Ministry of Education and Science of the Russian Federation
Federal Independent Educational Institution
"NATIONAL RESEARCH TOMSK POLYTECHNIC UNIVERSITY"

School of Engineering and Entrepreneurship
Direction of training (Specialty) *31.04.02 Management*
Division

MASTER'S THESIS

Topic of the work
Совершенствование информационной безопасности ИТ-решений для бизнеса (Improvement of data security of IT-solutions for a business)

UDC 334.012.64:004.056

Student

Group	Full name	Signature	Date
3АМ6Ф	**Victor Hohn**		

Scientific Supervisor

Position	Full name	Academic degree, academic rank	Signature	Date
Associate professor	**T.R. Rakhimov**	**PhD Economics**		

ADVISORS:

Section "Financial Management, Resource Efficiency and Resource Saving"

Position	Full name	Academic degree, academic rank	Signature	Date

Section "Social Responsibility"

Position	Full name	Academic degree, academic rank	Signature	Date
Associate professor	**N.V. Cherepanova**	**PhD Philosophy**		

ADMIT TO DEFENSE:

Head of the Division	Full name	Academic degree, academic rank	Signature	Date
Associate professor	**N.O. Chistyakova**	**PhD Economics**		

Tomsk – 2018

Ministry of Education and Science of the Russian Federation

Federal Independent Educational Institution
"NATIONAL RESEARCH TOMSK POLYTECHNIC UNIVERSITY"

School of Engineering and Entrepreneurship

Direction of training (Specialty) *31.04.02 Management*

APPROVED BY:
Head of the Program
_____ _____ _____ N.O. Chistyakova
(Signature) (Date) (Full name)

ASSIGNMENT
for the Master's Thesis completion

In the form:

Master's Thesis

For a student:

Group	**Full Name**
3AM6Ф	**Victor Hohn**

Topic of the work:

Improvement of data security in a company through IT-solutions	
Approved by the order of the Head (date, number)	**1982/C from 20.03.2018**

Deadline for completion of the Master's Thesis:	**June 9th 2018**

TERMS OF REFERENCE:

Initial data for work *(the name of the object of research or design; performance or load; mode of operation (continuous, periodic, cyclic, etc.); type of raw material or material of the product; requirements for the product, product or process; special requirements to the features of the operation of the object or product in terms of operational safety, environmental impact, energy costs; economic analysis, etc.).*	**Issues of Data Security** **Legal documents** **Dissertations on the topic** **Periodicals** **Interment sources** **Information on the company for internship.**
List of the issues to be investigated, designed and developed *(analytical review of literary sources in order to elucidate the achievements of world science and technology in the field under consideration, the formulation of the problem of research, design, construction, the content of the procedure of the research, design, construction, discussion of the performed work results, the name of additional sections to be developed; work conclusion).*	**1. THEORETICAL ASPECTS OF DATA SECURITY** **1.1. Definition, history of data security** **1.2. Data policy** **1.3. Risks and risk management of data security** **2. ANALYSIS OF THE OBJECT: HESUS COMPANY** **2.1. Description of Hesus company** **2.2. Company's IT System** **2.3. Challenges and potential evolution of IT tools** **2.4. Current implementation of data security** **3. PROBLEMS, DEVELOPMENT AND SOLUTIONS** **3.1. Problem areas** **3.2. Solutions for data protection** **3.3. Solutions and efficiency**
List of graphic material *(with an exact indication of mandatory drawings)*	

Advisors on the sections of the Master's Thesis *(with indication of sections)*	
Chapter	**Advisor**
Social responsibility	**N.V. Cherepanova**

Date of issuance of the assignment for Master's Thesis completion according to a line schedule	**November 15th 2017**

The task was issued by the Head:

Position	Full name	Academic degree, academic status	Signature	Date
Associate professor	**Timur R. Rakhimov**	**PhD Economics**		

The assignment was accepted for execution by the student:

Group	Full Name	Signature	Date
3АМ6Ф	**Victor Hohn**		

Ministry of Education and Science of the Russian Federation
Federal Independent Educational Institution
"NATIONAL RESEARCH TOMSK POLYTECHNIC UNIVERSITY"

School of **Engineering and Entrepreneurship**
Direction of training (Specialty) ***31.04.02 Management***
Level of education **Master's Program**
Division
Period of completion **spring semester 2017/2018**
Form of presenting the work:

Master's Thesis

Topic of the work
Improvement of data security in a company through IT-solutions

SCHEDULED COURSE ASSESSMENT CALENDAR

for the Master's Thesis completion

Deadline for completion of the Master's Thesis:	**June 09th 2018**

Assessment date	Title of section (module) / type of work (research)	Maximum score of the section (module)
24.04	*Chapter 1*	*…*
15.05	*Chapter 2*	*…*
01.06	*Chapter 3*	
05.06	*Social Responsibility section*	
06.06	*Internship Report*	
07.06	*Final control*	
11.06	*Downloading into the system*	

Made by professor:

Position	Full name	Academic degree, academic status	Signature	Date
Associate professor	**T.R. Rakhimov**	**PhD Economics**		

AGREED:

Head of the Division	Full name	Academic degree, academic status	Signature	Date
Associate professor	**N.O. Chistyakova**	**PhD Economics**		

ABSTRACT

Master's Thesis
_80_ pages,
_12_ fig.,
_00_ tabl.,
_64_ references,
_02_ appendices.

Key words: Data security, IT-system, threats, data, information, encryption, data protection

The object of the research is data security in Hesus company

The purpose of the work is to study connections between data security and IT-system of a given company

In the course of research theoretical aspects and topicality of data security were studied; Hesus company IT system was analyzed

As a result of research problems of IT system data security were pointed out and solutions for the problems were generated.

Basic structural, technological and technical-operational characteristics: doesn't apply

Degree of implementation: is not implemented yet

Application area: IT systems of any organization.

Economic efficiency/significance of research outcomes increase of data security and efficiency of IT systems.

Future prospects research results are going to be suggested to the company.

Table of Contents

Introduction

From the ages, humans needed to protect some information. It can be for many different reasons such as military aim, politic aim, or just to keep certain secrets. History had been built by stories, facts and events that weren't possible if all the information were free and available. In the modern history, the D-Day in Normandie, when American, Canadian and English soldiers landed on beaches, it was a huge operation, with a huge logistic. It could not have been possible if Nazis heard news about it. I England, they created fake tanks and other machines in balloon to make focus the enemies on the wrong target. Also, the decryption of the Enigma machine by Turing was another way to get information from intercepted communications. Those examples show how the information management gave decisive advantages on huge conflict.

Also, during the cold war, soviet airplane constructor successes to have the same innovations on the same period as the Concorde with the Tupolev 144. Finally, we discovered that plans were copied by soviet agent because they were hiring as cleaning agents and have accesses to the trashes where engineers throwed their paper plans. Those examples are sounding like there is only nations that are involved in this security information business, but it is not. This question affects much more entities than states.

Indeed, companies do not escape from this reality. We can take a lot of examples. One of the most famous is the Coca-Cola recipe. The company had been created in 1892, and since this time, nobody knows the recipe to recreate exactly the beverage. The legend say that a Coca employee get acknowledge about it, contacted Pepsi to sell the information, but those ones refused to know it and warned Coca about this employee. They manage to keep this information so secret that even if some competitors tried to imitate the original taste, none of them succeed. According to Coca, the original recipe stays in a safe box, and is known by only two person that never travel in the same time by the same transport way.

Nowadays, safe boxes are not enough to protect information. With the technology development, most of them had been digitalized and store on digital platform. Also, the vocabulary evolved. We started to talk about data, that is no more than information about something or someone stored on a computer or a server. The difference

is that digital environment offers new possibilities and creates new threats. The volume of exploited data dramatically increased the last years, and their content also been more complete.

The advantage of this new storage format is that the storage capacity increased, that allowed companies to offer new services, get more information, use more and more of it to be more and more competitive. It also developed a new convenient way to use it. If everything I stored in the same place, and you can organise this space as you want, needed information or reachable faster than with some paper files. Also, the risk of losing it is reduce by the necessity to take care of the computer, or more, just put them on a cloud to make them reachable from everywhere and make them depend only from a server.

Another point is the circulation. With all the new means, the information circulation potential also increased drastically. When in the past, the information was owned by one person, or worst, this person had to search into different files to find the good one, now, with the networks possibilities, needed information are available for the person that need it. The different information technology system allowed a very fast, well thought, and convenient information distribution. Each office can work more efficiently, for the same number of employees, the production possibility increased a lot, that permitted to companies to develop their activities, their competitively, and their horizons.

The individual employee potential is developed also. It allowed new models, small entities can generate very big turnover with few employees and have access to international markets very easily. They can make concurrence to giants on the place since many years and develop new ideas. Those solutions help the creativity potential to increase. A good idea now can be developed efficiency easier and faster than before.

But with those possibilities come new threats. In the past, to protect information or documents, the goal was to hide it, put it in a safe place (safe box for example), or just keep the idea or the knowledge in the mind. At least, paying attention with the brevet that were deposited and about the machine and the used process was the basic of the information security management. Nowadays, because information are data, they

are stored on server, with accesses. The past robbers have today equivalents as hackers. They have the same potential, but the difference is that technologies are less known by people generally. The fast evolution creates very new behaviours. The apparent complexity of those change makes them difficultly understandable.

In fact, because possibilities to connect to data are plenty, possibilities to attack them are also plenty, maybe even more. And because sensitive data are stored with other data, they need more protections. One of the reason that there is more and more sensitive data on networks is that basics information are used in more and more different ways. You can log in to the bank website, the cookies stored all your navigation traces, they can be used to know user tastes, user preferences, what user like, what interest the user, they can know him very well, maybe, according to some algorithm and determined patterns, more than himself.

Because the data usage increased and, the data value follows the same evolution, it is correlated. And because the amount of data available also exploded, protecting them began a veritable challenge. On one hand, it is very important because the company activities depend of the accuracy and the correctness of them. If they have some corrupted ones, it can make them make their job very badly, they can not complete their mission efficiently, with all the consequences that it includes (bad reputation, useless costs, bad products, decrease of productivity, new issues to treat, more work to do etc.)

On the other hand, consumers, especially since the European directive, called GRPD, are more aware that enterprises collect their data. They know that it can be sold to other companies with commercial aims. Even if they have the rights to manage it, ask for deleting, or transfer, anonymising etc., a company that protect in a bad way his data, that mean user data, will have a huge reputation deficit. Of course, with social networks, the reputation will be attacked from everywhere, some people that should not be involve will give their opinion, most of the reactions will be hostile. The traffic will decrease, user can delete their account, and if the movement is big enough, it can ruin an organisation.

So, the data security has a double challenge, ensure the society inside operations, to make it work in a good way, and ensure the consumer trust, to make them continued

to provide the requested data. But there is a third point, it is to keep the competitive advantage. Because data are stored online, competitive advantages are fundable in those data. For some businesses, it is applied prices, margin, purchase sources, or just the way that the structure is organized. But for some others, financial data and R&D data are also stored and make all their difference with their competitors. Because they are making research on a particular field, or because they want to invest in a special company to get a certain technology with the mind to develop a new function for their product, all that strategic information has to be protected. They can be guessed by simple deduction (what is called grey information in economic intelligence), with random data, such as flight tickets, phone call passed, the provide demand of new materials, hiring of new personal etc. The definition of sensitive data has a non-negligible role.

1. Theoretical aspects of data security

1.1. Definition, history of data security

1.1.1. *Key definitions*

The data security can wear plenty definitions:

“Data security refers to protective digital privacy measures that are applied to prevent unauthorized access to computers, databases and websites. Data security also protects data from corruption. Data security is an essential aspect of IT for organizations of every size and type.” From techopedia.com [17]

“Data security means protecting digital data, such as those in a database, from destructive forces and from the unwanted actions of unauthorized users, such as a cyberattack or a data breach.” From Wikipedia.org [18]

“Data Security refers to the process of protecting data from unauthorized access and data corruption throughout its lifecycle. Data security refers to data encryption, tokenization and key management practices that protect data across all applications and platforms.” From software.microfocus.com [9]

“Data Security concerns the protection of data from accidental or intentional but unauthorised modification, destruction or disclosure through the use of physical security, administrative controls, logical controls, and other safeguards to limit accessibility. Ways of securing your data include:

Data Encryption - converting the data into a code that cannot be easily read without a key that unlocks it.

Data Masking – masking certain areas of data so personnel without the required authorisation cannot look at it.

Data Erasure – ensuring that no longer used data is completely removed and cannot be recovered by unauthorised people.

Data Backup – creating copies of data so it can be recovered if the original copy is lost.

General good practice, however, goes beyond these methods. Stringent processes should be put in place to cover all areas such as Password Policies.” From edq.com [8]

All those definitions have in common the explanation of the main idea: data protection.

Considering all those definitions, we can define Data security as the idea to protect all data on database, on a special server, on a computer, on mobile, on the cloud etc. from different dangers, as cyberattack, data breaches, corruption or any different leaks, in media storage, mainly on hard drive, IT system, networks, devices, with technical solutions such as data encryption, tokenization, key management, data masking, data erasure, data backup for example, or also education about behavior to adopt.

1.1.2. *History of data security*

50’s – 60’s

At the beginning, confidential data transfers were made through dedicated phones lines, with lines specially prepared to link two data centers. Because internet was, basically, an American army creation, they had budget to allow special material

to create their network. Phone lines network and internet network were highly connected. At the beginning, it was even the same one. Early, in 1967, they founded that this way is very unsecure, and started to think how to put more steps between connexions, make them more complicated. The first act was to include different levels and passwords to have access at different files and have the equivalent of today administrator right. This solution was available until personal computer started to be famous, creating a new threat.

70's

Indeed, with first personal computers, in the 70's, came the first hacks. The first pirates discovered weakness in telephonic networks, and started what is calling "phreaking", what is pass free call everywhere, including long distance call (that was something very exceptional to be able to do one, and very expensive to pass this kind of call), from California to another state, until to Australia for example. It was like an achievement, and in the same time a crime. With this huge breach in their system, FBI became more aggressive with those hacker, and took really seriously the threat after the first attack who make money, around 70 million, against Chicago's first national bank in the 80's.

80's

The first virus, named "brain", was created in 1986, by two Pakistanis brothers. The aim of this innovation was just to block on a hard drive data and show on the user screen a message. Later, they explain that it was just to show their talents to IBM. The same year, a law, the computer fraud and abuse act, was promulged, in the idea to protect people against hacking, a new form of out of laws trend. Kevin Poulsen, a famous hacker who the figure of this move, was put on America's most wanted list before he was arrested in 1991.

90's

Years 90's are the true beginning of massive hacking, threat creation, like worms, viruses and other malwares. In 1989, the first computer worm is created by Robert Morris, and create the first big threat for internet. Because the machines weren't

an efficient as today, this malware had the potential to knock down the system. The reactions to protect users against it were, like the old adage advocates, “prevention is better than cure”. On the same time, a lot of computer security products appeared on the market. It’s also the beginning of viruses’ creation. The most famous was the Michelangelo virus. It was programming to attack on the 6th of March (birthday of the artist), by changing first seventeen lines of hard drives and so destructing some data, but not all of them. The 90’s are also the period when users started to put their personal data on internet. With this new use, appear anti-virus market, different from other security tools, because it pretends to defend your machine against unwanted malware. Another virus who started to make noise was Boza. He was targeting only windows system, but without mean effect. In fact, he just opened a dialog box every 31th of month.

2000’s

1999 is the beginning of a new period, kind of a golden age of viruses. Those who left a trace in cyber-security history are melissa virus, who was sent by David L smith in may 1996, and the ILOVEYOU virus. Both of them are working on the same model, they are sent by mail, but their new idea is to use the principle of social engineering to make the receiver open the mail and the attached file. Another revolution is that, with the new software box such as Office 97, Office 2000, and mainly with Word and Excel, who have some script language like visual basic, they can be executed automatically, without asking the user. All the purpose of the social engineering is to transform those viruses in macro-viruses, because they will touch a huge amount of people. When they infect one machine, they infest files, and contact list to send it self to all victim’s others contact. This combination was very new. Melissa was sending all documents on your device to your contacts, without considering the confidential state or not, in a random way (very dangerous for companies and governmental computer), while ILOVEYOU just modified first page and extension files. Other important malwares Nimda (could infest just by internet navigation and by mail, was replacing and renaming files), Blaster (who left sometimes to restart the computer before blowing with a lot worms), or Netsky (had a lot of variations, that’s why he was difficult to stop, could disactivated anti-virus and send himself by mails.).

It’s on this period that anti-viruses had a huge sold “boom”. They kept the same principle, the same adage, precaution is better than cure, but evolve their weapons. They focused their strengths on analysis of different programs, e-mails and on internet navigation, to prevent the user against malicious software. All this period had a good effect, it was a good education to internet surfing and which behaviour should you adopt with unknown mail or in front of strange program. It’s also on this period that Organized crime started to have interest in the field and thought that activities in relation with computer hacking could be a clever and efficient source of income. In the future, they will organize what nowadays we know as “dark web”. All those threat will not stop users to leave their data online. In 2003, more data had been left during one year than all human history on internet. The amount of accessible data created new interest.

Of course, when different organizations like big companies or government understood the potential of hacking, they started to implement their defence, but also to trick their competitors or foreign states. Since 1999, chinese colonel imagined and theorised a kind of war “out of limits”. This idea continued to grow and to be effective with the time, because of all the opportunities that it creates, not only in spying, but also in propaganda and in sabotage.

