MASTER'S DEGREE IN INDUSTRIAL ENGINEERING

MASTER THESIS

IMPLEMENTATION OF AN RPA TECHNOLOGY MAKING USE OF ARTIFICIAL INTELLIGENCE WITHIN THE STG WEB FLOW CHART

Author: Ángel Burgos Rastoll

Academic Advisor: David Patiño Mochales

Industrial Advisor: Javier Matanza Domingo

Madrid

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. **Ángel Burgos Rastoll** _____

DECLARA ser el titular de los derechos de propiedad intelectual de la obra:

Implementation of an RPA technology making use of Artificial Intelligence within the STG web flow chart, que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar "marcas de agua" o cualquier otro sistema de seguridad o de protección.
b) Reproducirla en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
e) Asignar por defecto a estos trabajos una licencia Creative Commons.
f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente)*.

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

a) Que la Universidad identifique claramente su nombre como autor de la misma
b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
c) Solicitar la retirada de la obra del repositorio por causa justificada.
d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e intereses a causa de la cesión.

d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusive del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- La Universidad adoptará las medidas necesarias para la preservación de la obra en un futuro.
- La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a 28 de Agosto de 2021

ACEPTA

Fdo. **Ángel Burgos Rastoll**

Acknowledgments

I would like to thank David Patiño, Jorge García and Javier Matanza for their help and advices that led to the conclusion of this report.

IMPLEMENTACIÓN DE UNA TECNOLOGÍA RPA HACIENDO USO DE INTELIGENCIA ARTIFICIAL DENTRO DEL SISTEMA STG WEB

Autor: Burgos Rastoll, Ángel.
Supervisor de la compañía: Patiño Mochales, David.
Supervisor académico: Matanza Domingo, Javier.
Entidad Colaboradora: Iberdrola

RESUMEN

Este informe tiene como objetivo capturar una imagen completa de la web de STG en Iberdrola con el fin de proporcionar suficiente información al proveedor de RPA para que el robot que automatizará todo el sistema pueda tener suficientes argumentos de entrada y salida para actuar correctamente.

Palabras clave: RPA, automatización, STG web

1. Introducción

STG es el sistema que tiene la función de leer los contadores eléctricos y detectar el mal funcionamiento y los errores en los dispositivos que conforman la red. Cuando algo falla, se genera una incidencia y las actuaciones y acciones que se deben realizar sobre ella es un proceso que Iberdrola quiere optimizar. Por ello, la introducción de una tecnología RPA que ayude a automatizar el proceso se plantea como una gran solución para aumentar la eficiencia de todo el sistema.

2. Definición del trabajo

La implantación del bot del proveedor de RPA debe realizarse tras establecer unos criterios precisos, por lo que se deben designar unos objetivos y una metodología para alcanzarlos. En primer lugar, se debe realizar un proceso de filtrado de datos para poder diseñar el robot adecuadamente, separar la información en grupos facilitará el estudio de los casos que sean realmente interesantes para el proyecto. También es necesaria una automatización de las tareas repetitivas para reducir el tiempo no valioso de los trabajadores, los cuales podrían estar haciendo cualquier otra cosa que produjera más beneficios para la empresa. Esta automatización estaría relacionada con la optimización de todo el sistema, que es el objetivo principal del presente informe.

La minimización de los errores y de las pérdidas de tiempo tiene una clara correlación con el aumento de la eficiencia, lo que a la larga se traduce en más beneficios para la empresa. Por ello, la automatización del sistema es esencial para alcanzar los objetivos propuestos, ya que proporciona herramientas que no podrían utilizarse con el modelo tradicional. Otro de los objetivos preestablecidos del proyecto es la preparación del sistema en caso de adopción de nuevas tecnologías y técnicas en el futuro, introduciendo los cambios necesarios.

El primer paso de la metodología propuesta consiste en recopilar toda la información posible sobre el sistema para conocer la infraestructura de la red y su funcionamiento. Tras esta investigación, se realizará un filtrado de datos mediante clustering y otras herramientas de análisis de datos y, una vez filtrados, se designarán los casos más relevantes a estudiar. Dichos casos deben ser analizados en profundidad para que la combinación de las tecnologías de Inteligencia Artificial y RPA pueda realizarse satisfactoriamente.

3. Descripción del modelo

A la hora de desarrollar el modelo, debe existir un camino claro a seguir, para que los pasos sean claros y fáciles de seguir hasta la obtención del resultado final. La primera etapa del proceso es el análisis del sistema, donde se organiza adecuadamente toda la información del mismo para su estudio. La web STG es el objeto de estudio de este informe, pero una mirada sobre el sistema en el que se incluye es imprescindible para entender su funcionamiento. Este sistema se llama TitaniumSTG y se encarga de conectar STGweb y otros sistemas a las bases de datos para que haya un intercambio fluido de información entre los datos generales de la empresa y los demás sistemas.

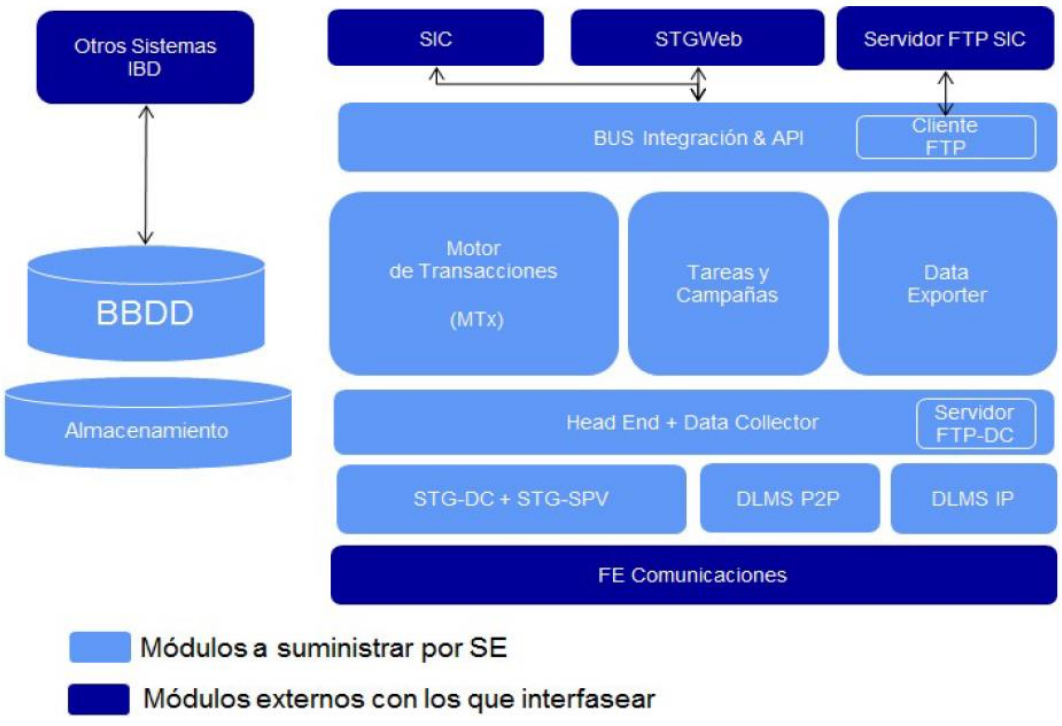

Ilustración 1. Visión general del sistema TitaniumSTG (Navarro, 2019)

Dentro de TitaniumSTG, STG web se centra en la red eléctrica, concretamente en los elementos de bajo nivel que la conforman y en su comunicación con el sistema y entre ellos. El proceso de generación, detección y actuación sobre las incidencias en el sistema es el que se pretende optimizar con el desarrollo de este proyecto. Las incidencias se pueden dividir en manuales o automáticas dependiendo de cómo se generen, si un agente observa un problema y quiere que esté en el sistema aparecerá una incidencia manual tras su inclusión o si el propio dispositivo detecta una avería en el sistema se genera sin la intervención de ningún trabajador. En este caso, sólo serán objeto de estudio las incidencias automáticas porque son las adecuadas para ser automatizadas por el robot una vez completado su diseño.

Durante el proceso de una incidencia en el sistema desde que se genera hasta que se concluye, son varios los pasos que puede seguir dependiendo de cómo cambie su estado y la actuación promovida sobre ellas.

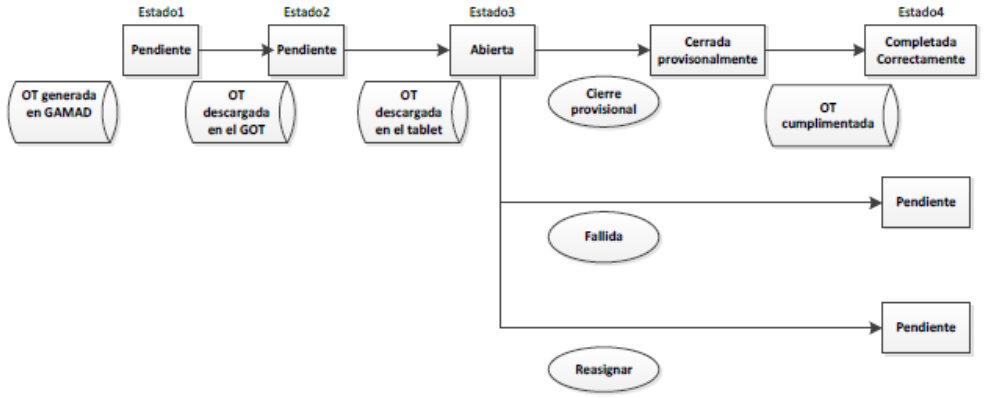

Ilustración 2. Diagrama de flujo de los cambios de estado en la web STG (Iberdrola, 2020)

El diseño del modelo se realiza atendiendo a la información recogida de todo el sistema y se llevará a cabo haciendo uso de técnicas de data analytics, por su capacidad para trabajar con grandes cargas de datos. En este caso, la metodología seleccionada es el clustering, que permitirá una división en grupos de los datos según sus características y el interés de los objetivos del proyecto.

Los diferentes casos de estudio que se realizarán son los diagnósticos más comunes, con el objetivo de abordar el tipo de incidencias más graves y repetitivas del sistema; las incidencias con mayor número de actuaciones, por ser las que necesitan un mayor recorrido antes de llegar a completarse y resolverse; la distribución de incidencias por años, para ver una imagen completa de la evolución del número de incidencias desde la creación de la web de STG; la distribución de incidencias por meses, para apreciar las posibles diferencias en la proporción de incidencias por épocas del año; y la distribución por zonas, donde podría haber un patrón que muestre que se generan más incidencias en algunos territorios y la distribución de las mismas dependiendo de si es verano o el resto del año.

Por último, la implementación de la metodología de clustering se realizará con Matlab 2020 e incluirá todos los casos de estudio propuestos.

4. Resultados

La metodología ejecutada para la implementación del modelo proporciona información realmente útil que puede ser sometida a un profundo análisis para que la futura implementación del bot en el sistema pueda realizarse con éxito.

Tras la revisión de los diagnósticos más comunes se obtienen algunas similitudes entre la mayoría de ellos, que suelen estar relacionados con las malas comunicaciones. En cuanto a las incidencias con mayor número de actuaciones, se repite el patrón a través de las mismas que apunta claramente a malentendidos y disputas entre los agentes implicados en el sistema. La distribución de incidencias por años ofrece una esclarecedora imagen de la evolución del funcionamiento de la web STG desde su lanzamiento en 2013.

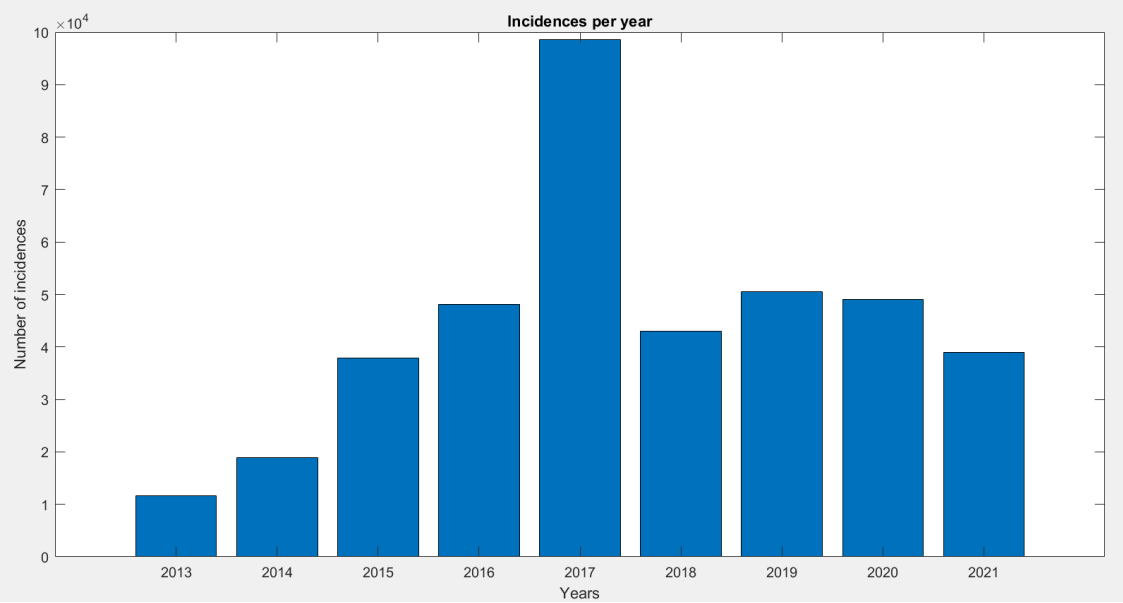

Ilustración 3. Número de incidencias por año

Si se vuelve a analizar este número de incidencias pero esta vez centrándose en los meses, se puede obtener otra perspectiva de los datos. Sin embargo, esto no es del todo interesante en el caso de la introducción de una tecnología RPA, ya que no aporta nada interesante al propósito del proyecto. Por último, la distribución por zonas muestra las consecuencias de instalar dispositivos de la red en zonas de alta densidad, que son las que más incidencias tienen por instalación debido a la alta concentración de elementos.

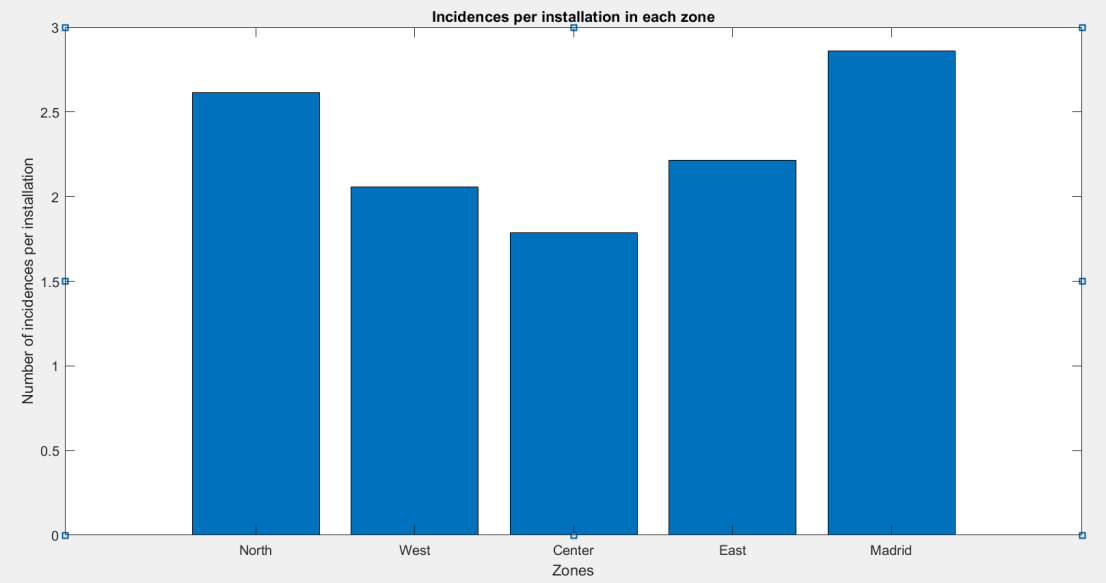

Ilustración 4. Número de incidencias por instalación en cada zona

5. Conclusiones

Tras la recopilación de resultados y su análisis, se pueden establecer algunas conclusiones sobre la influencia del presente informe en la posterior implantación de la tecnología RPA en la web STG. La optimización de todo el proceso mediante herramientas de automatización debe centrarse en los casos de estudio que resultaron interesantes tras el análisis profundo de los mismos. Las comunicaciones erróneas y la baja capacidad de envío y recepción de señales de los elementos de la red es uno de los principales problemas a resolver, e incluso su generación de incidencias debe ser modificada para que el sistema funcione mejor. Asimismo, las disputas entre agentes y organismos y el establecimiento de un criterio claro son esenciales para disminuir la pérdida de tiempo en ese tipo de incidencias. La inclusión del bot implicará que muchos de estos casos se resuelvan con mayor facilidad, ya que el robot sólo actúa según los criterios que se le dan cuando es entrenado.

