Data Science Machine Learning	LabVIEW Mathematica	
Flávia Rocha	Junior Data Scientist	MsC in Physics Engeeniring	Data Science Machine Learning Data Visualization	Python R SQL C Azure Machine Learning	Portuguese English
João Peixoto	Business Intelligence Data Science	BSc Technology and Information System	Advanced Analytics Business Intelligence Database architecture Data warehouse ETL processes Data management SAS SQL Oracle Teradata	SAS SQL, Oracle Teradata T-SQL PL-SQL SAS Macro SAS Dataset	Portuguese English
Ruben Moura	Team Leader Technical Leader	BSc Computer Engineering SCRUM Master	Functional analysis Software & prototype development Mobile development	Angular NodeJS MongoDB	Portuguese English
Tiago Gonçalves	Technical Leader	BSc in Computer Engineering	Web and Mobile Development Design and UX / UI Organization Delivery from idea to product	MongoDB Java NodeJS Mobile (Android and IOS)	Portuguese English

3. Cost description

In this section, we discriminate the costs of the proposed services. The presented costs do not include VAT. As the service, will be provided in Portugal, the VAT should add to the cost of the provided services. Currently the VAT rate in Portugal is set at 23%.

In the table below we present the cost by profile and according to the job titles defined by Swiss Mobiliar in the RfP.

Job title	Daily Rate[3] (€)	FTE	Monthly Cost (€)
Team Lead	550	0,5	5.775,00
Senior Data Scientist	400	1	8.400,00
Junior Data Scientist	300	4	25.200,00
Senior Actuarian	400	1	8.400,00
Junior Actuarian	300	2	12.600,00
Junior software Engineer	280	2	11.760,00
Administrative Assistant	185	0,5	1.942,50
Total		*11*	73.710,00

Table 1 – Proposed rates by the profiles defined by Swiss Mobiliar in the RfP. The daily rate assumes a standard 8 h allocation. For the purposed of defining a monthly cost estimation, we assume an average of 21 days.

[3] The standard rate assumes a daily 8 hour allocation.

Infrastructure

Baseline

The baseline suggestion comprises a virtual development and build environment in Azure Cloud. It includes the cost for a databases server on a virtual machine (VM), two VM for development and one VM for build purposes, as well as a general-purpose storage component, a VPN gateway, backups, and a site recovery option.

Service type	Region	Description	Estimated Cost
Virtual Machines SQL Server	North Europe	1 Standard VM D12v2 (4 cores, 28 GB RAM,, 200 GB disk) and 3 x 200 GB premium-P20 disks with SQL Server	€738,99
Virtual Machines	North Europe	3 Standard VM D12 v2 (4 cores, 28 GB RAM, 200 GB disk) tamanho: 744 horas	€698,31
VPN Gateway	North Europe	base tier, 1TB outgoing data volume	€97,35
Storage	North Europe	1 TB storage Block type blob, base tier, GRS redundancy, hot access scale, 1 TB IO, 1 TB data georeplication	€136,75
Site Recovery	North Europe	4 recovery instances	€84,33
Backup	West Europe	5 backup instances	€133,74
VSTS		Visual Studio Team Services	€30,00
Support		**Support**	€0,00
		Monthly Total	**€1919,20**
		Annual Total	**€23 030,42**

The total monthly cost for the proposed infrastructure is 1919,20 EUR.

Additional costs per VM and storage/transfer volume

The costs for additional virtual machines, storage options or network traffic is as follows:

Scenario	Description	Additional Cost	Total cost / month
baseline	Standard	-	1919,20 €
alt1	Standard + 500 GB disk storage for DB	80,13 €	1999,33 €
alt2	Standard + 1 TB network traffic and disk IO	136,25 €	2055,45 €
alt3	Standard + 1 TB Blob storage	130,15 €	2049,35 €
alt4	additional dev VM	266,33 €	2185,54 €

Note that additional costs not outlined here may incur. Specifically, very large data volumes or computation requirements will require a redesign of the architecture at a higher cost.

Additional Work

In case of additional work, Closer apply a correction factor to the hourly rates, as we describe following:

Additional work (time)	Correction Factor
From 6.00 pm to 7.00 pm	1.25
From 7.00 pm	1.375
Weekends/holidays	1.5

Example:

Job Title	Daily Rate €	Hourly Rate €	Correction Factor		
			1.25	1.375	1.5
Senior Data Scientist	400	50	62.5€	68.75€	75€

Naturally, these hours of additional work will be previously discussed between Closer and Swiss Mobiliar teams.

Assumptions

In our approach, the following assumptions are made:

- The team will be placed in Lisbon for the entire project duration, namely in Closer offices
- Traveling expenses will be charged for necessary trips separately, according to incurred and demonstrated costs.
- The specific training for the project will be charged, according to incurred and demonstrated costs.
- Closer will have access to a specific server from Swiss Mobiliar infrastructure

Financial Conditions

Closer will invoice the services of the project by the following conditions:

- 100 % invoicing monthly with the project
- Monthly, all the project expenses

Payment term of thirty (30) days from date of receipt of Invoice will apply.

General Conditions

Proposal Acceptance

Swiss Mobiliar can communicate the acceptance of these services through email or Fax mentioning the proposal reference.

4. Bidder Information

Company description

Closer was set up in 2006 with the mission "*to challenge complexity*". We gathered a team of seasoned IT professionals to offer our clients different approaches and new perspectives. We had the courage to combine hard work and knowledge by believing that bright people can make the difference.