Later in the 00’s, hackers changed their way and their target to focus more on financial, credit card, and bank information. One of the most active was Albert Gonzalez, who stole on average 45.7 million payment card data from an American retailer, TJX. These information are equivalent at more than 200 million dollar. At the same time, Apple, with the first iPhone, put internet in everybody’s pockets, creating a new way to consume it, to use it, but also new doors and way to attempt to steal personal data. Those new doors are calling applications. In fact, their development creates new weakness, so new potential breaches, while being true data’s vacuum cleaner. Android system follows next year Apple’s IOS, opening more gate, making every user data easier to steal. At least, it also opened the possibility to have different data from the same user, making data harvest more complete, and allowing those who use data to have more information on a larger public (juvenile, people who didn’t used internet or

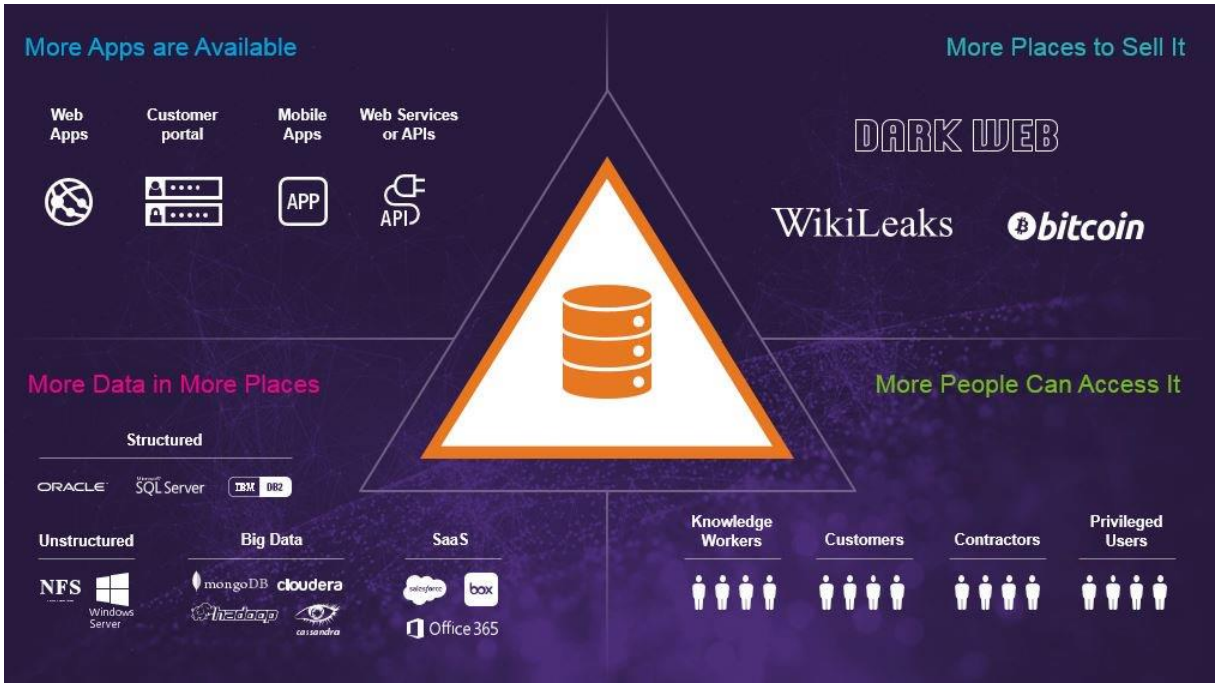
at least left as less traces as they could). That's means also more target, and more accurate profiles. It's on this period, with the GAFA (Google, Amazon, Facebook, Apple) or GAFAT (Twitter) affirmation, that the data started to have a value. With their business model (the user doesn't pay, he's the product, those websites sell space advertising to companies, the possibility to have an accurate targeting), the data's value grown and grown with stop.

2010's

That bring me to the next period, the actual one. In 2010, a small group reunifying bests computering/internet scientist across the world made a report with this conclusion:

"the cyber-universe is complex well beyond anyone's understanding and exhibits behavior that no one predicted, and sometimes can't even be explained well."

Internet is now so big and so important, technologies are very developed, and since the beginning, hackers always are more creative and have a step beyond. The Organized crime who started to have interest in hacking computer in the beginning of 00's, has now a true organization, with processes, better tools, more talents, they can organize for who ask for it, almost any attacks against any organization. They put their interest on this "market" because of the value of the data. The use of the word market is appropriate because there is a real demand and a real offer for those kinds of services (a two years old database of millions, maybe billion mail account, with mail address, passwords, IDs secret question and back-up mail can be sold for 300 00$ on the dark web). With the evolution of technologies, all the data theft activities found new ways to be done, and explored many new possibilities. Because data is the center of digital traffic, there is more and more apps containing user data, those data are stored in more and more different places, there is more and more people who put their data on different networks, and the things that allows the grown of the demand on the black market, there is more and more way to sell the thefts, in more and more secure ways.



Like the organized crime, the governments also improve their possibilities in this sector. The cyberwar, concept that born in 1999 in Chinese mind, is now so far from the beginning, governments are able to spy anybody (Edward Snowden, NSA, PRISM Project for example), they can break nuclear power plant (Iran), the used of data can influences elections and votes (2016 American presidential campaign, Brexit vote), they created literally cyber armies. Companies in each country are also weapons in this opposition. USA, with Google, Amazon, Facebook, Apple and Twitter, have five internet giants that open them possibilities for a lot of different intervention area (mainly in espionage and propaganda). But it's the same with Russian and Chinese internet giants. For example, Kaspersky researcher founded in 2016 a malware named ProjectSauron, who was so well developed, elaborated and so deep in systems, that their only conclusion was it had been sponsored by a state. According to their analyses, this program was in files since 2011, and was made to steal key encryption, passwords, plaintext and all other sensitive information from targeted computer in targeted organizations, even from devices without internet connection. Those organizations are from Russia, Belgium, Iran, China, and some other countries. There is a collaboration that give to state abilities to steal and use citizens data.

It is similar for dark web activities and companies, they have the same power. They have the same potential, they can do the same things as governments, particularly if they have the equipment. The thing who makes a difference is that today, the danger for hackers is not to fail on their attacks, but to be catch after this one. It's almost impossible to avoid all the offensive, ways are just too much, and very creative, complex, and there is more and more opportunities every days to invent new ones. The best dissuasive weapons against them is the justice pursuit. Effectively, after an attack, people in charge of security system can get the trace of the pirates, and find him in real life. That the biggest challenge. Basics anti-virus continue to implement the old adage, precaution is better than cure, but it still works only for a casual use of internet and computer, without sensitive data, to protect against general threats. More accurate dangers are almost impossible to prevent, because attacks are so sophisticated.

Those examples will be explained further with Yahoo! Case.

An important thing is to understand what imply to implement data security. It is a very complete idea, that is composed by a data policy (how and which data are harvest, stored and used, for which purpose) and a set of means implemented such as protection against attacks from outside (hacking attempt for example), attacks from inside (theft, espionage, etc.), data protection (how to protect them during an exchange, or just when they are stored), and processes to reduce as much as possible the risk of human error (wrong data sent, wrong receiver, a session not closed etc.).

The first step is to define the data policy. It is separate in three different areas : Data privacy, data stored, use of data.

1.2. Data policy

1.2.1. *Data Privacy:*

One of the most common mistake of people is to confuse data security and data privacy.

Data privacy, on internet, refer to the field of collecting/disseminating data, and which policy are used on different pages. You leave data every time you surf on internet. All request on Google, every page on Facebook, every time you buy something on

amazon, all private navigation, every website you'd visited, everything that you do on internet create data that will put in relation with you.

The most famous and knowing example of this process are internet cookies. Basically, when you go on a website, the owner can create, if he wants, and/or if you allowed it, a small file with the information of your navigation, what do you watched, what interest you, on which link did you clicked, that he will store on your computer. This file will be activated when the user will come back on the website to allowed the owner to get a historical and propose things who are correlated with his tastes. Cookies are just an example of data privacy, especially if website owner ask you or tell you that they use it.

The fields where data privacy is the most important are fields with very personal data, such as financial websites, or institutions, biological and medical records, justice case and investigations, public administrations etc. With those services, data privacy is very important, and they have to insure you that nobody who doesn't have the right will not be able to take a look on them, not only on internet, but also in real life. Therefore, by using all the possibilities to ensure it, you can use them with confidence, and they can respect their roles.

Around the world, most of the country adopted data privacy laws. USA, Canada, western and northern Europe, and Australia have laws that are considering as heavy, Argentina, China, Japan, and eastern Europe country are considered as robust, Russia, Mexico, turkey, Saudi Arabia, Egypt, south Africa are considered as moderate, other country as Pakistan, India or Indonesia have limited laws. But what does it mean? For example, Russian laws or considering as moderate. Indeed, most of the rules are coming directly from Russian constitution and Russian work code, but also from different international convention. Every personal data operators have to store, on databases or in server, all personal Russian data on Russian territory. If they don't respect this law, their website can be blocked. If it happens, they will be registrate on a special register. The problem is this law isn't old, and it still an unclear situation between who deserve to have the website blocked and who deserve to be in the register. In France, we have European laws mixed with national laws that we try to update as soon as it needs. We

also have a special organism who is dedicated to monitoring the respect of laws on internet.

Recently (the 25th of May 2018), the GPPR (General Data Protection Regulation) started to be obligatory. This European law has the aim to improve the security of users' personal data against companies who harvest it and use it for commercial activities or identifying people better. The law implies to create a permanent data monitoring service (most of the time it is one person who will be in charge of it). Users will have more rights, such as the forgiven right (delete the totality of data concerning the user), the right to manage his personal data, the possibility to transfer data from a point (website, company, organisation etc.) to another one without restrictions, the possibility to have the data anonymized after three years of inactivity and other measures that will allow to know more about the data that the user left after a web surfing session, those that are used about him, and help him to improve the management of them.

A major example of the difference between data privacy and data security is Facebook with the affair of Cambridge Analytica. Cambridge Analytica is a communication company who based their techniques on data exploration and analysis. In 2018, they are accused to have a very important influence on 2016's American presidential campaign and on the Brexit campaign, in United States of America and United Kingdom. Until here, a communication companies who's involved in political vote, it is not new. What is new, is, to designed the profiles of the voters they wanted to touch, they used Facebook data. For the American presidential election, the current numbers or talking about on average 87 million profiles who were used. It means that there is a leak of 87 million American profiles. The question is: how Cambridge Analytica uses an out of laws ways to get those data, who else can steal data from Facebook, and how are they protecting our data, so respecting their data policy about data privacy. In 2014, Cambridge Analytica already had relation with Facebook. They already obtain data from the giant by another app who's used by psychologists. This app is a personality quiz. It asked rights to have access to your profiles information and your friends' information too. With this way, with all those stolen data, a huge data base had been created with elector profiles. To reunify all those data, on this period, they used the same technique who looks to be used in 2016.

With Mark Zuckerberg auditions, American justice asks directly Facebook about his relation with data security, are they really able to ensure it, especially for a GAFA member. The actual chosen defence line is to assume that there is past mistake, the society is already investigating to understand where the breachs are, and they already took decisions to solve their old problems. Obviously, when question about user data harvest, for example about their navigation out of Facebook, or about the property of data, or else if there is a control of whom buy advertising space with which aim, or also if it is possible that other apps could, as Cambridge Analytica, stole huge amount of data, and if it is the case, does they know it.

On this point, we have to remind Facebook's business model: They have on average two billion users, with data about each one. It means they have one of the biggest databases in the world, about citizens from all the planet. Access to this social media is free, they make money by selling advertising space on their website, with the promise that the targeting will be very accurate, because gg

That's all the paradox of FB. On one hand, they give you the promise that you will be the master of your data, still have all control on them, but on the other hand, it's known that they use data the get from user to make money, with the promise towards company that, with their knowledge, they will have the best result, and they are the best if you want to invest in advertising. One of the point is that it's very easy to pay an advertising space, and there is a lot of modalities to pay, it's not only with an amount of money. You can pay according to the number of click, the success of your campaign, and you have many settings that you can define to target efficiency profiles that you want touch.

Therefore, the security of your data doesn't depend of data policy of the website where you'll leave them. As we see it with this case, data privacy is just a promise to don't use them with bad aim. One day, you can receive a publicity to vote for Trump without any justification, even if you didn't ask for anything, because your data would have been used by political service, that after analyse, choose to target you, because your profiles matches with potential voters. Of course, it is the same with companies.

1.2.2. *Storage of data:*

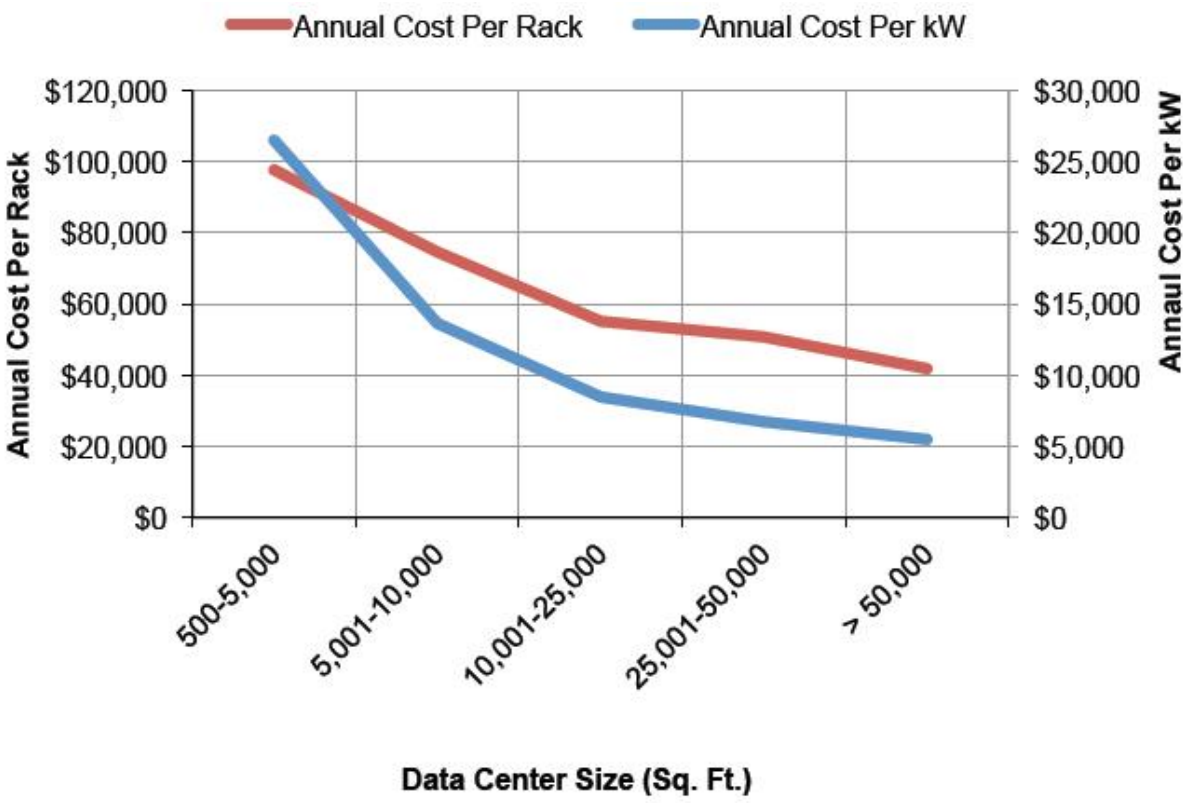
After harvest data, companies have to stored them to be able to reuse it later for their different applications. Nowadays, every data is stored on servers. There no organization that are doing it in another way. Only people, for personal use can stored their information on hard drives.

But what is a server? Basically, it is a huge hard drive that is totally dedicated to store data. The interface is as minim as possible, most important operation are doing by network connection. So, the challenge with servers is to keep their access to them safe to protect data against any theft, corruption or even spying. They are one of the main part of the network, a lot of devices can be connected to it / them, they can also be connected with other ones. If an unauthorized visitor success to be a part of this network, he will be able to get all the information he wants. The first point is to secure the connection to the device (we will see further how we can improve security, principally by the router's connection).

Another important point is to choose the way to implement the server in your network.

If the company has it own data-house, so you can rely it by cables. The connection is entirely configured by the Information Technology (IT) service, the enterprise has the entire control on all the network and the information circulation. But, there is a huge cost for it. Indeed, being the owner of a datacentre is very expensive. Data centers consumed 5% of the worl energy consummation in 2015. Half of it was just climatization to refresh those centers and make them working well. In France, this consummation, still in 2015, was higher than a 2 million people city consummation on the same time. According to those information, purchase enough energy to the data center to ensure his operating capacity has a certain cost (even if we exclude the buying, installing and maintaining cost). The best way to reduce this price is by making scales economies. As it is showed in the following scheme, economies can be very interesting if the capacity is increased a lot. For most of the companies, excluding the internet giants, reach this data storage needs is not possible. It can be a problem of

traffic (they don’t generate enough visits, they don’t have enough users), or their activity is to extensive, they have to store data into different centers, none of them can grow enough to pretend to have competitive price.



GAFAM members (the most known is amazon) found to this problem a market to explore. According to their statues and their size, whatever they needed huge data warehouse. They had to invest to buy and refresh new servers. Because of their big capacity, they started to offer to rent some space on it to other entities. It was a solution to fight the waste of storage capacity and monetize it. Moreover, it gives them some influence on other companies, or markets. This shape of storage is actually known as the “cloud”. Apple, Amazon, Google and Microsoft have developed their own solution for different markets. Amazon propose an offer to rent servers for companies in their data warehouse, with a complete solution to give it keys in hand to the developer to allow him to start the building and management of the platform as fast as he wants. Actually, they declare that they have as their cloud user famous such as Netflix, Air Bnb, NASA, Lamborghini for example. Apple has developed a solution only for his users (particularly for those who used portable devices such as phones, pads, watches

etc.). It allows the user to leave data on his own space (on the server, he has a dedicated space that is just the limit that apple allow him to use) and use it everywhere until he has the good connection IDs. Google Proposed to store data from his different services such a gmail (mail box services) or drive (online storage file) for example. Every information left on those platforms are keeping in google devices and keep accessible for the user under the IDs connection. The difference with apple is that you can accessed to your personal space from every devices, whereas for Apple cloud, you have to get an Apple product. Microsoft also get a position on cloud market, but they did it for companies. They specialized their offer not on individuals but on bigger structures. With Office 365, they developed a full set of software to make faster most of professional operations. Their goal is to make working the enterprises more efficiently. With those solutions, there is propositions of data storage that are complementor. They also buy Skype, and outlook, to improve communication possibilities. The only member of GAFAM that doesn't developed any offer of data storage is Facebook. The reason is obvious, their purpose is to collect data from their users. The first service to individual is a giant server where people leave data and can see these of other people.

Because of the cost, renting servers in data warehouse instead of getting his own is a very popular. Almost all companies use those services (not only with GAFAM, there is a huge market), because of the economics possibilities. As we will see further, even some governments chose this option.

The danger with those solutions, even if they are very convenient, and cheap, and fit to the society needs, is that those societies don't have the full control of their data process. So, the choice of supplier become critical to ensure data security.