Las zonas urbanas son también otro campo potencial donde las tecnologías RPA pueden mejorar el rendimiento del sistema. La detección de zonas donde las incidencias tienen un comportamiento similar puede ayudar a reagruparlas y tratarlas como iguales para que no haya problemas cuando la densidad de elementos en la red sea demasiado alta. Por último, no se debe romper el equilibrio entre el crecimiento de la tecnología y la inclusión de más elementos en la red para que no se produzca un colapso en el sistema.

6. Referencias

Iberdrola. (2020). *Procedimiento de gestión de incidencias de telegestión, telecontrol y telecomunicaciones.*

Navarro, J. G. (2019). *TitaniumSTG.* Schneider.

IMPLEMENTATION OF AN RPA TECHNOLOGY MAKING USE OF ARTIFICIAL INTELLIGENCE WITHIN THE STG WEB FLOW CHART

Author: Burgos Rastoll, Ángel.
Company Supervisor: Patiño Mochales, David.
Academic supervisor: Matanza Domingo, Javier.
Collaborating Entity: Iberdrola

ABSTRACT

This report aims to capture a full picture of STG web in Iberdrola in order to provide enough information to the RPA provider so that a bot that will automatize the whole system can have enough input and output arguments to perform correctly.

Keywords: RPA, automatization, STG web

1. Introduction

STG is the system that has the function of reading the electrical meters and detecting malfunction and errors in the devices that conform the grid. When something goes wrong, an incidence is generated and the actuations and actions that must be performed on it is a process that Iberdrola wants to optimize. For this reason, the introduction of an RPA technology that helps automatize the process is thought to be a great solution to increase the efficiency of the whole system.

2. Definition of the work

The implementation of the bot from the RPA provider must be done after establishing accurate criteria, so some objectives and the methodology to reach them must be appointed. First of all, a process of data filtering must be carried out so the robot can be designed properly, separating information into groups will make it easier to study the actual interesting cases for the project. An automatization of repetitive tasks is also necessary to reduce non-valuable time spent by workers who could be doing anything else causing more benefits for the company. This automatization would be related to the optimization of the whole system, which is the main goal of the present report.

The minimization of errors and waste of time has a clear correlation with the increase on the efficiency, which in the last term results in more benefits for the company. That is the reason why an automatization of the system is essential to achieve the proposed objectives, because it provides tools that could not be used with the traditional model. Another established goal of the project is the preparation of the system in case of the adoption of new technologies and techniques in the future, introducing the necessary changes.

The first step of the proposed methodology consists of collecting all the possible information about the system to understand the infrastructure of the grid and its functioning. After that research, a data filtering through clustering and other data analytics tools will be performed and, when filtered, it will appoint the most relevant cases to be studied. Those cases must be deeply analyzed so a combination of Artificial Intelligence and RPA technologies can be achieved satisfactorily.

3. Description of the model

At the time of developing the model, there must be a clear path to follow, so the steps are clear and are easy to follow until the obtention of the final result. The first stage of the process is the analysis of the system, where all the intel from it is properly organized for its study. STG web is the object of study of this report but a look on the system where it is included is essential to understand its functioning. This system is called TitaniumSTG, that is in charge of connecting STGweb and other systems to the database so there is a fluent exchange of information between the general data of the company and the other systems.

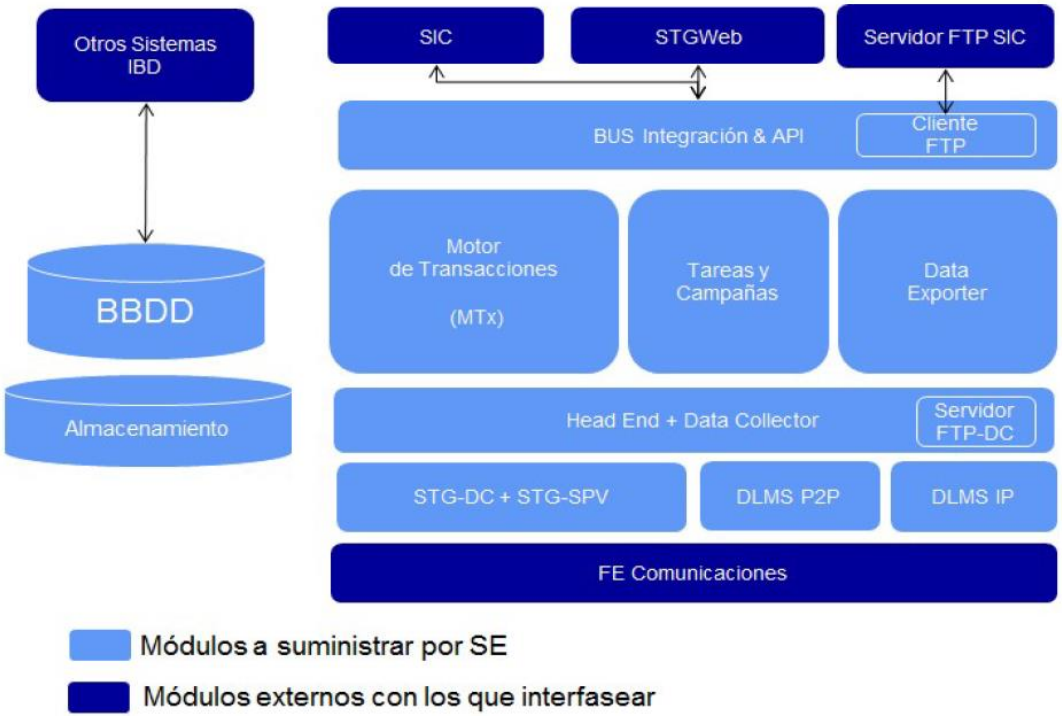

Illustration 5. General vision of TitaniumSTG (Navarro, 2019)

Inside TitaniumSTG, STG web focuses on the electrical grid, specifically the low-level elements that conform it and their communication to the system and between them. The process of generation, detection and actuations over the incidences in the system is the one that is aimed to optimize with the development of this project. Incidences can be divided into manual or automatic depending on how they get generated, if an agent

observes a problem and wants it to be in the system a manual incidence will appear after its inclusion or if the device itself detects a malfunction in the system it is generated without the intervention of any worker. In this case, only automatic incidences will be object of study because they are the right ones to be automatized by the robot once its design is completed.

During the process of an incidence in the system since they are generated and they are concluded, there are several steps that it can follow depending on how their status changes and the actuation promoted on them.

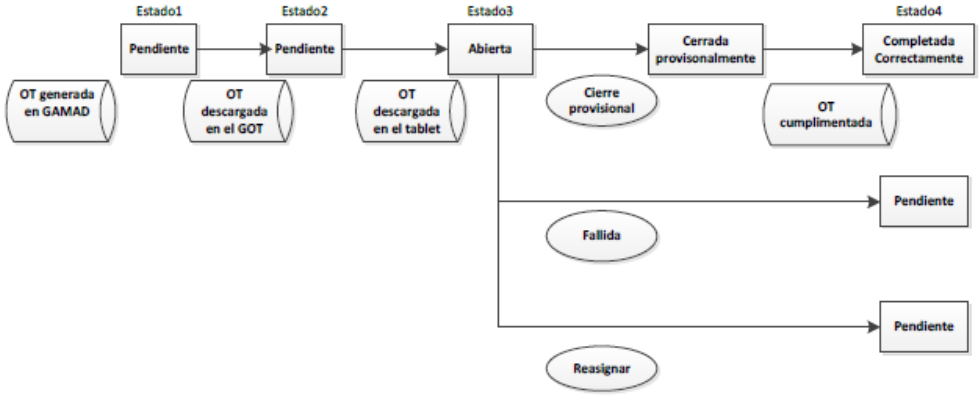

Illustration 6. Flow chart of changes of status in STG web (Iberdrola, 2020)

The design of the model is performed attending to the collected information of the whole system and will be carried out making use of data analytics techniques, because of their capacity to work with huge loads of data. In this case, clustering is the selected methodology and it will allow a division in groups of the data according to their characteristics and the interest of the objectives of the project.

The different cases of study that will be performed are the most common diagnoses, aiming to tackle the worst and most repetitive type of incidences in the system; the incidences with a higher number of actuations, because they are those that need a longer path before they get completed and solved; the distribution of incidences per year, to see a full picture of the evolution of the number of incidences since the creation of STG web; the distribution of incidences per month, in order to appreciate possible differences in the share of incidences per time of the year; and distribution per zones, where there might be a pattern that shows that more incidences are generated in some territories and the distribution of them depending if it is summer or the rest of the year.

At last, the implementation of the clustering methodology will be carried out with Matlab 2020 and will include all the proposed cases of study.

4. Results

The executed methodology for the implementation of the model provides really useful information that can be subjected to a deep analysis so that the future implementation of the bot in the system can be successfully done.

After reviewing the most common diagnoses, it is obtained some similarities between most of them which are usually related to bad communications. As for the incidences with more actuations, it occurs again the pattern through them that clearly appoints to misunderstandings and disputes between the agents involved in the system. The distribution of incidences per year provides a nice picture of the evolution of the performance of STG web since its launching in 2013.

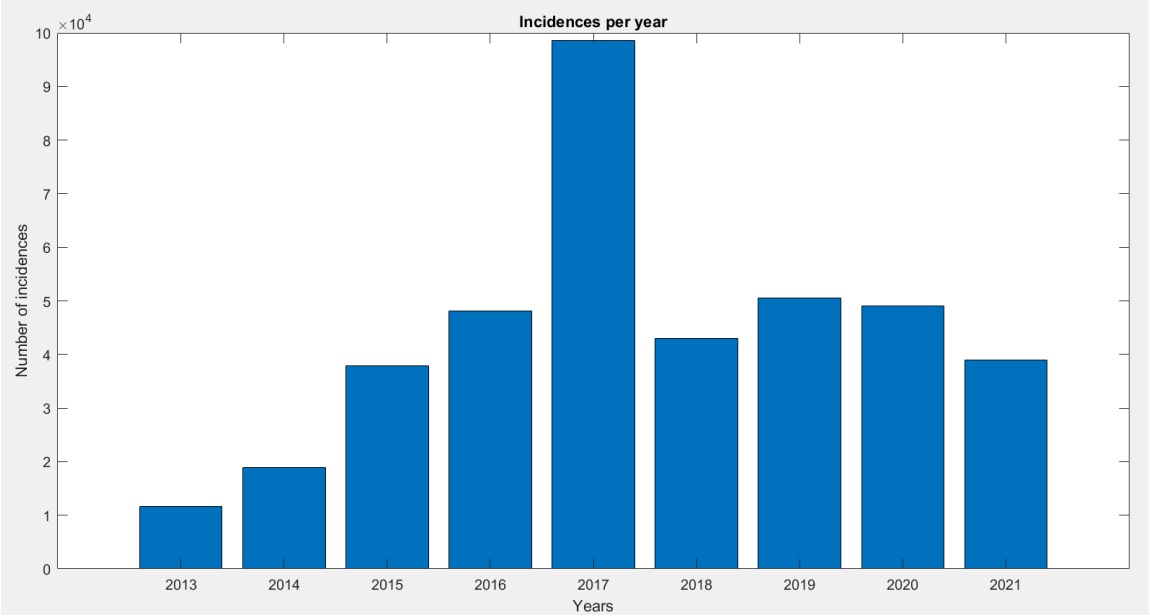

Illustration 7. Number of incidences per year

If this number of incidences is again analyzed but this time focusing on months, another perspective of the data can be obtained. However, this is not quite interesting in the case of introducing an RPA technology, because it does not add anything interesting to the purpose of the project. At last, the distribution per zones shows the consequences of installing devices of the grid in high density areas, that are the ones with more incidences per installation due to the high concentration of elements.

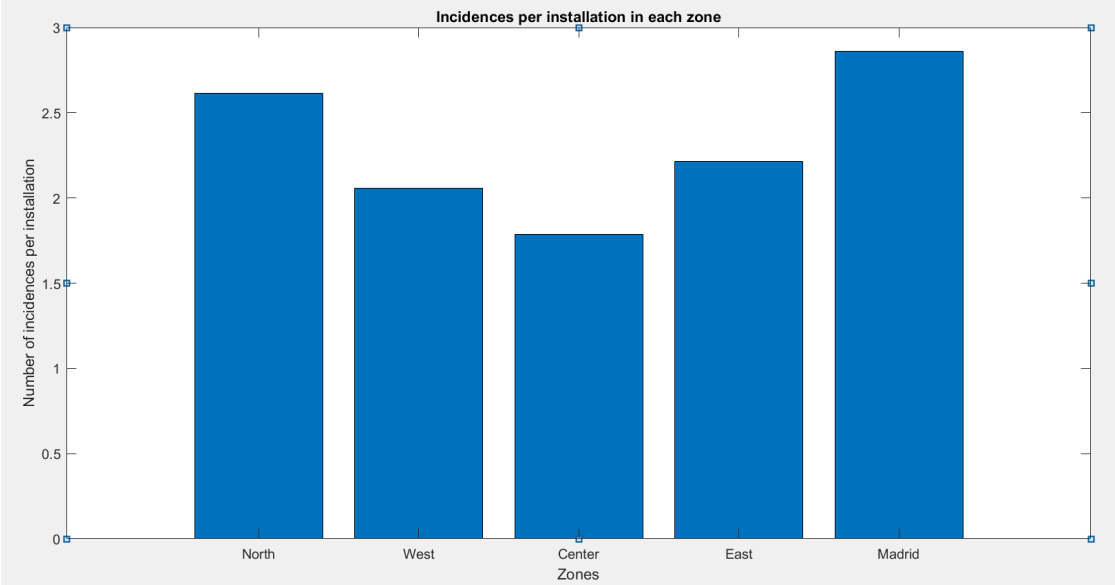

Illustration 8. Number of incidences per installation in each zone

5. Conclusions

After the collection of results and their analysis, some conclusions can be stated regarding the influence of the present report in the later introduction of the RPA technology in STG web. The optimization of the whole process using automatization tools must be focused on the cases of study that were interesting after the deep analysis of them. Wrong communications and low capacity to send and receive signals by the elements of the grid is one of the main issues to be solved, and even their generation of incidences should be modified so system performs better. Also, disputes between agents and organisms and the establishment of a clear criterion are essential to decrease the waste of time in that kind of incidence. The inclusion of the bot will imply that lots of these cases are more easily solved since the robot only performs according to the criteria that it is given when trained.

Urban areas are also another potential field where RPA technologies can improve the performance of the system. Detecting areas where incidences have a similar behavior can help regroup them and treat them as equal so there are no problems when the density of elements in the grid is too high. At last, a balance between the growth of technology and the inclusion of more elements in the grid must not be broken so there is not a collapse in the system.

6. References

Iberdrola. (2020). *Procedimiento de gestión de incidencias de telegestión, telecontrol y telecomunicaciones.*

Navarro, J. G. (2019). *TitaniumSTG.* Schneider.

Index of the report

Index of figures

CHAPTER 1. INTRODUCTION

Iberdrola is the leader of the energetic sector in Spain and one of the top companies in the world in this market. It has been almost a decade since they implemented STG (Sistema de Telegestión) to control the incidences that take place through the grid. STG is the remote management system that is in charge of detecting manual and automatic incidences in the grid and manage their control and supervision.

First of all, incidents are generated in the system due to the malfunction of an element and then a diagnosis is assigned to it. With this diagnosis, a series of actuations and actions are promoted to solve the incident and, if everything has worked well, its status is changed to solved and it will be closed so that it does not appear in the system as something to work with. However, this process of actions can be improved by optimizing and automating the process in some points where human intervention does not contribute. This is a point of improvement and economic profitability for the company.

For these reasons, Iberdrola is trying to adopt a new methodology which enables them to increase their efficiency and to be more competitive in their market. This methodology must be achieved by implementing some type of technology that helps moving the company towards that direction. After a long process of research, some ideas were proposed and, at the end, one of them was chosen as the most suitable for the system. This new solution would be the one that should take over the process to achieve an optimization of the whole system. Therefore, once the decision was made, it was a matter of time that one of these technologies was selected and the process of renovation started out.

CHAPTER 2. DESCRIPTION OF TECHNOLOGIES

As technology continue to grow, and the benefits they provide to companies become more tangible, automation technologies providers intend to offer a wider spectrum of products to their clients so each of them can adjust better to the certain needs of every customer. In this chapter, it will be analyzed the different solutions that the market is offering right now and the cases of use that are already implemented in Iberdrola.

2.1 TECHNOLOGICAL SOLUTIONS

2.1.1 RPA TECHNOLOGIES

Nowadays, digital processes are flooded with tons of data that multiply themselves over the years. RPA (Robotic Process Automation) refers to a bot or software that executes itself virtually. It aims to automatize repetitive tasks that do not require too much complexity to solve, like collecting and transfer of data, file management, scanning or gathering information (Urreta, 2019). It is a great tool for big scale processes where there is a huge volume of easy and standard tasks that can be optimized with this technology.

