

SMART Mail

A SMART Platform for Mail Management

Ricardo Raminhos¹, Eduardo Coutinho¹, Nuno Miranda¹, Maria Barbas², Paulo Branco²,
Teresa Gonçalves³ and Gil Palma³

¹VIATECLA SA, Estrada da Algazarra n°72, Almada, Portugal

²Instituto Politécnico de Santarém, Escola Superior de Educação, Complexo Andaluz, Apart.131, Santarém, Portugal

³Universidade de Évora, Largo dos Colegiais 2, Évora, Portugal

Keywords: SMART Mail, Analytics, Visualization, Exploratory Data Analysis.

Abstract: Email is a key communication format in a digital world, both for professional and/or personal usage. Exchanged messages (both human and automatically generated) have reached such a volume that processing them can be a great challenge for human users that try to do it on a daily basis and in an efficient manner. In fact, a significant amount of their time is spent searching and getting context information (normally historic information) in order to prepare a reply message or to take a decision/action, when compared to the actual time required for writing a reply. Therefore, it is of utmost importance for this process to use both automatic and semi-automatic mechanisms that allow to put email messages into context. Since context information is given, not only by historical email messages but also inferred from the relationship between contacts and/or organizations present in the messages, the existence of navigation mechanisms (and even exploration ones) between contacts and entities associated to email messages, is of fundamental importance. This is the main purpose of the SMART Mail prototype, which architecture, data visualization and exploration components and AI algorithms, are presented throughout this paper.

1 INTRODUCTION

In the universe of email management solutions, where the volume of data is continuously increasing, the existence of platforms/solutions that allow the treatment of messages in a graphical and intelligent way (supervised or not – i.e. using automatisms) is more necessity than an optional feature.

The email has been, throughout the years, the most ubiquitous way of digital communication, even taking into consideration the strong growth of instant messaging applications. The previous existence of an email account limits a large part of the online presence of an individual, from social network authentication, online shopping, access to web portals, as well as most forms of online communication.

According to the latest report “Email Market, 2015-2019” from “The Radicati Group, Inc.” from July 2015 (The Radicati Group, s.d.), there are 2.6 billion email accounts (in 2015) and this number is expected to grow above 2.9 billion by the end of

2019.

Table 1: Estimated market growth of email platforms for 2015-2019 (The Radicati Group, s.d.).

Year	Worldwide Email Users (M)	% var	Worldwide Email Market Revenues (\$M)	% var
2015	2 586		\$13 607	
2016	2 672	3%	\$19 353	42%
2017	2 760	3%	\$25 934	34%
2018	2 849	3%	\$32 592	26%
2019	2 943	3%	\$38 917	19%

Email solutions market will exceed \$13.6 billion in 2015, being expected to exceed \$38.9 billion by the end of 2019, which represents an average annual growth of around 30% - Table 1.

The volume of emails generated, including both business and personal emails is estimated to be around 205 billion per day during 2015, increasing to above 246 billion per day by 2019 –Table 2.

Table 2: Estimated email traffic for 2015-2019 (The Radicati Group, s.d.).

Year	Total Worldwide Emails Per Day (B)	% var	Business Emails Per Day (B)	% var
2015	205.6		112.5	
2016	215.3	5%	116.4	3%
2017	225.3	5%	120.4	3%
2018	235.6	5%	124.5	3%
2019	246.5	5%	128.8	3%

While the vast majority of email management platforms are focused on the technical component of the messages exchanged, all “intelligence” needed to organize, prioritise and discover historical content regarding email messages is currently depending on the actual end user intelligence.

Although there are some solutions (please refer to the State of the Art section) that are capable of presenting the user some data/email metrics in a more analytical way, this is still considered a secondary feature.

VIATECLA understands the need for R&D effort on the creation of instruments that offer mechanisms to support the user, including a change of paradigm focused on the presentation and exploration of email data versus the current focus – to send and receive email messages.

That is the purpose of the SMART Mail R&D project (SMART Mail webpage, s.d.), a project developed by VIATECLA (VIATECLA, 2015) and supported by Universidade de Évora (Universidade de Évora, 2015), Instituto Politécnico de Santarém (Instituto Politécnico de Santarém, 2015) and GTE Consultores (GTE Consultores, 2015), and co-financed by QREN (Quadro de Referência Estratégico Nacional) (National Strategic Reference Framework (NSRF), s.d.).

2 THE SMART MAIL PROJECT

With the SMART Mail project, VIATECLA aims to potentiate the importance of email, thus contributing to an increase of productivity to its users, through a software functional prototype that allows the current standard/static system of email messages to have a certain degree of intelligence.

Using SMART Mail it is intended for users to reduce information overload, becoming information more visual by representing email patterns as well as propose priorities for processing email.