Another question about server security is the access. Who have access to the data? How are they going out of the database?

1.2.3. *Use of data:*

This question is very important. Every businesses implement the way that they use their data differently. Even for enterprises who are working on the same market, with similar position, offer, and structure, the data distribution will be different. The choice of needed information will not be the same, even if it can be pretty similar. Moreover, different data will go to different offices that have different missions. For

example, a company who is producing and selling phones. They will collect a very big amount of data, mainly about the users, and the devices. Those information will not be used by the same departments. The commercial office will not have the same interest as the R&D office, that will not have the same as the production office. But, data will be stored in the same database. Moreover, different IT systems can be implemented in the same society, such as an internet website, a CRM, an ERP etc. All those informatic structures need data, if possible from the same source (what can imply some other problems in the way to save them, and need a strong and very well build database). All those objects are connected and have access to the server. There is as much connections as accesses to define and secure.

Today, for every type of structures, IT companies developed offers that include main services that enterprises need in general through CRM (salesforces, Sellsy, etc.) and ERP (SAP, JD Edwards, Oracle etc.). Their aim is to facilitate the information circulation from the storage place to the person who need it. By facilitating the circulation, they open the accesses to more and more people. So, the problematic in this case is who can see which data. For some uses, a limited amount of data is needed. To reuse the example of a telephone company, commercial department doesn't need to know which component is weaker and create duration life battery problems. In the same idea, both marketing and R&D (developers) should know which function is the most used, how and why, to potentially improve this point and continue to innovate and grow as company. Decision to distribute data to departments is taken after consultation by heads of departments and IT department, who can give advices and find appropriate solutions. For bigger societies, each branch has it own IT system fitting to it needs, linked to the appropriate database (who are also located in data warehouse specialists). Information are then aggregated and joined under one readable form by some ERP aggregator, that will translate different data from different software. The process is the same for CRMs.

Those software are, as server purchasers, well protected with a high encryption level. When hackers want to attack it, they try to enter in the system, so they try to steal passwords to get some rights. Those entrances are particularly targeted because it opens the possibility to get information of all the system, especially with administrator rights.

As it is showed further, the admin right policy is very important, less people have access to important rights, better the system can be protected. To be able to reduce admin number, the data steam has to be accurate. When it is well designed, different user have access only to information they need to manage their mission, not useless ones. The purpose of the IT team is to be sure that everybody gets the good data, and so minimize the number of people that can get access to all the database. This action cover two fields, make sure that other departments can focus on their mission and don't waste time to get good conditions, and secure the different exchanges.

Respecting those rules can avoid a lot of risks. Of course, there is always things that are unpredictable. But those simple settings are basics of data security policy. Having a clear data policy, mastering the storage of his data, know how his own data are used in his own organization is essential, is a very important part to improve the data safety. An intrusion can have devastating effects. The amount of data that can be stolen can reach amazing rates, costs of leaks can be very high, some consequences can important further than simple economic costs.

1.3. Risks and risk management of data security

1.3.1. *Biggest data breaches*

Experian (2013)

A young Vietnamese man (24 years old), named Hieu Minh Ngo, create a company, NGO'ID, in Vietnam, and had a contract with an American company, Court Ventures. This Company has access to the huge database US Info Search, as his clients. The purpose of this company was to allowed the access against money. Hieu Minh Ngo had a contract as a private instigator operating out of Singapore, and was paying regularly through a Singapore's bank, according to the contract. US Info Search has data about on average 200 million American people, like name, date of birth, social security and other records. He was siphoning data through Court Venture, until Experian, one of the three US major credit's bureaux, bought them. This big company had much more data about American citizens, as number of driving license, bank account and credit card data etc, data that mister Ngo also stole. The final aim was, obviously, to resell it. Finally, he was catched by U.S. secret service, who create a fake meeting to invite him

on American's territory. In 2014, American justice estimated that 3.1 million of people get their data used because of him, a small amount comparing to the 200 million records that he had in his possession. [20]

U.S. Military

What is funny, it is that, unlike company, U.S military have a lot of data breach, quiet regularly, with sensitive data, sometimes even classified. The three following stories are not because of hacking, it's just due to human error.

The first really huge data breach happened in 2009. At this time, an agency, we don't know wich one, sent back a hard drive to it vendor, GRMI, to get repairing and recycling. The problem? Sensitives data, as veteran's information, including medical information, weren't erased. The agency thought that it wasn't a matter because they had a contract with signed privacy promises. But GRMI was unable to fix it, so then send it to another company for recycling. They didn't know that all devices with those kinds of information should be destroyed by the National Archives and Records Administration. [14]

There is also more recent data breach, like for example one who is dating from 2017. A huge amount of data, who were stored on a private cloud server (Amazon Web Server 3), because of an update, started to be totally free and accessible to anybody. Those leak files were extremely sensitive, it was data from operating commandment in middle east, Africa, Asia, south Est Asia and Pacific. Those regions are, obviously, the most critical region of U.S. foreign policy. They contained, according to a renowned security researcher, Chris Vickery , at least dozens of terabytes, full of spying social media files, with key word that U.S secret services use, as Coral, the most known, to identify some dangerous people, and/or networks, but also programs like the outpost program, that should influence and monitored on social media young target overseas "steer them away from terrorism". Of course, as soon as they noticed this leak, the server adds many programs to protect and recreate previous privacy. [23]

Another story is about someone who sent by mail a lot of personal and private information about reservist marines. The mistake? It was sent through the wrong mail list. It's on this way that people who shouldn't have access to those data, even civils,

received it on their mailbox. This wrong mail contained information about social security numbers, credit card information and bank routing numbers of, at least, 21 426 marines and people in relation with marine. [24]

To conclude about U.S. security, Chris Vickery and Bob Diachenko, another security researcher, found out, on the beginning of 2017, after plenty of other data breach as OPM data breach, U.S. voter's data breach, or also a data breach about national printing chain PIP Printing, who involve celebrities, that their security system I very bad, particularly with portals like website recruitment and other pages that all public can access.

JP Morgan Chase (2014)

Late in July 2014, the security team discovered that the bank is under attack, and finally success to stop it during august. After investigations, it appears that the hacking started during June 2014. Hackers, apparently by spear phishing an employee, got administrative rights, that why the system didn't react by itself. They were able to get information from on average 76 million household and 7 million small businesses. Majority of them were mobile application users. But it was information like name, address, postal code, social number etc, not the most sensitives. By the way, still according the investigation, it seems that they were no data corrupted. The interest lies in the fact that hackers had also access to the complete list that JP Morgan Chase use on their computer. This attack looks more as a spying intrusion than a theft. It can be explained due to the huge importance of this bank in world financial system. Even if basically, Russian or eastern Europeans hackers were suspected, finally it appears that it was from Israel, were four people were arrested. [28][27]

Yahoo! (2013-2014)

Data breach does not happen only to other internet entities, but also to giants. In this case, Yahoo! Was the victim of two different attacks, one in 2013, the second in 2014. Both of them were discovered in 2016, when the Verizon group was trading to buy Yahoo!. Strangely, that's the 2014's one that had been discovered or announced first. The subject of this breach was to steal user data. After investigations, the security

team came to the conclusion that it attacker(s) was state-sponsored, in the environment of cyberwar as it is described earlier in the text. The result was that on average 500 million user accounts had been hacked on this time. The stolen data were names, email addresses, telephone numbers, dates of birth, hashed passwords, and in some cases, encrypted or unencrypted security questions and answers. Investigators found that attackers made it by using a cookie-based attack. It was well enough organized that it broke the bcrypt algorithm, a defender one that had the reputation to be difficult to pass. On this way, they could have access to information of some users and maybe get right accesses to manage this huge rob. We can easily imagine that hacking such a company to get information about citizen in the world is much more easy and fast than making the database by itself.

The second attack on Yahoo!, temporally the first, was targeting the same type of data. It is not impossible that some data had been stolen two time during those attacks. Contrarily to the previous aggression, Yahoo! Didn't communicate on what broke their defence, but their first estimation was about one billion hacked user accounts. They hired a firm, InfoArmor, to help them in their investigation to discover from where it came, and who asked for it. Finally, after some research, some conclusion appeared. They found, on the dark web, that a list of one billion users was available for 300 000$ in 2015, 200 00$ in 2016, maybe after the announce of the breach. It was not a state-sponsored attack, but a third-party hacker that provides the attack. Other evidences showed that it was a command. There were three buyers, two big spammers, and a third one that had a special demand: be sure that there is at least ten names of united states and foreign government official in the list, with information about them. This buyer is suspected to work for a foreign intelligence agency. The result was that on average 150 000 people who were working for government and military were on the list, in additional to other accounts linked to officials from other countries such as Japan, Australia, Canada or even European union. Later in 2017, Yahoo! officialised that this data breach didn't affected one billion accounts, but three billion. This data breach is considered as the biggest in the internet history. [65]

1.3.2. *Hacking techniques*

Most famous hacking techniques: [31]

Phishing

The principle of this technic is to reproduce an official page (from company, banks, administrations, police etc), generally by mail, and ask for your private information as password, ID data, Bank account, credit card number etc. Once this page is sending to random people, hackers are just waiting for result. It can be more sophisticated if they reproduce an entire website.

Spear phishing

This is the main idea of phishing, but on this case, the target is not random people who will believe in your fake mail/website and give you what you want, but some specific people, with certain posts, in bank for example. The aim will be to appear like someone that the target knows, and start discussion. It required a lot of preparation, a huge time of spying, stalking all social networks, to know the maximum about the victim. At this point, it's very close to social engineering, some hackers are able to use exactly the same mail vocabulary when they're phishing. Generally, targets have a good place in a company or administration and access to interesting administrators right or sensitive information.

Key logger

Key logger is a software who'll record the way that you use your keyboard, with for example, sequences that you use letters. It saves those data in a small log file on your computer. It also can be a hardware, who, in this case, will target not only your keyboard but all the machine emissions, as electronics ones, smartphone sensors etc. They can be very dangerous, that why, generally, online banks have a virtual keyboard to offer you instead of yours.

Denial of Service (Dos\DDoS)

DoS is not a hacking technique to stole data. The aim is different, it's made for crashing down site or server. The idea is to create a lot of request, and simultaneously, send it on the target. After that, usual user cannot use the service anymore during

maintenance period. Most of the time, because only one device is not enough to generate a huge amount for the attacks, they use botnets and/or zombies computers or devices (all connected devices can be used, from the connected washing machine to the connected hoover, or pad, or even monitoring cameras) that the only mission is to send a request. They successes to attack and crash down OVH, one of the world leader host. The volume of the attack was bigger than 1Tbps (Tera bytes per second).

Waterhole attacks / Fake WAP

On both of those cases, the idea is the same. The purpose is to create a fake wireless access point, make you log on, by a fake name or just open this Wi-Fi point for free, change your most visited websites, divert your connexion, to pages that he already controls, and then then just serves himself to your data. Once the hacker has access to your data, it's very difficult to expulse him and stop the process. The most efficient ways to protect yourself is to use a good VPN service, and keep your software updated. Those technics can be efficient against one specific target or random one.

Eavesdropping

All technics aren't focus on data stealing, there is also passive hacking, such as this way. The principle is quite simple, it's to monitoring the way that you are in contact with your environment, generally through your mails, your instant messaging services, your phone calls etc. Because your system is not harmed, it will not react, that why it's one of the most discrete solutions to get information from target(s). The most famous demonstration of it efficiency come from the NSA. According to Edward Snowden, they were/are monitoring the whole world, until Angela Merkel's cellphone.

Malware

There is a plenty of different sort of Malware, all with different actions on your device. Here, I'll talk about famous ones.

Worms: This malware will not steal your data or spy your activities, it will only replicate himself to make your machine slower, your hard drive full faster, and get immobilize as much as it can any devices. It hides all their action from the user, you can have a device full of them and don't imagine it.

Viruses: They are like worms with the difference that they attack your files, they can delete it for example, and put damages on your computer. To be active, they need to be supported by a host program. Otherwise, they're ineffective. They can infest files, video, mails, links, programs etc.

Trojan: the origin of the name is directly inspired by "trojan horse", this subterfuge that Greek people used to enter in the Troy city. The idea is similar, it will not attack anything, but create a backdoor that the user will not hear about, to allowed the creator to get access to all your data, include the most sensitives.

Adware: This malware will not cause any damages to you, and generally is coming when you download and install free-to-use software. It will only display advertising, but it can be interpreted as a signal of a breach in your security system.

Spyware: The principle is quite similar to the adware, at the difference that it will send to the remote user your browser habits. Also, it can facilitate installation of unwanted program. They are operating like worms or trojan, the user doesn't have to notice their activities.

Bot: It's a process designed to automatically work without human hands. It can interact through internet. Good or bad, it can be considered as a worm evolution, with abilities to steal passwords, log keystrokes, analyse network traffic, relay spam, launch DoS attack, or open backdoors on infected hosts. The main danger is that it can infest many machines, take information on all of them, be a tool to coordinate massive attacks, and be controlled by only one user. Create this kind of program need a lot of time and a hard work. The potential of this system is huge, as we saw previously, it can be used to infect monitoring camera and other electronic devices, devices that can hardly be protect.

Ransomware: This malware encrypts devices data with the aim to block partially or completely the using. It generally shows only a message that may explain the purpose of this ransomware, but also and mainly the amount of money that pirates want to free the machine. If the ransom is not paid, they have the possibilities to definitively delete all data and "clean" the machine from any files as there.

Clickjacking attacks

The clickjacking is a common technique, that happen while an app is downloading, before watching a streaming movie or on torrent websites. The hacker tries to make the user click on a hidden link on an advertisement or because a page opens suddenly, to drive him to a page where the hacker wants. Most of the time, this technique is used for unwanted advertising, but can also be used to steal personal data.

Cookie theft

Cookies contain your navigation's data, like browsing history, username, passwords for different websites etc. they are files who are store in your computer by the website owner, file with information that they can have access when you connect on internet. The problem if the hacker stole those data? He can use your digital ID and log in as yourself. By this way, he can have access to your personal data. The best way to be protected from this kind of attack is to the less that you can SSL protocol (HTTPS).

Bait and Switch

It's a technique based on a investment: on this example, the hacker buy advertising space on websites. He makes it attractive with the aim to get a click from visitors. The link brings the user to a page who's infected by malwares. On this way, the owner of the advertising space can install malware and adware on the machine. As he wants, the hacker can download and use specific program, who looks authentic to the casual user (so he will give his agreement to install it), to get (unprivileged or privileged) access to your computer, and do what he want. This technique is the most malicious cause the user has less chances to guess what happen than other techniques. The best way to avoid these risks is to refuse strange program that you don't know, who ask to be installed, specially those who ask for administrators right.

Most of the threats can be avoid by just prudency in your behaviour, and at least a good antivirus. Those threats or the more famous and don't target someone specifically (except the DDoS attack), so they aren't the most dangerous and the most difficult to defend against, even if when they infect your computer, it can be difficult to delete it. If you have to protect sensitive data (from your company, government or just your

personal bank data (I strongly advised against this idea, it's literally a lot of risk for no benefits)), you'll need other protection against accurate attacks.

But there is also new trend in common threat:

Cryptojacking

To be cryptojacked, you just need to surf on internet, maybe you are presently. It is not a threat that will still your data, crash down your system or something else with a potential direct impact. It is a new way that website use to make money, instead of advertising and / or storing data. Basically, a software opens the gates of your devices and use the power of the processer to create crypto money (such as bitcoin for example). BY creating crypto value through this way, it engages less costs, and remuneration is good. That's why many websites turn their attention to this solution. Their creation need a lot of energize and can be split between many devices. For instance, the only action will be to make your machine running slower than it should. So, the consequence is that it will annoy the user. To noticed that the computer is used by a third person, user would have open a code monitor that check continually which code is used.

Stenography

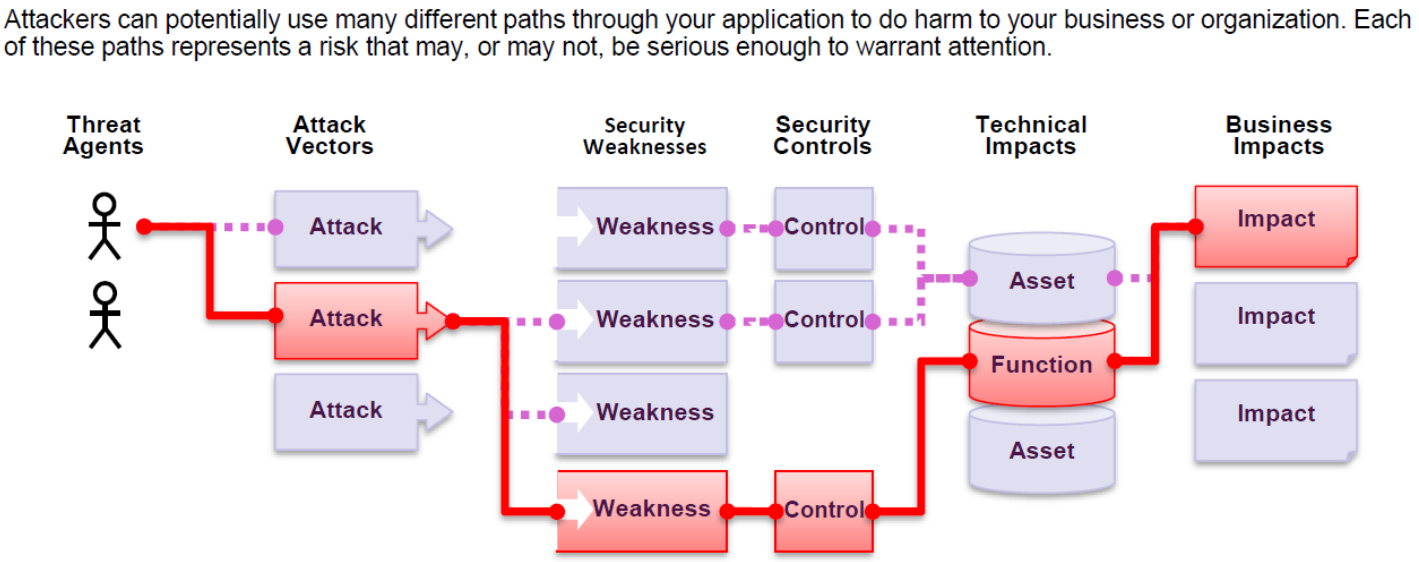
It is not a new attack, it is quite old, but the appearance of the use of meme and the habit of sharing pictures makes it dangerous and offer a second youth to this malware shape. Effectively, those malwares are hiding in images. The danger exists since decades but is really to consider since only few years. Such as first virus, it is hidden in an attachment file. Just open it make the malware enter in your computer. Classical anti-virus can protect the user against it if their analysis is complete and deep enough.