Back-office departments are some of the most beneficiated by RPA since human resources, finances, accounting or client service among others need to deal with tons of information that can be standardized and managed as a robotic process taking away load of work from the worker. This could be seen as a threat by employees, but it is actually a tool that they can take advantage from so they can focus on other tasks that really need the supervision of a human. This way, a better use of time and optimization is achieved only by removing simple and repetitive tasks from the requirements of workers.

This technology is being already applied worldwide and is expanding its range of market. Its software is not invasive, performs fast and does not need much supervision once it its installed and working. It can be easily integrated in IT environments and complements the

rest of processes involved in the system of the company enhancing a great collaboration between them. Its configuration is designed to be treated transversally so it does not stick to certain kind of departments, and it can be adapted to methods that were not originally supposed to be used.

The robot is designed to work in the user interface (UI) as if it was a human worker, so it would move the mouse or use the keyboard like any other user would (Deloitte, 2020). Before it gets implemented in the system it must be trained so that it gets used to the configuration and environment where it will perform. This training is considered as an education of the robot that is fundamental in structured systems with specified rules that it must follow and must adapt its methodologies to them. It is important to state that these bots are not trained to learn from experience and to act wisely when they find an exception in the system, then it is the human expert the one who must solve the issue reducing his area of influence to really needed tasks.

2.1.1.1 Benefits of RPAs

RPA technologies provide a big number of benefits to the companies that adopt them and usually leave a great footprint wherever they get included. Probably, the most characteristic feature of them would be the decrease of staffing costs and human errors (Boulton, 2018). The automation of a system always implies the suppression of human failures and mistakes, which can be provoked by fatigue or distractions that will never take place in a robotic process. The redesign of the workload and the rescheduling of the system lead to an increase of the capacity of the whole company, thanks to them it can be achieved a better efficiency in the time that workers spend doing their tasks.

Another advantage is the low cost it takes to implement the bots. It does not take too much money or time to implement this technology in the whole system what is fundamental if a not too big company wants to adapt to it. Also, the implementation is relatively easy to make so every enterprise who shows interest in embracing this new methodology can be able to do it without having to invest too much effort in it.

The computed working hours have no competence with other technologies, automation in this case means that bots can work without interruptions 24 hours a day, 7 days a week. Considering these measurements, it is not hard to notice the increase of efficiency and reduce of resting time that it implies. Furthermore, amortization of RPAs will not take long if the system is well designed and the configuration of the process and its methodology are introduced correctly in the company (Urreta, 2019).

2.1.1.2 Drawbacks of RPAs

The benefits of RPAs in companies also have a counterpart in some of them that need to be studied in case of adopting this technology. The main drawback is related to human resources and the readjustment that there is to do in order to adapt everything right. The inclusion of RPAs, and any kind of automation methodology, involves the reduction of jobs, which can be seen as good for the company, but it is obviously a big obstacle for workers, who then need to find the way to be irreplaceable in their environment. Bots can do easy and repetitive tasks in much less time than humans so it is this area where people will not have a very bright future in their enterprise, instead, it can also be seen as an opportunity to get a training about how to manage these robots and how to maintain and fix them. In conclusion, lots of job positions will be replaced by bots but, at the same time, other tasks and other needs will be created so a balance between them must be achieved by the company.

This transition between traditional and automatized work is not always easy, sometimes technology or resources are insufficient and changes fall bad in the performing of the system. RPAs can be considered as a brand-new technology that is not totally developed so it is not perfect at this time, that means that errors are common and a good efficiency is not always achieved. According to a Mckinsey report of 2017: “Several robotics programs have been put on hold, or CIOs have flatly refused to install new bots” (Edlich & Sohoni, 2017), what was stated by Alex Edlich and Vik Sohoni, senior partners of the consultant enterprise. This is a real problem for companies because they must take into account the work and training that must be done with bots so failures in the performance bots decreases to a point where it can be profitable to invest in it.

However, there is one more inconvenient, and it is the scalability of the bots. This is one of the most commented issues around RPA technologies by customers who find it too difficult to install a higher number of bots in their system. Platforms have not achieved yet a proper structure where bots can complement each other and enrich the whole environment. On the contrary, it is commonly known among clients that when it comes to a gathering of more bots and to add them to other tasks, failures spread all over the system and it gets impossible to control them. Another report about this was published by Deloitte and authors claimed that: "Only three percent of organizations have managed to scale RPA to a level of 50 or more robots," (Deloitte, 2017).

2.1.2 INTELLIGENT AUTOMATION

Intelligent automation is a concept that covers a set of automation tools based on RPA technologies combined with Machine Learning (ML) and Artificial Intelligence (AI) (Penteo, 2021). It is defined as any kind of tool that can improve their own processes and results, optimizing IT resources and efficiency through data analytics.

Intelligent automation could be considered as a combination of different technologies that conform a higher-level concept with a bigger scope. In the next figure, it can be appreciated the structured data processes, based on rules, combined with non-structure data processes, based on judgements:

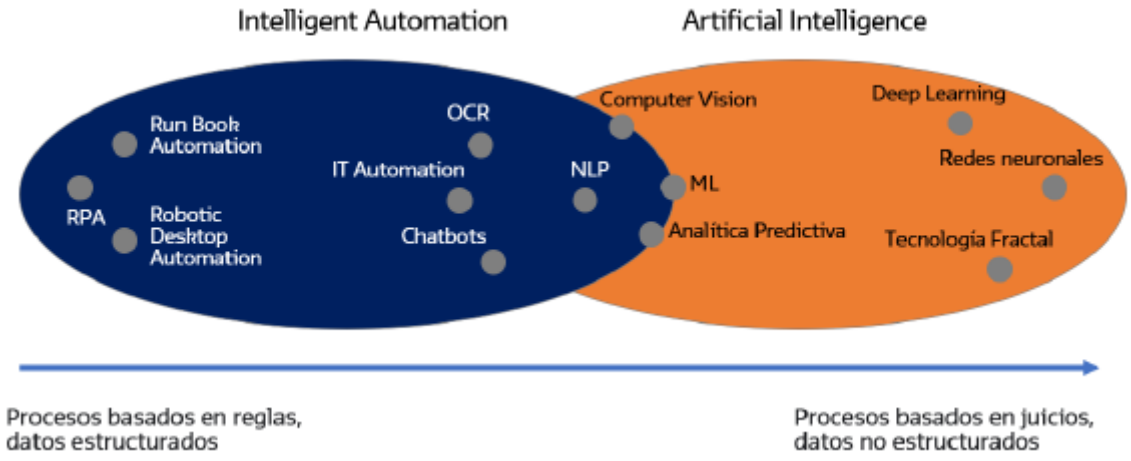

Figure 1. Intelligent automation as a combination of AI and RPA technologies (Penteo, 2021)

2.1.2.1 Key tendencies in Intelligent Automation

In the last years, some tendencies have come up around this concept and are getting a good welcome by users, being the most remarkable ones:

- **Hyperautomatization and based on platform**

The orientation of RPA industry is moving towards hyperautomatization, which consists of covering every last piece of the system and automatize everything making use of Machine Learning (ML) and Natural Language Processing (NLP). The goal is to achieve the full cycle automatization so RPA providers are investing time and resources in evolving through this path. Thanks to this, it can all be structured regrouping workflows in the process with the discovery of them and adding collecting and predicting models for training of the bots. They also aim to cover data analytics, document processing and file management or unified dashboards.

- **Self-service and Low Code Platforms**

Low Code allows easier and simpler implementations of the bots in the different platforms, so this is a trend right now among companies who want to make their processes in a way that does not take too long to understand and use. This is done through simple parametrizations that enable the user to work with this technology in any situation and its

development is growing exponentially because of the emergence of pre-design bots, connectors, scripts or IT skills that continue to impulse this sector. Predefined workflows and emerging self-service capacities like self-reparation, self-scalability, self-learning or self-installation will encourage a better and faster implementation of these bots in the platforms.

- **Scalability, Flexibility and Prices**

Since the beginning of RPA commercialization, it has been usual the flexibility in prices like prices per process or bot and costs based on obtained return. Nowadays, the orientation is directed to reduce the Total Cost of Ownership (TCO) and entry and exit barriers, this is the reason why automation ecosystems are growing because of the inclusion of hyperscallers and software providers who collaborate to put together all of their solutions and achieving like this a great scalability. Also, models like SaaS, Bring Your Own License and automatization packages are contributing to the expansion of these methods.

- **Consolidation as key tools for the continuity of business**

Business stability is usually very volatile and suffers from changes all the time. With these bots, providers try to help companies cope with fluctuations of demand and interactions with clients thanks to specialized solutions, free capacitation programs and "hackatons" to identify new application areas. During COVID 19 pandemic, there has been the need of working from home, online, and to solve mistakes remotely. This has been possible because of the use of assisted automatization that consists of a environment where Help Service Desk agents can work online and help guarantee the productivity and well-functioning of the whole system.

2.1.2.2 Impact of Intelligent Automatization

The great scope of Intelligent Automatization allows to differentiate into the techniques and tools that conform it so we can observe how each one of them contributes to the generation of competitivity in companies. In the next figure provided by the PENTEO report (Penteo,

2021), we will be able to confirm the areas where the impact of IA is more noticeable and the ones that do not seem to be affected by changes in this field.

Potencial para ... (*)	Experiencia Inmersiva	Blockchain	IoT	Ciberseguridad	Chatbots	AI (ML/DL)	Automatización inteligente (RPA)	Analítica Avanzada
Incrementar ingresos	8%	13%	21%	15%	45%	38%	40%	96%
Reducir los costes	26%	20%	60%	25%	50%	43%	45%	43%
Crear nuevos modelos de negocio	26%	48%	19%	5%	48%	48%	30%	52%
Transformar o eficientar el marketing	19%	11%	4%	6%	60%	57%	55%	75%
Transformar o eficientar las ventas	17%	16%	9%	10%	55%	45%	50%	73%
Transformar o eficientar las finanzas	0%	23%	2%	2%	34%	34%	60%	81%
Transformar o eficientar la gestión del talento (RRHH)	24%	9%	7%	9%	50%	43%	35%	73%
Transformar o eficientar la producción o logística	41%	16%	80%	22%	48%	53%	35%	47%

Figure 2. Percentage of success of AI tools in certain tasks (Penteo, 2021)

Rows show:

- Increasing incomes
- Reducing costs
- Creating new models of business
- Transforming marketing
- Transforming sales
- Transforming financing
- Transforming talent management (Human Resources)
- Transforming production or logistics

Columns show:

- Immersive experience
- Blockchain

- IoT
- Cybersecurity
- Chatbots
- AI (ML/DL)
- Intelligent Automation (RPA)
- Advance data analytics

Besides, after a comparation between companies that were successful in their digital transformation and the rest, the results show that successful companies had a bigger use of digital platforms:

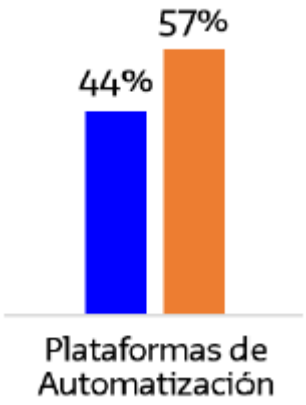

Figure 3. Use of automatization platforms in succesful vs general groups (Penteo, 2021)

2.2 INDUSTRY USE CASES FOR RPA

RPA technologies are already adopted in a lot of different sectors of activity, being more useful in big industries where competition is high and management of data is key. These are some of the most remarkable use cases in existing industries:

2.2.1 RETAIL

Retail industry is a fantastic fit for RPA technologies, the market is constantly changing and orders and shipping information floods the system. Status of orders or client directions, along with prices of products conform an ideal environment to introduce these bots that can make the whole process much easier and fast. The great boom of e-commerce (online market) has provoked that many retail companies have been forced to be reorganized and to adapt their business to face (Amatech Group, 2020):

- Increase of production costs
- Low reliability in the deliveries
- Slow market growth
- New technologies, like payment platforms
- Need to focus on the client's culture

All of them are included in the offer of RPA providers in a bigger or smaller scale so that is why this technology is having a strong impact in the sector.

2.2.2 BANKING

Banking processes are widely known for being especially slow and long, but they also involve transactions of lots of data with a huge number of inputs and outputs. For this reason, RPAs has found a perfect market to be spread in many different aspects. One of the most criticized departments in the banking sector is customer support, however, these bots can be used to design a system where appointments are established following an automation protocol that will be much more efficient than what it used to be. There are other interesting cases of use like incorporating Machine Learning to learn what kind of product are more suitable for clients or to cut the time that it usually takes to finish a loan application.

2.2.3 INSURANCE

In the insurance sector, it happens something that is similar to banking, processes are really long and include multiple stages in every one of them, what again suits RPAs perfectly.

There are lots of operations in the industry that are likely to be automatized like renewals, processing of claims or the reports including information of clients. Automating these processes can lead to a decrease of time and an increase of efficiency.

2.2.4 COUNCILS AND LOCAL AUTHORITIES

Council and local authorities have the duty to control and manage a lot of different tasks and the collaboration between them and their configuration is not always easy. With RPA technologies, they can be organized reducing the difficulty of managing that quantity of data and information.

2.2.5 HUMAN RESOURCES

Selection processes are another field where automation has found expansion, profile filtering and selection make this industry a prime use case for RPA (Foundry, 2020). In this case, it is more noticeable its influence over higher scale companies, since they have a bigger human resources department and deal with more employees and more job positions. In this area, it is not only useful to determine the perfect candidate in recruitment processes but, it is also a great tool for calculation job positions that are no longer needed because of the time and effort wasted there with poor results.

Nevertheless, there are a lot of options for this technology, its potential has no limits, so other sectors may be included in this top in the following years. The expansion of robotics and the development of technology create a new scenario where almost every aspect of any business can be improved, and its efficiency maximized so it is not hard to imagine the impact that RPAs might have in the future. Besides, big companies are investing lots of money in it so there is no doubt that everyone will have to adapt to this new actor and its benefits, so a fast adoption of its methodologies is truly advisable.

2.3 CASES OF USE IN IBERDROLA

Iberdrola decided on adapting RPA technologies for an automatization of STG web after studying the good results that they had obtained from these technologies in other departments where they had already been implemented. These are some cases where RPA technologies were adopted in the company.

2.3.1 GEM

Global Energy Management was one of the first departments in Iberdrola that considered an evolution through the automatization of their processes. Their system has some similarities with STG web and the use of a bot to optimize the different steps to follow has led to an increase in their profits. The selected RPA provider in their case was Uipath, which had to adapt their methodologies to the methodologies in BluePrism, since this is the one that was originally implemented in Iberdrola. The implemented automatization was performed in order to enter international websites and compare prices with multiple organisms. Reports from the project show the good results that this RPA technology has obtained in this area of the company and, besides, workers from that department give a positive feedback from their job and the tools that they have there.

2.3.2 SIROCO

Siroco Inteligente is an application that provides documentation to the companies that work for Iberdrola and receives documents from them. Once the file is submitted to Siroco it must be approved or rejected according to some criteria. UHIS is the name of the service of automatization, that now has changed to Dolffia, which does not only stick to an RPA technology, but it also combines Artificial Intelligence techniques with Machine Learning processes. The bot from Dolffia is trained to automatize every document that gets to Siroco and to validate it only if it fits the predefined requirements, being very important the quality of the files. If this quality is not good enough, it does not matter the content of the document since the program will not be able to read and will discard it. Right now, there are five processes that are automatized with this application:

- Personal Identification Document (DNI, NIF)
- Up-to-date payments with the Tax Agency
- Up-to-date payments with Social Security
- Technical Inspection of Vehicles (ITV)
- Medical history

CHAPTER 3. STATE OF THE ART

In this chapter, it will be analyzed the market of RPA providers that are currently top leaders in their sector.

3.1 UIPATH

Uipath is an RPA tool that is mainly used in Windows and aims to automatize repetitive tasks so human intervention can be minimized. It tries to add value to all activity sectors since it eliminates meaningless handwork and greatly increases the productivity of the user.

Figure 4. UiPath (UiPath, 2021)

This RPA tool is currently established as one of the leaders of this sector due to its huge investment on the expansion of the range of products that it offers to the user. Innovation is one the main pillars of the philosophy of Uipath, they continuously try to find the most updated solutions for the clients so they can achieve a digital transformation that suits all the required needs.

They offer three key factors that make them one of the most powerful tools in the market:

- The first one is the high capacity that it has to automatize complex processes keeping the risks at a low level. It enables the choice between having their robots under surveillance of an operator or on their own.
- The second one is the agility, the ability to develop new software with high velocity. Also, this agility is a key factor to quickly adapt to new environments or priorities.
- The last one is the scalability. It shows great capability to be modular and extensible to any circumstance.