Therefore, SMART Mail goals are based on three pillars:

- Enable a set of charts and exploration controllers which applied to email information will allow showing statistical data about email usage (e.g. reception, sending and classification of emails), through the application of statistical processes;

- Data normalization for both Organizations and Contacts, which will make them more suitable to be shared within the business context, not only its specific information but also the meta-information related to its relevance within the organizational context;

- Application of knowledge and rule models, directly to each email message data or to the statistical numeric aggregated values from its history and the definition of reading and reply messages’ priorities.

In this context, SMART Mail aims to be a very valuable tool, improving the productivity of its users and helping them in the decision making process. The knowledge resulted from this investigation has been materialized in a prototype for a generic platform for email visualization and interaction.

This article presents the architecture, data visualization and exploration components, Artificial Intelligence and alert detection algorithms, implemented in the SMART Mail. The article “Email solutions – state-of-the-art and possible evolutions” (Raminhos, et al., 2015), also written in the context of the SMART Mail R&D activities, presents the state of the art for email management platforms in a greater detail.

3 STATE OF THE ART

Generally speaking, although email management platforms have evolved during the execution period of SMART Mail (2013-2015), no disruptive changes have occurred, either at the internal data models, interaction layer, graphic representation or AI Algorithms. The main changes occurring in this area have been driven by research and development initiatives (and later creation of product) sponsored by businesses to be applied to the corporative environment, with very low contribution from the academic sector.

Therefore, initiatives like Sidekick (Sidekick, 2015), Google Inbox (Inbox, 2015) and Verse from IBM (Verse, 2015), end up contributing with advances in the area of email management, through the inclusion of graphical mechanisms and some

intelligence on the comprehension and navigation of emails and contacts. Each one of these initiatives is currently a market leader, specifically:

- Under the form of plugin/add-on, using existing email management platforms, and providing a layer of intelligence on those platforms – Sidekick approach;

- Reinventing the entire email experience, natively integrated in the email client application – Google Inbox approach;

- Under the form of an autonomous platform based on the integration of email management platforms and business analytics/documental platforms – Verse (from IBM) approach;

On the other hand, Xobni plugin, one of the main reference platforms in a recent past, having been acquired by Yahoo! (Xobni Support Homepage, s.d.), has since been withdrawn from the market. The reason behind its acquisition (supposedly to be integrated within Yahoo!'s email offer) is not clear at the moment, and a practical outcome of any integration with Yahoo!'s email has yet to be observed.

From the state of the art analysis, it is observed that there are not many tools that address the email management and optimization of time spent using email. The number of applications is even lower when it comes to solutions available that are capable of visually representing results and perform its analysis.

From the technological point of view, there is a large dependence of Microsoft Outlook under the form of plugin creation for processing, visualization and interaction with the information. In relation to the approaches on the web visualization of email indicators, these are also limited to a specific client (Gmail in the case of Gmail Meter) or by using internal resources (GetResponse (GetResponse Homepage, s.d.)) that do not allow to work with services or clients of external emails.

Based on these conclusions, some trends/possible evolutions foreseen on this domain are presented next:

1. Change of paradigm – from the current one, which is focused on an isolated email message, to an email message “in context”. Each message should be put into historical context, in case it exists (the history), especially with regard to other Contacts and Organizations also present in the conversation;
2. Integrated interface – following the suggested change of paradigm, the way these implementations are put in place cannot be seen as something that is secondary/optional, but as a

new area that is always present and visible in the user interface;

3. Higher intelligence – through reflection on which artificial intelligence capabilities can be applied to emails, either on the proposal of relevance levels, or on the correlation of those levels;
4. Exploratory and Interactive – making graphic controls that go beyond the graphical representation of information, and enable the user to have a certain degree of interaction through: (i) filtering, (ii) search, (iii) definition of temporal scope of the search, (iv) drilldown/roll up mechanisms on the universe of data selected to be analysed;
5. Collaborative – through the construction of repositories (e.g. of Contacts, Organizations) where information can be constructed and consulted in a collaborative way in the context of one entity (e.g. a company) or even in a general/global way;

Considering the aforementioned points, it is noted that there is urgency and concern with these subjects. As it is an element that, for businesses and organizations, involves many working hours of their employees, so the possibility of optimizing that use is entirely relevant, knowing that at any moment, through metric analysis, whatever is being made (whether well or not), and what can be corrected.

4 ARCHITECTURE

The Figure 1 shows a global vision for the SMART Mail architecture. Following the client/server paradigm, the solution is decomposed in four main functional areas that are interrelated, namely (i) SMART Mail Plugin, (ii) SMART Mail Core Server, (iii) SMART Mail Catalogue and (iv) SMART Mail Back office.

Being a prototype directly conditioned by a standard email management platform that deals with the email exchange component, Microsoft Exchange/Outlook was selected as the most adequate for the creation of a pilot capable of demonstrating the capabilities of the developed prototype.