From the beginning, we chose to invest in R&D and focused on quick-win solutions in BI, Risk Management, Compliance, Cloud Computing, Mobility, Big Data and Social Networking. We focus on Data & Intelligence and use Advanced Analytics techniques (like simulated annealing, economic variable transformations in machine learning, math models and algorithms) to guarantee that our clients make the smarter decision for any real-world challenge.

Experience

Our experience, can be illustrated by the references in the Table below. These describe briefly some of our main projects. For further detail, we please ask you to consult the appendix III.

Project Name	Main Features	Tools
Impairment Alternative Approach	• Estimation of the credit portfolio risk is unaffected by the expansion/contraction of the credit portfolio; • A lower impairment level is obtained in scenarios of contraction, therefore minimizing its recessive feedback; • Compliant with IFRS 9 regarding the impairment computation as it looks at potential losses considering the full lifetime of the contract (unlike the 12 months currently used).	SQL Server 2014, R, Azure Machine Learning
Credit underwriting in subprime market	• Machine learning (Random Forest algorithms proved to be the best option); • All the several hundred variables from a credit bureau with which the institution interacted were considered (plus several other derived variables); • Back-testing showed a reduction of 25-30% in FPD occurrence, while keeping the underwriting acceptance rate; • Automation of the checking of the classifier accuracy and the need to retrain.	SQL Server 2014, R, Python
Churn mitigation in Banking	• Machine learning and time series analysis; • Loyalty drivers were identified; • Time to action was identified based on the monthly analysis of the customer behavior	SQL Server 2014, R, Azure Machine Learning
Car Buying Likelihood and Profile	• Machine learning and time series analysis; • Use of internal and external data (regarding the car market demographics); • Profiling of the existing customers in terms of car buying likelihood (in a defined period) and car segment propensity	SQL Server 2014, R, Azure Machine Learning
Credit card campaign design	• Machine learning and time series analysis; • Identification of customers for the three different campaigns;	SQL Server 2014, R, Azure Machine Learning

and churn risk identification	• Test campaigns currently ongoing	
Segmentation & Churn Modeling	• Customer value management analytical repository; • Churn, segmentation, X-Sell and Auto-churn modeling; • Web based front-end for analytical engines	BD Oracle 10g,Unix,SAS Enterprise Miner; SAS Enterprise Guide, Java.
Non-maturity deposits projection based on interest rate curves	• Simulation of balances with external factors • Liquidity Forecasting • Reporting through SAS Web Report Studio.	BD Oracle 10g,Unix,SAS Enterprise Miner; SAS Enterprise Guide, Java.
Social Media Presence and Impact Monitoring	• Pre-defined dictionary, contexts and languages (editable in the administration area); • Contextualization of key words by subject (association can be parameterized in the administration area); • Data Visualization (Dashboards/Cloud Tags); • Drill to detail of the specific feed whenever desired	Connectors for the main social media (flickr, tumblr, google, facebook, linkedIn, twitter); Compatible with cloud based or on premise infra-structure.
Prediction of Hourly Energy Consumption for Low Tension Installations	• Estimation of the hourly consumption profile for an installation where the hourly consumption is unknown; • Prediction of hour profiles for all of tension level installations; • Allows comparison to be made for the energy cost, between a profile using only network energy and the same profile using a combination of network and solar energy.	PostgreSQL, R, Python,SQL Server
Chat.Bot	• Performs an automated and real-time conversation using natural language processing, machine-learning and AI algorithms. • Implements user-defined workflows and automated actions	MS Cognitive Services, NLTK, Python, SQL Server
Email contextualization	• Intelligent, large-scale Email contextualization • Seamlessly integrated with a complex back-office solution	Python, SQL Server, NLTK

References

These are some of the companies we work with

5. Legal statement

Closer understands the contract supporting a future agreement for providing the services within the scope of the current proposal will be governed by Swiss law. The exclusive place of jurisdiction will Bern.

We assume that the general terms of the service will be aligned with the "Oustsourcing Agreement" that Swiss Mobiliar shared along with the RfP for "Project Tagus". The specific and final agreement will be defined upon the contract formalization.

The current proposal is valid until the 31st January 2018.

6. Contacts

Av. Eng. Duarte Pacheco,

Torre 1, 15º

1070-101 - Lisbon, Portugal

Tel: (+351) 217 957 426

Any request for additional information may be directed to:

João Jarego
Email : joao.jarego@closer.pt
Tel: + 351 91 103 21 35

Ricardo Cruz
Email : ricardo.cruz@closer.pt
Tel : + 351 91 359 10 49

7. APPENDIXES

Appendix I – Human resources policy

Control	Timing	Scope	Control Description	Why is it important
Identity Verification	Offer	All roles	Verify the identity of an individual by checking and retaining valid, original photographic evidence, and retaining a copy as evidence.	To prove that the individual is who they say they are
Legal Right to Work in Assignment Country	Before Offer	All roles	Check that the individual is legally entitled to work in the relevant jurisdiction by obtaining the original appropriate government-issued documentation, and retaining a copy as evidence.	To verify that the individual is legally entitled to work in the relevant jurisdiction(s).
Address Verification	Offer	All roles	Check that an individual currently resides at a fixed abode, by obtaining a suitable and recent document addressed to the individual that bears their home address and retain a copy as evidence.	To verify the physical address of the individual to corroborate identity checks.
Career and Company Verification	Recruitment process	All roles (excluding roles related to facilities and building domestic maintenance)	Verify: ▪ employment history for the last three years (five years if it is a regulated role); ▪ reference checks to confirm that the employment history was without incident and that career gaps can be evidenced; and ▪ via independent means that the contact details for references from previous employers are bona fide	To confirm: ▪ the suitability and integrity of the person; ▪ that career gaps greater than three months are investigated and assessed to ensure that all information on previous employment is accurate; and ▪ that references are genuine
Criminal Record Checks	Before the offer	All roles	Carry out, via legal agencies, a check for criminal convictions, and retain evidence of such checks.	Checks that the individual is of good character, and helps guard against inappropriate disclosure of information by individuals with criminal or malicious intent.
Changes to Personal Circumstances and Failure to Disclose Information	As required	All roles	Supplier must have a process in place: ▪ for individuals to notify them of relevant changes in circumstances (e.g. subsequent convictions); ▪ to ensure that subsequent discovery of adverse information is investigated and action taken if appropriate (e.g. a criminal charge that was not disclosed during screening).	Verifies that: ▪ changes in circumstances that come to light which may change an initial screening decision are reviewed; ▪ persons failing to disclose material information are removed from assignment;