1.3.3. *Application security*

Previous threats concern web pages and computer. But as it is said upper, data are not only on websites, or stored in hard drive. The access to databases are not only through computer, but also from applications. Those objects have their own threats, different from computer's one, but they are not less dangerous. Because they have more

user, have to be easier to use and use the data exchange in their processes, their architecture can be weaker, and security might be not as worked and strong is it should.

A non-lucrative organisation, Open Web Application Security Project (OWASP), published every four years a document that classify and warn about what they call “top 10 most critical web application security risks”. It does not sum up different weakness that an application can have and which danger there is because of these weaknesses only, but also explain which evolutions there is since pasts years and try to educate people in charge of application development (developers, architects, managers, designers, etc.) about risks that their product can be victim, as the following scheme will show.



By publishing such a public guide (2), OWASP allows people that want to create and develop their own application to be aware about the potential dangers and risks that they will meet. Also, their publication can show which trend application security sector is following, and with which evolution.

The dangers that can aggress an application and are the most common / dangerous for the data security are the following: [57]

1) Injection

The hacker tries to inject some code in the existing code to trick the interpreter by sending untrusted data as a part of a query or a command. So, this one will read those new commands and execute unwanted command. On this way, the attacker can get some accesses that he should not.

To prevent it, it is advised to separate data from command and queries.

2) Broken authentication

For incorrectly implemented applications, some attackers are able to comprise passwords, keys, sessions tokens etc. It makes them able to impersonate some users, possibly permanently.

To prevent it, it is advised to implement a multi-factor authentication, to do not ship and deploy any default credentials, do not implement weak passwords check, neither align passwords length, complexity and rotation policy, ensure registration, limit failed attempt etc.

3) Sensitive data exposure

Some application does not protect well some sensitive data. They do not implement extra security, such as encryption at rest or in transit, tokenization, etc. The pirate can get possession, use or modify this kind of weakly protected data. It gives to him the possibility to theft identity, conduct a fraud card, or other actions, depending of the data type.

To prevent it, it is advised to classify data, apply controls as per the classification, don’t store sensitive data unnecessarily, encrypt it at rest, ensure the encryption settings, encrypt data traffic with secure protocols, disable caching for responses that contain sensitive data, store passwords using strong adaptative and salted hashing functions, verify independently the effectiveness of the different implementations.

4) XML external entities (XXE)

XML processors that are old or badly configured XML evaluate external entity references within XML documents. By doing it, the XML processors will allow some programs to be executed, have an action that they should not be able to have. This function can be used to make some action, that be piloting to open private files and open the gates for a hacking action.

To prevent it, it is advised to have trained developers, use less complex data format, patching or upgraded all XML processors, implement a lot of controls about use of XML functions.

5) Broken access control

If the rights policy is not well build or well implemented, some users can have access to functions that they should not be able to use, such as admin function, and so they might see, modify, delete data, change other accesses rights, edit the content etc. because user restrictions, according to his identity, are not enforced enough. It is a basic principle of data that must be respected, whatever the device.

To prevent it, it is advised to enforce the access control, through a trusted server-side code, or server-less API.

6) Security misconfiguration

This is the issue the frequently meted. It can come from bad configurations (by default, incomplete, ad hoc etc.) for example, but also from other sources such as open cloud storage, misconfigured HTTP headers, verbose error messages containing sensitive information etc. A lot of files have to be firstly well configured and secured, and they have also to be patched and upgraded in a timely fashion.

To prevent it, it is advised to implement secure installation processes by including, a repeatable hardening process that makes it fast and easy to deploy another environment that is properly locked down, a minimal platform without any unnecessary features, a segmented application architecture, to ensure secure separation between different objects, an automated process to ensure the solidity of the configuration.

7) Cross-site scripting (XSS)

When this kind of breach happen on a website, the attacker can modify some part of the script, in the code, to change some actions on the website.

For example, if there is a field that ask for your IDs, the attacker can change the place where that information is stored. With this way, it can collect many data from many different user, without any signals, except for the owner of the platform.

To prevent it, it is needed to separate untrusted data and active browser content. It is possible by the use of a framework that automatically escape XSS by design for example.

8) Insecure deserialization

When a deserialization is executed badly, it can implement a lot of trouble in the code, in the execution of programs, can create some breaches for attacks. This is kind of information corruption.

To prevent it, it is advised to implement an architectural pattern that does not accept serialized objects from untrusted sources or use serialization mediums that only permit primitive data types.

9) Using component with known vulnerabilities

If a component (such as libraries, frameworks, software module generally) that vulnerabilities or weakness or known from everybody, so it will be a target for pirates that want to attack the system. Moreover, it will be difficult to defend because it is known, so exploited. It open gates for a huge variety of dangers.

To prevent it, it is advised to use a patch management process to improve the component. The most efficient way is to be sure that there is an ongoing plan for monitoring, triaging and applying different updates.

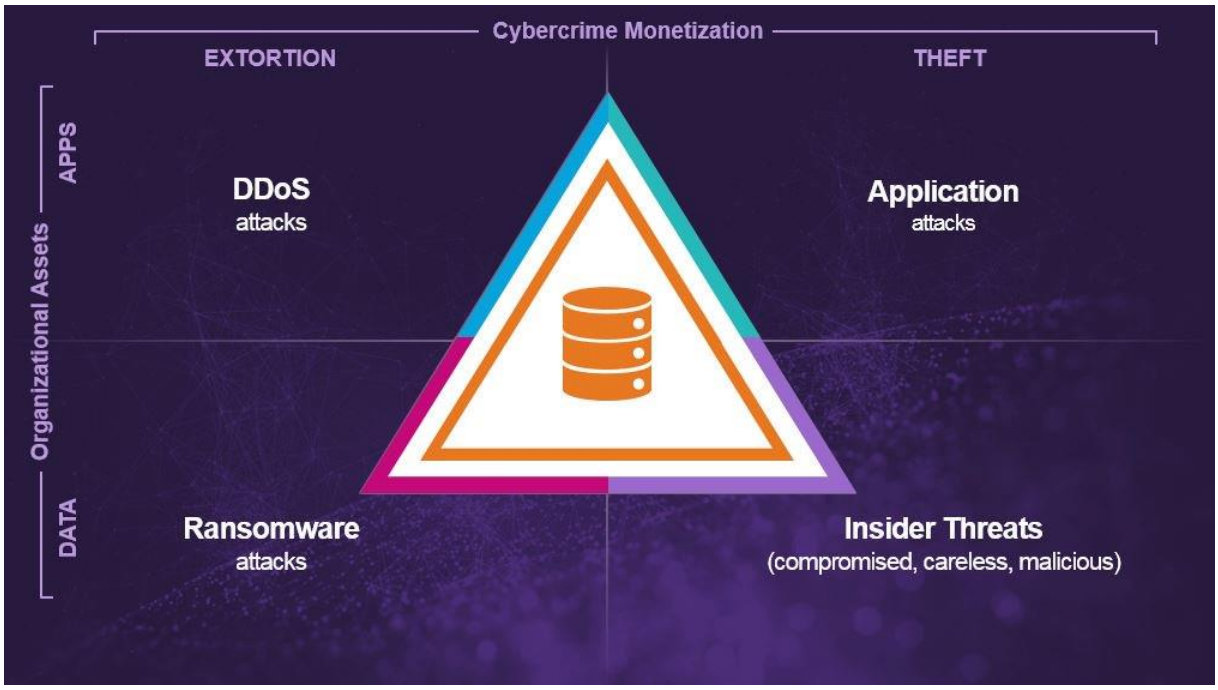
10) Insufficient logging and monitoring

Don’t log and monitor enough can create a lot of danger, especially during a strong a well prepared. Do not have enough back log can block the owner of the application and the users to use again the functions and can stop

every exploitation. Do not monitor enough can drive to spying actions against the application. Such as JP Morgan data leaks, it is possible to be under an attack during months without having any information or clue about it.

1.3.4. *Current threats classification*

When a company is directly targeted, there still only few methods to drive the aggression. Attacks are divided in two categories: the theft and the extortion. But in these two ideas, possibilities are huge and without limits, except creativity. The following scheme show how targets and way of attack are connected and related according to the way that the hacker wants to gets the wanted data. In each solution, there is plenty shape to adopt.



Hopefully, there exist efficient way to protect data. If nowadays data are the favourite target, and have access to them is one of the main question, even if it still complicated, very complicated in certain cases, there is another problematic, how translate and understand those data. There is another level to protect them, it's to make them unreadable. There is also a technique that consist to hide them. And still some others.

But there are also new trends in the common threat field. Due to new technologies, and new opportunities, threats also evolved. Those two new ones can show you new dangers (strong or not) that can target your computer.

6) Application threats

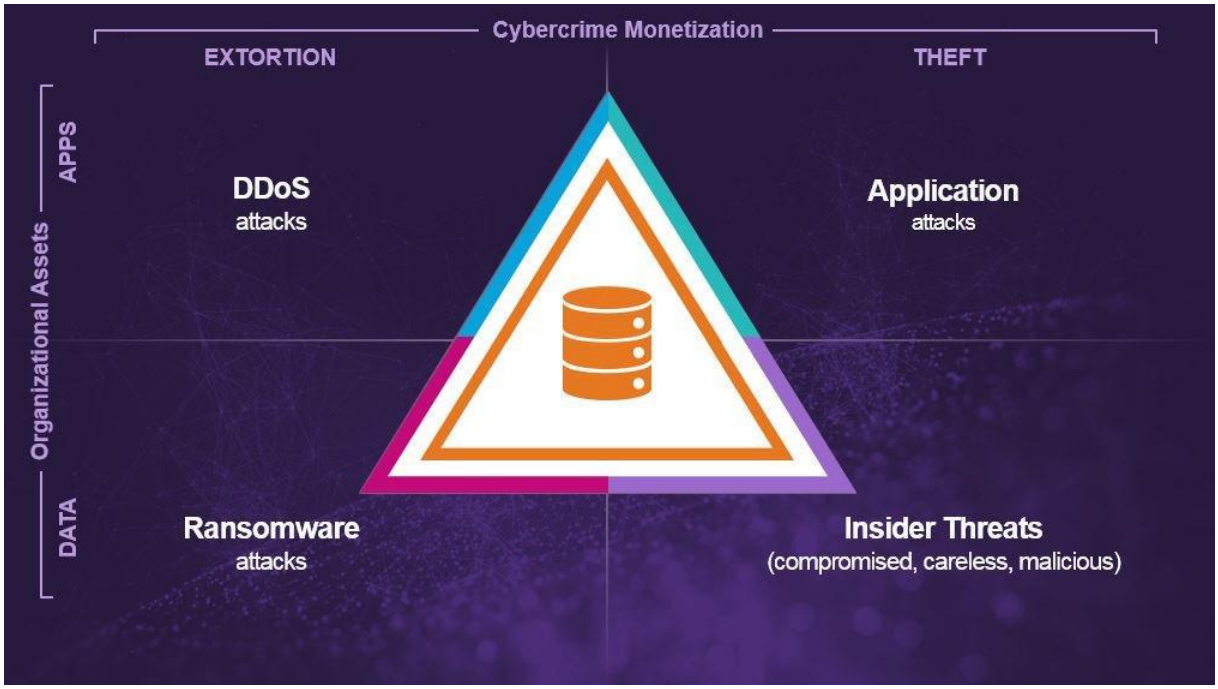
To prevent it, it is advised to ensure all login, access control failures, server-side input. Also, it is recommended to ensure that logs that are generated in an appropriate format (one that can be easily consumed by a centralized log management solution for example), high value transactions have an audit trail. There is a need to implement an effective monitoring (that alert when suspicious activities are detected, with a respond in a timely fashion. Also, establish a plan to have some responses and a recovery solution, to be sure to do not loose important data.

Those different risk have different level of dangerousness. Owasp made a table that sum up them with a rate of the different damages they can cause, their difficulty to exploit, their prevalence and their detectability. The classification takes the shape of a table, each risk getting a score for each factor, with a final one.

RISK	Threat Agents	Attack Vectors: Exploitability	Security Weakness: Prevalence	Security Weakness: Detectability	Impacts: Technical	Impacts: Business	Score
A1:2017-Injection	App Specific	EASY: 3	COMMON: 2	EASY: 3	SEVERE: 3	App Specific	8.0
A2:2017-Authentication	App Specific	EASY: 3	COMMON: 2	AVERAGE: 2	SEVERE: 3	App Specific	7.0
A3:2017-Sens. Data Exposure	App Specific	AVERAGE: 2	WIDESPREAD: 3	AVERAGE: 2	SEVERE: 3	App Specific	7.0
A4:2017-XML External Entities (XXE)	App Specific	AVERAGE: 2	COMMON: 2	EASY: 3	SEVERE: 3	App Specific	7.0
A5:2017-Broken Access Control	App Specific	AVERAGE: 2	COMMON: 2	AVERAGE: 2	SEVERE: 3	App Specific	6.0
A6:2017-Security Misconfiguration	App Specific	EASY: 3	WIDESPREAD: 3	EASY: 3	MODERATE: 2	App Specific	6.0
A7:2017-Cross-Site Scripting (XSS)	App Specific	EASY: 3	WIDESPREAD: 3	EASY: 3	MODERATE: 2	App Specific	6.0
A8:2017-Insecure Deserialization	App Specific	DIFFICULT: 1	COMMON: 2	AVERAGE: 2	SEVERE: 3	App Specific	5.0
A9:2017-Vulnerable Components	App Specific	AVERAGE: 2	WIDESPREAD: 3	AVERAGE: 2	MODERATE: 2	App Specific	4.7
A10:2017-Insufficient Logging&Monitoring	App Specific	AVERAGE: 2	WIDESPREAD: 3	DIFFICULT: 1	MODERATE: 2	App Specific	4.0

7) Present trends

When a company is directly targeted, there still only few methods to conduct the aggression. Attacks are divided in two categories: the theft and the extortion. But in these two ideas, possibilities are huge and without limits, except creativity, as we saw previously with the different threats and risk that platforms are supporting. The following scheme show how targets and way of attack are connected and related according to the way that the hacker wants to gets the wanted data. In each solution, there is plenty shape to adopt.



If the users need to protect very sensitive data (from your company, government or just user personal bank data (I strongly advised against this idea, it's literally a lot of risk for no benefits)), he will need other protections against accurate attacks. The first decision to implement is to create a department dedicated to do it. Even if there is a lot of companies that built their business on IT systems and security assistance, and even if they have a true knowledge, it is necessary to have someone totally dedicated to these problematics.

The actual way to store data in external server and get access to them through cloud, being popularised by some providers, such as amazon, but also OVH, or 1&1 for example, is the most common one. Those structures implement security of their

servers by recruiting the best developers they can. Their main threat is the attack by DDOS, because they cannot continue to provide their services if it happens. But for other part of connections and use of data, it is up to companies to secure their network by themselves. Basically, they can get the access of basics function of SSL protocol, but also have more settings by buying more securities. The sellers of it are server providers. Amazon or 1&1 have different offers about the data security, the treatment data, the way the secure the different transfers etc. It is up to the department in charge of the company data security policy to determine it and define what they want by discussions with the different actors. Today, a company need to call outside actors to implement the most complete security system fitting to their organisation.

2. Analysis of the Object: Hesus Company

2.1. Description of Hesus company

2.1.1. *Internal environment*

Hesus is a company with a construction site soils specialisation. Their main activity is to orientate soils to the appropriate center according to pollution rates in the ground and transport price. The idea behind this action is to reduce the waste and offer more ecological solutions. They are in relation between jobsites, transporters and centers to ensure the best offer and travel to those material. They are considered as waste, even if most of the time, this consideration is wrong. Effectively, many tons of grounds were directing to the wrong center with the wrong treatment. Generally, it was because the knowledge about retreatment network was far from efficient and the easiest solution was chosen to allow the building activity starting in time (construction site have commonly a lot of late, so executers try to save time as often and as much as they can). By putting in contact different actors from the same activity, Hesus was able to propose new solutions, offer to the companies to save some money, make the ground recycling in the best way, and sometimes make jobsites making money by connecting them who have corresponding demand and offer. They are the only one in France with this trade. So, they have a unique expertise on this field. According to this competitive advantage, even if some construction site directors have their own solutions, they are able to purchase a complete and independent analyse. The offer is not only interpretation of laboratories analysis and waste orientation, they also have enough knowledge of laws (in France, waste treatment regulation is really complex, and change often), ground recycling market and recycling center network (not only in France but also in Belgium and Netherland, countries with a different skills level in this sector) to propose to the construction site decision maker a plan to care about the different soils from the excavation to the final treatment in the appropriate structure. Hesus does not only ensure the chosen center, but also the transport of different type of products from the jobsite to the recycling, or storage center. This second activity is also a strength because it makes all the process much easy for the jobsite header and bring him some serenity because he does not have to manage this part, except asking for how may trucks he needs ach days to evacuate his construction site. The final aim is to create a closed circuit from the excavation of the materials and/or soils to the reuse after, until the recycling step and his second life.

The company is young, has been created ten years ago, in 2008. The creators are two brothers, Pierre and Emmanuel CAZENEUVE. Only the second one still at the head of the society. It grows up since the beginning, until making turnover calculating in million of euros today. The company structure evolved to be, currently composed by one commercial service, on studies office service, one supply chain service, one transport service, one product service, one financial and HR service, one innovation service and one Information technology service.

After ten years of existence, Hesus created partnership with very big companies and s recognize for his expertise. The uniqueness of his mission imply that they could not copy other organisations. They developed an operation way based on the competences of his employees to find the best solutions and present options that answer the best to their client needs. According to their experience, they also developed a unique network center that help them to find different and more efficient solutions to issues that primary look obvious with one solution.

Currently, Hesus looks like continuing his growth by creating some partnership with other companies to develop solutions over offer. Also, they integrated a start-up company, named pick my waste, to complete their activities directly connected to the jobsites, and enrich the offer to be the only interlocutor that construction site decision maker has to meet to treat his wastes.