Therefore, the **MS Exchange Mail Server** component (external to the SMART Mail) is responsible for the exchange of messages between users, agenda management and message synchronization with client components (e.g. desktop, mobile). A plugin was developed for the **Microsoft Outlook Client** – the **SMART Mail Plugin** which deals with accessing email messages

locally present (and previously synchronized with the **MS Exchange Mail Server** component) processing them with the purpose of directly obtaining data and metadata and indirectly present in the messages through the **Email Data Access** component.

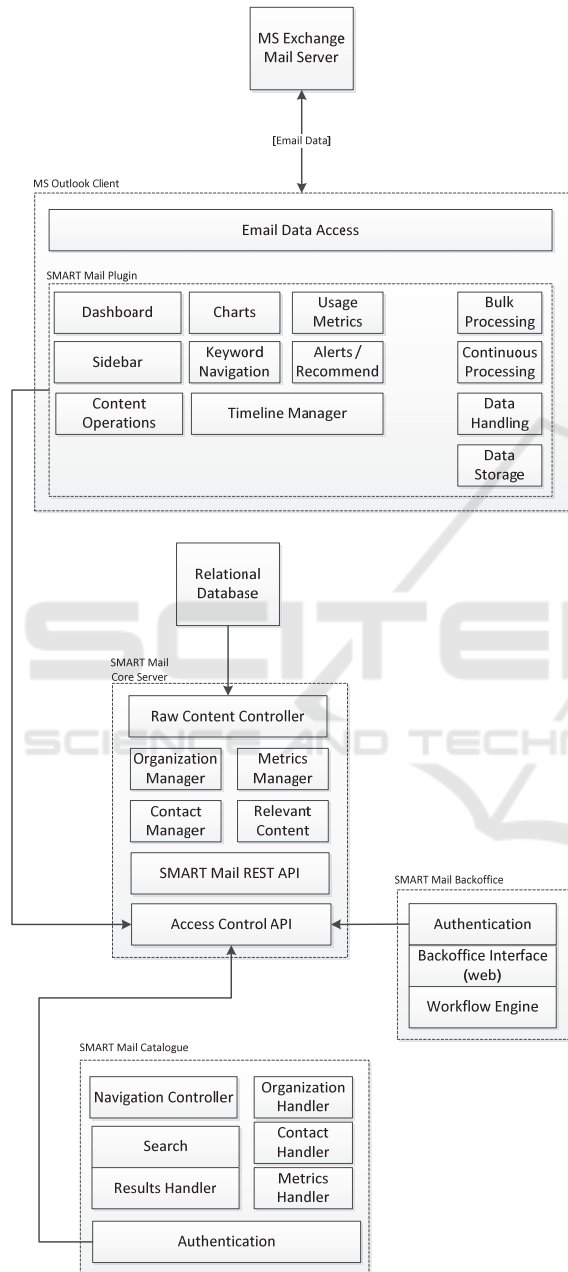


Figure 1: General diagram for the SMART Mail platform architecture.

The data access component is used by two processing modules of the SMART Mail – **Bulk Processing** and **Continuous Processing**. The **Bulk**

Processing component is used in the initial phase of the **SMART Mail Plugin** and in case the user has a high volume of historic emails that need to be progressively processed in background. It is intended that the historical volume be processed as quick as possible, so this module guarantees that this happens with transparency and in a non-intrusive manner to the user’s daily operation – especially during the peaks of information processing. Once the historical information is processed, the **Continuous Processing** module takes the **Bulk Processing’s** place, by processing emails incrementally as they are received. The data obtained from the email messages – regardless of the processing type – are kept in data optimized files within the private area of the user’s file system (i.e. **Data Storage** component), being the **Data Handling** component in charge of the management of all low-level data access.

Visually, two main graphical components exist – **Dashboard** and **Sidebar** - integrated within the SMART Mail Plugin. The **Dashboard** component presents a global vision for email management process – (i) presentation and navigation in emails categorized as high priority items, (ii) navigation and exploration of information from Organizations and Contacts, (iii) possibility to create and update Organizations and Contacts through the **Content Operations** sub-component, (iv) overview of Contact’s history and (v) intelligent interfaces for search and identification of emails using multi-filters.

On the other hand, the **Sidebar** component always shows information in context according to the specific email selected. As a result, it is possible to visualize and navigate on Organizations and Contacts present in the email, as well as to access a set of metrics related to the same email, aggregated by Organizations and Contacts or, in the case of a conversation thread which extends in time, to graphically represent its temporal iterations.

Both **Dashboard** and **Sidebar** include four main components throughout the process, although with a few variants according to the space available for the presentation of the information, (i) presentation and exploration charts (i.e. **Chart** component), (ii) aggregated or non-aggregated numeric values – e.g. by organization, contact, or temporarily (i.e. **usage Metrics** component), (iii) application and extraction of relevant words from the email message, and possible relationships with other messages that share the same set of contents (i.e. **Keyword Navigation** component) and (v) presentation and management of Alerts and suggestions (i.e. **Alert/Recommendation**

Engine component).