Appendix II – General training plan for Closer's Data Science team

		Data Science			Open source Technologies				Microsoft			SAS		
	Month	IST Machine Learning School	Text Mining/NLP	Data Science Specialization (Coursera)	Haddop	Spark	R	Python	BI Suite	Azure Machine Learning	Cognitive Services	SAS Base	SAS VDMML	Leadership & Communication Skills
2017	1st Quarter		X	X			X	X	X			X		
	2nd Quarter	X		X			X	X		X	X			X
	3rd Quarter				X	X	X	X						X
	4th Quarter			X			X	X		X	X	X	X	X
2018	1st Quarter		X	X			X	X	X			X		
	2nd Quarter	X		X			X	X		X	X			X
	3rd Quarter				X	X	X	X						X
	4th Quarter			X			X	X		X	X	X	X	X

Appendix III – Closer Project Portfolio

This appendix is a non-exhaustive description of the project portfolio covering the following areas:

Data Warehouse and Business Intelligence – Implementation of DW, Datamarts and BI oriented solutions;
Risk Management – Risk related projects developed in costumers to ensure proper risk management and compliance with regulatory demands;
Forecast and Analytics – Projects with emphasis on forecasting, optimization and analytics;
Research – Our research and original production made to build internal knowledge.

Client-Project relations are not disclosed but can be provided for at justified request. Our references are:

Data Warehouse & Business Intelligence

Bank Corporate DW

Goal	Platform
Full implementation, from scratch, of a bank corporate Data Warehouse together with specific Datamarts and Reports.	• SQL Server; • Integration Services; • Analysis Services; • Business Objects.
Main Features	
• Full business scope, from investment to retail banking; • Full banking regulatory support, including accounting; • Full Data mart support, including commercial, regulatory and management control.	

Bank Corporate DW Re-engineering

Goal	Platform
Replacement of an existing DW with a new structure delivering state-of-the art performance, while ensuring that it can be run with minimal HR.	• Oracle; • Informatica Powercenter; • SAS.
Main Features	
• Architecture and Technical design that ensures full scalability, maintenance simplicity and high performance; • Implementation of application-like release control;	

- Definition of Key Operational Procedures for the DW concerning its interface (in/out), operation, development and nomenclature standards, testing, production rollout and documentation (functional, design and technical);
- Definition of maintenance/support services with corresponding SLA;
- Full support to external applications and datamarts.

Full ETL re-engineering of a bank DW

Goal	Platform
Full optimization of ETL process in a bank Data Warehouse. These were spread between different servers (making the monitoring more time consuming and less clear) and taking an excessive amount of time to run. An execution engine whose parameterization is centralized on control/dependence tables, has cut the execution time by 2/3.	• SQL Server; • Integration Services; • Analysis Services; • Business Objects.

Main Features

- Modular ETL, based on logical blocks with external parameters;
- Centralized control structures for statistical process measurement;
- Parallelized pumping modules to optimize dependences.

Data Warehouse Strategic Definition

Goal	Platform
Review, define and optimize organizational, methodological and documental processes for a corporate Data Warehouse evolution and provide an application to manage it.	• Power Designer; • Power Centre; • Oracle; • Business Objects.

Main Features

- To organize logically all tools that supports the Data Warehouse to provide support to the development life cycle and indentify the managers for development and maintenance.
- To provide rules and methods that establish the sequence of actions that each developer must comply in order that full information consistency is insured.

- To insure the documentation associated with ETL processes and data structures that comply exactly with the production schemas.

Multi General Ledger Model

Goal	Platform
Multidimensional model intended to provide from single business facts the coexistence of several accounting plans and rules.	• SQL Server; • Integration Services; • Business Objects.

Main Features

- Production of balance sheets and accounting maps according to several accounting realities;
- Drill-to-detail on every accounting figure;
- Ability to measure the contribution of each position on several accounting plans;
- Four-eyes adjustment capabilities;
- Auditing capabilities to the minor positions.

Bank Profitability Model

Goal	Platform
Model which gathers the profits/costs associated with any single deal in the portfolio and allows its analysis along the several relevant dimensions (product hierarchy, commercial network et. al.).	• SQL Server; • Business Objects.

Main Features

- Financial margin calculation;
- Commission calculation to each contract;
- Online front-end for process control and administration;
- Reporting.

Bank Funding Modeling

Goal	Platform
Internal funding calculation, for each of the contracts in the portfolio (assets and liabilities) and Kondor+ deal generation. Allows the financial margin calculation and its analysis along the several	• SQL Server; • VB.NET; • Business Objects.

relevant dimensions (product hierarchy, commercial network et. al.).	

Main Features

- Average Outstanding calculations (period and YTD);
- Average refunding rate (period and YTD);
- Refunding interest calculations (oustanding and pastdue based);
- Exception and Alert report;
- Cash flow deal generation to Kondor+;
- Online parameterization and process control front-end.