2.1.2. *External environment*

Hesus is located in Paris area, but has activities in all the country. The activity is a part of the construction sector. Their partners are in majority very big construction companies. But not only. They are also in business with many different transporter, due to the surface they need to cover. Jobsites with whom they are working can be commanded by the same company that has the property of the center. This position, the only one in France, is quite uncomfortable. On some point, the partners are also the competitors. The capital gain has to be enough important to convince jobsite decision maker to choose the proposed solution than his usual centers. On the other hand, transporters can also be in competition with some centers. Some of them can propose the transport option in their offer. The Hesus advantage in this position is to offer time and

save money to jobsite. With this opportunity, they can focus on different point and reorient their attention on other obligations. It is some comfort that is proposed.

One Hesus problematic is the legislation. Because the activity is in relation with wastes (polluted soils are considered like that), the legislation is very strict, there is a lot of laws to respect. The waste treatment is submitted to very accurate rules, according to the present type of pollution. There is security measure to ensure for transportation, different for each type of pollution. Not every transporter is able to ensure it, also not every centers are able to take in charge to product. There are some documents that are proving the effectiveness of the transport in good conditions and the delivery. If the dedicated authority asks for it, they must be able to show them to prove that they respected the law. If they cannot purchase it, they can be condemned to pay penalties. The advantage of this strict legislation, in addition to be complicated so knowing it is a true skill, they will force the jobsites to recycle 70% of their waste in 2020. Soils are considered as jobsite wastes. So, the work supervisors have to find a valuable solution to reach this rate if they do not want to be out of the law. This measure is one of the few that is stable (for instance). Others, such as taxes can change every year, according to politics decisions and needs. They can take the decision to create new ones, to change the value of them, to make it less constraining for construction actors, or change their mind and change the ecological policy of the government. Every center based their pollution rates on government official threshold. They make the national difference between polluted soil and non-polluted soil. To define the difference between different pollution level, it is the different prefectures that decide which pollution rates are in their region. So, every region has different rates of pollution and acceptation level. A ground can be considered as very polluted and toxic one area, but on the other side of the border, it is counted as a normal waste. There is the same problem with foreign country rules and uses. Hesus is working with centers in Belgium. Their center recycling skills are far better than the French ones. But the legislation is different, and centers are looking closer certain information that French ones do not consider that important.

The fact they integrated the start-up “pick my waste” will not make it easier. In addition to have to manage polluted soils laws, they will manage other kind of wastes. They will be more various, less common, potentially more dangerous.

The good side of this action is that they will be able to help the jobsite during a bigger time. Also, some stuffs considered as waste have value. They can be sell to center. The best example is the copper (even if it is very rare because everybody knows this material has value). But, those product does not have a stable value. It is dependant of the markets, the balance between the offer and the demand. Because the company margin is basically similar on every offer, the mission to find the best supplier offer is central. So, an important side of the mission is to stay aware of market evolutions and keep good relations with centers directors. Some external environment can have important repercussions on channels. For example, currently in France, the wooden one is full because of this winter, wood has not been sold, so between wooden factories and recycling wooden factories, there is a huge wood offer, that make difficult for jobsites to recycle them by selling.

Polluted soils and materials are more taxed than non-polluted soils and materials. The difference is very important. The solution to this problem is to depollute the soils to down the price. The obstacle of this method is the technology, and his cost when it exists. There are bio solutions not very expensive, but the default is that the treatment time is very long, some the economizes come very late. This is the reason why this type of solutions is not the first chosen by jobsites. Belgium and Netherlands centers have a more advanced technology and a better one. According to this point, their prices are inferior to French prices. Moreover, because their return is faster, their performance is higher. It gives them the possibility to enter on French market with a very strong position. Their technology is more developed, so they have more opportunities. When a soil is considered as very polluted in France, it is only considered as polluted abroad, so the treatment price is sensitively different.

Hesus use all those settings to find the best offer to his client and put in competition the different centers. To continue to grow, they decided to implement some IT system to become more efficient and spend less time on unproductive tasks.

2.2. Company's IT System

Today, there is an existing platform. It has been developed previously by another team. It name is Hesus-store.com. Data, as a lot of different websites and platforms, are stored in a database that is hosted by amazon servers. This is one of the most convenient and suitable solution for this type of entity. The project being at the beginning, this choice of host fit well, because it permits the platform to be created, and then developed more and more. We can implement the functions that we want, and it leave us an interesting grow potential. The server structure will continue to follow our evolution and ensuring operations and access to database while we will implement changes. Amazon has the reputation to provide good development environment, while it also ensures protection against outside attacks. They also give us the possibility to adjust our security policy to our needs, mainly by the upgrade of the SSL protocol that we are benefiting.

They also provide us a virtual database, that is paired with the development version of the platform. With this option, we can analyse every implementation, every new features, and watch all the implications in the process, to understand how it is running and which setting do we have to change to improve the use of the platform. It also allows new person that are working on it to get use of it faster.

Currently, the platform fulfils some basic functions. On the user side, after creating a jobsite, this one can make a demand for excavations or backfilling of different type of materials. It can be done with polluted soils (according to the rate of pollution to redirect it to an appropriate center to give it the appropriate treatment), non-polluted soils (to be stored and reused on the future and be sure that those soils are not mixed with polluted soil and so do not waste ground) or some materials (such as specific kind of soils, or some waste such as rubbles, pebbles etc.). There is an algorithm that, according to attached documents, and soil/material properties, will determine which centers are compatible and which one is the cheapest for the user by treatment price and transport price. The strength and the other advantage of this platform is that it also proposed to put in relation two jobsites, one that have an excavation different than polluted soils, and the other that have a backfilling to do. The algorithm compares settings to be sure that construction site demands are compatible according to dates and

quantities. The user has to create an account to connect and use the functions. Additionally, the user can also decide to accept or refuse the result of his demand, but also, once the jobsite is closed, he can have access to a service that show him the performance of the different services he used on the platform, under the figure shape with statistic such as number of trucks needed theoretically and number of truck needed effectively, or how many percent of his soils had been retreated (a French law that will enter in application in 2020 will force jobsites to retreat at least 70% of their waste, soils are considering as wastes.)

About the partner side, the platform had been built with the idea to propose them to readjust an offer before it is proposed to the client. When a demand is created and validated by Hesus, the algorithm chooses the most suitable center for each partner, according to specificities of the demand. It displays it on their screen, and propose them to make this proposition directly, to modify it if they want, or to refuse the case. Once they answered, the file still displayed in another tab that separate non-answered request than answered requests. Partners can also manage their centers. They can create a new one, or edit an existing one, by change basics information (location, contacts, accepted quantity, catchment area), and adjust pollution rates accepted or material type received. They can also delete a center if they want.

The platform admin are members of Hesus. They have the rights to reject any demand that has not validity, according to different analyses purchased by the demand creator, the jobsite owner. The admin can also follow all the process of the algorithm, see which option were chosen and which not, and why. He also has access to every centers, jobsites, demands and answers of demands. Moreover, he can create new user, on user platform (those that create jobsite and make demands for excavation and backfilling), and on partner platform. He can give some restricted rights, or at the opposite, can create new admin with full accesses. The admin has also another role, it is him that define the role of the user. When a new one registrate by the website, he can be a classic user (director of a jobsite) but also the representant of a company, or a holding. Those distinctions are used to define different rights according to the role.

The platform fulfils also another function, it makes in relation consumer and other Hesus partners. As it is write before, Hesus has some partnerships with companies that are not competitors, but that propose offer that are compatible with hesus activities. The idea of those agreements is to create win-win relations, to make a network of companies able to purchase the longest services chain. Hesus-store.com host links that bring to those partners, so do partner platforms host hesus-store links.

Currently, Hesus is also working with a Costumer Relation Management (CRM) that manage all interaction information between clients and salesmen. Every client information is stored in this CRM. Some of them are imported from hesus-store.com, but most of the data are enter directly through this software. It runs in pair with a cloud to store different files. Salesmen, through this tool, can have a fast access to client information, the state of the relation, the project, on which step is it, who is involved in it, which relation have different actors between them etc. As for a lot of companies, it became necessary to implement this kind of solution to increase the efficiency and reduce the administrative time spent. Also, it facilitates the arrival of new comers, they can have access to the needed information in an easier way if they have the appropriate rights. They do not have to ask to others to be aware of the different cases. This solution also encouraged the information circulation. As an IT tool, it helps for as important part of the management, the knowledge management. Because the information is circulating, and because the person that need it have access to it, the general knowledge is spread in the company and so, it is not the property of one individual. The betterment I high for all the enterprise. In this case, the departure of one person will not cause the loose of his knowledge and his information. The data sharing principle in the company provide against this kind of risks.

Another IT tools is ongoing installation. Because there is not only salesman that can see their function optimized, but also other departments, an ERP is planned to be installed. It will interact mainly with financial department and supply chain, to reduce the number of paper documents, and increase the speed of exchanges, limit the risk of errors, and make situations clearer. This ERP will be linked with the actual CRM, but

not with hesus-store.com to connect all services faster and more efficiently. The betterment of this implementation will be the number of case that each team can manage, and the clearness of those ones.

Presently, Hesus split his data usage between two systems, the CRM and the hesus-store.com platform. Both have their utility, in different way even if they have synergies, by exploiting common data with different aims. A third one will be implemented, the ERP, that will use different data, but will be linked to the same cloud as the CRM is. So, it means that there will be three different sources that will provide data. The difficulty will be to be sure the different data streams are fulfilling their purpose, without bogues, and be under appropriate shape to be used by different users. Make those systems be linked, communicate together, without wrong interferences, or data duplicates will be the main challenge, and also the principle interest.

2.3. Challenges and potential evolution of IT tools

The IT pole is focused on the development of an internet platform that will recreate and make faster all the different processes that are working separately. The deep idea is to change the actual general process to a new one who will make the working team able to manage more files, in a more efficient way. The user will be able to enter his jobsite information, add the needed files for analysis, and then will automatically have an optimized answer to his demand with the more appropriate center. This center will have to ensure two important conditions: it must be able to treat soils and material send by the jobsite, and it will be the cheaper comparing to the other competitors. The model is tinder operation. The system is simple, accessible, and have to be fast. The second idea is to add new activities, not only stay on construction site soils sector, but treat other construction site wastes, and continue the grow by developing a general solution to be able to offer to their partners, at the end, one global offer for all their wastes. It will be facilitating their job and ensure a good way to reuse or recycle it. It is a win-win relation.

To develop such a tool, an efficient platform is necessary. It represents a huge IT system connecting many different services, which implies the same amount of different job, all of them related to the same database. The first challenge is to connect all of them, with the same platform shape, even if they have various needs. So, developing

different functions is the first step, with the question of the different rights, access, who need which access, to which information. Those questions are answering by understanding the job processes. To do it, there is a fundamental requirement: understand all different services processes, very accurately, step by step. Once this part is complete, another side of this IT system building is to make the architecture of the platform, defining which part is connected with which one, for which goal. The term architecture is appropriate because it's a complex construction, due to all the different interferences that functions can have together. Also, to execute an algorithm, it goes to use data in the database, data that can be modify by another part of the same algorithm. Everything is in relation, and it has to result as a very intuitive and efficient model that can follow future evolutions of the company. According to those conditions, the second challenge is to distribute the right information on the good time to the right person. The one who will need and use the information can be the construction site decision maker, the partner centers, or the Hesus team. There is lot of different type of information to deliver, to different users, with different sensitivity. The security policy has to be precise and don't forget any cases. The third challenge is to make this platform efficient and well-constructed enough to evolve with the company and follows it strategic goals. The next ambitions are to propose news services, with new partners, for the same jobsite, and maybe try to work not only on French territory but also on foreign country's market. The platform has to fit with those hypothetical evolution, without needing to be change, or redesigned.

Developing such a platform is very interesting, because it interweaves many questions. Between the interface that new user has to be able to understand, to use, also about the different functions to implement to make it convenient also for Hesus team (one of the aim of the platform is to make their job easier, to help them to be more effective), but also anticipate future strategic developments, and those actions have to be doable through a database well-built, offering huge navigation and new service implementations possibilities. Obviously, because this platform is going to grow, to have more and more users, and the structure will be more and more complex, the security questions will have to be treated efficiently. The main question for instance is how the

right distribution has to be done. Which information have to be available, and which one do not have to. Because it is not a technological company, and because financial data are not linked to this platform, the most sensitive information are not candidate to be subject to strong attacks.

Another part of the mission will be to merge the platform with in CRM in place. The main challenge with it will be the data circulation between the entities, especially to do not lose any of them during exchanges, but also do not corrupt information contained, and finally, send the data from one place in the platform to the equivalent place in the CRM. This action will allow much more functions to be implemented. The other challenge will be to do not lose any data during the merge, and defining how exchanges are running, following which process. It is important because the time that it can make win can be huge, permitting salesmen to be more efficient on their task, by having faster the information about the client and the project they need.

2.4. Current implementation of data security

2.4.1. *Database security*

The data security is ensured by amazon services. Since their accident with the US army during one of their update (as it is explained in the "data breaches" part), amazon pay much more attention to their security policy, and has improve it. For database connection, there is the possibility to create user IDs that have accesses, give them some rights, and defines conditions. A user ID can be defined to be available for a certain period or without limitation, to be sure that connections possibility still limited. For Hesus case, accesses are parameterized by developers. They manage to give rights to users or not, and the conditions of those rights.

After that we get the accesses rights to reach the database, we have to be sure that connect on it is safe, and also use data from it to the platform is without risks. On this point, amazon services ensure the different connections. They are encrypted following the SSL protocol (an internet security protocol that secure server access, confidentiality and the fact that data still uncorrupted, it is easy to use it with HTTP protocol). This SSL protocol is used by most of the host providers to implement the security. Every requests and exchanges are treated and secured on the same way, following the

basic requirements that it implies. Those settings are the less protective, according to the free amazon formula. Even if it is the lowest level of protection that the host proposed, it still safe enough to allow a use without trouble, since there are not strong attacks targeting accuracy the Hesus dedicated server. It is possible to have better versions of this protocol, more secured, with different options, with more insurances, but it is another offer with different costs.

Contrary to other data, passwords used for different identification on the platform are more encrypted. Considering that the most common attacks are by stealing user IDs, the password benefits of two different level of protection. The first one is the same as other data, the SSL protocol encrypts the plaintext with a unique key, to allow the use of this one and the data transfers that are concerning this information, that only the server has the same key, so only it can decrypt the object of the encryption. But also, when an admin is in the database and he is in the passwords table, the passwords are unreadable, they are displayed already encrypted. Indeed, passwords are submitted to a second encryption. What is displayed on the database is the ciphertext of the passwords. It prevents threats that can come from different sources, such as stolen accounts, inside threat, eavesdropping etc. and this measure reduces the potential leak damages. Without the key to decrypt it, you can have every hashed code in the world, it will be very difficult to restore the original information, especially if there is a double encryption that was applied on it.

To protect the server, another possibility, instead of implementing the classic IDs connection with logs-in and passwords, is to define the way to access differently. A solution can be, additionally to the first option, to block every IP address to go on, except the authorized user one. In other words, it is to define the connection rights according to the IP address. It means that only few of them can have access to the database, even if log-in and passwords are known and used from another device. It is considered as safer because it prevents against a lot of threats that or focusing on user IDs to get rights and infiltrated the targeted network. As it is explained during the previous part, most of the danger and the threats tries to get log-in information to be able to use another account. By this filter way, not only the information about identity are important, but also the machine used, that can be identify by only one item, an item

that is much more difficult to faked. Moreover, a connection attempt from another device can give the alert that someone stole or at least copied data to get connection rights and allowed the IT department to take appropriate measures and start an investigation.

Another threat that amazon servers are protected against is the DDoS attack. Because it is possible to have unexpected huge number of requests during a short period, there is a mechanism to prevent this case. The principle is to adapt the response capacity of the server to the request demand by allocating more calculating power to the session, in a temporally way. It is called auto-scaling. The server capacity is automatically adjusted to be able to not crash down when the traffic increase dramatically. This defending solution is possible because amazon has a lot of different server, with huge storage capacities, and most of the time, the dedicated space allocated to a platform or a website is bigger that they need. So, there is in emergency free space to use, in case of attack. It is one of the security measure that a host provider has to implement to guarantee so performance to his users.

There is also another possibility to provide this danger. When the user rent servers to a very implanted provider, he can ask to have a different structure than the simple database. This idea is to implement multi-database. This solution is design to be convenient to very big structures that have a lot of account and data to protect. The design will be different from the simple database model. Instead of having one database for every partners, every client, every user, the cunning point is to create a different database per partner or user. With this shape, it will split the number of connection per server, so it will help to decrease the request number per database and make the request amount easier to manage. This technique, coupled with the auto-scale, should make the DDoS attacks less effective. But with the proliferation of new connected objects, it more and more potential request senders. To be protected against the bigger attacks, there is just one solution, you at Hesus, we have implemented this one almost from the beginning of the platform. The solution looks very easy but is not that obvious at the beginning.

Fortunately, our database is not subject to this kind of aggression. According to the structure and the main idea of the service, the connection is not free, and the access to other pages is not available for every internet user. For each user, unlike some other

platform, the only way to access to the services is to connect on a session. It is needed before get access to the functions. This authentication part as the same role as a firewall.

If the attacker successes to get accesses to the servers and the database, he could access to every data, from the more useless to the most sensitive. He can see them, copy them, make exports, but also modify it, corrupt it, or even delete it. And, as we saw in JP Morgan case, this kind of pirate can stay hidden during weeks, or months, or maybe even more.

So, the question of the traffic monitoring has some value. Currently, there is only the developers monitoring that is applied. They have the charge to verify if the connections are usual, if IPs are looking normal, if they know it, or at least if there is a reason because this new IP is connected (new comer for example). But if there is no reason to an unnormal connection, no news on this way from the IT department, no valuable explanation, developers are those who have and the only ones who can react to this anomaly.

Because the structure is not that big, they do not have developed tools to fulfil their mission. There is not monitoring system implemented, to help them to keep an eye on all the connections, and every action made during those connections. They neither can separate activities per blocks. Because the database is one structure, they cannot limit access to certain zones, certain tables. To forbid access from a part of the server but allow it to another one, it required more means, and a different organization. Some systems make able the database administrators to separate the rights, to prevent the fact that a hacker enter in the system. In this case, his action power is very limited, as well as the damages he can do. Pass from one zone to the other is very difficult. It needs to have some rights that are generally very well protected, and very safely stored.