The **SMART Mail Core Server** represents the server layer responsible for the persistence of Contacts and Organizations' information, as well as some general metrics related to the relationship between a user and their specific Contacts and Organizations. Due to privacy issues, no information regarding specific email messages is kept in this repository – this shall be responsibility of **MS Exchange Mail Server**. Therefore, all accesses to the server layer (based in human action – e.g. back office authentication request, or based in software programs/automatizms) is performed through the **Access Control API**. Upon successful authentication, access to the **SMART Mail Rest API** is made available on the form of a web service REST for CRUD Organizations management (i.e. **Organization Manager** module), Metrics (i.e. **Metrics Manager** module), Contacts (i.e. **Contact Manager** module) and keyword extraction (i.e. **Relevant Content** module). On a lower level, all the data is kept in persistence in a **Relational Database**, being the **Raw Content Controller's** responsibility to access and manipulate them.

Being possible to manipulate the information from Organizations and Contacts through the **Dashboard** and **Sidebar** components included in the SMART Mail Plugin, the **SMART Mail Back office** enables access to all contents created and their management and administration in a web browser environment. Apart from the consultation and manipulation processes, the administrator has access to the contents being able to decide if it should be removed or whether the access should be or not be restricted to some users. The **Authentication** component is responsible for the user's authentication in the back office, where, if the authentication is successful, a **Web Interface** is made available. Through the **Workflow Engine**, which implements a set of access policies to contents, it is possible to remove/restrict the access to these contents.

Finally, the **SMART Mail Catalogue** provides a web access to the Organizations and Contacts catalogue, from which information is used by the **Dashboard** and **Sidebar** components. Thus, after a successful **Authentication** by the user, it is possible to search for Organizations and Contacts through the **Search** module in natural language terms. Whilst the **Results Handler** manages the results obtained in an intelligent manner, the sub-modules **Organization Handler**, **Contact Handler** and **Metrics Handler** deal with the visualization of Organizations, Contacts and metrics associated to the user that has

performed the search. The **Navigation Controller** module guarantees the navigation process between all entities involved.

5 SMART VISUAL COMPONENTS – DASHBOARD

SMART Mail, as previously presented in the “Architecture” section, has three main graphical interfaces to communicate and interact with the end user regarding statistics on Emails, Contacts and Organizations, namely (i) SMART Dashboard, (ii) SMART Sidebar e (iii) SMART Catalogue. Due to the extent required for describing in detail each of these interfaces, a decision was made to present on extent only the SMART Dashboard interface in the current article.

The SMART Dashboard area can be accessed via a ribbon present at Microsoft Outlook top toolbar section. While the Dashboard is activated only “on request” by the user via ribbon, the Sidebar graphical component is always visible and is located on the right side pane of Microsoft Outlook application (Figure 2).

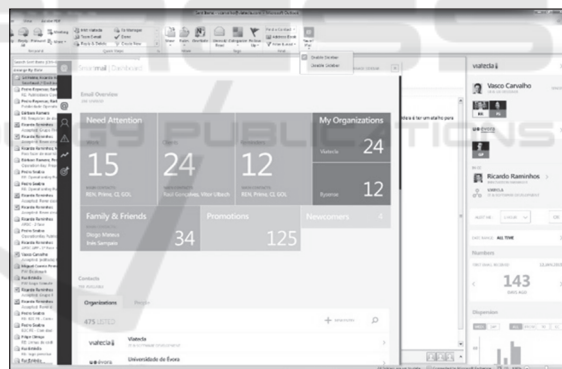


Figure 2: SMART Dashboard – global vision while integrated on Microsoft Outlook client.

The central Dashboard area comprises metrics, graphics and global actions applied to all received email and contacts associated to these emails and the lateral Sidebar area displays content always in context with the selected email (just one email), presenting the related information (e.g. organization, contacts, metrics, charts) accordingly.

In a general way Dashboard works as a “memory panel” which never forgets its previous content/usage, keeping historic behaviour for the latest actions performed on its five main areas, namely (i) SMART Email Overview, (ii) SMART

Contacts, (iii) SMART Alerts, (iv) SMART Overview and (v) SMART Search.

Below the description and presentation of each of these functional areas is presented.

5.1 SMART Email Overview

The Email Overview area (Figure 3) is the first on display at the Dashboard component. It presents an area chart, where the volume of email messages is proportional to the area used by each of the email classifiers, namely:

(i) **Need Attention:** Main focus for the user regarding received messages, and not yet processed. Since SMART Mail is used mainly in corporate context this area is sub-divided in three: “Work”, “Clients” and “Reminders”;

(ii) **My Organizations:** A dynamic area changes according to the number of Organizations the user works for. This separation is mainly relevant in multi-company environments as displayed in the example;

(iii) **Family & Friends:** Even if the SMART Mail focus is mainly corporate it is not possible to isolate familiar/friends message interchange during work time, thus a specific classifier exists for that purpose;

(iv) **Promotions:** Promotion / advertisement emails which although secondary are not considered SPAM messages;

(v) **Newcomers:** Reference to emails for which there is not enough classification information, thus being classified in this area (rather any of the previous ones).