Bank Objective Control Model

Goal	Platform
Evaluation and reporting of commercial network across its different hierarchic levels.	• SQL Server; • Business Objects.

Main Features

- Definition of objectives per product and agency;
- Definition of bonus/malus policy;
- Definition of weights;
- Web front-end;
- Reporting.

Bank Portfolio Profitability Model

Goal	Platform
Implementation a portfolio profitability system by customer, counter and manager.	• Integration Studio • SAS 9.1

Scope

- Provide customer portfolio performance for commercial purposes;
- Full Drill-to-Detail on customer, manager and counter dimensions.

Financial Institution DW and Budgeting Control

Goal	Platform
Development and maintenance of a Data Warehouse and Budgeting System.	• SAP BW; • SAP SEM BPS.
Scope	
• Insurance group Data Warehouse. • Budgeting System.	

Financial Group HR DW

Goal	Platform
Implementation of a Human Resources Data Warehouse and Tableau de Board for management.	• SAP BI 7.
Main Features	
• Data Warehouse with full HR process information • Designing and production of the Tableau de Bord.	

Insurance Group Financial Control Model

Goal	Platform
Development of a decision support system for financial control for all insurance business areas.	• SAP Neatweaver 2004s – R/3, EDW.
Main Features	
• Revenue assurance of all business areas and accounting reconciliation.	

Process Workflow System

Goal	Platform
To implement a generic workflow system capable of supporting all organization process.	• Oracle; • ASP .NET; • C#; • Analysis Services; • Reporting Services.

Main Features
• Process, status, actions parameterization; • Workflow rules parameterization; • Process control • Automatic decision process parameterization; • Status control; • Alert parameterization; • SLA management and control; • Multidimensional statistical reporting; • Audit Trail; • Process Search.

Risk Management

Bank Credit Risk Management Model

Goal	Platform
Implement a system for regulatory calculations under Basel II Accord.	• SAS Credit Risk Studio; • SAS Data Integration Studio; • SAS 9.1; • Web Report Studio.
Main Features	
• Minimum Capital requirement calculation for Basel II Standard and IRB Advanced methods; • Support information report for regulatory reporting; • Risk management support reporting.	

Market Risk / ALM

Goal	Platform
Implement a system for evaluation e quantification of Market Risk/ ALM.	• ETL Data Integration Studio (SAS); • Oracle; • Total Commander (FTP); • Adaptiv 360 (Sungard).
Main Features	
• Manage the periodic upload of the portfolio; • Manage the upload of risk factors and benchmarks; • Risk engines to evaluate and quantify portfolio risk; • Ad-hoc reporting and regulatory reporting; • Multidimensional reporting and querying; • Scenario drawing for stress testing.	

EAD's & LGD's Evaluation

Goal	Platform
Implementation of a EAD's e LGD's calculation engine. Full historic information upload and information quality analysis.	• Oracle 9i • MicroStrategy • C# .net

Main Features

- Historic data upload to a risk datamart;
- EAD/LGD calculation engine development;
- Detail reporting and recovery system feeding.

Credit Limits Control System

Goal	Platform
Implementation of a management system capable of control exposures and limits in several dimensions and control levels. Data sources: Kondor+, KVaR and IBS.	• Sql Server (DW e ETL em DTS's); • Analysis Services; • Reporting Services; • .Net C#

Main Features

- Exposure proposal management system;
- Multidimensional Corporate exposure management system;
- Exposure control analytics, crossing exposure and limits in all dimensional hierarchies;
- Multidimensional report.

Rating Cross Validation System

Goal	Platform
Provide rating discrepancies between several sources.	• Oracle; • SAS Base; • SAS WEB Reports.
Main Features	
• Based on parameterized sources, the system checks for rating discrepancies that can induce credit process mistakes.	

Forecast and Analytical

Impairment Alternative Approach

Goal	Platform
Since the 2007/2008 crisis, the banks have been under pressure from the regulators to increase their impairment levels based on more conservative assumptions. Besides that, they have been forced to deleverage their loan to deposit ratio, therefore shrinking the credit on their balance sheet. While this happened, non-performing loans remained on the portfolio, therefore affecting the quality of the portfolio and leading to further increase in the impairment level and reduction of new credit concession. This happens because the observed default ratio is considered as a probability of default (which is not a solid assumption when the portfolio is changing). To deal with this we developed a survival analysis approach which relates the PD with the age of the contract. This calibration is done for several types of credit (consumer credit and mortgage have very different maturities and default behaviours). A lower and more stable impairment level is obtained.	• SQL Server 2014 • R • Azure Machine Learning

Main Features

- Estimation of the credit portfolio risk is unaffected by the expansion/contraction of the credit portfolio;
- A lower impairment level is obtained in scenarios of contraction, therefore minimizing its recessive feedback;
- Compliant with IFRS 9 regarding the impairment computation as it looks at potential losses considering the full lifetime of the contract (unlike the 12 months currently used).

Credit underwriting in subprime market

Goal	Platform
Provide a subprime credit card financial institution with an improved underwriting process. The main concern lied on the FPD (First Payment Defaults) which account for most of the losses as this behaviour is seldom recovered.	• SQL Server 2014 • R • Python

Main Features

- Machine learning (Random Forest algorithms proved to be the best option);

- All the several hundred variables from a credit bureau with which the institution interacted were considered (plus several other derived variables);
- Back-testing showed a reduction of 25-30% in FPD occurrence, while keeping the underwriting acceptance rate;
- Automation of the checking of the classifier accuracy and the need to retrain.