Automation Cloud is one of the main features of Uipath. It oversees a hybrid cloud of both public cloud and on-premises infrastructure; with Automation Hub, it launched a platform for automation centers of excellence. The open platform integrates easily with enterprise applications like Google Workplace, Microsoft Office, Oracle, Salesforce, ServiceNow, and Workday; UiPath builds and supports native integrations along with technology partners. UiPath offers an enterprise-grade and innovative RPA solution augmented by a large ecosystem of partners, making it a good fit for large, global enterprises with demanding needs for support and governance.

Task mining isn't yet part of Automation Hub, but process mining features, collaborative discovery, assessment tools, and graphical visualizations provide all you need to analyze and track automation ROI and its realization. StudioX requires no programming skills, while Studio targets advanced developers. Embedded AI functionalities classify and extract data from unstructured, semistructured, and structured documents and reliably scan at high speed. Bot-triggering capabilities are limited, with workarounds. Security, access control, and authentication features are state-of-the-art. The open platform integrates easily with enterprise applications like Google Workplace, Microsoft Office, Oracle, Salesforce, ServiceNow, and Workday.

3.2 BLUE PRISM

Blue Prism offers an RPA with high security and great compliance that has been the head of the RPA industry for the last decade. It is focusing right now on the artificial intelligent field

in order to offer a wider spectrum of services to enterprises. Their investments on R&D have grown 100% in 2020 and is expected to continue through this path in the following years. Blue Prism covers a pretty wide ecosystem that include several partners in a lot of countries around the world, which makes them a global company with a standard license.

Figure 5. Blueprism (Paradavisual, 2021)

Automation Lifecycle Management is a set of tools that provides the user an assistance from the first stage, consisting of automation ideation, to the final deployment of the project. The bot-building process is quite intuitive, and it offers more than 1800 prebuilt automations which can be downloaded from its website anytime. The bots are designed with a high capability to be easily automated reducing the complexity of its coding.

Blue prism has several functionalities and techniques like OCR, optical character recognition, that are very useful at the time of classifying data from documents. This data is required to be structured so it can be recognized by OCR so it must be standardized. In terms of security, there are strong restrictions for accessing the data in order to avoid cyberattacks from unauthorized people. The company has the goal of making its ecosystem safe for the system and users when the automation of the system is or is not attended by them.

3.3 AUTOMATION ANYWHERE

Automation Anywhere intends to encourage a collaboration between the human and the bot. This RPA stalwart completely rebuilt its product to turn it into a cloud-native, web-based,

AI-infused platform, with capabilities from discovery to automation delivered via an integrated and seamless experience. It aims to widen the spectrum of use for RPA by integrating business matters into their scope. Automation Anywhere has created alliances in more than ninety countries with more than two thousand partners, including some technological partnerships with Amazon Web Services (AWS), BMC, Citrix, Google, IBM, Microsoft, MuleSoft, Oracle, Salesforce, and SAP (Schaffrik, 2021).

The universal recorder can be used by any user to generate any kind of bot, therefore, they can develop their own Automation Anywhere bot with a set of tools which may not be too hard to use. This bot mixes machine learning techniques with different tools to manage data, what makes a great combination when there is a big amount of documents to be analyzed in order to extract information from them. Automation Anywhere has developed a program called Automation 360, originally known as A2019, that intends to open their field of use in many aspects. One of them is the intention to take advantage from text analytics or from natural language processing with IQ Bot and make them easy to implement for the partners which form their alliances like Google, Microsoft, IBM and other companies. Another feature to be considered is the capability to define their opinions by the user thanks to Bot Insight, getting them from data operations or audit trails. Automation Anywhere also provides a digital assistant by the name of AARI which helps completing the automation lifecycle and enriches clients experience providing them with tools to export their work to other agents. This trade of information and data must have a great network of cybersecurity so the RPA platform can be implemented in complex industries that could suffer from cyberattacks or hacking. On the other side, there can be problems at the time of transmitting bots with great complexity or with past versions. Nevertheless, the platform has been developed in the last years to make it competitive in the current RPA market.

3.4 NICE

NICE is a technological platform which aims to automatize different cases of use for the client. It is specialized in voice recognition applications and now it is introducing in the market a higher speed system that checks the inventory, analyzes the user speech on real

time and finds recommendations related to contact centers. The automation of this center is one of its main goals so there is a lot of effort invested in improving this issue. Another characteristic feature of this RPA technology is text analytics. Nice Automation Finder is in charge of collecting and data from the employees and designing algorithms to automate some processes, so the system works more efficiently. Click to Automate is another tool which enables the generation of automations and is intimately joined to AI ethics, as well as all the tools in NICE. This is a key aspect in the culture of this RPA, they intend to get every client and customer to understand the consequences and the opportunities that this technology offers.

In this moment NICE can count on a big number of customers, around 700 in 2021. Its model is being improved every year and the solutions that they offer include new versions of object controls and screen management by augmenting the range of application connectors or OS. Both voice recognition and text analytics characteristics share the need of having a contact center which gives them the opportunity to classify, process and release the collected information. For this reason, their bots must be properly visualized so they can be monitored and errors can be quickly detected. These bots are usually not supervised by a human so this makes it difficult to test the quality of them. It is one of the main drawbacks of this technology and the most problematic for users, however, it is being improved in the last versions and it is expected to reduce mistakes in the future.

As for highly regulated industries like electric companies' environments, its performance is great and it can be rapidly adopted by any actor to make use of it. It must not envy its competitors in security issues, including user's authentication or their access control to the system. This RPA technology would not only cover contact center automation but also voice and text recognition making use of not supervised bots but also some attended ones, which are under development.

3.5 KRYON

Kryon has a different perspective in the way of dealing with RPA designs, it focuses on the discovery of processes that can decrease the time that it takes to produce something or accelerate business. Unlike some of their competitors, in Kryon they make an approximation to data by applying Full Cycle Automation to the whole process. It consists of implementing automated steps in every stage of the cycle so there is a methodology of continuous improvement and data mining from the beginning. This solution is also adaptable to other platforms like Software AG and Verint so it can be utilized with other purposes.

The Full Automation Cycle is combined with AI to enable more features for data mining and for data training so that algorithms can be designed to give the best solution when there is the need of dealing with tons of information. Business process discovery is the field where Kryon leads the market so they intend to create an environment where users from business companies can get a great experience since they launch their enterprise or only to design new processes once the company is already performing. It is very useful to design Canvas or to introduce methodologies like Agile, which are now very demanded worldwide.

At the time of finding failures, Kryon allows fallback and other procedures so users can avoid severe errors in bot performances. Nevertheless, as some clients state, this performance is required to be low so the errors detection and handling procedures are actually effective, if it is not the case, their resilience decreases more than it could be expected. To sum up, if a company is interested in approaching business from the RPA perspective, Kryon is the platform that can fulfill best all the requirements from the company including the design and implementation of the bot.

3.6 PEGASYSTEMS

Pegasystems is another leader of the sector of RPA technologies. They do not only stick to RPA itself, but they also look for a complete automation adding new features like advances in recording that allow saving complex processes for any user. As skill development and

training is one of Pegasystems' strategic investment pillars, it has initiated countless university collaboration programs and takes pride in its 300,000-strong global user community. They are committed to the good use of their system, so it is included in Pega Infinity a variety of webinars and video tutorials which are destinated to users who are not familiar with RPA technologies.

When approaching automations under human supervision they include X ray vision in their offer to make it easier for customers to create and maintain these bots. X ray vision program, included in the package Infinity, complements RPA features with intelligent automation, data mining and the development of not very high programming code. Thanks to this, it is more intuitive to detect better automation candidates with an established collaboration and priority system. Businesspeople are capable of selecting candidates with a more clear visualization of their skills and it can be used a first step to manage an Agile methodology in the company.

Pegasystems users usually show fidelity to the company and stay for many years with it. This is a fact that proves the durability and good performance of their technology. According to their main customers, priority and collaboration program is truly effective and reduces great amounts of time when talking about picking the right candidate. However, they also comment about the difficulty they find to use every offered tool in the set and to combine them to get a full experience of the system. As a summary, Pegasystems is leader in RPA market and they continue to grow thanks to the confidence they create in their clients.

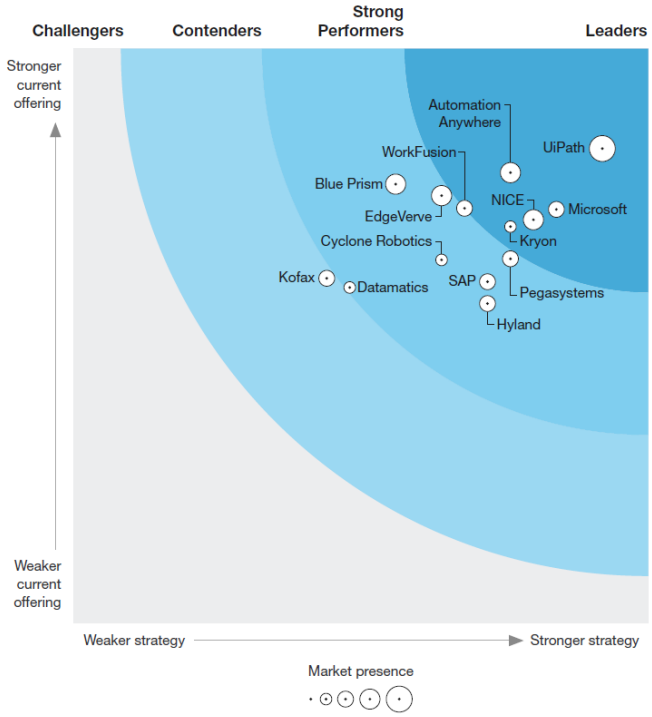

Figure 6. Market share of top RPA providers (Schaffrik, 2021)

CHAPTER 4. DEFINITION OF THE WORK

4.1 JUSTIFICATION

RPA technologies have become a great tool for the adaptation of established systems to the new ages. Nowadays, companies are trying to update their models of work so that they can be competitive against the rest. Iberdrola wants to follow this path and get a better efficiency in their departments. These technologies are already implemented in some departments but, in this report, we will focus on the adaptation of RPAs to STG (Sistema de Telegestión), and the preparation of intel of the system and different cases of use so once the bot is implemented it can be properly trained and work in the direction we need.

The later implementation of the bot requires a previous work that provides enough information to design the automatization of the system properly so an optimization in the whole grid is achieved. This automatization will be performed attending to specific criteria that will be proposed according to the needs and the potential improvement areas in the process. Once the system has been totally reviewed and conclusions have been made, it can be approached by the robot to improve the system and avoid waste of resources that could be used in other segments of the system.

4.2 OBJECTIVES

This project intends to achieve different goals that will be exposed in this chapter.

4.2.1 DIVISION OF DATA IN SPECIFIC GROUPS TO BE STUDIED SEPARATELY

There are tons of data in the STG that make it really hard to understand what is going on in every moment and the reasons why error occur. One of our goals is to determine what is really interesting in the system and which data can give us useful information that can be utilized to design the robot. A division of data in groups makes it much easier to discriminate

important and interesting information from the rest and to regroup those cases with characteristics in common to create a bigger picture of the system and allows the user to focus on the specific data that will help advance the project.

4.2.2 AUTOMATIZATION OF REPETITIVE TASKS

There are multiple tasks in STG that are performed again and again and are not complex at all. These are the ones that RPAs can help manage, because by automating them the system gains efficiency and saves time to focus on more relevant issues. Therefore, the localization of these tasks and the calculation of their time, delays and costs is one of the main objectives of this report.

4.2.3 OPTIMIZATION OF THE WHOLE PROCESS

This is the main objective of this project; it is the reason why this study is being performed and why Iberdrola wants to adapt robotic automation in STG. The system of incidences is far from being perfect and a lot of processes and configuration can be modified so that work is truly effective and there is no waste of time. There is a need to identify those processes and the reason behind their low performance and, if they are correctly selected and analyzed, suitable modifications can be implemented that will make the system more efficient and will increase the degree of optimization of the whole grid.

4.2.4 INTRODUCING IMPROVEMENTS IN THE SYSTEM TO EMBRACE NEW TECHNOLOGIES IN THE FUTURE

The process of automatization does not only cover the short-term goals that we are trying to reach in this report, data will also continue to grow, there will be more items and issues involved in the system in the future, some of them we cannot even think now about. For this reason, it is mandatory to prepare the system for future innovations and technologies that will be able to get included more easily if we already take that chance into account.

4.3 METHODOLOGY

4.3.1 COLLECTION OF INTEL ABOUT THE SYSTEM

Before attempting to filter data and to analyze the steps in the process to be optimized, we need a deep understanding of the system and how the grid and its infrastructure is organized. In order to do so, there must be an exhaustive research on the documentation about STG web and the electrical and telecommunications grid that conform it. The collection of documents and files that explain this system should provide tools and items to be used at the time of designing the robot and will help implementing it by a fluent integration. The following steps in the methodology are based on this research so sources must be official, and the information must be proved and useful.

4.3.2 DATA FILTERING: CLUSTERING

At the time of dealing with so much information, it is unavoidable the use of data analytics in order to filter and to gather different elements into groups that will be organized through some characteristics that will be defined according to our preferences. Clustering is a great tool that gathers data in groups (Unir, 2021) according to some predefined criteria that must be established according to the research of information that was performed in the previous step. When specific criteria are selected, the construction of the analytical model must be developed. This model might suffer from changes depending on the direction that may take the project in certain moments but, at the end, it will be flexible enough to cover all the possible cases that are interesting to analyze for the bot design. After the model is completed, all the results must be noted so the next step can be performed right.

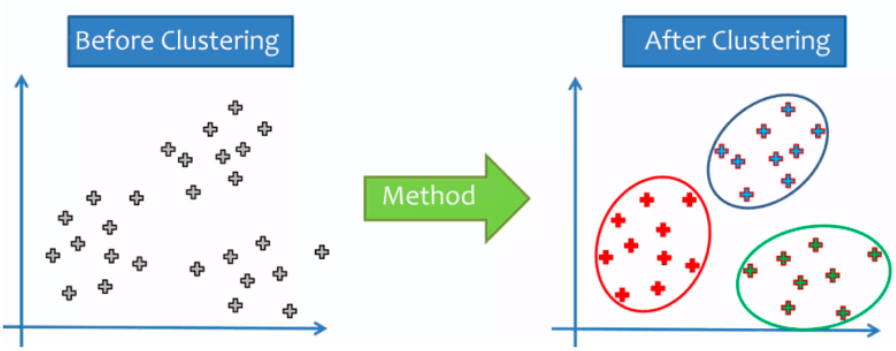

Figure 7. Clustering methodology (Jacobsoft, 2020)

4.3.3 STUDY OF THE MOST RELEVANT CASES

Once the information is divided into groups, there is to decide which of them will be more interesting for our interest. Results from the clustering methodology will offer such a variety of potential cases of study that a analysis of their interest for the project and their difficulty to be executed later becomes truly essential. Some cases could turn out to be not too useful at the end but even that confirmation might be necessary to complete the project. Before choosing the right criteria to pick some cases or others, a conversation with expert on the topic in Iberdrola is always advisable not to make mistakes in such an important stage of the process.

4.3.4 ANALYZE THE POSSIBLE CHANGES TO BE INTRODUCED

With all the collected data, it is time to study how the most relevant cases can be modified so the process is optimized and an increase of the efficiency can be achieved. Analyzing the most relevant cases will answer the questions about their performances and their potentials, if they provide enough information, a scheduling of improvements can be established and modifications that adapt to these cases can be designed properly.

4.3.5 COMBINING AI METHODOLOGY WITH RPA TECHNOLOGIES

When applying clustering methodology, it is necessary to make use of data analytics techniques and other artificial intelligence methods. However, the intention of our project is to include all this previous work in the implementation of the bot from the RPA provider,

that means that all the work must be done to be complementary with RPA technologies. It is important to combine languages and tools so every step of the process can be managed from both sides and there can be a fluent communication between them.

4.4 PLANIFICATION AND ECONOMICAL ESTIMATION

4.4.1 PLANIFICATION OF TASKS

When the project was launched, an initial time planification was elaborated through a Gantt diagram. This diagram contains the tasks and their duration during the development of the project and, with it, the intention was to divide the project into different stages assigning a number of days to each of them so we could get a picture of how the work would be distributed.