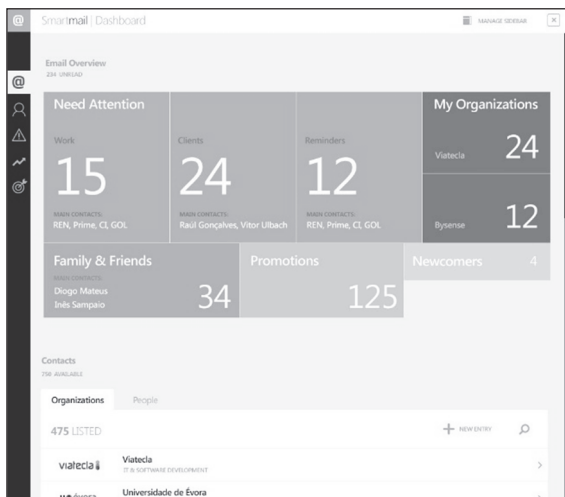


Figure 3: SMART Dashboard panel.

Besides the number of received emails messages associated to any classifier area, the main related Contacts and Organizations are also presented in the case of the areas “Need Attention” and “Family & Friends”. In the classifier “My Organizations”, each Organization is identified by its name in the specific area.

If the user selects any of the “Need Attention” internal areas, he is positioned in its details represented as a two dimension matrix with all Organizations involved in sent email messages (horizontally) and two columns with “New” and “Follow up” classifications (vertically).

5.2 SMART Contacts and Organizations

The “SMART Contacts and Organizations” area is activated any time the user (i) selects a Contact or Organization within any Dashboard area and intends to see its detail, (ii) navigates specifically to the area – using the left-side anchor button (Figure 3) or by scrolling on the Dashboard panel.

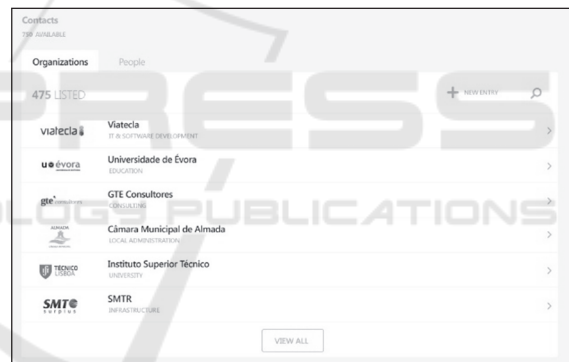


Figure 4: SMART Dashboard – Organization listing.

While in the first case the detailed record is presented immediately for the selected Contact or Organization, in the second a list with Contacts and Organizations available is shown.

As an example, Figure 4 presents a list with the Organizations previously created, while pressing the “People” tab will present, in an equivalent way, a Contact listing.

While in listing mode each Organization/Contact is characterized by its logo/photo, name and business area/business role, access to its detail is performed by clicking on the full extent on the listing entry. Also, in this panel, it is possible to create a new Organization/Contact (according to the selected tab) or to identify a content using textual search.

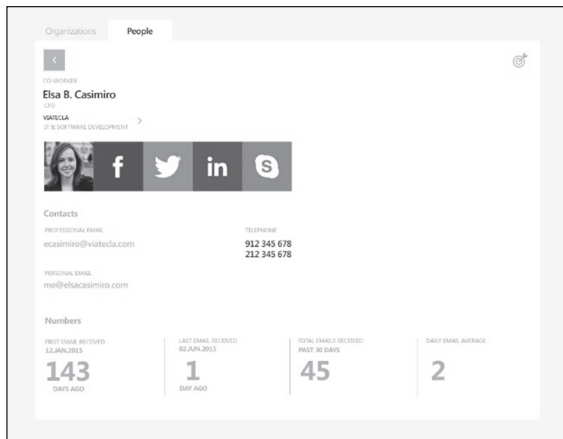


Figure 5: SMART Dashboard – Contact detail.

Figure 5 presents the detail for a Contact where, besides its associated information, a set of statistics regarding previous email interaction history between the user and the selected Contact is shown.

5.3 SMART Alerts

SMART Mail presents two types of alarms in the Dashboard area.

The “Losing Contact” (Figure 6) alarm refers to personal Contacts which, according to its previous email exchange record, detects that the frequency of exchange emails has been progressively decreasing (e.g. in the last days, weeks, months) accordingly to the message volume exchanged in the past.