The only way that our developers are able stop any intrusion is to suspend rights of the user, temporally or not, to make this gate close for a next attack. Also, to be sure to do not be attacked by the same user, or at least by the same IP address, this one can be blocked, also temporally or definitively, to be sure that the user on the other side will not try again from the same point. The problem with this solution is that it does not prevent any attack. It is just in reaction of an event. It needs to wait that something

happened to start to be used. Another problem is that if the attacker is, at least, a little bit experimented, he will be able to connect through another IP address. So, if breaches are not repaired during this time, he will be able to reconduct his attack again.

Fortunately, in addition to do not be the owner of very sensitive and critical data, that imply to not be the first target of spies, hackers, and other dangerous person, our database connection process is ensure by amazon. They are our supplier and have the responsibility to protect us against the threat that we can be victim.

On the other hand, there still be a way to penetrate the system, a way that does not make amazon responsible for this lack of security. It is the company network.

2.4.2. *The network security*

To improve data security, improving data defence is good, but it's not enough. These securities are good in case of a successful attack (or an unfortunate leak/breach) but feel safe with it is not enough. There is another side to improve to insure data safety. It is the defence against pirate attacks. To defend your information, you have to defend your database. To defend your database, you have to defend servers. To defend servers, you have to defend your network. To defend the network, you have to defend the accesses. You can hack a website, get the rights from it, and try to visit it, but it will not be enough to have access everywhere. To do it, you have to pass through routers. This object makes the connection between different devices, is a node, is essential to the network and is the biggest door. So, it is a strategical point to ensure.

The first and the basic protection is the password. Generally, it is a key, that can have different form. The is different kind of key, the WEP one, the WAP, the WPA2 or WPA2-PSK. The two lasts ones are the newest versions, and the most efficient ways to protect the network password to individual users. Also, it is possible to personalize it. The problem and the danger with this option is that the user will chose some words, that are findable in dictionary, to make it easier to remember or to write. It can be connected to his life or not, maybe containing some personal information, such as some birth date or kid names etc. Another precaution that may not be applied is the implementation of special shape of writing. For example, the classic user will not think to add majuscules, numbers or special prints sur as "@" or "!".

Hesus, that has an informatic school below, has improved his password until students from this school are not able anymore to broke it and enjoy the wi-fi. If the device connects to the wi-fi, it means that the connected machine is on the same network as other machine that use the same wi-fi. Considering that even informatic student, at least, have great difficult to break the password, we can consider the strength of it as good.

Another problem concerned connection ports. They are one of the breaches that can use an attacker to enter the network. Every devices have ports and use it on the same principle. They permit the connection between different machine by networking and allowed the data to enter or not in the inside system. Most of the port are every time close, because open only when the appropriate protocol arrives, and then close when the transfer is over. There are few exceptions, it concerns those who are allowing specific connections such as protocols FTP (port 21) or HTTP (port 80). But, because they are sensitive, they are very well secured.

To use those kind of doors, pirates use some software to scan which port are open and which are not. Those software send TCP and UDP packets (some protocols) to different machines, and according to the returns, their algorithm can define if there are ports open and which ones. Because of the type of the protocols, entrances or easy. This is the reason why those scanners are using those type of packets.

In the case of the hacker success to get an access to a device, he has another barrier. He needs to get access to the admin session to get the rights to access to the wanted data. Once he does it, he can have an activity on the network, visit the places that his statue allows him.

To respond to this problematic, Hesus hire a supplier that is expert in IT systems and security to implement some processes and securities. For example, they block the access to admin session to every computer that is using. There is no possibility to download and install software without the approve of supplier. This company ensures the network security. They have the different IDs to connect to every admin session and manage to do the actions that, as a simple user, it is not possible to do. Also, they provide a support to help the team in case of problem.

The CRM is another part of the network. The system comes from a supplier. So, it is his role to ensure the security. This IT tools is reachable by internet. That implies an ID connection. The main danger come from here. The connection moment is the most sensitive. It is protectable from some spying or robbery, but not from all of them, even less if the IDs are known by the hacker. A solution that had been adopted is to define different rights for users. They have access to different parts of the platform, they can watch different information, but they cannot access to others part if the admin does not allow it. In this case, as for the network, the admin function has to be protected more than the other because of the rights and his action field.

The CRM supplier provide us also a cloud service. This cloud is reachable on the same way as the CRM, by internet access. It is the same conditions to connect on it, by IDs connection. But the internal right structure is a little bit different. There are admins so do as networks and CRM, but on this platform, the content creator is the admin of his content. For example, if a user creates a file, he can give access to this file to everybody, only few persons, or nobody. It makes things more interesting in the idea to protect access, because every file creator can block the entrance and the content. He can protect the information from his colleagues, but not from the admins. It permits to create information structure only with people involved in the project, separate different document. This cloud is also used as storage for the documents loaded on the CRM. They are connected.

Another danger gate can be the mailbox. For this reason, it is very important to define a clear policy of mail consulting. Implement a list of non-desirable mails, a blacklist of different addresses, and do not use the professional mail address are few of the basics of the safety. Nowadays, people are more aware of the different dangers that can be contained in unknown mail. That’s the reason why some techniques, such as spear phishing, based on social engineering appeared. Using a mailbox from Microsoft is not dangerous if the information that are communicate do not interest them. They can have access to every mail send, so they have to do not be so serious and critical for the business and the enterprise.

3. Problems, development and solutions

3.1. Problem areas

After analysing the organisation and the structure of the different systems, and according to the different threats, from different natures and types that we saw previously, with the different dangers and consequences that they could imply if they are exploited by a skilled attacker, we can conclude that there are two potential weaknesses that can be identified.

3.1.1. *Internet navigation*

The first one is the **free connection or navigation on different internet browsers**. The fact that there is the possibility to go freely everywhere means that there is the possibility to be tricked by any casual threat. The visit of an unprofessional website can be dangerous, especially if the device is not protected by an efficient system. The only precaution is too block the admin function, to prevent any unwanted software installation. The problem is that malwares, or trojan for example can take a place on the computer without needing any admin authorization. For this kind of program, it could be easy to steal copy IDs and passwords already entered, or just to keep the information about the touches that are pressed during the session. Also, an eavesdropping use is possible. Once the hacker got those accesses, he will be able to use the victim's rights. There is not constant monitoring system installed to control the activity. Because almost none of the users are high skilled in computer science, to notice the fact that the session is occupied and used by someone else, particularly if this person does not change anything, or at least nothing viewable will be very difficult. It can take a very long time, allowing the hacker to monitor the hacked device, maybe try to get the admins rights, but principally to copy and / or steal the targeted data.

3.1.2. *Work at home*

The other potential danger is the possibility to connect to the different platforms from the user personal device. This function can do not be that dangerous if the machine is surely protected, and if the user pay attention and adopt a very careful behaviour. But there is also the possibility that the computer, or the pad, or even the

smartphone, is already contaminated by some programs. If the machine is accessible for the pirate from before the first connection to the different IT tools (CRM, mail box, internet platform linked to database), he will keep the access to those places, even if then, the access from the computer deleted. If the user is not yet under attack, and makes some connections to those applications, traces (such as cookies for example) will be findable. The problematic in this case is not to protect the different gates in a very strong way, it is to defend the computer. But personal user is rarely as protected as a company's computer. The consequence is that it makes them weak targets. It is advised from every cybersecurity professional to do not connect to sensitive data with an unprotected machine.

3.1.3. *Mailbox*

The mailbox can also be a target of a pirate. This item is full of information, can contain everything the hacker desire, and even more. Every internal discussions can happen and deliver some secrets. Because employees think this is safe and without danger, they can exchange more information than they should. Also, spying a mailbox can be very useful to choose a better target. Different job in the hierarchy are displayed, and some positions have more rights, power or influence than other. Hacking the CEO's mailbox is more useful than the intern marketing's one. The content of the messages is the main point. If they are not readable, it is useless to attack a mailbox.

Of course, no system is unattackable, even Apple's cloud had been forced in august 2014. Hackers found a breach and make leaks a celebrity's pictures. No server is totally safe, even if the company has its own one. The quality of the reputation can change from a day to the other (yahoo!). The access to data cannot be one hundred per cent sure.

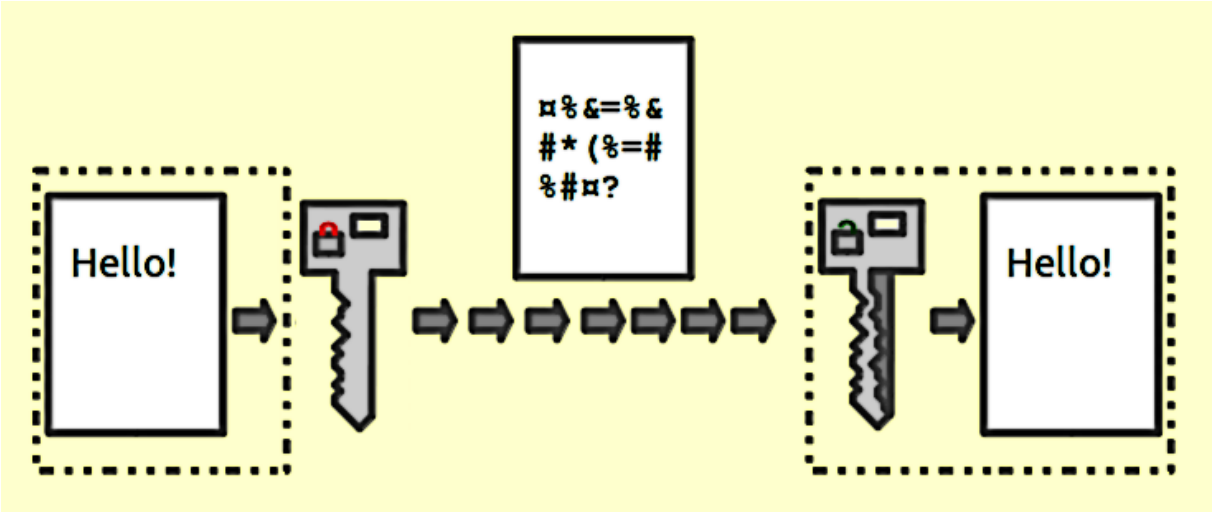
3.2. Solutions for data protection

There are different ways to protect your data:

3.2.1. *Encryption*

Encryption is a technique used since ages. Encoded messages have been created by generals in war time to exchange secret and sensitives information without interferences or spying. Encryption follow the same principle. The idea is to protect the information against unwanted third parties. Theoretically, there is only two parts who can understand the message, the encoder, the part who encrypted the message, and the receiver the other part, who knows the way to interpret the code, and transcript the original message. The data, following encryption process, become a cyphertext, that can be understand only by those who have the good key to interpret it. During the process, there is one similar key for encrypted the information, and for decrypted it. Generally, this key is randomly generated by an algorithm to be sure that a similar previous key doesn't exist already. This way is calling a symmetric key (or a private key), because only some users, the less possible, have to communicate on this way, in opposition with the other practical, that provide and different use of the encryption.

The second way to use encryption is calling asymmetric key (or public key). Where we have one key to encrypt and decrypt in the symmetric way, there is two different in the asymmetric one. The reason why it is called public key is because the encryption key is public, a lot of people can use it to encrypt messages. The decryption key is different, and only few users are able to use it, so few users are able to understand messages. It's very convenient in system where communication is going in majority in one way, such as credit card payment system. All information are going from the merchant to the bank, and most of data in this kind of transaction are encrypted on this way, allowing a better general performance.



Even if encryption ask some calculation power, who slows data exchanges, the advantage is that you can encrypt data without using them. You don't have to communicate, are make an exchange, a communication, a transaction or something else to encrypt information, it's possible to encrypt them and store them like that. Encrypted data at rest is a security against data theft, eavesdropfping, copy or even reverse engineering.

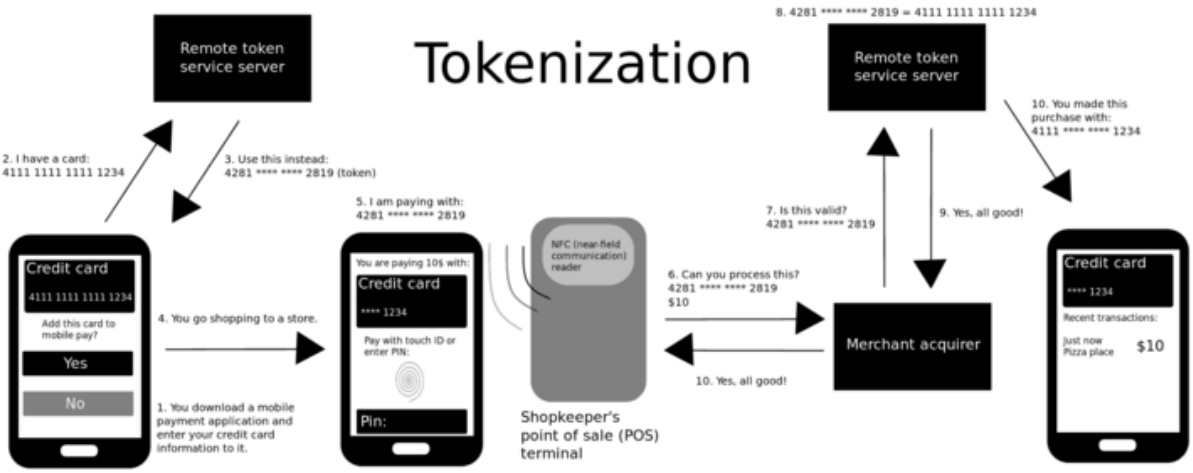
But Encryption can also have other uses. For example, the crypto-shredding. It can be used with if the user wants, or by a hacker. The principle is to delete all decryption key, to make encrypted key unreadable. Apple use this technique when the function "erase all content and settings" is activated. There are solutions in case of that happen and the user wants to recover his data. The first solution is to try a brute-force-attack, to crack the encryption, in the case of this one wasn't that strong finally. It's easier to do it on public key than private key. Another solution is to find flaws, but if encryption was doing well, there is a very small amount of chance that happen. The last solution is to wait that computer get faster, specially with next generation of quantum computer, that will have an incredible speed of calculation.

To broke encryption, pirates are very creative and imagined a lot of different way and ideas. The most known principles are the brutal-attack, where the encryption key will be attacked directly, frontally, and it will be just a question a calculation power and the attacking software (it's not the most efficient), but also known plaintext attack, when the plaintext is known and the attacker find the relation between the plaintext and

the ciphertext, the chosen plaintext, when a cryptanalyst analyse the result of encryption of a known plaintext, generally used for public key. Another form of attack is called adaptive chosen plaintext / ciphertext attacks. This shape of offensive follows the idea of adapting and analysing offensive based on prior result

3.2.2. *Tokenization*

Not very convenient to exchange data with third parties, tokenization is a technique that substitute your data with another object (another type of data or something with the same shape). It uses the token principle. The original data has a shape, is identify through a protocol, is stored under a certain form. The tokenization will happen when the original data have to be used. Generally, this solution is implemented for sensitive data (such as financial data, credit card number, bank account etc.) to allowed the user to have same properties without the risk that it been intercepted. Instead of using the sensitive data, the user will use the token to do what he need. With this idea, if this token is catch, hacked, stolen or maybe deleted, no information will be lost or corrupted in the database. Only the token. To use this principle, a tokenization system is required. It will generate in a random way a token that will fit with the sensitive data to replace it. This system will make the link, the relation, between the original data, and the object that is used instead of. Every time that the token will have to send a message, an information, it will pass through the tokenization system, who will refer, through the using protocol to the original data. By this way, tokenization in a supplementary step in communication, but it makes it much safer and avoid unwanted action during this time.



The critical point with this solution is, the tokenization system. It has to be sure. The first security measure that has to be taken is to isolate and segment it from the data process that already stored the data that you want replace by a token. An important point is, only the same system can tokenize or detokenize data, to insure the security, be sure that the tokenization way is unique and does not exist in another system, to certificate the efficiency of the system.

The aim of the token is to replace the data for exactly the same use, for the same purpose giving the same needing authorisations and access. Application has to be able to use it. Only the minimal needing applications should be able to use the original data.

There are two types of tokenization: The High-value one (HVT), and the low-value one (LVT).

The HVT looks similar to the data it replaces. Multiple HVT can be attached to one data and can also have restrictions, like specific networks/merchant, and/or specific device, or region, or blocked for a certain activity. A good example is the primary account number. It is bound to one or more credit card, and can wear multiple HVT, depend if one credit card has one or more HVT and how many credit cards are bounded. But it doesn't work on the other way, one token can be link to only one data.

The LVT is not enough to complete a transaction, that's why it is less secured, because it needs the former data, so it is possible with it to find back the original data, but only under a strict control. Using LVT implies a huge trust in the tokenization system security.

One of the strength of this technique is it energize consumption. Effectively, it consumption is lower than encryption's one, and it has, as a result, a continuous performance in the database. Calculation speed almost does not change, so it saves a lot of time.

Tokenization is only used while the data has to go out of the database, during a communication phase. In the former location, the information still has the same state, the same shape.

There is no perfect solution, and this one isn't an exception, there is limits. For the first generations, tokens needed a lot space in database because they weren't deleted, making live transaction more complicated a slower. Another point is the level of security. Indeed, there is not standards. The comparison between two systems is difficult. There is also the question of the strength of the random number generators, if there is pattern that repeat themselves, maybe can be predictable, if it has bias etc.

New technologies, such as vaultless tokenization and stateless tokenization, that represent the future of tokenization, gad been independently validated to be used in the Payment Card Industry Data Security Standard, proof of their effectiveness. Moreover, stateless tokenization discovered a solution to don't need anymore storage in database, while retaining the isolation properties of tokenization.

