PROSEGUR RESEARCH

Highlights

ChatGPT:
Security risks
and opportunities

PROSEGUR

Index

01

Introduction: technology as a tool for the future

Highlights

We are living in exciting times; there is no doubt about it. And we are possibly **witnessing the beginning of a new revolution** that will affect every aspect of our lives: education, communication, work, personal relationships, and the man-machine relationship.

Adapting to new technological developments requires a mindset of growth which embraces **the only constant the world is experiencing: change.** However, the world in which we live requires increasingly complex and varied decision making, complicated by various international factors such as economic crises, health emergencies or technological disruptions of which quantum computing stands out.

Technology plays an important role in this process of change, being a lever for development. As a result, it is considered one of the great future game changers, having an affect on all areas of our lives and almost every company worldwide.

This accelerated growth, coupled with technological convergence, increases the complexity of everything around us, while allowing us to do the seemingly impossible.

PROSEGUR RESEARCH
Highlights

02

ChatGPT: security risks and opportunities

ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.

PROSEGUR RESEARCH
Highlights

ChatGPT, popularized in recent months, is one of the generative intelligence developments made available to the public. This refers to a **form of machine learning, based on language models and capable of generating new content in any format** (text, image, video, voice, code). In the past, applications such as Dall-e, Midjourney or Stable Diffusion have amazed society with their ability to produce high quality images from text. A few weeks ago, Microsoft presented its proposal for Bing with AI, whose potential is much greater than that of ChatGPT. Additionally, Google's solution, Bard, is expected to be launched soon. In the meantime, a multitude of proposals have proliferated, offering more than just textual results based on language models, but also the generation and combination of any type of content.

As previously mentioned, there is no doubt that **technology supports progress.** Technological breakthroughs have improved fundamental aspects for human beings such as communication, education and healthcare. But we also know, especially in the field of security, that **every opportunity generates new threats,** just as opportunities can be detected from many threats.

That is why at Prosegur Research, we point out the **ten main security risks of this technological development:**

PROSEGUR RESEARCH

Highlights

ChatGPT security risks

1

Geopolitical and geoeconomic power struggle

Within the context of unclear and fragmented powers, leadership can no longer only be measured by economic, diplomatic, and military capabilities. In 2017, Vladamir Putin pointed out that whoever mastered artificial intelligence (AI) would dominate the world. Geopolitics and geoeconomics present new risks, but also opportunities for states and companies that know how to predict the future. Data, as well as technologies, are at the heart of power, generating asymmetry between those who have it and those who do not.

2

Social polarization

Generative AIs, given their ability to produce multimedia content, can be used to spread hate or discrimination, as well as radical or extremist messages.

3

Disinformation

The objective of generating false news is to influence public opinion, damage social cohesion or affect electoral processes. Disinformation is a clear issue that affects national security, harming social cohesion and democratic principles.

4

Doxing

Disinformation is also likely to affect companies and organizations with the spread of hoaxes, biased information, the creation of false job profiles, or the manipulation of documents to damage the credibility of organizations. Their purpose can range from imitation to reputational attacks and influencing the markets.

PROSEGUR RESEARCH
Highlights

5

The creation of malicious chatbots with criminal objectives

Highlights the interactions between individuals with the intention of gaining sensitive information or with illicit economic purposes.

6

Identity theft

Achievable through the use of so-called “deep fakes” and AI’s ability to generate texts, images, videos and even simulate voices. It supports the creation of avatars that integrate all these and thus increasing the authenticity of their identity.

7

Creation of malicious codes

Such as viruses, Trojans, malware, ransomware or spyware, among others, with the aim of committing all kinds of cybercrimes.

8

Phising

Automatic generation of emails which appear real are used to deceive users in order to gain access to confidential information or IT systems. Generative AIs produce high quality content which invalidate suspicions usually triggered by low-quality phishing.