Churn mitigation in Banking

Goal	Platform
Identify customers at the risk of churning and the key factors to ensure its retention.	• SQL Server 2014 • R • Azure Machine Learning

Main Features

- Machine learning and time series analysis;
- Loyalty drivers were identified;
- Time to action was identified based on the monthly analysis of the customer behaviour.

Car Buying Likelihood and Profile

Goal	Platform
Identify which existing customers of the bank are considering buying a car (either their first car or by trading in their existing car) in the near future.	• SQL Server 2014 • R • Azure Machine Learning

Main Features

- Machine learning and time series analysis;
- Use of internal and external data (regarding the car market demographics);
- Profiling of the existing customers in terms of car buying likelihood (in a defined period) and car segment propensity.

Credit card campaign design and churn risk identification

Goal	Platform
Provide a credit card company with classifiers which identify three different targets: Customers at risk of stopping using the card; Customers with potential for uplift in the use of the card on dinning; Customers with potential for uplift in the use of the card overseas.	• SQL Server 2014 • R • Azure Machine Learning
Main Features	
• Machine learning and time series analysis; • Identification of customers for the three different campaigns; • Test campaigns currently ongoing.	

Segmentation & Churn Modeling

Goal	Platform
Implementation of an information repository to support behavior prediction models and segmentation.	• BD Oracle 10g; • Unix; • SAS Enterprise Miner; • SAS Enterprise Guide; • Java.
Main Features	
• Customer value management analytical repository; • Churn, segmentation, X-Sell and Auto-churn modeling; • Web based front-end for analytical engines.	

Non maturity deposits projection based on interest rate curves

Goal	Platform
Bank liquidity depends in a big extent from the will of depositors to leave their money on deposit accounts with no established maturity. We modeled the dynamics of these deposits as Ornstein-Uhlenbeck stochastic processes dependent on the interest rate curves for the bank to be able to project liquidity levels.	• BD Oracle 10g; • Unix; • SAS Enterprise Guide; • SAS Web Report Studio; • Java.
Main Features	

- Simulation of balances with external factors
- Liquidity Forecasting
- Reporting through SAS Web Report Studio.

Prepayment Modeling

Goal	Platform
Institutions that actively hedge their risks tend to make derivative deals based on active contracts on their portfolio. But, usually, costumers have the option to prepay or anticipate the end of the contract leaving the institution with open positions on derivatives.	• BD Oracle 10g; • Unix; • SAS Enterprise Guide; • SAS Web Report Studio; • Java.
Main features	
• Modeling of interest rate curves based prepayment model; • Forecasting of options exercise; • Simulation of open position risk; • Reporting through SAS Web Report Studio.	

Curve Modelling

Goal	Platform
Yield curve construction is a fundamental feature for all market/liquidity/interest rate risk systems, including derivative valuation and ALM.	• BD Oracle; • Unix; • SAS Enterprise Guide; • SAS Web Report Studio; • Java.
Main features	
• Cubic splines smoothing techniques; • Bootstrapping; • Forward rate extrapolation algorithms; • Reporting through SAS Web Report Studio.	

Social Media Presence and Impact Monitoring

Goal	Platform
EVALYZE SOCIAL MEDIA PRESENCE AND IMPACT, more than a project, is a productized solution that will guide through a learning process on how to best interact with social media to get the most of it, minimizing negative exposure. With our data collection algorithms we gather information about your presence and your competitor's presence in the social networks giving you the ability to define strategies and actions to capture the best opinion leaders, make successful viral marketing campaigns and put the word of mouth working for you, your products and your services.	- Connectors for the main social media (flickr, tumblr, google, facebook, linkedIn, twitter); - Compatible with cloud based or on premise infra-structure.

Main features

- Pre-defined dictionary, contexts and languages (editable in the administration area);
- Contextualization of key words by subject (association can be parameterized in the administration area);
- Data Visualization (Dashboards/Cloud Tags);
- Drill to detail of the specific feed whenever desired;
- Permanent monitoring;
- Ability to contact the originator of the feed (with proven results in avoiding churn).

Portfolio Management by Random Matrices Filtration

Goal	Platform
Markowitz theory on portfolio management using asset correlation and allocation to minimize risk at constant return or maximize return at constant risk has a problematic flaw. Correlations are the most volatile parameters in the problem. To address this flaw, we developed a random matrices filtration system to remove volatility from correlations combined with an asset allocation algorithm based on simulated annealing.	- BD Oracle; - SAS Enterprise Guide; - SAS Data Integration Studio; - SAS Web Report Studio; - Java.
Main features	
Upload of market time series; Portfolio aggregation/evaluation/optimization; What-if scenarios (back-testing, forecasting); Front-end for visualization; Back-office for parameterization.	

Solar thermal forecasting system

Goal	Platform
To capture information produced in solar thermal panels to allow small energy investors to forecast the equipment production in several time horizons. Energy production forecast is financially critical since small energy producers make future contracts with distributors for short time horizons. Huge energy producers have the means to hedge their contracts by making use of several forms of energy (wind, hydro, solar, thermal, etc…). Small producers do not have that capability and they have a considerable loss due to the lack of a forecast system. For that reason the Portuguese Energy Agency decided to build a platform for small producers where they register their equipment and provide interfaces getting in return the forecast for their plants.	C++; SAS Enterprise Guide; SAS web report studio.
Main features	
Equipment registration;	

Data collection interface (for production time series);

Fokker-Plank time series analysis;

Scalability to other energy sources;

Correlations with weather numeric models output.