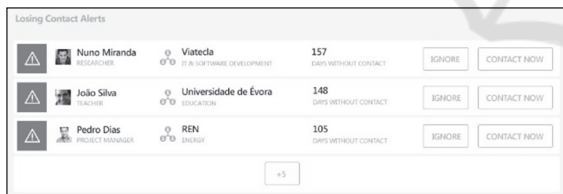


Figure 6: SMART Dashboard – “losing contact” area.

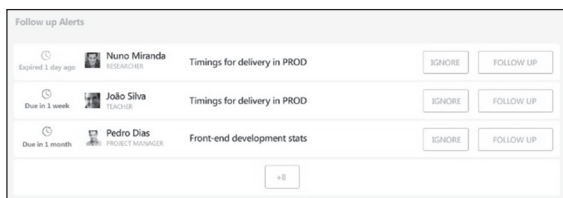


Figure 7: SMART Dashboard – “follow up” area.

Thus, in order not to “lose contact”, an alert is issued (initially in the form of a popup message) where the user is invited to “reconnect” with the Contact (or ignore this proposal). This feature can be of relevant

importance especially in commercial context or when associated with networking actions.

Consulting the Dashboard the user can see which alerts have been issued (and are active/not ignored), with indication on the Contact and associated Organization (if existing) and the period extension in which no contact has occurred. The “Ignore” action will cancel explicitly the alarm while the “Contact Now” action will create a new email message for the selected Contact.

The second type of alarm – “Follow up” (Figure 7), is always defined explicitly by the user. Through the Sidebar, and having selected an email, it is possible to associate a future moment/date to perform a follow up action. Once this period is attained a popup message is presented to the user alerting that is time to act. On the other hand, the user can, by accessing Dashboard, consult the scheduled follow up actions, ordered chronologically, and decide to cancel them or provide an early reply. A special highlight is presented for those alarms for which the initial date/time to reply has been exceeded.

5.4 SMART Overview

The SMART Overview control (Figure 8) enables a graphical display using an area representation, where received email messages are mapped to the associated Organization within a time frame. The area size is proportional to the volume of received emails within the defined time frame.

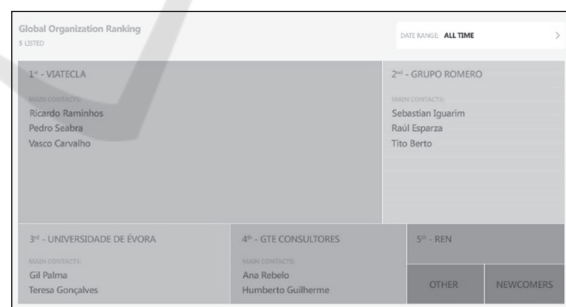


Figure 8: Graphical overview for the received emails associated with Organizations in a time frame.

The time frame definition is performed on the right most control where the user can select the full extent of time (from the time he/she received the first email message until current time) or define a specific time interval.

The user, when selecting an area associated to an Organization, visualizes a list for those emails (ordered chronologically).

5.5 SMART Search

The last visual control present on the Dashboard is the SMART Search which enables the user to identify email messages according to the relations between Organizations, Contacts and Conversation Threads.

The usage of this control can be triggered in one of two ways: (i) in context – when selecting an Organization, Contact or Conversation Thread, the control is invoked, making the selected value to be automatically pre-selected in the dynamic search filters or (ii) without context – directly accessing the control via Dashboard where none of the dynamic search filters present pre-selected values.

Thus, by defining values to the Organization, Contact or Conversation Thread filters, associated emails are dynamically filtered. Associated to each filter area, there's the possibility to perform a textual search in order to faster determine the filter value to apply. Further, each filter area can be expanded/collapsed in order to gain further space area when the number of search results/available filter options is high.

When the user selects a specific email the Dashboard area is hidden and the email is selected in the Microsoft Outlook client, being this information completed by Sidebar that presents the associated metadata regarding the email.

6 SMART RELEVANT CONTENT

The recognition of keywords is a subarea of knowledge from information extraction that intends to identify and classify relevant elements on the text – e.g. which mention pre-defined categories as people's names, organizations, locations, time expressions, monetary values and percentages. For those detected keywords (retrieved from the email's *corpora*) which are found to be transversal to multiple / different users, can be proposed as a dynamic classification attribute.

There are several approaches to the problem (Nadeau & Sekine, 2007), from systems that use rules defined by people (Aberdeen, et al., 1995) to systems based in automatic learning techniques (Mitchell, 1997) through the use of classification algorithms as decision trees (Baluja, Mittal, & Sukthankar, 2000) or support vector machines (Takeuchi & Collier, 2002).

The process implemented for SMART Mail for the automatic classification uses a generic inference process (commonly designated by "learning") from

which a classifier is automatically built.

During the learning phase, several examples of email text with keywords manually classified are provided, from which the algorithm concludes the features that define each one of those entities. This conclusion is reached using a set of attributes that characterize each word that can be found in the email text. As a result, a model of knowledge that summarizes the rules of identification of keywords is obtained.