PROSEGUR RESEARCH
Highlights

9

Information leaks and data theft

Companies such as Amazon and Google have warned employees about the risks of sharing company information on ChatGPT and similar applications, as it may later be disclosed in the in the answers it provides to users.

10

Fraud and scams

These criminal offences have increased significantly in recent years. Traditional fraud cases, existing in all economic sectors, are further enhanced by the internet, social media and new technology. Generative AIs can contribute by developing high quality scams, in addition to profiling targets.

However, **technology, as a major driver of change,** is not generally created to be used maliciously, but instead is determined and enhanced by human action: suffice to recall that, according to IBM, ***up to 95% of cyber-attacks are due to human error.***

Therefore, it would be neither fair nor proportional to highlight the potential risks without similarly pointing out the opportunities that ChatGPT can generate in the field of security.

From Prosegur Research, we have identified **ten main opportunities ChatGPT presents for security matters:**

1

Automation of routine tasks for safety features

This, in turn, would enhance the more human capabilities and facilitate employee moral by eliminating repetitive and tedious tasks.

2

Creation of appealing chatbots

A more attractive profile for the consumer would improve customer interaction and other humans alike, promoting the customization that is currently required by companies.

PROSEGUR RESEARCH

Highlights

3

Access to vast amounts of security-relevant information, in a structured way by use of language

The capabilities of open-source intelligence (OSINT) can be of great benefit, while remaining conscious of source reliability and the credibility of the information.

4

Risk analysis

Support for risk detection and classification within organizations in various sectors.

5

Pattern identification

The value is not only in the pattern, but in its discrepancies: that which is out of the norm and could generate a weak signal or an early warning in an in a data-intensive and volatile situation.

6

Intelligence analysis support

It can assist in formulating hypotheses, whereby trends are identified and potential outcomes are drawn up. Although AI cannot substitute human creativity it can be an interesting accessory to think outside the box.

Highlights

7

Structuring recommendations on security issues

From how to defend yourself against a cyber-attack to what security measures to adopt before or during a trip. However, this work cannot replace the work of an international security analyst, who has purely human abilities which enrich the analysis.

8

Predictive analytics

It can assist with certain predictions, along with their associated probabilities, based on a vast amount of information.

9

Preventing and detecting cybercrimes

Within cybersecurity it can be used to support the detection of phishing, trial and test of codes, identify vulnerabilities, generate secure passwords, simulate conversations with adversaries and even with potential targets, thus being able to anticipate their actions among other effects.