Prediction of Hourly Energy Consumption For Low Tension Installations

Goal	Platform
Environment awareness has been growing for about a century. This awareness has been promoting sustainable policies in companies' bussiness plan. Investing in environmental friendly solutions will also improve companies' competitivity by allowing them to have great energy savings, while paying the investment by itself. Allowing customers to visualize how they would improve savings if they change their energy source for a greener one would promote a better adherence for this proposal. For instance, by showing a client how much energy they are spending on a hourly basis everyday and how much they would be saving if they implemented solar panels is a great approach in order to improve market acceptance.	PostgreSQL R Python SQL Server
Main Features	
Estimation of the hourly consumption profile for an installation where the hourly consumption is unknown; Prediction of hour profiles for all of tension level installations; Allows comparison to be made for the energy cost, between a profile using only network energy and the same profile using a combination of network and solar energy.	

Bots for Workflow Automation

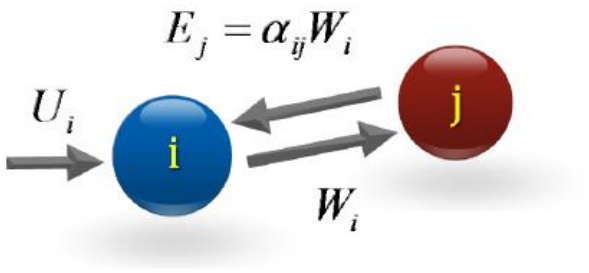
Goal	Platform
Bots are rapidly gaining ground in enterprise-customer interaction, as they can create interest and provide additional information with a human touch, thus creating customer empathy, more leads and improved CRM. Most bots, however. are restricted to performing simple keyword searches that then redirect the customer to help or FAQ documents, with low added value besides an initial interaction. We, on the other hand, have added a new level of customer experience by including a series of innovative workflows, which can guide the customer through complex menus while giving the impression of a natural conversation. Based on layers of 3rd party AI interfaces and machine learning solutions that understand customer intent, our solution learns from past interactions, improving customer satisfaction. On top of that, we have added a general conversational component.	MS Cognitive Services NLTK Python SQL Server
Main Features	
Performs an automated and real-time conversation using natural language processing, machine-learning and AI algorithms. Implements user-defined workflows and automated actions	

Research

Criticality in Economic Processes

(J.P. da Cruz)

The occurrence of huge drops on financial markets that are not explained by the general accepted models is a well reported phenomenon. With the purpose of explaining it, we based ourselves on the reasons for an economic relation between agents to establish the microscopic interactions within the system. Assuming the existence of this mechanism, we can drop all the assumptions on which the stochastic processes that are used in the current models are based, namely that the economy is equivalent to a system in thermodynamic equilibrium, and assume that it is an open system where the different agents interact between them.

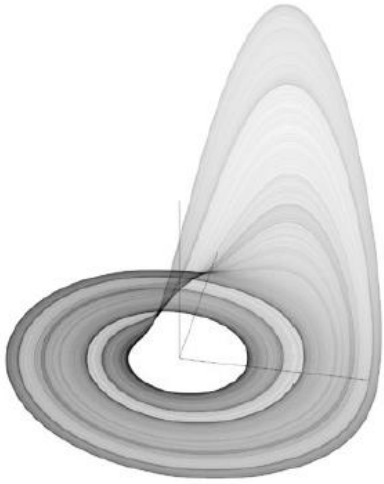


Assuming that the economic agents are organized in complex networks, we show that a system with those interactions only exists in a critical state. Consequently, extreme events with large magnitude have a significant probability of occurrence. We also show that using economic first principles, our model reproduces power laws identical to those that are observed on the real markets and the observed exponents depend linearly from the exponent of the underlying complex network. We argue that, on opposition to some economic theories, the economy is not a closed system on thermodynamic equilibrium but as open system on stationary self-organized critical state.

Stochastic Differential Equations in Finance Modeling

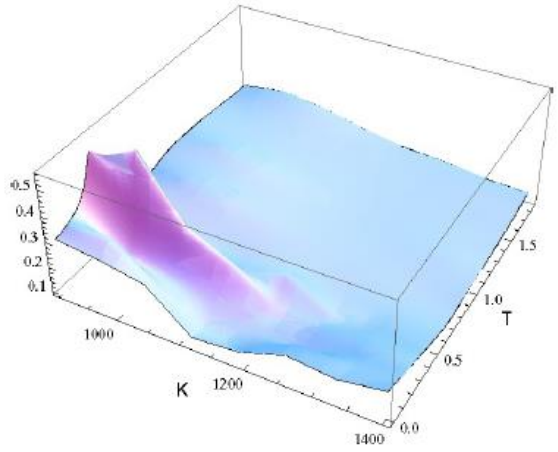
(C. Vaz)

The pricing of derivatives is a useful tool for anyone involved in the derivatives market. Thus, good models and numerical tools for the valuation of these financial instruments are crucial. The classic model for this valuation is the Black-Scholes model. However, due to instability of the market, new models had to arise in the financial landscape, as it is the case of models with jumps and models with stochastic volatility. While the new models have fewer inconsistencies than Black-Scholes model, they also require a careful calibration of its parameters. Here, using the models proposed by Merton and Heston, a calibration is performed against the values observed in the market. After performing the calibration the parameters will be applied in the numerical approximation, aiming to find an option value as close as possible to the given market values.