Later in the classifying phase, the model obtained is used to identify entities on new email text.

There are two main phases during the learning and classifying process. One is responsible for the extraction of spelling and morphological attributes from each word; the other one is responsible for identifying and classifying keywords within the text.

The morphological attributes are obtained from SVMTool (Gimenez & Marquez, 2004), a word labeller that together with the spelling attributes is used in a second stage for the construction of the entities classifying model.

The SVMTool is a sequential label generator based on support vector machines (Cristianini & Shawe-Taylor, 2000) (the same automatic classifying algorithm already mentioned above). As such, for it to be used, it is necessary to have a model for the language in which the email messages are written in. As the Portuguese model is not included in the available language models (i.e. Spanish, Catalan and English) it has been necessary to develop it.

On an operational level, the two stages have a component in common that is based on automatic classification algorithms. The entity recognizer, like the labeller, uses support vector machines.

The choice of this algorithm was due to a set of tests made initially in order to assess the performance of some algorithms recommended by the scientific and academic literature as the most suitable (Quinlan, 1993), (Zhang, 2004), (Caruana & Niculescu-Mizil, 2006), (Keerthi, Shevade, Bhattacharyya, & Murthy, 2001), (Witten & Frank, 2005).

Table 3: Spelling Attributes.

Spelling Attributes			
!	?	Uppercase	Capitalization
()	Lowercase	Alphanumeric
“	‘	Unique character	Numeric
;	:	Upper and lowercase	Letters
+	-	Initial on its own	
«	»	hyphenated words	

Table 4: Morphological Attributes.

Morphological Attributes		
Determiners	Common noun	Verbs
Proper noun	Contraction (determiner)	Prefix
Conjunction	Contraction (adverb)	Adverbs
Adjectives	Contraction (pronoun)	Pronoun
Interjection	Punctuation	Preposition

Following, the attributes used in the classification of words as point of entry for the keyword extractor, are described.

The spelling attributes, in a total of 25 binary attributes, are extracted from the words orthographic characteristics. Table 3 lists all attributes considered. The morphological attributes, in a total of 15 binary attributes, indicate the morpho-syntactic class of each word in the text. Table 4 lists all attributes considered.

To improve the efficiency of the process and to add information from the context of the word analysed, a “window of context” is created, which consists in joining the describing attributes of the neighbour words (before and after the word that is being analysed).

As an example, for a window of context of size 5, the analysed word is present at a central position and the two previous and two posterior words are also considered in the classification process, as depicted bellow.

word - 2	word - 1	analysed word	word + 1	word + 2
----------	----------	---------------	----------	----------

Therefore, the base pillars for the operation of the keyword extraction mechanism have been listed: the window of context, extraction of spelling attributes, extraction of morphological attributes via support vector machines and finally, the final classification of keywords using support vector machines, and consuming all available resources obtained from the stages previously mentioned.

7 EVALUATION AND FUTURE WORK

For the evaluation of the SMART Mail prototype two validation test pilots have been performed – one conceptual and another functional. The SMART Mail conceptual pilot was performed by creating a test case simulation/ scenario /story using a set of images (resulting from the design mockups) for the different interface screens. Using Microsoft Sketchflow technology (Corporation, 2015) it was possible to create interaction areas in the design

mockup images and simulate a navigation flow, making it a dynamic and living experience. This way it is possible to attain a conceptual prototype (even if static and image based) for testing the concept, organization, perception and navigation flow.

The test case environment took place during 3 days (with 1 week interval between each) where new prototype versions were presented according to the collected feedback in the previous version. In the first session both project, concept, objectives and the simulation test scenario itself were presented to the tester team involved in the process.

The main feedback recoiled from these sessions focused the following points:

- Layout issues – both comprehension related and ergonomics – i.e. positioning, size, colours and graphical elements;

- Discussion, proposal and conception of the “Losing Contact” area (integrated in the Dashboard component);

- Inclusion of social networks information associated to both Contacts and Organizations;

- Proposal on removal the “Attachments” visual control (integrated in the Sidebar component) which would aggregate links for the attachment files present in an email message.

The functional pilot testing had the duration of one month and was performed on a continuous way using the SMART Mail prototype software which was available at the time. Also, during this period, feedback (of a more functional nature) was collected, namely:

- Layout issues – both on understanding and ergonomics – i.e. positioning, size, colours and graphical elements;

- Issue identification and solving;

- Discussion, proposal and conception of the “SMART Search” area (integrated in the Dashboard component);

- Proposal and integration of “Take me there” functionalities which would potentiate georeferenced information (mostly Organization addresses) integrated with external map applications for both Desktop, tablet and mobile environments;

- Proposal and inclusion on the “SMART Catalogue” interface of high level statistics regarding the communication between the end user and other Contact or Organization;

- Inclusion of a “Clients” section in the Dashboard’s “Need Attention” area (previously the proposal was focused on “Work” and “Reminders” sections).