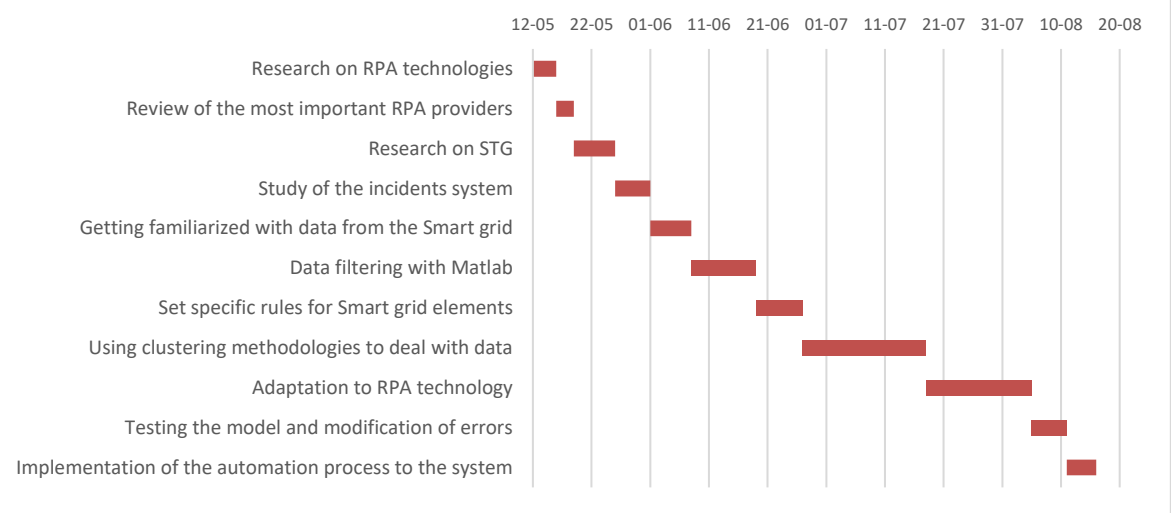

Figure 8. Gantt diagram of the predicted schedule for the project

However, this planning suffered from several changes during the development of the project. The most relevant ones are related to the delay of some of them because of inconvenients

that were out of our hand and depended on external factors, as it could be holidays from people in the project in certain moments. For this reason, we were forced to change priorities and to modify preferences at the time of approaching the process.

4.4.2 ECONOMICAL ESTIMATION

This report in included in a much bigger project, so there are steps like the development of the bot that are not studied here. The duration of the work that is stated in this document only lasts from May to August, however, the final results of the whole project will not be available until the end of the year. The economical estimation has been carried out with the help of consultant advisors from Capgemini.

4.4.2.1 Scope of the services

After the delivery of this report, next steps will be the design of the robot based on the information provided and predefined requirements, and then, the construction of it, including unitary and integrated tests and support of the operation areas in the production of the robot.

4.4.2.2 Price of services

The economic prediction of the project has been developed through an exhaustive study of the system and the information that is going to be required and has been divided into the design and the construction stages. In the next figure, it is included the service, the complexity and the price of each of them as well as the total price.

Servicio	Complejidad	Precio
Robot Design	High/Alta	1,311.87 €
Robot Construction	High/Alta	6,559.35 €
Total		**7,871.22 €**

Figure 9. Budget for the project from Capgemini (CapGemini, 2021)

4.4.2.3 Functional scope and justification of evaluation

The complete breakdown of how the project and every stage of it has been evaluated and calculated is included in the annex. It can be found there also the sizing and complexity that has been assigned to every use case and functionality.

CHAPTER 5. DEVELOPED MODEL OF SYSTEM

In this chapter, it will be described the implemented system and the models that have been designed to optimize the process of transfer of data in STG web.

5.1 ANALYSIS OF THE SYSTEM

First of all, we must carry out an exhaustive analysis of the system that provides us enough information to design the process in a proper way. This analysis will be executed over the infrastructure of the system and over the elements of the grid that play a part in the exchange of information in STG web. This flow of data is bidirectional and goes from the element that is not functioning correctly to the different units that are in charge of detecting and fixing the issue.

5.1.1 TITANIUMSTG

The system of incidences that belongs to STGweb is the one we are going to focus on but, it is included in a bigger system called TitaniumSTG that comprehends more actors that are present in Iberdrola. STG web is in the same level as SIC (Client Information System) and FTP SIC server. They connect with databases through an integrated bus and an API (application programming interface) which allow the flow of data between them and the whole system. Information provided by STG and SIC is gathered in these databases and once it is required by them, it flows in the opposite direction.

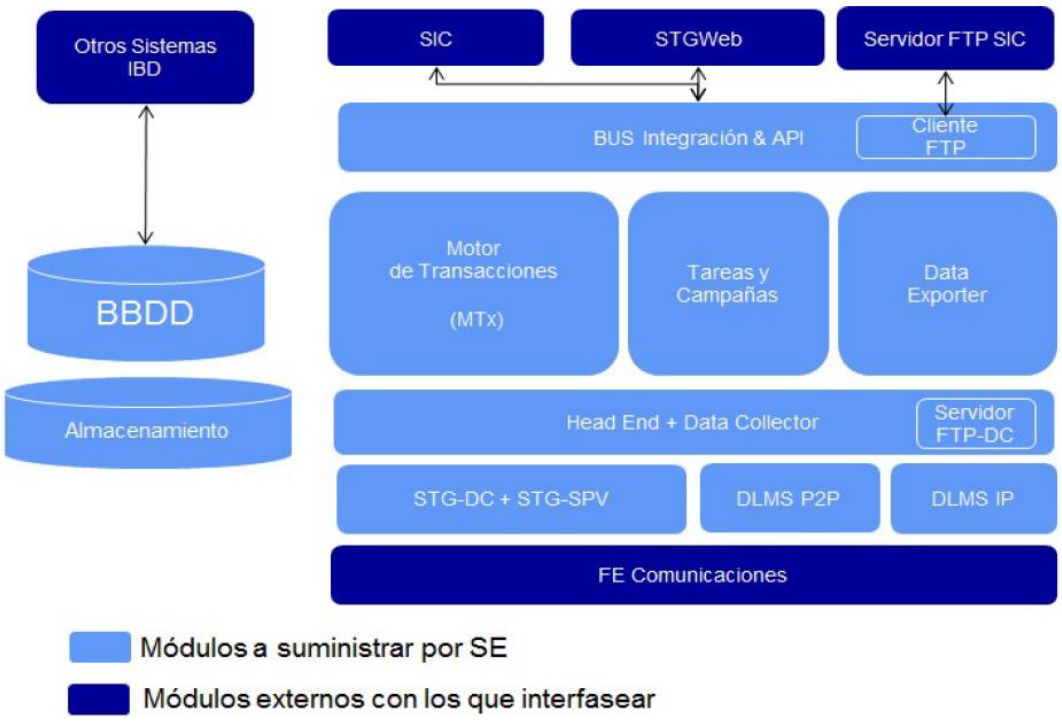

Figure 10. General vision of the system (Navarro, 2019)

5.1.2 STG WEB

Inside TitaniumSTG we can find STG web, which is more focused on the lowest part of the electrical grid, the one regarding DC, meters and concentrators among others. In STG web, diagnoses differs between works and incidences, a work is a task to be performed on an element of the grid where the solution is related to inventory, punctual need of work or old equipment to be replaced. On the other side, an incidence has its origin in a problem or a fault which has been detected in an installation or an equipment. This problem can be detected manually or through algorithms that have been developed in STG web through some parameters.

5.1.2.1 Automatic incidences

Automatic incidences are generated automatically if they fit the predefined parameters that have been previously defined. These are the ones that are going to be studied in this project

since are those that can be automatized and then, object of study by the robot. Although they fit predefined parameters, they will only be generated if:

- There is not a previous incidence of the same group already open.
- There is an open incidence in remote management, there will not be anyone for CN group.
- The transformer is not in the status "being tested".
- The transformer must be in the stage of "Exploitation" or "Finished".

An incidence is discarded automatically as long as it does not fit the parameters it did when it was generated and it is not under treatment yet. In that case, it is assigned to CSD (Center of service and diagnosis) and it will be reassigned from there.

5.1.2.2 Manual incidences

An incidence of an equipment can be generated by the person in charge of the device. CSD will open an incidence manually if:

- The automatic process has not generated an incidence when it should. In that case, it would warn SISCO (Systems of Control) for its analysis.
- There is an incidence without automatic generation where CSD can identify the problem.
- Incidences that suppose an actuation on devices or equipments.

5.1.2.3 Incidence status

Incidences are divided according to their status:

- Pending: the incidence is waiting to be treated by the assigned group.
- Pending visa: the incidence is waiting to be visa by CSD once it has been generated by another group.
- Open: the incidence is in process of management.
- Correctly completed: the incidence has been solved satisfactorily.
- Discarded: the incidence has been removed.

- Excluded: the incidence gets excluded so it does not generate automatically again while it is in this status.
- Provisionally closed: the incidence is closed until closing information from GAMAD (Management of Distribution Files) arrives.
- Failed at closing: this happens when it is closed wrongly.
- Closed reassigned: this occurs when the user reassigns the incidence at the time of managing it, creating a new one to be reassigned.
- Cancel: this status is requested when the incidence is assigned to UTD (Territorial Unity of Distribution) and from STG web, they want to cancel it.

5.1.2.4 Changes of status

- Step to open: the incidence changes to open when the user starts to manage it. Once open, it will not be discarded except for certain cases.
- Step to completed correctly: when CSD is assigned a pending visa incidence, they will have to get sure remotely that the incidence is solved indeed. If this is the case, the incidence will be closed correctly and, if not, it will be managed as failed.
- Step to discarded: the incidence moves to this status in case that it does not fit the parameters that were responsible for its assignment.
- Step to excluded: it will be established by the unit in charge, that will have to perform these exclusions only temporarily and when they get solved, they will be reopened again.
- Step to provisionally closed: this change only occurs in the incidences assigned to UTD where CSD calls them to state that is has been closed correctly.

In the next figure, it will be shown the flow chart of the process where:

- Pendiente: pending
- Abierta: open
- Pendiente visar: pending visa
- Completada correctamente: correctly completed
- Excluida: excluded

- Descartada: dicarded
- Anulada: cancel
- Gestionar: manage
- Cerrar correctamente: close correctly
- Completar correctamente: complete correctly
- Cerrar fallida: failed at closing
- Excluir: exclude
- Descartar: discard
- Reasignar: reassign
- Cerrada reasignar: close reassigned
- Anular: cancel

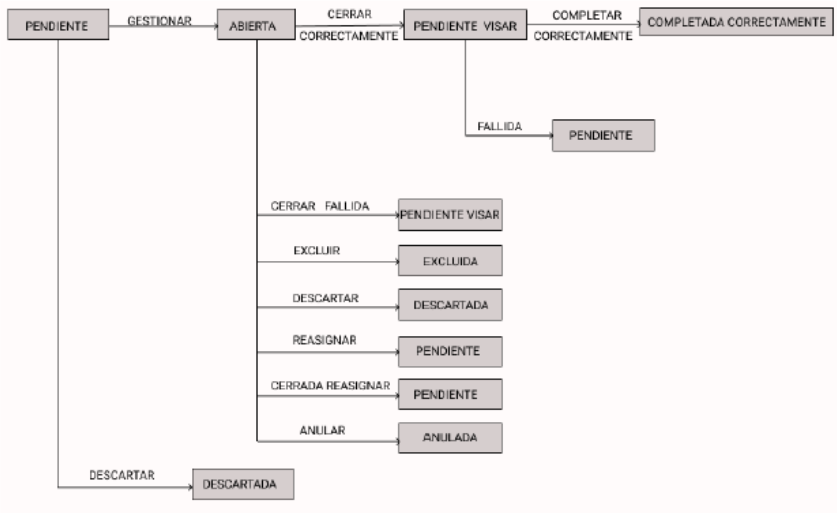

Figure 11. FLow chart of incidences process (Iberdrola, 2020)

5.2 DESIGN

5.2.1 DEFINITION OF THE DESIGN METHODOLOGY

The design of the model is performed following the collected information about the system and the processes that data follow since an incident is detected until it is closed. However,

before implementing any change in the system, we need to filter all the information embedded in the system and to gather it in groups that will have the same characteristics. As we want to automatize the whole system, the adopted solution for the design of the project is based on a data analytics technique called clustering.

Clustering is an important process within Machine Learning that allows automatic learning algorithms to train and adequately know the relevant data that they will need to develop certain activities (Grapheverywhere, 2020). This process helps bots to generate analysis skills minimizing the number of errors in every iteration. Its main purpose is to manage a group of sets of unlabeled objects so that smaller groups of data known as clusters can be created and controlled more easily. These clusters are applied to non-supervised machine learning models and are conformed by objects that have some similar characteristics between them and their organization is done by a predefined criteria established by the user.

5.2.2 CASES OF STUDY

In our project, the discernment of the clusters is crucial to properly analyze the different use cases that can be optimized in the system.

5.2.2.1 Most common diagnoses

The first case of study is focused on the most repeated diagnoses in STG web. Filtering data through this criterion will give us information about what kind of problems are occurring more often in the elements of the grid and, with that information, there could be a minimization of them. Tackling the worst and most repetitive cases enables a huge increase on the efficiency of the whole system by removing those problems that have a higher frequency all over the grid.

Nevertheless, this will give us an information that is not entirely useful since, when filtering the most common diagnoses in the whole system, there will be some of them that are not important at the time of designing the bot. Another filter to be implemented over the most common diagnoses is the way incidences are generated. In our case, only automatic

incidences will be object of study since manual ones have no interest at the time of automatizing them and, in consequence, will have no use when the bot works on them.

5.2.2.2 Incidences with more actuations

One of the biggest improvements areas in STG web is the number of steps that an incidence needs to take until it is finally closed or discarded. For this reason, if we study those incidences that have the longest path, which means the highest number of assigned actuations, we will discover how to optimize the rest of them. After being generated, Incidences are always under the responsibility of an organism and, there they are treated and analyzed. Once the process is over in each department, their status changes and they are moved to another one until they are completely finished. However, these series of actuations are not always efficient and there are useless steps that do not add anything valuable to the process. If that is the case, efficiency falls dramatically, so those are the ones that we will study in this chapter.

5.2.2.3 Distribution of incidences per years

STG web is working since 2013, so there is available data since that year to present. A full study of how the number of incidences has been changing since the implementation of the system can give us a bigger picture of what has happened and when.

5.2.2.4 Distribution of incidences per months

Another case of study will be the distribution of the incidences through the year, so it can be appreciated if there is a balance between months or there are some of them when the highest number of incidences concentrate.

5.2.2.5 Distribution of incidences per zones

Iberdrola divides their territory in the country in five zones: North, West, Center, East and Madrid. This division of zones can be interesting to study and observe if there is a pattern that shows a higher number of incidences in any of them. If that is the case, there will be some reasons that can provide us information about the whole system. Also, we want to see

if there is a great difference between the number of incidences in summer and the rest of years in these territories.

5.2.2.6 Evolution of resolution time

One of the main indicators of the efficiency and performance of the process of managing incidences is the time that it takes to solve them. If this process is too long, it means that there are some aspects that can be improved so the efficiency can rise. Evaluating the resolution time in the last three years will provide a good perspective of the impact of the implemented changes on the grid and their degree of influence in the performance of the system.

5.3 IMPLEMENTATION

Once different cases of study have been defined according to the preferences of the project, we are ready to implement the clustering methodology to get results from the system. Data analytics provides us with a set of tools that allow us to make division among data and to regroup it in the most interesting way for our interest. The implementation of this clustering methodology has been carried out with Matlab 2020 and the source code of every case of study that was defined in the previous chapter is included in the Annex.

CHAPTER 6. ANALYSIS OF RESULTS

After implementing the proposed methodology making use of data analytics tools, we are going to proceed to analyze the obtained results in each of the cases of study previously defined. Some of them will provide more useful information than others and the intention is that everything helps the future implementation of the bot in the system.