Finite Differences in European and American Option Valuation

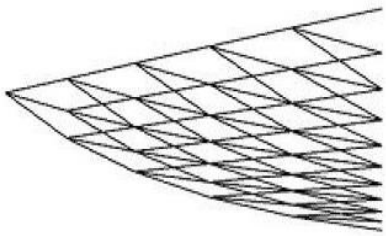
(M.M. Fernandes)



Starting with the Black-Scholes model, we used finite differences schemes to value European and American options. Two approaches will be considered: the transformation of the Black-Scholes equation in the heat equation and the logarithmic transformation of the asset value. The two approaches provide a good approximation to the analytical solution of the Black-Scholes equation. However, it is not possible to adapt the approach using the heat equation to the real world - as the volatility is assumed to be constant with time and asset value, in the standard Black-Scholes model. In real world models but the truth is that it changes with strike price and maturity date of the options - thus, the second approach presents is more general. An application to real data, using the Local Volatility Model to determine a volatility surface consistent with the Black-Scholes equation will be presented. Applying this local volatility surface to the finite difference method, it is possible to calibrate the Black-Scholes model, obtaining a reasonable approximation of the real market values.

Interest Rate Derivative Valuation Using Hull-White Model

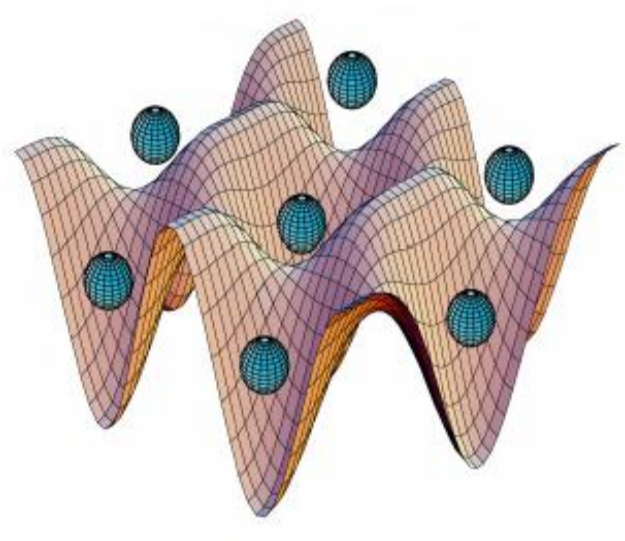
(P. Teixeira)



In this work we developed numerical methods for the valuation of interest rate derivatives using trinomial trees consistent with thee Hull-White model. It is a popular way of modeling the term structure of the instantaneous interest rate. These methods are tested using the available analytical solutions for the most simple financial derivatives. The calibration of the model it's a way of finding the model's volatility parameters that allow us to correctly price the available interest rate derivatives. Specifically, we used the Simulated Annealing algorithm in order to minimize the difference between the model prices and the available calibration instruments. It explores the controlled cooling process of a metal in order to obtain a crystal and it's used, in general, as a way of finding the global minimum of a given function in a large space. To correctly price any financial derivative, the knowledge of the models and techniques available it's a requirement. In the future, the main objective is the integration of these methods in a system capable of valuing a large range of financial derivatives.

Quantum Mechanics Framework to Minimize the Lack of Stationary Properties in Markov-like Credit Risk Models

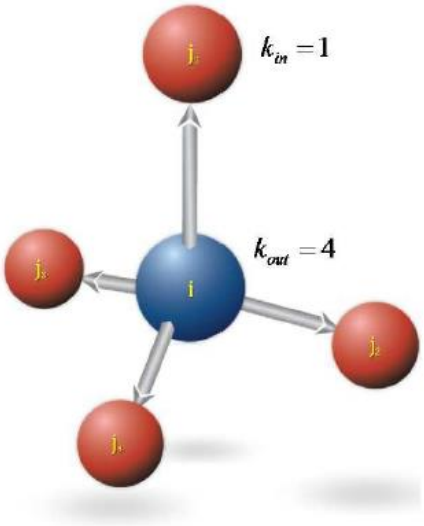
(J.P. da Cruz, J. Jarego, A.C. dos Santos)



The recent crises in financial markets demonstrated that the lack of stationary economic properties in stochastic processes can be a limitation when it propagates into credit risk modeling. To incorporate uncertainty in stochastic metrics, we take advantage of a quantum mechanics framework to redefine the existing credit risk models in order to forecast a final risk distribution. In this work, we define the obligor as a wave function with intrinsically uncertain properties that will produce different risk measures with different probabilities. These probabilities can be derived in order to obtain a complete probability density function that incorporates the lack of stationarity of the variables. This can be achieved without the need to significantly adjust the information infrastructure that underpins the data and models.

On the emergence of critical behavior in evolving financial networks

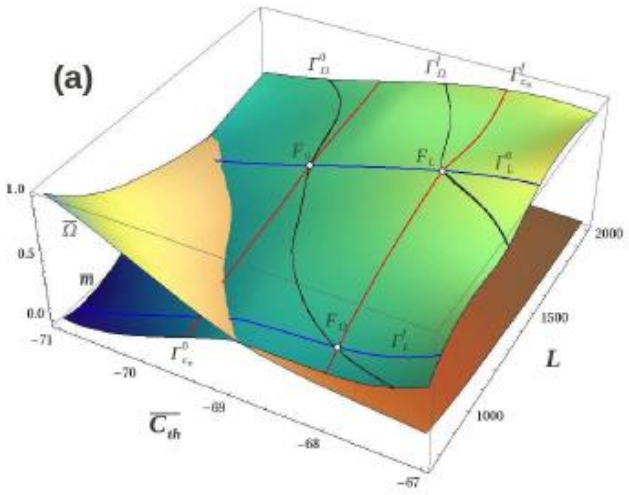
(J.P. da Cruz, P.G.Lind)



We introduce a simple model for addressing the controversy in the study of financial systems, sometimes taken as brownian-like processes and other as critical systems with fluctuations of arbitrary magnitude. The model considers a collection of economical agents which establish trade connections among them according to basic economical principles properly translated into physical properties and interaction. With our model we are able to reproduce the evolution of macroscopic quantities (indices) and to correctly retrieve the common exponent value characterizing several indices in financial markets, relating it to the underlying topology of connections.