In a complementary way, also during the functional prototype testing, some ideas/feedback

regarding the possible commercial positioning of SMART Mail were addressed/discussed:

- Development of a 100% web interface which would not be dependent on a Desktop based email client (such as Microsoft Outlook which was used as a first testing environment);
- Possible SaaS business model (“Software as a Service”) where the end user would refund the SMART Mail promotor according to the actual platform usage;
- Possible unification of both Contact and Organization repositories at a higher macro level which could be used “openly” by different independent entities;

During the operational prototype testing, in order to access the usefulness of SMART Mail, a set of tasks (in the format of a script) was provided to user subjects. Group A performed tasks supported with SMART Mail while Group B used a standard desktop email client. Users using SMART Mail, and after the initial learning curve, proved to be 5% - 15% more productive (i.e. time per task) than Group B users.

Both pilots strongly contributed for testing, validation and evolution of SMART Mail.

Being available at the current development stage a first functional version of SMART Mail, future work will mainly be directed to the promotion and support of “live” clients in real environments. According to the feedback collected from both users and enterprises, future SMART Mail clients, a roadmap (both technological and business oriented) will be defined in order to contribute to the platform further refinement and evolution.

REFERENCES

- Aberdeen, J., Burger, J., Day, D., Hirschman, L., Robinson, P., & Vilain, M. (1995). {MITRE}: description of the Alembic system used for {MUC-6}. *MUC6 '95: Proceedings of the 6th conference on Message understanding* (pp. 141-155). Morristown, NJ, USA: Association for Computational Linguistics.
- Baluja, S., Mittal, V., & Sukthankar, R. (2000). Applying Machine Learning For High Performance Named-Entity Extraction. In *Proceedings of the Conference of the Pacific Association for Computational Linguistics*, (pp. 365-378).
- Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. *ICML '06: Proceedings of the 23rd international conference on Machine learning* (pp. 161-168). New York, NY, USA: ACM.
- Corporation, M. (2015). *Sketch Flow*. Retrieved from Cristianini, N., & Shawe-Taylor, J. (2000). *{An Introduction to Support Vector Machines}*. Cambridge University Press.
- GetResponse Homepage*. (n.d.). Retrieved 2015, from GetResponse: <http://www.getresponse.com/>
- Gimenez, J., & Marquez, L. (2004). {SVMTool: A general POS tagger generator based on Support Vector Machines}. *Proceedings of the 4th LREC*.
- GTE Consultores*. (2015). Retrieved from <http://www.gte.pt/>
- Inbox*. (2015). Retrieved from <https://inbox.google.com>.
- Instituto Politécnico de Santarém*. (2015). Retrieved from <http://www.ipsantarem.pt/>
- Keerthi, S., Shevade, S., Bhattacharyya, C., & Murthy, K. (2001). {Improvements to Platt's SMO Algorithm for SVM Classifier Design}. *Neural Comput.*, 13(3), 637-649.
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1), 3-26.
- National Strategic Reference Framework (NSRF)*. (n.d.). Retrieved 2015, from <http://www.qren.pt/np4/home>.
- Quinlan, R. (1993). *C4.5: Programs for Machine Learning*. San Mateo, US: Morgan Kaufmann.
- Raminhos, R., Coutinho, E., Miranda, N., Barbas, M., Branco, P., Gonçalves, T., & Palma, G. (2015). Email solutions – state-of-the-art and possible evolutions. *Sidekick*. (2015). Retrieved from <http://www.getsidekick.com/>
- SMART Mail webpage*. (n.d.). Retrieved 2015, from <http://www.viatecla.com/inovacao/smart-mail>.
- Takeuchi, K., & Collier, N. (2002). Use of support vector machines in extended named entity recognition. *COLING-02: proceedings of the 6th conference on Natural language learning* (pp. 1-7). Morristown, NJ, USA: Association for Computational Linguistics.
- The Radicati Group, I. (n.d.). Retrieved Julho 2015, from <http://www.radicati.com/wp/wp-content/uploads/2015/07/Email-Market-2015-2019-Executive-Summary.pdf>.
- Universidade de Évora*. (2015). Retrieved from <http://www.uevora.pt/>
- Verse*. (2015). (IBM) Retrieved from <http://www.ibm.com/social-business/us/en/newway/>
- VIATECLA*. (2015). Retrieved from <http://www.viatecla.com>.
- Witten, I., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques* (2nd ed.). San Francisco, US: Morgan Kaufmann.
- Xobni Support Homepage*. (n.d.). Retrieved 2015, from Xobni Support: <https://support.xobni.com/home>.
- Zhang, H. (2004). The optimality of naive Bayes. *Proceedings of the 17th International FLAIRS conference*. AAAI Press.