3.2.3. *Data Masking*

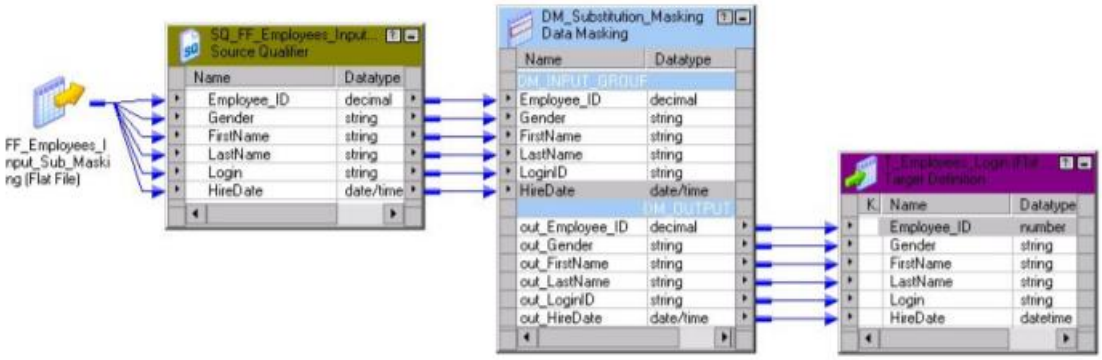
To ensure data security, and protect them, another way than protect them from reading is to mask them, and specifically hide or change their value to make the data in the table or in the database that have no sense all together. There are some techniques more or less known, more or less easy to implement, with different advantages, targeting different type of data, that can fit better on certain system. Most of the following ideas look simple, maybe obvious, but they still efficient to protect data and make the hacker task much more complicated.

3.2.4. *Substitution*

Substitution follow the idea of protecting data by masking them in the database. Here, it is another solution than encryption or tokenization, another way with different advantages. As encryption, Substitution can be used only on the data consisting the database. In a classic one with, for example name and surname of customers, and their gender, if we apply this method, the general look will be similar to the original database. The structure will be the same, with a column name, the second for the surnames and the last one for genders. But, the subtilty is that none of them will be true. It will be just fake ones, none of the names will be an effective customer, the same for the surname and maybe also for the gender. The point is that there is the possibility to choose what is substituted and what is not. The huge advantage is that it looks like a

real one, with the same structure. It can take time to the pirate to realize that what he steals is fake and he can do nothing with it. To realize it, a large substitution dataset is required, otherwise, some patterns will appear and be noticed easily.

The following scheme show how the source data, by using tables, are substituted in different data, and how the originals ones are stored and saved until the user need to use them again.



To use this system, the database has to store initial data and faked data in two different table, and make sure they don’t have any direct relations.

3.2.5. *Shuffling*

Such as substitution technique, shuffling is similar in the principle, but it complicates more the hacker task. Indeed, where substitution masks data by replacing those by new ones that looked pretty similar, the shuffling method, as it name suggests, shuffled them into the same column. The result is that, instead of don’t knowing if the data are true, and specially in financial sectors, when numbers should be related to name or account or other type of data that can own money, the values aren’t connected as it has to be. It means that it also can protect against internal threat, such as an intentional leak, or spying. Even if the person is familiar to those data, it’s very difficult to notice that the order is wrong, and that all the copy or theft of data is useless cause there is no sense with data under this shape. As for an encryption, the ciphered can be broke and the process can be reversed. It can be associated to substitution method, to increase efficiency.

3.2.6. *Number and date variance*

Use the variance is efficient only on numeric value (percentages, rates, amount of money, date, account, zip code etc.). As the mathematic principle, the variance defines a gap from a value. Generally, the value is the average of statistical data, and is used to define the trust that you can accord to a survey or a study, or something different with percentage that should represent a huge number but doesn't. If we refer to surveys, the different institutes or newspaper that give us those studies tell that the value is not totally sure, there is a percentage doubt. This percentage is calling the variance. It's above and below the average, generally with the same value on both sides. In data security, this gap is used to mask data. Let's say that a data has a value of 10, and we configure a 20% variance. According to this percentage, and in a random way, the original value, 10, will appear in the database under a value between 8 and 12. What make it difficult to crack is that it's generated randomly. So, for each value, or date (it also works with days, less or more 50 days for example), in a tab, none of them will be true, with the same maximal gap. Of course, guessing the gap is almost impossible, and in this gap, depending on how deep or decimal (if it can be only 10, or 10.1, 10.01, 10.001 etc.), there is potentially a huge multitude of possibility. Moreover, all those value don't have relation between their masking, except that they can be more or less close from the original value.

3.2.7. *Nulling out or deletion*

In opposition to the previous ways of masking data, that have the goal to change and protect the data without showing to the pirate that they are, doing it in a discreet way, nulling out or deletion indicate directly that data aren't those that he expects. The principle is to replace values by zero. By nulling value, none of logic application can work (it's a clear indication). Furthermore, any human who check stolen data, and find out that there is null value will notice immediately that they have been corrupted. So, what's the advantage? A null value doesn't have value. It's used to hide real value of a data, there is no way to find out what is the original one, there is no reverse way. As previous technics, this one is more efficient on numeric data, such as financial are date.

3.2.8. *Masking out*

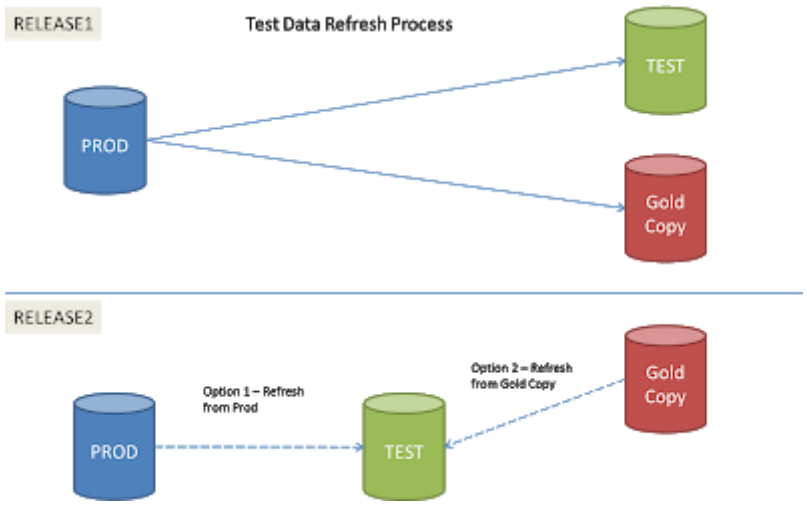
Masking out is another technique very useful to hide and protect public financial, bank and numeric data in general. As nulling out, it affects the visual of information. But contrary to the previous one, it doesn't change anything in the data (even if it's obvious that data are protected), it just replaces them by a new sign, generally a "X".

It is not the most efficient one, but it is very convenient for casual user (like to hide it during an internet payment) and for billing in many different situations.

Those various solutions were created for different cases and are not used on for the same purpose.

For static data, data that are not use, just stored, can be used as a golden copy (dataset that can be used for tests, implement changes, or just as a backup), the encryption and the substitution can be applied.

Golden copy principle



There is also a type of data exchanging that is called one-the-fly data. This kind of transfer is a transfer from one environment to another one. This is used most of the time by companies who have continual sending data needs, for example to run heavy applications. This type of company does not have the time to save it object on database or in golden copy, most of the time because it is small packets that are send continually. For example, the time of the day is an on-the-fly type of data. To this mode, techniques as encryption, tokenization, substitution can be used to protect data from theft during the transfer. If a hacker tries to intercept them, he will not have any key or way to decrypt information. The only hope for him is to be the proprietary of a strong brutal-attack program and a powerful computer to run it.

Also, there is dynamic data exchange. It is pretty similar to the on-the-fly one, but the difference is that data that are send by data exchange are going to another database then to be shared. Generally, it is a relation sever client. It is the most common way to organize efficient transfers. Because it is the most common, there is the most way to improve the data protection. All the previous and seen way to protect information can be used with this protocol. There many different types of information that can be concerned by a transfer between server and client, this situation multiplies the possibilities to attack the data, database, the website, and the object of the transfer. So, because there is not only one way to attack, and hacker are so creative, there is not only one perfect solution. That is the reason why the number of solution for those problems increase.

To improve data security, improving data defence is good, but it's not enough. These securities are good in case of a successful attack (or an unfortunate leak/breach) but feel safe with it is not enough. There is another side to improve to insure data safety. It is the defence against pirate attacks. To defend your information, you have to defend your database. To defend your database, you have to defend servers. To defend servers, you have to defend your network. To defend the network, you have to defend the accesses. You can hack a website, get the rights from it, and try to visit it, but it will not be enough to have access everywhere. To do it, you have to pass through routers. This object makes the connection between different devices, is a node, is essential to the network and is the biggest door. So, it is a strategical point to ensure.

3.3. Solutions and efficiency

Nowadays, there exist some tools to implement those different solutions. Some software offers those services without telling you, such as rented servers (encryption), or banks (tokenization) for example. But also, implementing those solutions together, merging with some other offer can improve a lot the security.

An easy way to implement those tools to solve the problems that hesus can be victim is the Virtual Private Network (VPN). This system is a something comparable to a private connection. There are only authorized machines that can connect on it. The advantage is that it can creates just a network used by few and chosen devices. So, there is only those machines that can communicate together, exchange and make transfers. So, it means that connections cannot be corrupted by external object. It ensures a safer way to navigate on internet, by isolating the interior of the computer from the browser. It creates a kind of funnel totally encrypted with only keys owned by the connected objects. By being isolated, the VPN connection is protecting the IDs connection on different platform, particularly for the CRM, the cloud, and hesus-store. This solution is easy to implement and do not disturb the different operations in progress. The different processes will not either be impacted. There will be only on step more, that will be the connection to the VPN. This part of the new process can be protected by a tokenization system, that provide the safety of the action. By protecting every different moment, different steps, the connection to sensitive tools will be improve, more than what is proposed by the provider. This solution is effective and used

by many companies. A Local Area's Network can be built only with a VPN, according to the society size. On this way, the probability that an external intrudes can connect without authorization is very low.

Another solution, this one to protect the user that wants to work from his personal computer, at home or somewhere else, is to install a software to recreate a virtual office on the computer. It is not every time possible to join the office, there can a true blizzard outside, or as in France recently, some periods of trouble (strikes) in public transportation. For those various cases, a solution is needed. It is working at home. To allow it, a sage solution is imperative. Every devices, independently of the OS used (Linux, IOS, Windows) has a welcome office. This function is used to stored files, create shorter ways to access to some files. Because it is just a part of a software (and not the most difficult to copy), some solutions propose virtual office. It allows to do the same things as a normal office, it permits to store exactly on the same way as a normal office, it is possible to use exactly on the same way the different software that are installed. The internet navigation is also the same, except that it comes from a virtual origin. So, it is empty of malwares, trojan, spying programs etc. also there is nothing to observe or to steal. The connection will come from a software, through the computer to a browser. With this security, none of the actual spying malwares can have action through this system. It is not protecting from every threats, such as ransomwares for example, but it protects at least against programs that would spying or steal data stored on the device. Avoiding this threat is equal to eliminate the principle danger.

Using those two solutions can improve a lot the security of the connection between device and network. Compilating those securities is not bad, they are complementary. Adding it to the actual solutions can make the system much stronger and safer. It will be a second protection more suitable against the threat.

A solution to protect the messages from the mailbox is to encrypt them. But to do it, it is necessary to add an encryption to the mailbox encryption. To do it, Edward SNOWDEN gave some advises during an interview. He talked about security, and especially the different ways that he defended his data. When he was in contact with journalists, before the scandal, his used to communicate by mail. He added many different tools. But one of them has better results. According to the movie/documentary

Citizenfour, even the NSA was enable (in 2015) to break this protection. The name of the software is PGP (for Pretty Good Privacy). It is a freeware accessible from everybody, in open source.

Against the attack of the server, the company cannot do anything. The servers are out of their influence area and out of their skills area. The only thing that is possible is to corrupt the own societies data, in a way that the society will know how to restore it. They have to be the only ones able to do it. As it is described above, it is easier with financial data. But it still possible for other kind of data. The attacker can do not pay attention to the data format, their correspondence, or their logics, especially if he does not have enough time. Those systems can be implemented to improve the efficiency of the database security. It can trick any pirate that does not pay attention enough, and not reveal every information from the data.

4. Social responsibility

ASSIGNMENT FOR THE SECTION "SOCIAL RESPONSIBILITY"

For student

Group	Full name

School		**Division**	
Level of education		**Direction / specialty**	

Initial data to the section "Social responsibility"	
1. Description of the workplace (working area, technological process, equipment used) for the case of occurrence of: *- harmful manifestations of factors in the production environment (meteorological conditions, harmful substances, lighting, noise, vibration, electromagnetic fields, ionizing radiation)* *- dangerous manifestations of factors in the production environment (mechanical nature, thermal nature, electrical, fire nature)* *- negative impact on the environment (atmosphere, hydrosphere, lithosphere)* *- emergency situations (man-made, spontaneous, ecological and social)*	Not applicable
2. List of legislative and normative documents on the topic	
List of issues to be investigated, designed and developed:	
1. Analysis of factors of internal social responsibility: *- the principles of the organization corporate culture;* *- the system of labor organization and its security;* *- development of human resources through learning programs and training and development programs;* *- system of social guarantees of the organization;* *- assistance to workers in critical situations.*	I study the internal environment, I noticed some important point: the atmosphere, the ambiance, and the way that it is maintain.

2. Analysis of external social responsibility factors: *- assistance in environmental protection;* *- interaction with the local community and local authorities;* *- sponsorship and corporate charity;* *- preparedness to participate in crisis situations, etc.*	The ecological responsibility is important. By helping to the recycling process, they participate to reduce the amount of wastes.
3. Legal and organizational issues of ensuring social responsibility: *- Analysis of legal norms of labor legislation;* *- analysis of special (characteristic for the investigated field of activity) legal and regulatory legislative acts;* *- Analysis of internal regulatory documents and regulations of the organization in the field of research activities.*	Law n°28-00 « gestion des déchets et leur élimination »
List of graphic material:	

Date of issuance of the assignment according to a line schedule	

The task was issued by the Advisor:

Position	Full name	Academic degree, academic status	Signature	Date

The assignment was accepted for execution by the student:

Croup	Full name	Signature	Date

4.1. Social responsibility with external stakeholders

The different stakeholders that are interfering with Hesus are the jobsites, the centers, and transporters. The activity needs to make them working together. On a certain way, the company bring them work and help them to be more efficient. The risk is to be victim of untruthful jobsites. There is the possibility from some of them that they cheat on the content of their waste. They can announce a certain type of soil, put it in the trucks and send it to the center. Once the center receive the ground, he notices that it does not respect the agreement that had been made. On this case, Hesus is not responsible, it is the jobsite. It is also the same with transporters. If they have the wrong information and transport a very polluted type of soil in an unappropriated way, instead of some inert once, the responsibility of the mistake is assumed by the jobsite. By being just an intermediate, those responsibility questions are not directly concerning Hesus.

But, since few years, they are engaged in a cause: The improvement of the trackability of the wastes, from the jobsites, to the centers, to the transporters. The idea is to being able to know o every time where the sent soils are, and control that they arrived on destination as it planned. Actually, this kind of following exist for the polluted soils, under the shape of letters. Each truck has to fulfil a letter when they come on the construction site, and when they leave the product in the center. After that, they have to send it back to Hesus. The first problem of this process is that it is very long. It can take a month or even more to get some letters from a case. Because it is a law, they have to be in possession of those letters as soon as possible. Secondly, they can be lost during the process. Generally, the bigger risk of loosing is from the postal service. This one is not one hundred per center guaranteed. Also, the political environment can have an influence. If there are some strikes, it can make all the process slower. The idea is to make the process automatic. It can take the shape of an application, where truck drivers will fil the needed information on each step and send it to Hesus by mail. It will be more storable, faster and with more trust. The obstacles are the transporters. Implementing this kind of solution means for them to buy for every trucks a mobile phone with this application already installed. The cost would be high. The other one is the truck drivers. They do not want to have the feeling to be monitored and spied. They

are showing this disapprobation. Their mentality has to evolve on this question, it is on the challenges that the company has to brave.

4.2. Ecological responsibility

The Hesus mission is to sort the different types of wastes. The aim is to redirect it to the suitable center and do not pollute proper places. Their main expertise is about soils. For ten years, the society participate to find the best solution in term of valorisation center, and also in distance term. Their action does not reduce only pollution that should not be there or make some valorisation of product that was sending to another type of center. One of the high value of the service is to find also the closest center answering to the different criteria. The commercial advantage and strength to this value is that, because it is close, trucks can make more passages, so extract more ground from the construction site per day, so the process can be faster and less expensive. This idea is important for jobsites, because they have some different dates to respect, and if they can save time, or maybe even win a little bit, they will take this possibility to get some margin in their different prevision, according to the construction they are making.

By offering this prestation, they are helping the fight against savages' pollution, solutions that does not respect the rules or the laws. They encourage different societies to call them to find a solution to the different wastes. By managing all the process, and ensuring a strict laws' respect, they motive the jobsite decision maker to use their offer. It is a societal problem, because construction companies have the reputation to do not follow the rules and hide everything can be hide. Because some legislations are not clear and accurate, there is some lack in the laws that allow some behaviours that can be considered as on the limit of the moral. Fortunately, they are not all like this, the majority is in the law. The bigger groups try to show the example for those recycling questions, because it is beneficed to their image and their reputation.

Even if soils were their first expertise field, currently they also treat material, and develop their activities (for example by integrated companies to make synergies and improve the efficiency of the different services, such as Pick my waste for example) to integrate other type of waste in their offer. By this strategy, the aim is to be present during every step of the construction site. It means that their offer will include

more and more wastes. Those wastes, instead of be randomly treated, sometimes maybe even do not be, they will finish in the suitable and appropriate center for them. It means that there will be less and less type of pollution rejected from the different constructions sites.

4.3. Social responsibility in the company

Taking care of his employee looks very important to Hesus. The ambiance at work is one of the priority. Because it is a small structure, every employee knows them each other. Links are easier to create. The people are not separated in private and small office. They are working in open spaces. This configuration makes the communication easier and much more comfortable. It allows the different discussion, to take news about each other and staying at the office. This point is very important. Having some exchanges and a fluent communication is good to improve the information circulation. If a salesman needs a technical precision, their office is not far from the studies service. By this way, they don't have to send mail, use official communication ways such as mails, or even to move from a closed office to another one to ask for the needed inquiry. They can be on a call with clients and without moving from their desk, having what they need to transfer. Because the communication is easy and fast, the information circulation is done better, more efficiently.