6.1 MOST COMMON DIAGNOSES

At first, we intended to discover the most repeated diagnoses in absolute terms and this was the result of the data filtering:

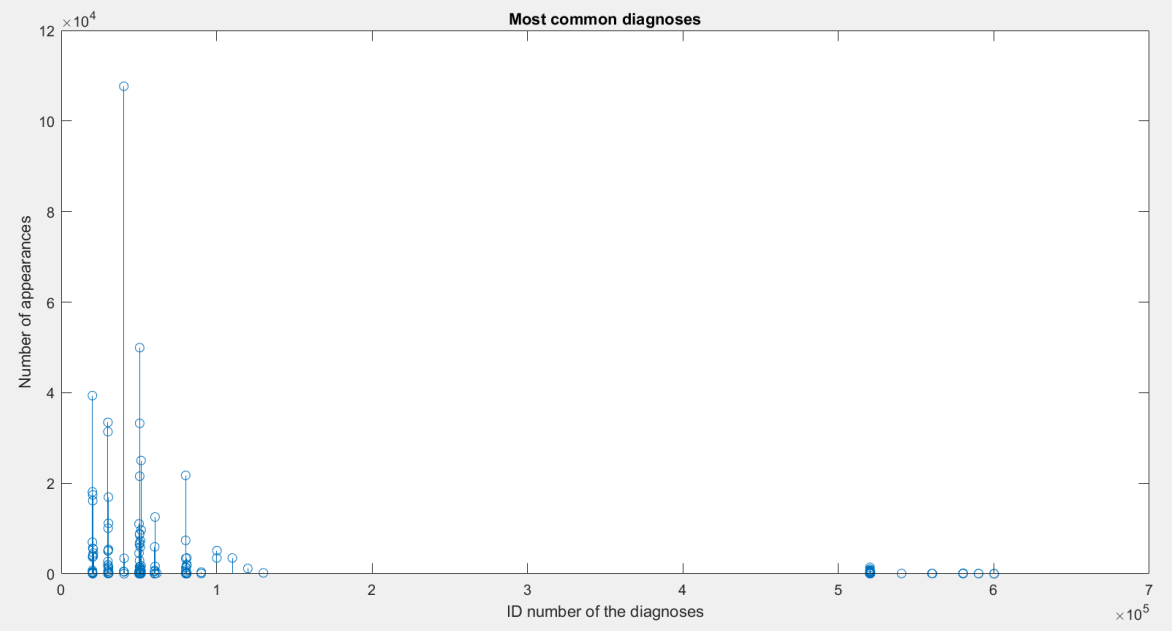

Figure 12. Most common diagnoses

In this figure, it can be appreciated the id number of the diagnoses in the x-axis and the number of times that it has occurred in the y-axis. However, at the time of analyzing the most repeated ones, we could observe that most of these diagnoses were related to manual incidences, those that were generated intentionally by an agent of the system. In our case,

this has no real interest for us since this kind of incidences are not likely to be automatized because of their nature, so we need to find those that are generated automatically so the robot can predict and optimize their process. After applying another filter, this time only for automatic incidences, we got the diagnoses that appeared the most. The selected diagnoses had to form a group between 10 and 15 samples after a discussion about the most relevant ones, so it was only picked those that had appeared more than 7000 times since 2013. These were the results:

- 20101: Not accessible remote. It is generated when the communication of the device works bad. However, it recovers quickly and then it is removed. It is typical from transformers.
- 30101: Not accessible concentrator. It is generated like the remote, but it only applies concentrators.
- 30103: Accessible concentrator without readings. In this case, the concentrator can be accessed by communications, but it does not get a clear reading. It can be caused by problems in Prime system or in the configuration of the concentrator itself.
- 30201: Not accessible node. It gets generated like the previous ones, due to a lack of communication.
- 30401: Not accessible TGB. It is generated the same way but it refers to a concentrator with an integrated router.
- 40301: Low functioning / bad quality RTU. It gets generated by noise in PRIME network, that is why it cannot be read correctly.
- 50101: Not accessible PLC system. It is generated when a switch or a PLC system has no access to the equipment situated downstream due to a cut of the communications. After some research, we have discovered that this used to happen more often during the first years of STG web and then, it did not occur so frequently. The reason behind this is because it used to happen to all the devices downstream that they all created an incidence, but now it only affects the first one in the line.
- 50501: Not accessible Router GPRS/3G/3GTP. If this router does not provide any signal in 36 hours an incidence is created.

- 50503: COM ACCES NOK/COM TELEG OK. This is similar to the previous one but it only generates the incidence if it has been 48 hours since the last provided signal.
- 51001: Not accessible PLC-Switch. It is generated if the switch or PLC system does not provide any communications in 12 hours.
- 51401: Not accessible GTP. It affects an auxiliar node with a router that communicates with a centralized concentrator.
- 51404: Accessible GTP without readings. It has the same process than the incidence for the concentrator. GTPs are actually quite problematic, their functioning is far from being optimized.
- 80101: Not accessible remote. It is generated as 20101 but it refers to OCRs, that equipment situated in the standings of the devices. They open and close MT current so their influence area is intended to be minimized.
- 80109: Low functioning / bad quality reading CN. It gets generated by noise in PRIME network and their accessibility percentage is smaller than 50%.

After the collection of results, we can state that most common diagnoses are related to devices that do not answer back to communications. This might happen for a lot of reasons, but the most frequent one is the malfunction of the communication system. This system is continuously under review aiming to be perfectioned while more agents enter the system and new devices are integrated there. If we want to reduce these kinds of diagnoses, it could be done an increase on the margin that these devices have to create the incidence, the more flexibility you give the system, the less incidences are generated.

6.2 INCIDENCES WITH MORE ACTUATIONS

In this case, we were trying to see why some incidences were redirected too many times until they got solved. That is the reason why we searched for those that had a bigger number of actuations. According to the predefined criterion, we would need between 10 and 15 samples

to make remarkable results so we searched for those that had more than 18 actuations, and these were the results:

- 53186: this incidence was generated in 2015, and in 2021 it is still generating actuations. It is related to a GTP unit and it has not been fixed in any of the actuations. According to the observations it looks like it has been wrongly programmed and there is always too much noise to reach it.
- 79216: in this case, the incidence affects a transformer and all the actuations were done in 2015. Apparently, the transformer was situated in a place with low coverage and even though the technician was fixing the issue every time it was sent to him, it was not until the problem of the coverage was faced that this transformer stopped giving problems.
- 162535: this incidence takes place in 2017 and the problem is the number of iterations that is has to run until it gets to the organism that is in charge of solving it. CSD and UTD send and receive the order until they get to a conclusion.
- 196302: this is carried out between 2017 and 2018. Concentrators in this unit are causing too many problems so there is an exchange of information between CSD and UTD to make a lot of revisions of the units. This exchange of information is what leads to this high number of actuations.
- 201735: this incidence affects a concentrator that needs multiple reviews and agents do not agree about how to solve it. There is not an agreement on how the issue should be solved, remotely or presential so each of them sends the order to the other one.
- 296824: this incidence involves the noise in a transformer whose origin is not detected easily so the process takes too many orders until it is finally completed.
- 357226: this incidence involves a problem in the saturation of a concentrator and has so many actuations because it is not done what is originally ordered, according to the observations, it is assigned to a group that has not proceeded to do the job and another one has done it instead. It looks like there is a failure in communication between them.

- 730990: a transformer is not performing right and there is not an understanding on how the problem should be solved. CSD is not happy with the monitoring carried out by the agents, so they order to repeat operation multiple times.
- 733902: a transformer is not performing right and the orders given by SISCO and the specialist are not carried out properly so it takes too long until it is solved.
- 737853: a router is not communicating properly and there is an exchange of information between UTD and the specialist about the monitoring and the operation to be done, since they do not seem to work.
- 874339: there is no coverage in the transformer so UTD says that they cannot monitor it anymore because it does not answer to the applied modifications that they were told to implement. This exchange of information takes too many steps in the process.

When looking at the incidences that have more actuations, there is a clear pattern through them that is repeated in most of the cases. The main reason behind them is the lack of understanding between the different agents or the dispute for having the best solution in every case. However, there is another reason that covers many of them that is the mistakes in solving an issue in the way that the specialist of the device has said it should be. This can be due to the tons of incidences happening every day that can cause misunderstandings and confusion to the different agents.

6.3 DISTRIBUTION OF INCIDENCES PER YEARS

A full picture of STG web performing can be achieved by observing the number of incidences that have occurred since it was implemented in Iberdrola. For this reason, an algorithm to calculate them has been developed and the results are:

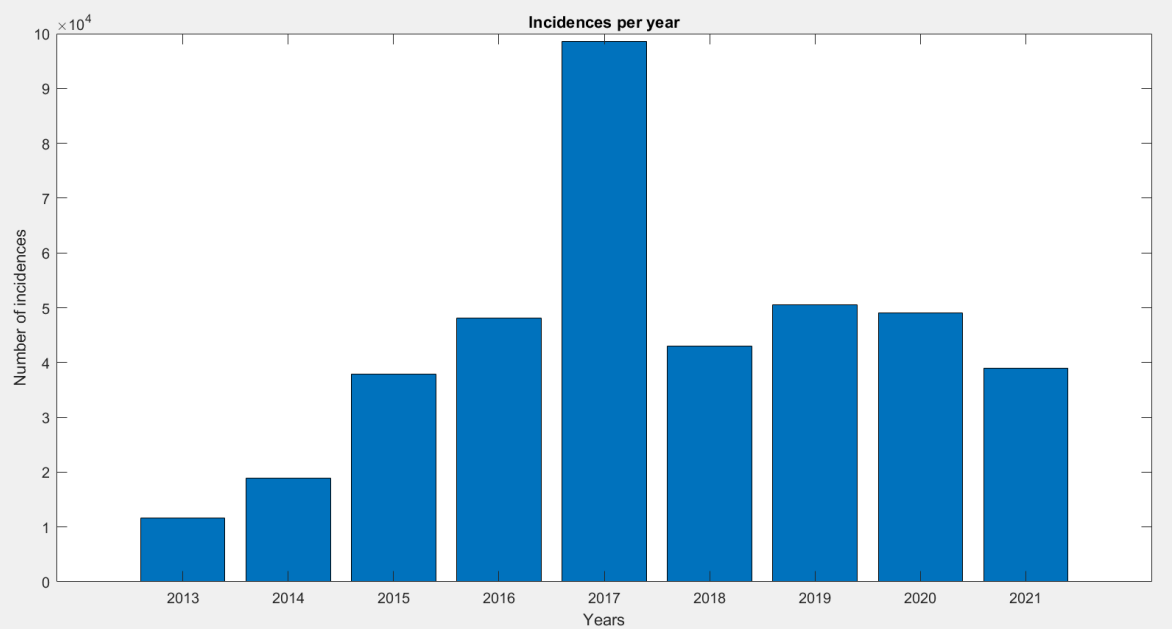

Figure 13. Number of incidences per year

As it can be appreciated, incidences show a curve that grows from 2013 to 2017 and then decreases until today. There are multiple reasons to analyze this effect, one of them is the inclusion on the grid of devices and their entry to STG web. In 2013, it was a new system with only not many smart devices so less incidences were created, but then lots of them were being included during the years and that obviously caused an increase in this aspect. However, the adoption of 3G technology instead of 2G has been also a progressive process that along with the inclusion of new technologies and tools has enhanced a decrease in the number of incidences as it can be appreciated since 2017. Since that year, a stability has been reached because of the balance of these new implemented technologies and the addition to the system of new devices that generate more incidences.

6.4 DISTRIBUTION OF INCIDENCES PER MONTHS

Just like it was done with years, the same analysis is performed with months in order to see if there is some pattern in the incidences during a year. In the next figure, it will be shown

the results of the number of incidences of every month since the implementation of STG web.

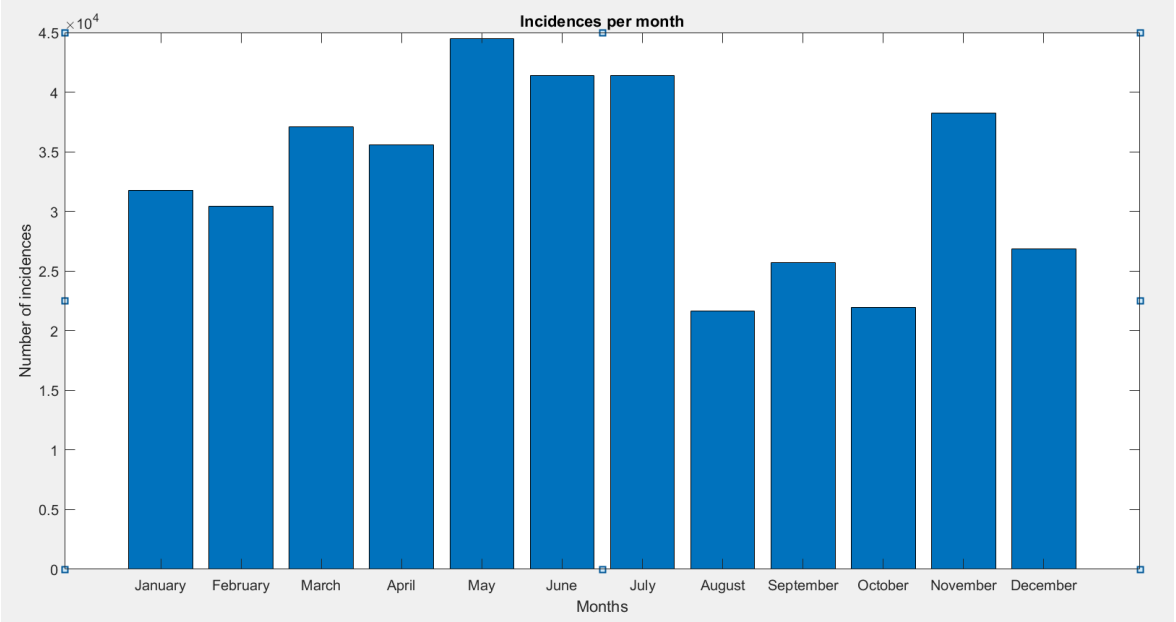

Figure 14. Number of incidences per month

At first sight, it can be observed that it is more or less linear during the year except for the months of August, September and October. However, after matching these results with the graphic of years, we found out that this is more similar to year 2017 and, as this is the year with more incidences, it increases the results on the rest of the months. Knowing that, there is still a little decrease on those months but apparently, after some research, there is not a clear reason behind it, so this case is not too relevant for our purpose of training the bot.

6.5 DISTRIBUTION OF INCIDENCES PER ZONES

At last, we wanted to observe if there was a tendency in the territories to have more or less incidences and if there was a big difference between this proportion in summer and the rest of the year. This is the general distribution:

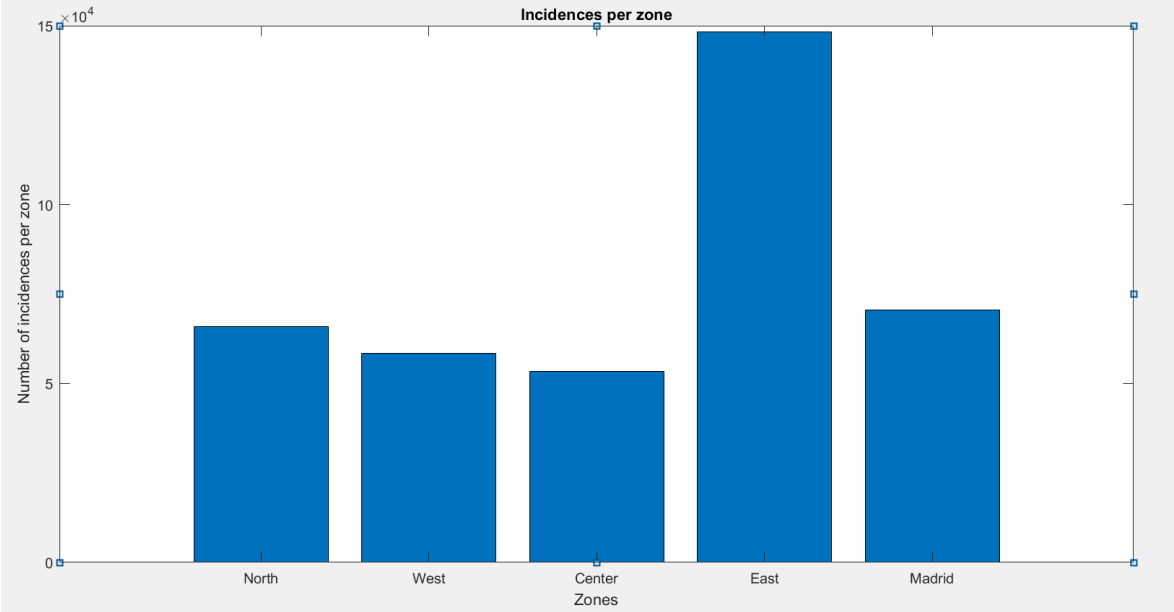

Figure 15. Number of incidences per zone

According to these results, it could be stated that there is a balance between the number of incidences in all the territories except for the East zone, that has more than the double of them than the others. Our first conclusion was that this zone had higher temperatures so that could be a possible reason. However, after a deeper research, it was found out that East zone was much bigger than the rest and, because of that, there was a higher number of installations in that territory. Therefore, the more installations there are, the more incidences there will be in that zone. To tackle this issue and have a more accurate vision of the system, it was calculated the number of installations in every territory and then, the number of incidences per installation.