The Dynamics of Financial Stability

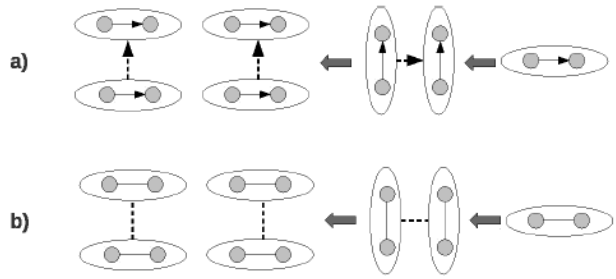
(J.P. da Cruz, P.G.Lind)



The social role of any company is to get the maximum profitability with the less risk. Due to Basel III, banks should now raise their minimum capital levels on an individual basis, with the aim of lowering the probability for a large crash to occur. Such implementation assumes that with higher minimum capital levels it becomes more probable that the value of the assets drop bellow the minimum level and consequently expects the number of bank defaults to drop also. We present evidence that in such new financial reality large crashes are avoid only if one assumes that banks will accept quietly the drop of business levels, which is counter nature. Our perspective steams from statistical physics and gives hints for improving bank system resilience. Stock markets exhibit critical behavior and scaling features, showing a power-law for the amplitude of financial crisis. By modeling a financial network where critical behavior naturally emerges it is possible to show that bank system resilience is not favored by raising the levels of capital. Due to the complex nature of the financial network, only the probability of bank default is affected and not the magnitude of a money market crisis. Further, assuming that banks will try to restore business levels, raising diversification and lowering their individual risk, the dimension of the entire financial network will increase, which has the natural consequence of raising the probability of large crisis.

Bounding heavy-tailed return distributions to measure model risk

(J.P. da Cruz, P.G.Lind)

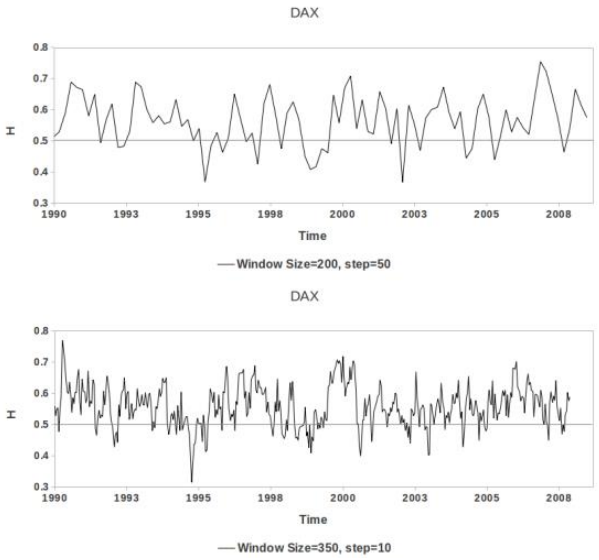


This article makes use of the apparent indifference that the market has been devoting to the developments made on the fundamentals of quantitative finance, to introduce novel insight for better understanding market evolution. We show how these drops and crises emerge as a natural result of local economical principles ruling trades between economical agents and present evidence that heavy-tails of the return distributions are bounded by constraints associated with the topology of agent relations. Finally, we discuss how these constraints may be helpful for properly evaluate model risk.

Fractional Brownian Motion in Finance

(Susana de Matos Neves)

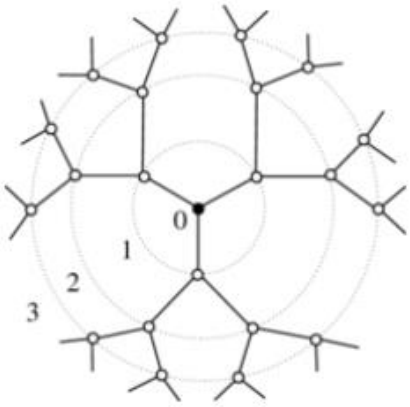
Some of the statistical properties of the financial data are common to a wide variety of markets: long-range dependence properties, heavy tails, skewness (gain/loss asymmetry), jumps, volatility clustering, etc. The need to seek new models for financial products has increased in recent decades due to the inability of current models to explain some of these facts. One of these models is fractional Brownian motion.



This work aims to give an overview of some studies that were done on the financial applications of fractional Brownian motion, in particular the work of Paolo Guasoni and Patrick Cheridito which shows that, if we assume certain restrictions, we can eliminate arbitrage opportunities. Moreover, we also present empirical studies with market data, in order to show how to obtain an estimator for the Hurst index (the fractional Brownian motion parameter). To this end, we used two methods, the Rescaled Range Analysis and the modified Rescaled Range Analysis. This study allows us to discuss the effect of memory on the time series of some market indices.

Percolation in simulated financial systems

(Natália Navin)



It is well known that the market price curves display non-finite variances and do not follow a Brownian motion. This behavior challenges the standard mathematical tools used by the financial mathematics. This work studies the economic variables which improve the predictability of models described by stochastic differential equations regarding the Brownian motion. Accordingly, other market variables, which display finite variations, are selected since these are compatible with the application of stochastic differential equations with Brownian motion.

Since economy results from the interactions of economical agents, we model it as a cluster of connections in which each node represents and economical agent and each connection represents the interactions between agents.

The fore mentioned entities will replace the market price on the modeling of the financial markets. The cluster formation will be analyzed under the complex network theory and its metrics are studied applying percolation and self-organized criticality theories.

At last, we intend to validate if the cluster metrics follow a Brownian motion, which will allow the market evaluation based on the cluster metrics and not on the financial ones.