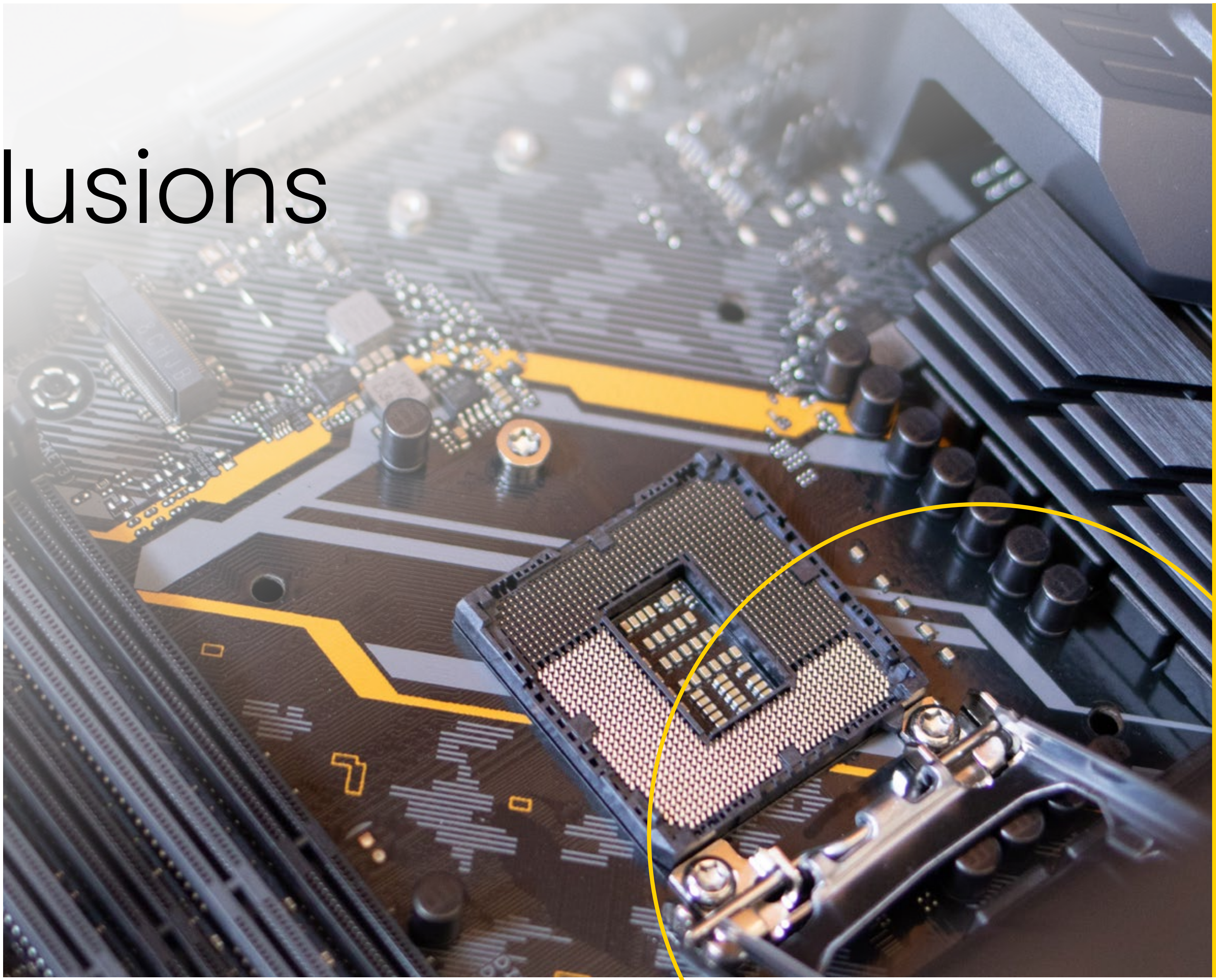
10

Learning

Generative AI can be a starting point for learning about issues related to security, particularly those linked to technology or risk.

03

Conclusions



PROSEGUR RESEARCH
Highlights

As Xavier Marcet rightly points out, “technology empowers us if we don’t stop thinking, but if we stop thinking, they only hypnotize us”.

Every security professional must continually refine his or her skills to face the present and the future with new instruments. This will require modifying our mental patterns and thinking differently in order to deal with change in a favorable manner, and to know how to take advantage of the opportunities presented to us.

Only empowered citizens and purposeful companies can achieve the desired results, guided by a defined and ambitious objective: to contribute to the global prosperity. Leading this task is a challenge but having experts with capabilities powered by disruptive technologies makes it possible.

And, of course, ChatGPT reminds us that the future is uncertain, while some positive developments remain, such as improvements in the quality of life and social and environmental efforts.

We are all protagonists of human, technological and social progress, and therefore we must be responsible for the legacy we leave to those who will succeed us. Because without progress, there is no future.

The future is uncertain and can be influenced by a variety of factors.
It is important to work together to create a better and fairer future for all, which may require international collaboration and the adoption of creative solutions. It is also important to consider the impact of our actions on the environment and society and look for ways to reduce the negative impacts and promote a more sustainable world.

Source: ChatGPT, 2023

PROSEGUR

We guarantee the safety of people, businesses and society as a whole.

research@prosegur.com PROSEGUR RESEARCH www.prosegurresearch.com