Another advantage of this open space and this friendly mood tried to be installed, it is that some actions can be organized to maintain it. Every year, there is a seminary that is organized, to enforced links between different teams. If there are open spaces, it does not mean that every department are always in relation together. Some of them do not have a lot of communication between them. This event is the good moment and occasion to create links and discover people out of the work prism. It can reveal some personalities, create affinities.

When relations are made, the organization of small events inside the company is much more convenient. The different participants are more in the mood to participate and be invested. It can be the occasion to install some competitions. It can stimulate the creativity and the efficiency. Also, it is a way to make the life of the office more interesting. This is also a way to offer a reward. Because the competition is an occasion

to put in concurrency different services, it can create an emulsion. Finally, attribute a reward to the winner or the winning team is a way to legitimise the fact to give some present to the worthiest people, while others are not feeling jealous.

The start-up mood is trying to be keep even if the company grows. An example of this whish is the different certification they are eligible for. This year (2018), they have been certificate from the website choosemycompany.com. This platform is a center where employees can give a mark to their company according to different criteria. The scale is from zero to five. Five is the best mark possible. Hesus get a score of 4,26/5.

According to many studies, the fact to be happy to go to work is very healthy, and one of the best ways to improve the efficiency. Having good relations with co-workers is one of the best way to do not be sad to got to the office.

4.4. Global consistency

The different actions drove by Hesus are looking in accordance with the policy and the ideas they are supporting. They try to bring transparency in his sector, and in the same time, they improve the waste retreatment, while they are taking care about their employees. They are invested in each field they are a part and are trying to do their best for every cases. Even in the office, there is special measure taken to recycle more products.

Conclusion

What about the future?

One of the possibilities is that governments and companies continue to become closer and closer. Government, with different reason like the struggle against “fake news” or to many data breach, or bad influence etc. will ask for an internet regulation, through internet operators. (It is a debate in USA). Those new regulations will give them more influence on giant internet companies, so on data that they have. Also, innovative companies will be targeting more and more by foreign big companies for their discoveries. Foreign big companies will have the support of their government, in this idea of cyberwar, so innovative companies will have to protect themselves. Government will provide their help to keep those innovation on their territory. Country who will not have enough strong knowledge on this field will undergo the influence of other powerful neighbours. To insure data security, the technology race will continue to grow, the Artificial Intelligence technology mastering will be decisive, especially for very sensitive data. Those data (and the companies who host them) will be defend more and more by governments, leaving other societies insure by themselves their information security. If they are considering not so important, there is a lot of chances that it’s because their field is not that crucial, so they will not be under strong attack from competitors, specially competitors supporting by another country. The most important risk they will run is data breach that a competitor will exploit.

Also, some researchers are looking for a new way to store the information. Currently, it is stored in kind of hard drive, trying to make devices with bigger and bigger capacity. But there is one natural structure that can stock more information than actual human devices. This solution is called DNA. Because it can store so many data that we are unable to sock DNA’s data from one human, the DNA structure, or at least model, can be future in term of data storage. It can solve many problems, due to the energize and the required space. Also, it can open new possibilities, in term of protection or just in term of technologies. For example, quantic computer requests a huge calculation power, but also, they will need some support to run correctly. And actual technology is not ready for it.

References

1. BUTTLER, Peter, 5 Ways to Enhance Data Security, 23 June 2017, https://www.globalsign.com/en/blog/5-ways-to-enhance-data-security/
2. Data Security, Detect and Protect for Digital Transformation, *informatica*, https://www.informatica.com/products/data-security.html#fbid=MRoot-mMuJ0S
3. Jagadeesha, Swetha Mysore, User Activity Monitoring and User Behavior Analytics for Entreprise Security, 15th november 2017, https://blogs.informatica.com/2017/11/15/user-activity-monitoring-and-user-behavior-analytics-for-enterprise-security/#fbid=sS3fJ2kYLOt
4. Udumula, Dhana Lakshmi, Choose Informatica Cloud Test Data Management for your Salesforce Security Model, 8th february 2018, https://blogs.informatica.com/2018/02/08/choose-informatica-cloud-test-data-management-for-your-salesforce-security-model/#fbid=sS3fJ2kYLOt
5. Thippanna, Anaga, AI and Automation – A Combo to Manage Cyber Security Threats, 18th January 2018, https://blogs.informatica.com/2018/01/18/ai-and-automation-a-combo-to-manage-cyber-security-threats/#fbid=sS3fJ2kYLOt
6. IBM, *IBM security guardim family*, https://www.ibm.com/security/data-security/guardium
7. ANSSI, *STORMSHIELD Data Security (SDS)*, https://www.ssi.gouv.fr/entreprise/qualification/stormshield-data-security-sds/
8. EXPERIAN, *What is Data Security?* https://www.edq.com/uk/glossary/data-security/
9. Micro Focus, *What is Data Security ?* https://software.microfocus.com/en-us/what-is/data-security
10. NG, Cindy, *The Difference Between Data Security and Privacy*, 21 december 2017 https://blog.varonis.com/the-difference-between-data-security-and-privacy/
11. DUPAUL, Neil, *Ultimate Data Security Guide*, https://www.veracode.com/security/data-security
12. Datasecuritty, homepage, http://datasecurityinc.com/
13. LORD, Nate, *The History of Data Breaches*, 21th April 2015, https://digitalguardian.com/blog/history-data-breaches
14. *A brief history of internet security*, 24 September 2009, https://www.scmagazine.com/a-brief-history-of-internet-security/article+/556389/
15. http://www.slate.com/content/dam/slate/blogs/future_tense/2013/08/01/130801_FT_DataBreachFullsize.png
16. CHAITIN, Daniel, *Edward Snowden: Facebook is a surveillance company rebranded as 'social media'*, 17th March 2018, https://www.washingtonexaminer.com/news/edward-snowden-facebook-is-a-surveillance-company-rebranded-as-social-media

17. Techopedia, *Data Security*, https://www.techopedia.com/definition/26464/data-security
18. Wikipedia, *Data security*, https://en.wikipedia.org/wiki/Data_security
19. KREBS, Brian, *Experian Lapse Allowed ID Theft Service Access to 200 Million Consumer Records*, 10th March 2014, https://krebsonsecurity.com/2014/03/experian-lapse-allowed-id-theft-service-to-access-200-million-consumer-records/
20. KREBS, Brian, *Experian Sold Consumer Data to ID Theft Service*, 20th October 2013, https://krebsonsecurity.com/2013/10/experian-sold-consumer-data-to-id-theft-service/
21. RAYWOOD, Dan, *Lost hard drive could affect 70 million US military veterans*, 05th October 2009, https://www.scmagazineuk.com/lost-hard-drive-could-affect-70-million-us-military-veterans/article/565201/
22. AISHWARYA, Srivari, *US DoD suffers massive data breach*, 1st December 2017, https://www.army-technology.com/news/us-dod-suffers-massive-data-breach/
23. THOMSON, Iain, *Massive US military social media spying archive left wide open in AWS S3 buckets*, 17th November 2017, https://www.theregister.co.uk/2017/11/17/us_military_spying_archive_exposed/
24. EGNASH, Martin, *Defense Travel System Data Breach Leaves Thousands Open to ID Theft*, 1st March 2018, https://www.military.com/daily-news/2018/03/01/defense-travel-system-data-breach-leaves-thousands-open-id-theft.html
25. https://www.upguard.com/blog/us-airforce-suffers-massive-data-breach
26. Upguard, *US Air Force Suffers Massive Data Breach*, 26th December 2017, https://www.upguard.com/blog/the-opm-data-breach-and-compromised-nuclear-data
27. Wikipedia, *2014 JPMorgan chase data breach*, https://en.wikipedia.org/wiki/2014_JPMorgan_Chase_data_breach
28. ROMAN, Jeffrey, *JPMorgan Chase Confirms Cyber-Attack*, 15th September 2014, https://www.bankinfosecurity.com/jpmorgan-a-7319
29. Wikipedia, *Hameçonnage* https://fr.wikipedia.org/wiki/Hame%C3%A7onnage
30. Wikipedia, *Spear phishing*, https://fr.wikipedia.org/wiki/Spear_phishing
31. SHEKHAR, Amar, *Top 10 Common Hacking Techniques You Should Know About*, 30th November 2017, https://fossbytes.com/hacking-techniques/
32. KARAYAN, Raphaële, *Les objets connectés, nouveaux relais des attaques informatiques*, 28th September 2016, https://lexpansion.lexpress.fr/high-tech/les-objets-connectes-nouveaux-relais-des-attaques-informatiques_1835475.html
33. Regards Connectés, *Data, emploi, éthique, politique : Regards sur l'intelligence artificielle*, published on 16th March 2016, video, https://www.youtube.com/watch?v=tL7ojiOTQho

34. TIWARI, Aditya, *What Is The Difference: Viruses, Worms, Ransomware, Trojans, Bots, Malware, Spyware, Etc?*, 4th October 2015, http://fossbytes.com/difference-viruses-worms-ransomware-trojans-bots-malware-spyware-etc/
35. CONFESSORE, Nicholas, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*,4 of April 2018, The New your times, https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html
36. Elodie, *5 choses à retenir de l'audition de Mark Zuckerberg devant le Congrès*, Journal du Geek, th of April 2018, https://www.journaldugeek.com/2018/04/12/5-choses-retenir-audition-mark-zuckerberg-congres/
37. @becket, *Photo of Zuck's notes, by AP's @andyharnik*, twitter, https://twitter.com/becket/status/983846618263891968/photo/1?tfw_creator=journaldugeek&tfw_site=JournalDuGeek&ref_src=twsrc%5Etfw&ref_url=https%3A%2F%2Fwww.journaldugeek.com%2F2018%2F04%2F12%2F5-choses-retenir-audition-mark-zuckerberg-congres%2F
38. Elodie, *Comment une société a siphonné les données personnelles de 50 millions d'utilisateurs Facebook pour la campagne de Trump*, 20th of March 2018, Journal du Geek, https://www.journaldugeek.com/2018/03/20/societe-a-utilise-illegalement-millions-de-donnees-dutilisateurs-facebook-cibler-electeurs-pro-trump/
39. *Compare data protection laws around the world*, https://www.dlapiperdataprotection.com/
40. *Data protection law of the world*, *law*, comparison between France and Russia, https://www.dlapiperdataprotection.com/index.html?c=FR&c2=RU&go-button=GO&t=law
41. Bisk, *The History of Information Security*, Villanova university, https://www.villanovau.com/resources/iss/history-of-information-security/#.WtdU6YhuZPY
42. JULIAN, Ted, *Defining Moments in the History of Cyber-Security and the Rise of Incident Response*, 4th of December 2014, https://www.infosecurity-magazine.com/opinions/the-history-of-cybersecurity/
43. *10 virus marquants*,6th of April 2006, Journal du Net, http://www.journaldunet.com/solutions/0604/diaporama/10-virus-marquants/1.shtml
44. The Washington post, *A history of Internet security*, 30th of May 2015, https://www.washingtonpost.com/graphics/national/security-of-the-internet/history/
45. *Cyberguerre*, Wikipedia, https://fr.wikipedia.org/wiki/Cyberguerre#Types_d'attaques
46. *Tokenization (data security)*, Wikipedia, https://en.wikipedia.org/wiki/Tokenization_(data_security)
47. *Tokenization vs Encryption*, skyhigh, https://www.skyhighnetworks.com/cloud-security-university/tokenization-vs-encryption/

48. *Encryption*, Wikipedia, https://en.wikipedia.org/wiki/Encryption
49. *Crypto-shedding*, Wikipedia, https://en.wikipedia.org/wiki/Crypto-shredding
50. Cryptomove, https://www.cryptomove.com/
51. GERHART, Morgan, *The Evolution of Cybercrime and What It Means for Data Security*, 27th of June 2017, https://www.imperva.com/blog/2017/06/the-evolution-of-cybercrime-and-what-it-means-for-data-security/
52. GOODIN, Dan, *Researchers crack open unusually advanced malware that hid for 5 years*, arstechnica.com, 09 of August 2016, https://arstechnica.com/information-technology/2016/08/researchers-crack-open-unusually-advanced-malware-that-hid-for-5-years/
53. Conrad, Eric, *Types of Cryptographic Attacks*, academia.edu, http://www.academia.edu/4739047/Types_of_Cryptographic_Attacks
54. *Data masking*, Wikipedia, https://en.wikipedia.org/wiki/Data_masking
55. CHANDLER, Ellen, 2009, *Data Masking with Substitution*, Informatica, https://kb.informatica.com/h2l/HowTo%20Library/1/0102_Data_Masking_Substitution.pdf
56. RAJARAMAN, Raghuraman, *Gold Copy in Test Data Management (TDM)*, 9th of March 2013, tdminsight.blogspot.fr, http://tdminsights.blogspot.fr/2013/03/gold-copy-in-test-data-management-tdm.html
57. OWASP, *OWASP Top 10 – 2017 - The Ten Most Critical Web Application Security Risks*, 2017, owasp.org, https://www.owasp.org/images/7/72/OWASP_Top_10-2017_%28en%29.pdf.pdf
58. *Crosse-site scripting*, Wikipedia, https://fr.wikipedia.org/wiki/Cross-site_scripting
59. *XML – Processors*, tutorialspoint/com, https://www.tutorialspoint.com/xml/xml_processors.htm
60. *Réseau ad hoc*, wikipédia.org, https://fr.wikipedia.org/wiki/R%C3%A9seau_ad_hoc
61. *Sérialisation*, wikpédia.org, https://fr.wikipedia.org/wiki/S%C3%A9rialisation
62. *Library (computing)*, wikipedia.org, https://en.wikipedia.org/wiki/Library_(computing)
63. OWASP, *Input Validation Cheat Sheet*, 11th of September 2017, owasp.org, https://www.owasp.org/index.php/Input_Validation_Cheat_Sheet
64. GRONDIN, Anaelle, *Vie privée : 5 outils utilisés et approuvés par Snowden*, the fifth of November 2016, lesechos.fr, https://www.lesechos.fr/05/11/2016/lesechos.fr/0211453241629_vie-privee---5-outils-utilises-et-approuves-par-snowden.htm
65. *Yahoo ! data breaches*, Wikipédia, https://en.wikipedia.org/wiki/Yahoo!_data_breaches

Thanks for their precious help to:

Charles (hesus developer)

Manoj (hesus developer)

Pawandeep (hesus developer)

Vincent Bignalet (hesus-store responsible)

Jean-luc Faisans (Global Infrastructure project manager)

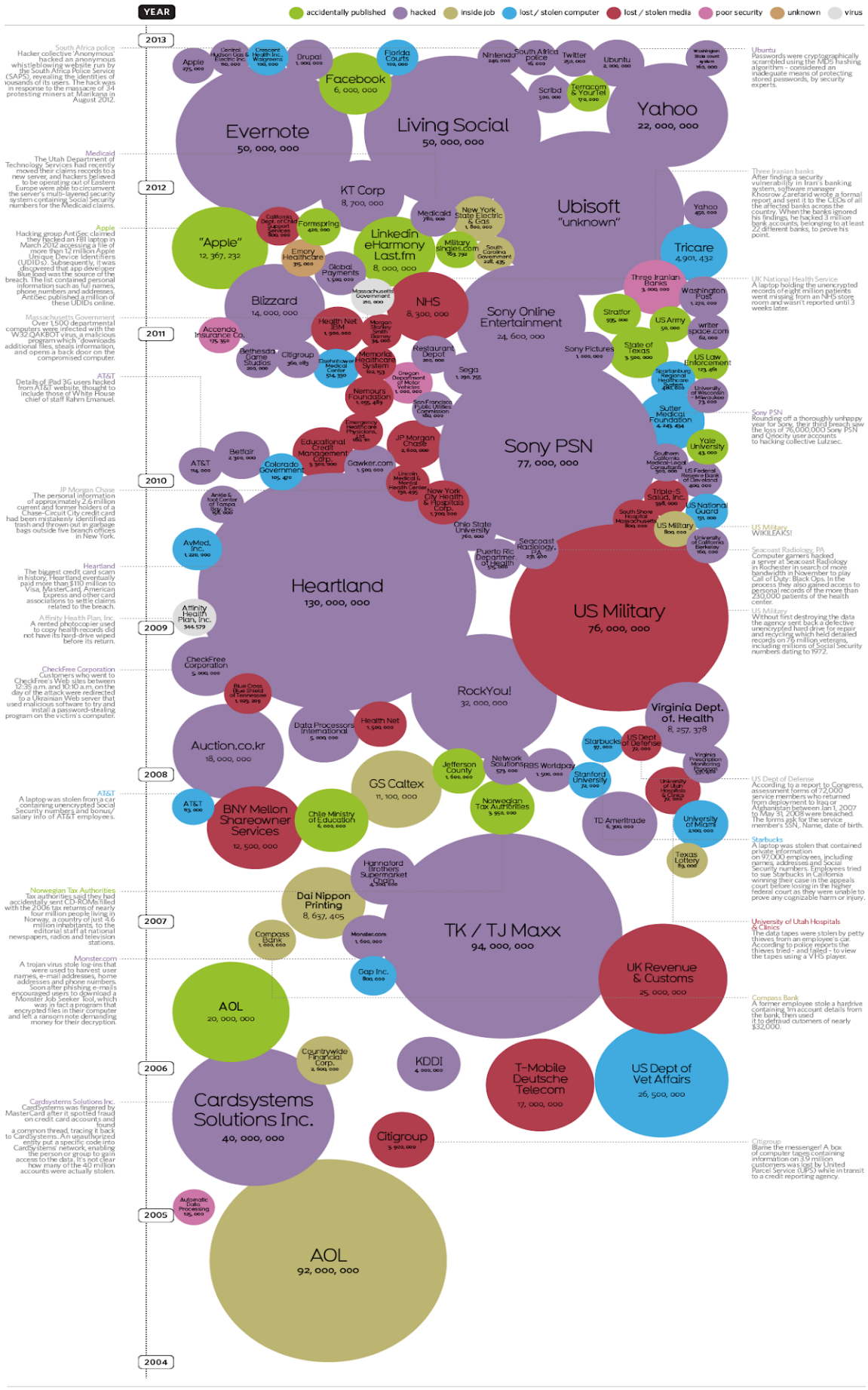
Xavier Clergeaud (consultant senior SI)

Annexes

(1) Different type of data breaches

World's Biggest Data Breaches

Selected losses greater than 30,000 records



(2) OWASP guide: https://www.owasp.org/images/7/72/OWASP_Top_10-2017_%28en%29.pdf.pdf