ZONES	Nº OF INSTALLATIONS
NORTH	25208
WEST	28420
CENTER	29865
EAST	66983
MADRID	24622

Figure 16. Number of installations per zone

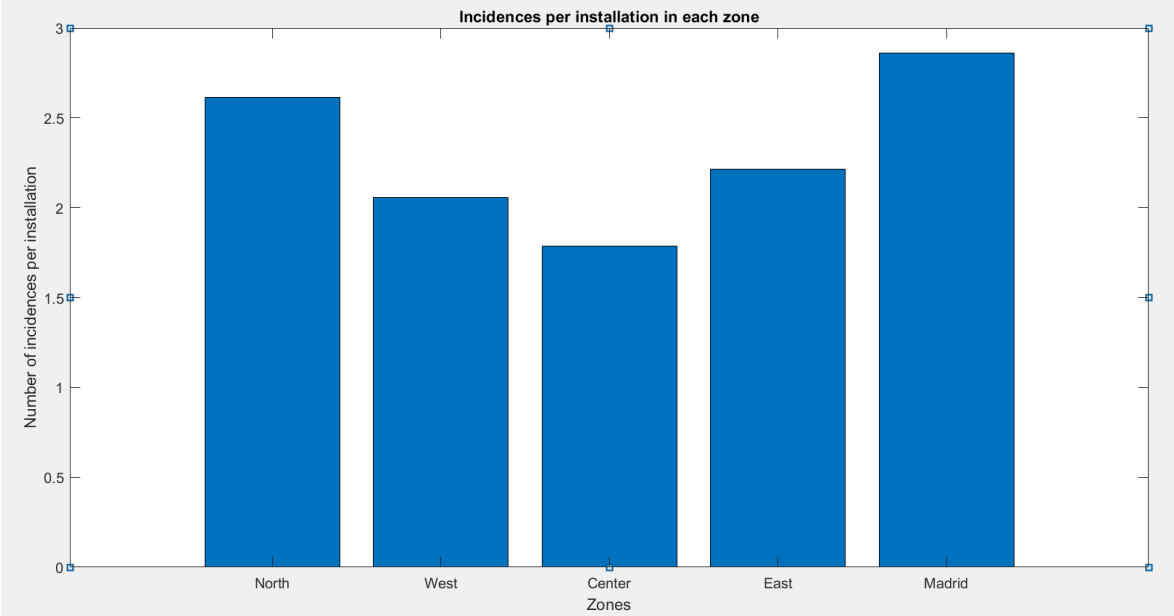

Figure 17. Number of incidences per installation in each zone

Now, it can be seen a more real picture of the situation of incidences in Spain. The area of Madrid is actually the one with more incidences per installation and the Center zone is in the opposite side. The explanation for this is the concentration of devices in a smaller zone, as it happens in urban areas like Madrid, instead of huge territories with a lower density of installation, as it happens in Center zone. A lot of incidences occur because of bad communication of devices that are harder to access when they are too concentrated, so that is the main reason why this is distribution has that shape.

Now, it would be interesting to watch if there are any changes if we filter data between summer and the rest of the year.

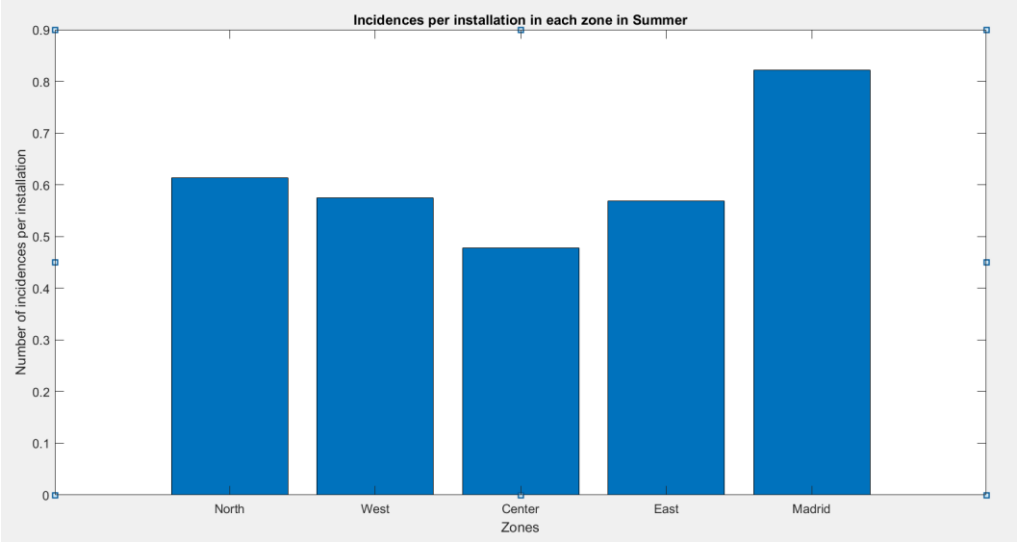

Figure 18. Number of incidences per installation in each zone in Summer

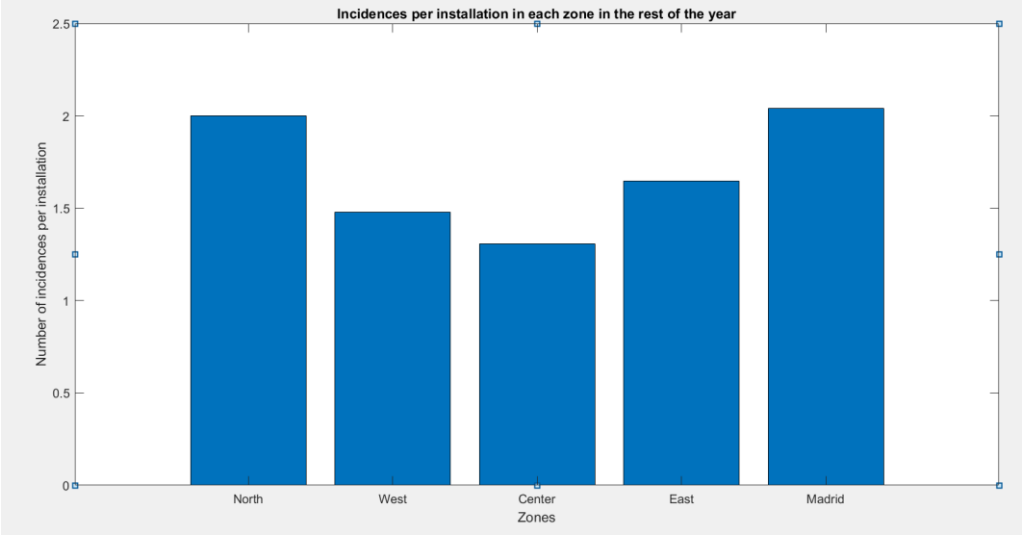

Figure 19. Number of incidences per installation in each zone in the rest of the year

There are no big variations in both situations except for a couple of cases. The proportion only varies because of a little decrease in Madrid and West zone and a bigger increase in North zone. This increase of incidences in the North is due to floods in the rivers of the zone that usually cause malfunctions in the devices nearby. Also, those floods make it difficult to repair the installations and usually the incidences need more actuations until they get completed.

6.6 EVOLUTION OF RESOLUTION TIME

At the time of analyzing the resolution time for the incidences in the system, a "boxplot" graphic was developed in each of the last three year to observe if there are interesting changes and the impact that they had on the grid. With this tool, it is shown the 75 percentile (top blue line), the 25 percentile (bottom blue line) and the average (red line) of the samples, apart from several cases between the 5% and the 25% of appearance.

Evolution of resolution time between 2019 and 2020

Studying the evolution from 2019 and 2020, it can be appreciated the decrease of the resolution time from one year to another in general terms. The average does not change too much due to the high number of incidences in the system, but the percentiles show this reduction of time clearly.

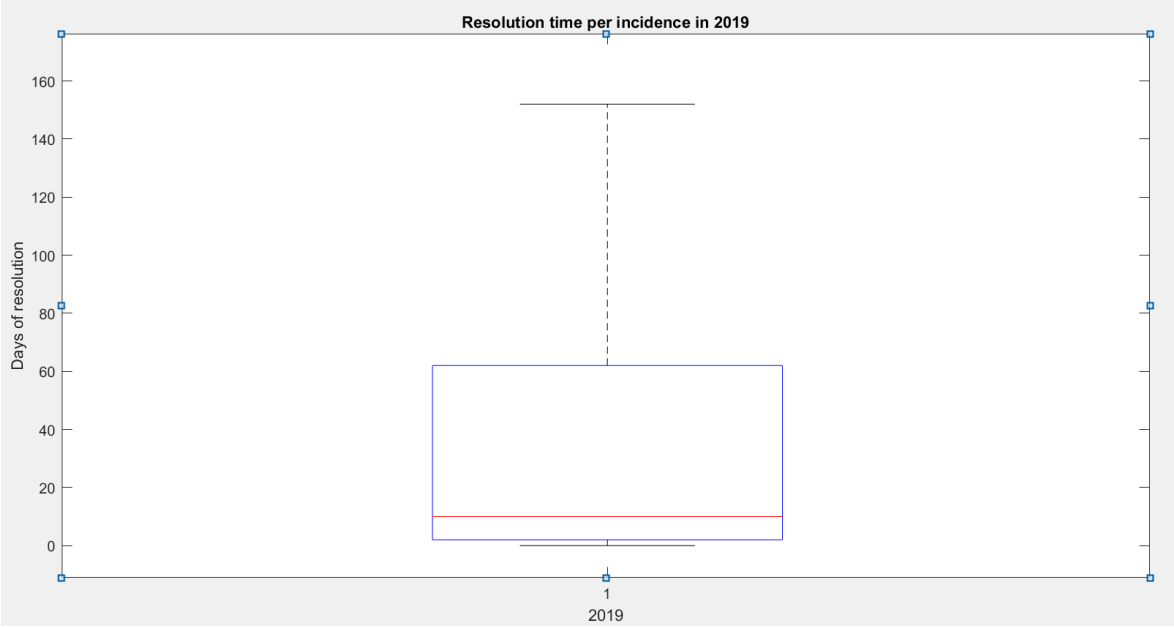

Figure 20. Resolution time per incidence in 2019

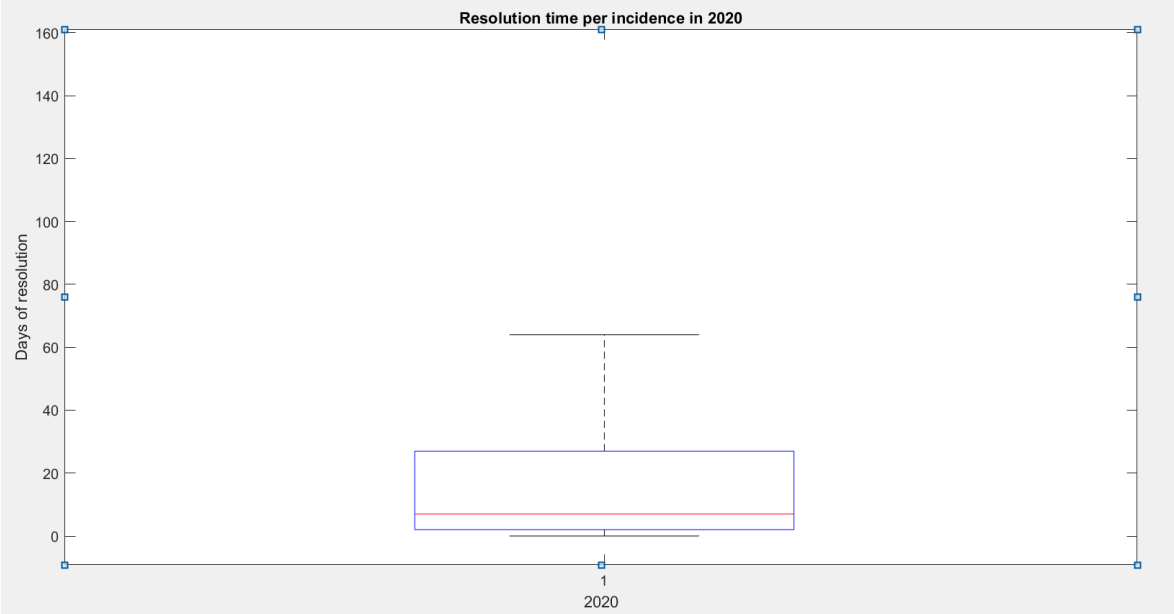

Figure 21. Resolution time per incidence in 2020 (1)

The main reason behind this is the improvement on the resolution processes with better reports, processes and equipment. Also, the experience of operators and digital coordinators is continuously growing so less mistakes and a higher efficiency is achieved.

Evolution of resolution time between 2020 and 2021

It is also interesting to study this evolution of the resolution time in the last year, so it can be analyzed the current situation of the system. The modifications of the processes that have been implemented during this period will be quite similar to those promoted in the next year along with the process of automatization so it can provide interesting information.

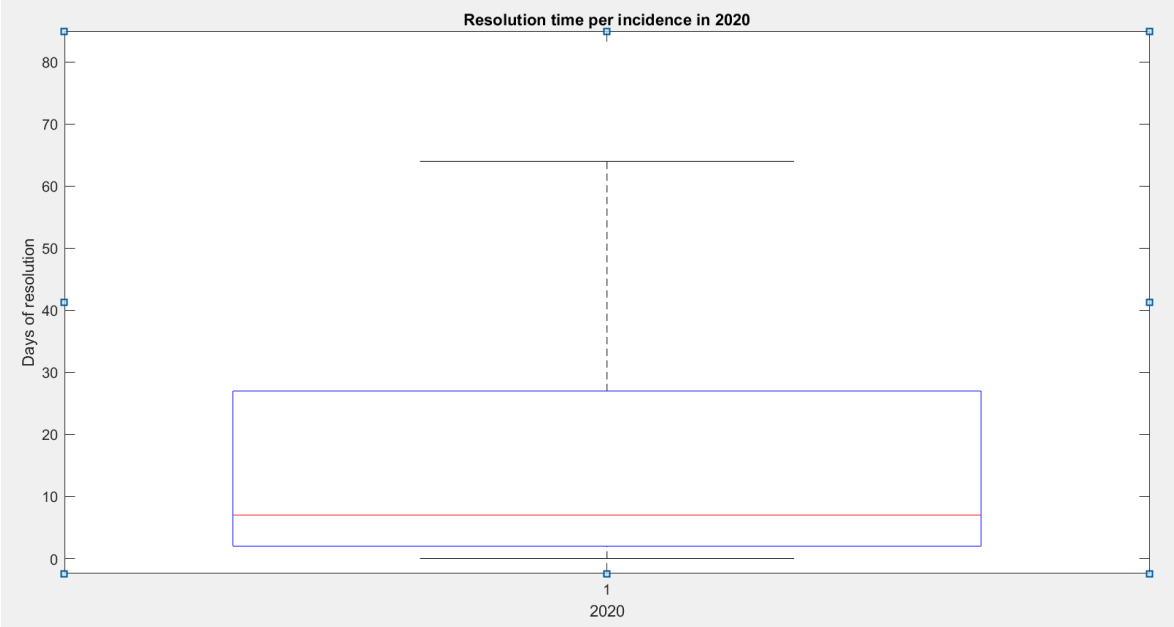

Figure 22. Resolution time per incidence in 2020 (2)

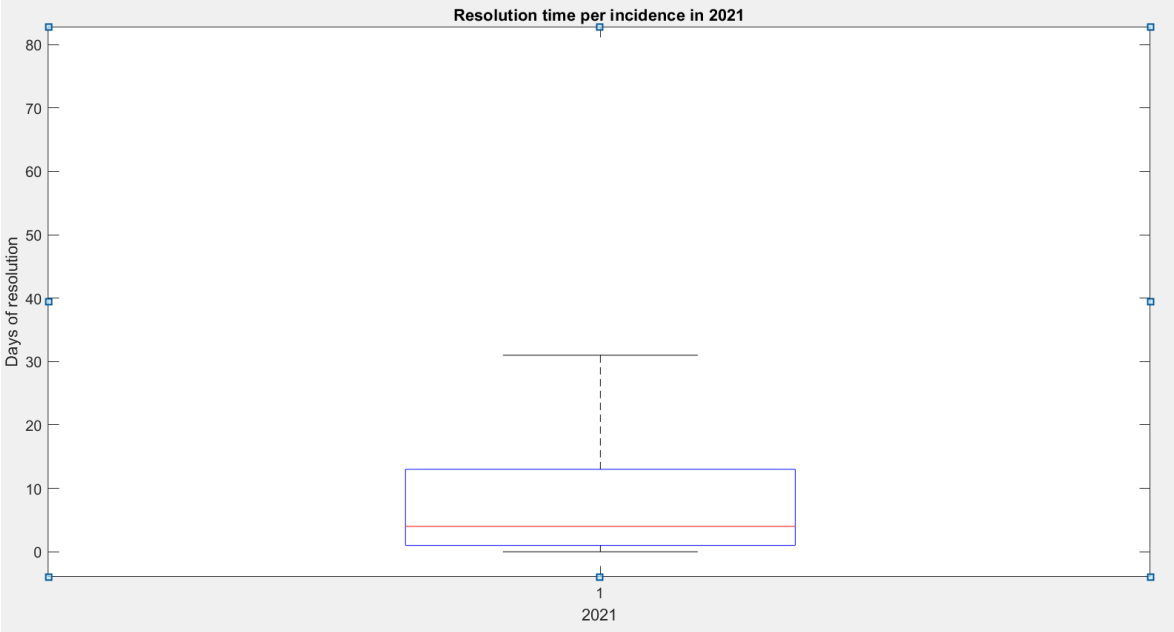

Figure 23. Resolution time per incidence in 2021

There is again a decrease on the resolution time per incidence, which means that the improvements implemented on the different processes have been working correctly. The reasons that explain this evolution are related to modifications and adjustments of the process of generation and discard of incidences. It has been improved so the generation values and the incidences per equipment become more accurate and more realistic compared to the actual situation. The discard process used to create many more incidences than it should because sometimes those discarded incidences showed up again automatically due to a slow update of their status. The system has been optimized so they do not generate again if they are already solved.

CHAPTER 7. CONCLUSIONS AND FUTURE WORKS

After the collection of results and their analysis, there are some conclusions that can be stated regarding the influence of the present report in the later introduction of the RPA technology in STG web. At the time of introducing an automatization methodology in the system, it is crucial to start implementing it in the situations and places where it will be easier to adjust the parts of the system and the results are more favorable. In this case, the project aimed to do a deeper research on the whole STG web to obtain great improvement areas and work methodologies that did not reach the expected performance.

According to the obtained results, there are interesting conclusions that can be assumed. On the one hand, there is a clear pattern in the diagnoses that appear the most in STG web, bad communication between the elements of the grid is the most repeated issue in the system. However, there are multiple cases where these issues are immediately solved and do not require an actuation on them. This is a step in the process that could be optimized by the robot by automatizing that kind of incidences and saving time and effort for the agents, also by adjusting the flexibility of the devices before they call for an incidence, that may be modified to only appear when it is really necessary.

Another interesting aspect that can be applied in the design of the robot is the optimization of the process of actuations of the incidences. This process has the biggest potential to be optimized, but it needs to be analyzed from different perspectives. Attending at those incidences that have more actuations until they are solved, it can be observed that sometimes there is a lack of understanding between the different organisms involved in the process. Disputes between them at the time of choosing the right criterion in every case and misunderstanding caused by a wrong communication make the process quite longer than it should be resulting in incidences that take months to be solved and multiple efforts and people involved that could be somewhere else doing other tasks.

STG web involves lots of different organisms and agents who theoretically would need to try to find the best solution for the common welfare. There is a clear scheme in STG web that shows the steps to be followed every time an incidence pop ups, however, when humans enter the equation, there can be discussions and differences of criterion that lead to a waste of resources. For this reason, the bot will have to follow a predefined path that, unlike people in the process, will always go on the same direction, without space for misunderstandings. Designing the robot taking into account every possible scenario will lead to a huge optimization of time and cost for STG web.

Urban areas are also another potential field where the system can be improved by the RPA technology. High concentrations of devices usually result in more incidences and warnings, so the robot must help relieving these areas of work for the agents. After the developed study of the system, it can be concluded that there are a lot of similarities between the incidences that are generated in theses areas, they usually include actuations of reviews and modifications that are not efficient until they get to the conclusion that there are too many devices concentrated there and there is the need of applying another measure. If the bot is given this information, it will be able to identify those incidences that are similar between them and create a standard automatized methodology for all of them.

Technology grows exponentially and the number of elements in the grid is going to be increased in the following years, for this reason, the robot must be designed to predict this kind of changes in STG web and to be capable of adapting to the new elements or methodologies that could be implemented in the future. As it could be deduced from the performed analysis in the report, there must be a balance between the growth of the present technologies and elements in the grid and the tools to manage them because if it is not the case, it might lead to a huge increase of incidences in the system that might provoke a collapse in the system.

The implementation of the proposed improvements in the system will not be appreciated until the end of the year 2021 at least. After the delivery of this report, the project will continue and the next step is using all the information and requirements provided here to design and implement the robot in the system. Due to the deadlines of this project, it was not possible to see the changes in the system and to observe if the work is actually well performed, however, according to the feedback of the company it can be stated that the project is truly valuable for Iberdrola and they will keep it for future implementations of automatizations in the system.

BIBLIOGRAPHY

Amatech Group. (2020). *Amatech Group.* Retrieved from https://www.rpasolutions.es/rpa-sector-retail/

Boulton, C. (3 de September de 2018). Obtenido de https://www.cio.com/article/3236451/what-is-rpa-robotic-process-automation-explained.html

CapGemini. (2021). *Gestión de incidencias de concentrador y CN.* CapGemini.

Deloitte. (2017). *The robots are ready. Are you?* Deloitte.

Deloitte. (2020). *Deloitte.* Retrieved from https://www2.deloitte.com/es/es/pages/operations/articles/que-es-robotic-process-automation.html

Edlich, A., & Sohoni, V. (24 de May de 2017). *Mckinsey.* Obtenido de https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-blog/burned-by-the-bots-why-robotic-automation-is-stumbling

Foundry. (2020). *Foundry.* Retrieved from https://foundry4.com/5-industry-applications-of-rpa

Grapheverywhere. (2020). *Grapheverywhere.* Retrieved from https://www.grapheverywhere.com/que-es-el-clustering/

Iberdrola. (2020). *Procedimiento de gestión de incidencias de telegestión, telecontrol y telecomunicaciones.*

Jacobsoft. (2020). *Jacobsoft.* Retrieved from https://www.jacobsoft.com.mx/es_mx/clustering-analysis/

Navarro, J. G. (2019). *TitaniumSTG.* Schneider.

Paradavisual. (2021). *Paradavisual*. Retrieved from https://www.paradavisual.com/entrevista-al-country-manager-de-blue-prism-para-espana-y-portugal-jose-manuel-garcia-gavilan/

Penteo. (2021). *Intelligent Automation Tools.*

Schaffrik, B. (2021). *The Forrester Wave™: Robotic Process Automation.*

UiPath. (2021). *UiPath*. Retrieved from https://www.uipath.com/es/

Unir. (2021). Clustering: qué es y cuál es su uso en Big Data. *Unir*.

Urreta, P. (2019, September). *Ricoh.* Retrieved from https://digital.ricoh.es/que-es-rpa-automatizar-negocio/

ANNEX I. ALIGNMENT WITH SUSTAINABLE AND DEVELOPMENT GOALS

The Sustainable Development Goals have been established as a wake-up call for all countries in the world, no matter level of richness to keep prospering without leaving aside the protection of our planet. Improving the level of welfare of countries and reducing poverty to the minimum by promoting strategies that build economic growth and address a range of social needs including education, health, social protection, and job opportunities, must not be an excuse to stop paying attention to climate change and caring about the environment. Covid 19 pandemic has increased the need, and it was already urgent, to take all the factors into account so every country can grow protecting the planet.

In this project we will try to adequate our proposal to some of these goals. In our case, these are the ones that we will try to adopt and to promote with the obtained results:

- **Decent work and economic growth**

Automatization of the process implies a huge change in the configuration of job positions. Those that were destinated to simple and repetitive tasks will disappear in favor of other ones with more responsibility and that will need a special training and education. This process will allow a fast economic growth in the area.

- **Industry, innovation and infrastructure**

New methodologies in the system introducing RPA technologies are included with the goal of improving the whole electrical infrastructure as well as the system of incidences of Iberdrola. Furthermore, innovative skills and techniques have an impact on the energetic industry pushing competitors to introduce new methodologies if they want to keep being competitive and having a strong position in the market.

- **Sustainable cities and communities**

The increase of efficiency that is pointed in this report as one of the main objectives will have an immediate effect on the sustainability of the whole system. The electrical grid will be improved and optimized reducing waste of time and resources in tasks that do not need to be cover that way.

ANNEX II. SOURCE CODES

In this annex, source codes from the implementation of the clustering methodology in Matlab 2020 are included.

```
b = importfile_CT_INCID_ACTU_DATA_TABLE("STG_INC_CT_INCID_ACTU_DATA_TABLE.dsv",
[2, Inf]);
c = importfile_CT_INCIDENCIA_DATA_TABLE("STG_INC_CT_INCIDENCIA_DATA_TABLE.dsv",
[2, Inf]);
d = importfile_DIAGNOSTICO_DATA_TABLE("STG_INC_DIAGNOSTICO_DATA_TABLE.dsv", [2,
Inf]);
```

The first case of study, most common diagnoses was coded this way:

```
inc = b.ID_DIAGNOSTICO;
[GR,GC]=groupcounts(inc);
figure
stem(GC,GR);
ylabel ('Number of appearances')
xlabel ('ID number of the diagnoses')
title ('Most common diagnoses')

[maxi,posi]=max(GR);
GC(posi);
GR(posi);
j=1;
for i=1:length(GC)
    if GR(i)>7000
        incid(j)=GC(i); %repite los valores
        j=j+1;
    end

end
%%
ind = find(GR > 7000);
tabla=table();
for iter = 1:length(ind)

    ind_tabla = find(b.ID_DIAGNOSTICO==GC(ind(iter)));

tabla = [tabla; b(ind_tabla,:)];
end

%%
```

```
ind_auto = find(d.TIP_INCIDENCIA=='A' & d.ID_DIAGNOSTICO==GC(ind(1))); %solo las
automaticas



%%
tabla = table();
ind_auto = find(d.TIP_INCIDENCIA=='A');
id_auto = d.ID_DIAGNOSTICO(ind_auto);

for iter = 1:length(id_auto)

    ind_tabla = find(b.ID_DIAGNOSTICO==id_auto(iter));

tabla = [tabla; b(ind_tabla,:)];
end


inc = tabla.ID_DIAGNOSTICO;
[GR,GC]=groupcounts(inc);
% figure
% stem(GC,GR);


ind = find(GR > 7000);
tabla_maximos = table();
for iter = 1:length(ind)

    ind_tabla = find(tabla.ID_DIAGNOSTICO==GC(ind(iter)));

tabla_maximos = [tabla_maximos; tabla(ind_tabla,:)];  % ESTA ES LA BUENA
end
```

Incidences with more actuations

```
inc2 = b.ID_INCIDENCIA_CT;
[GR2,GC2]=groupcounts(inc2);
% figure
% stem(GC2,GR2);

j2=1;
for i2=1:length(GC2)
    if GR2(i2)>18
        incid2(j2)=GC2(i2); %repite los valores
        j2=j2+1;
    end

end
%%
ind2 = find(GR2 > 18);
tabla2=table();
```

```
for iter2 = 1:length(ind2)

    ind_tabla2 = find(b.ID_INCIDENCIA_CT==GC2(ind2(iter2)));

tabla2 = [tabla2; b(ind_tabla2,:)];
end
```

Number of incidences per years

```
inc6 = c.FEC_CREACION;
y = year(inc6);


[GR6,GC6]=groupcounts(y);
figure
bar(GC6, GR6) %muestra las incidencias por meses
ylabel ('Number of incidences')
xlabel ('Years')
title ('Incidences per year')
```

Number of incidences per months

```
inc3 = c.FEC_CREACION;

m = month(inc3);
[GR3,GC3]=groupcounts(m);
figure
bar(GC3, GR3) %muestra las incidencias por meses
[GR3,GC3]=groupcounts(inc3);
ylabel ('Number of incidences')
xlabel ('Months')
title ('Incidences per month')
```

Number of incidences per zones

```
%verano
ind_verano =
find((month(c.FEC_CREACION)==6)|(month(c.FEC_CREACION)==7)|(month(c.FEC_CREACION)
==8));
id_verano = c.COD_REGION(ind_verano);
[GR4,GC4]=groupcounts(id_verano);

%resto del año
```

```
ind_inv =
find((month(c.FEC_CREACION)==1)|(month(c.FEC_CREACION)==2)|(month(c.FEC_CREACION)
==3)|(month(c.FEC_CREACION)==4)|(month(c.FEC_CREACION)==5)|(month(c.FEC_CREACION)
==9)|(month(c.FEC_CREACION)==10)|(month(c.FEC_CREACION)==11)|(month(c.FEC_CREACIO
N)==12));
id_inv = c.COD_REGION(ind_inv);
[GR4,GC4]=groupcounts(id_inv);

%
GC5=[1,2,3,4,5];
GR5=[0,0,0,0,0];
for i3=1:length(5)
    GC5(i3)=i3;
end

GR5(1)=GR4(1)/25208;
GR5(2)=GR4(2)/28420;
GR5(3)=GR4(3)/29865;
GR5(4)=GR4(4)/66983;
GR5(5)=GR4(5)/24622;
% GR5(6)=GR4(6)/29865;
figure
bar(GC4, GR4) %muestra las incidencias por zonas
ylabel ('Number of incidences')
xlabel ('Zones')
title ('Incidences per zone')

bar(GC5, GR5) %muestra las incidencias por zonas en verano o invierno
ylabel ('Number of incidences per installation')
xlabel ('Zones')
title ('Incidences per installation in each zone in the rest of the year')

tabla3=table();

for iter3 = 1:length(id_verano)
   iter3
    ind_tabla3 = find(c.COD_REGION==GC4(id_verano(iter3)));

tabla3 = [tabla3; c(ind_tabla3,:)];
end
```

Resolution time per incidence

```
[filas, columnas]=size(c);

c.dias_resolucion=repmat (duration (0,0,0),filas,1);

c.dias_resolucion=day(c.FEC_CIERRE-c.FEC_CREACION);

% 2017
```

```
dia_limite1=datetime(2017,1,1);
dia_limite2=datetime(2018,1,1);

% ind=find(c.FEC_CREACION<dia_limite1);
ind=find((c.FEC_CREACION>dia_limite1)&(c.FEC_CREACION<dia_limite2));

figure
boxplot(c.dias_resolucion(ind),'OutlierSize',0)

ylabel ('Days per resolution')
xlabel ('Zones')
title ('Incidences per zone')

%2018
dia_limite1=datetime(2018,1,1);
dia_limite2=datetime(2019,1,1);

% ind=find(c.FEC_CREACION<dia_limite1);
ind=find((c.FEC_CREACION>dia_limite1)&(c.FEC_CREACION<dia_limite2));

figure
boxplot(c.dias_resolucion(ind),'OutlierSize',0)

ylabel ('Days of resolution')
xlabel ('Year')
title ('Resolution time per incidence in 2018')

%2019
dia_limite1=datetime(2019,1,1);
dia_limite2=datetime(2020,1,1);

% ind=find(c.FEC_CREACION<dia_limite1);
ind=find((c.FEC_CREACION>dia_limite1)&(c.FEC_CREACION<dia_limite2));

figure
boxplot(c.dias_resolucion(ind),'OutlierSize',0)

%2020
dia_limite1=datetime(2020,1,1);
dia_limite2=datetime(2021,1,1);

% ind=find(c.FEC_CREACION<dia_limite1);
ind=find((c.FEC_CREACION>dia_limite1)&(c.FEC_CREACION<dia_limite2));

figure
boxplot(c.dias_resolucion(ind),'OutlierSize',0)

%2021
dia_limite1=datetime(2021,1,1);
dia_limite2=datetime(2022,1,1);

% ind=find(c.FEC_CREACION<dia_limite1);
ind=find((c.FEC_CREACION>dia_limite1)&(c.FEC_CREACION<dia_limite2));
```

```
figure
boxplot(c.dias_resolucion(ind),'OutlierSize',0)
```

ANNEX III. ECONOMIC ESTIMATION

The breakdown of the economic estimation is the next one:

Use cases detail and criteria evaluation per feature			
Use case and functionality	**Common Operations**	**Application**	**Screens**
1. Carga datos maestros			
1.1. Carga datos maestros (estructuras flujos y configuraciones por defecto), assessts y excepciones	Excel / Lectura (Simple)	MSExcel	1
1.1.1. Error Acceso/lectura datos maestros	Mail / Enviar Correo (+Adjuntos)	MSOutlook	1
2. I-DE Applex		Web	1
2.1. Activación Firewal		Web	1
2.2. Acceso STGWeb		Web	3
2.2.1. DIAGNOSTICO		Web	1
2.2.1.1. Panel de búsqueda --> Madrid --> CSD		Web	1
2.2.1.2. Tratamiento pendientes		Web	1
2.2.2. Proceso BAJAS PRESTACIONES/MALA CALIDAD LECTURA CN		Web	16
2.2.3. CONCENTRADOR ACCESIBLE SIN LECTURAS		Web	7
2.2.3.1. GESTIÓN EQUIPOS		Web	8
2.2.4. CONCENTRADOR NO ACCESIBLE		Web	8
2.3. Error tratamiento STGWeb	Mail / Enviar Correo (+Adjuntos)	MSOutlook	1
3. Tratamiento log			
3.1. Escritura log	Excel / Generar (Simple)	MSExcel	1
3.2. Envio de correo gestión de planta/usuarios	Mail / Enviar Correo (+Adjuntos)	MSOutlook	1

Screens	Data Set Management (0 - Simple) (1 - Complex)	Business Rule (0 - Simple) (1 - Complex)	Comments
1	0	0	
1			
1		0	
1			
3			
1		0	
1			
1	0		
16	1	1	Complejidad validación estructura flujo, reglas a tener en cuenta y recuperación/tratamiento datos en pantallas
7	1	1	Complejidad validación estructura flujo, reglas a tener en cuenta y recuperación/tratamiento datos en pantallas
8	1		
8	1	1	Complejidad validación estructura flujo, reglas a tener en cuenta y recuperación/tratamiento datos en pantallas
1			
1	0		
1			