Second National Conference on Internet of Things : Solution for Societal Needs
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology | ISSN : 2456-3307 (www.ijsrcseit.com)

Data Privacy in Artificial Intelligence

Ms. Mounica B[1], Dakshata N. Ramteke[2]

[1]Sr. Asst. Professor, ISE Department, New Horizon College of Engineering, Bengaluru, Karnataka, India

[2]M. Tech. Scholar, Cyber Forensics and Information Security, ISE Department, New Horizon College of Engineering, Bengaluru, Karnataka, India

ABSTRACT

Information, data privacy and security concerns are a persistent trend that we've been reporting on nearly every year since computers started booting up. As artificial intelligence is emerging with dynamic changes ,it works with both pro-cons in field of data privacy . Artificial intelligence magnifies the ability to use personal information in ways that can intrude on privacy interests by raising analysis of personal information to new levels of power and speed while artificial intelligence also helps in analyzing and detecting patterns in large volume of data called as big data which can be beneficial for some business as well as threat to data privacy.

Keywords: Data Privacy, Artificial Intelligence, Big Data

I. INTRODUCTION

Recent news about data leaks, the lack of control over content, and significant influence of social networks has provided an increasing awareness of how social media platforms misuse personal data, which in turn has had an effect on the level of trust users have in such platforms and digital services . Many social media platforms get their (economic) value from capturing user's behaviour either directly (via services offered) or indirectly (by tracking users' online activities). With the migration from laptop- or PC-based browsing via web browsers, to consuming media on mobile devices and via dedicated apps, it has become possible to collect far more different types of data surrounding this behaviour in a far more targeted manner; now it is possible even in near real time . Combining places where people go digitally and physically offers many possibilities, but also brings about many new privacy risks. Although location data are explicitly categorized as personal data in the General Data Protection Regulation, it is not always clear what kinds of risks such kind of data poses, specifically in combination with other types of personal and non-personal data. Discussions on what personal data exactly means and how to apply personal data protection. In the context of large-scale data analytics are even more pressurizing the current landscape of data protection regulation . Slowly but definitely, companies and governments deploying big data analytics and process personal data are applying (and complying to) the GDPR[1]. Beyond the growing awareness of the need to comply. there is a wider societal need for trust in digital environments. As we know one of the most important reasons business, especially consumer facing business, wants to have lots of data is to know as much about the market, us, as possible. Artificial intelligence (AI) has

made that focus on customers more and more easier and accurate. While business has been becoming more invasive, government are focusing more into this and passing regulations that begin to provide certain limits. Privacy matters to the all, and smart business looks at how to use data to find out information while remaining in compliance with regulatory rules.

Information, data privacy and security concerns are a persistent trend that we've been reporting on nearly every year since computers started rapidly increasing. And now, the economic stakes, social consequences and technology get more serious and complex[2]. Privacy issues used to be centered around evading online activity trackers as they follow you around with ads for things you don't want or want. Now exposed as centre to many political and ethical scandals, data privacy has become one of the defining social and cultural issues of our era. Our need to control what we hide and what we shares extends from our very person and lead to our homes, businesses, communities, and governments. And because of the pervasive nature of technology, the data it creates and carries has burrowed into our lives in ways that we now take for granted.

II. LITERATURE SURVEY

Privacy and AI are terms that are emotive and misused. AI is associated with reduce the cost of prediction. The outcome on privacy is ever more platforms of data will be utilized to predict thousands of outcomes. However, this paper will compare, contrast and critique contextualized examples to highlight the Privacy and artificial intelligence [3]. The privacy is that most of the information created and produced in our digital universe, is created and consumed by the consumers themselves: social media networks, digital TV, sending and receiving mobile images between devices connected to the internet. However it is companies that that have 80% of the responsibility and therefore the liability of privacy, copyright and compliance. The biggest media company in the world creates zero content – Facebook, the individual users create it for them and continue to use the platform despite Facebook recently paying a $5 billion dollar fine for the infamous Cambridge Analytical data breaches, as documented in the article 'How Cambridge Analytical used your Facebook data to help the Donald Trump campaign in the 2016 election' . Psychometrics, based on the five personality traits known as OCEAN: openness, conscientiousness, extroversion, agreeableness, neuroticism can assess what sort of person we are and predict how we will behave. The issue however, was always the data collection – this all changed with the combination of the internet, the emergence of Facebook, Michal Kosinski and Cambridge Analytica [4]. The predictive model that Kosinski and his team had created by 2012 could prove that on an average of sixty eight Facebook likes by the user that they could predict: the colour of your skin (95% accuracy), sexual orientation (88% accuracy), affiliation to Republican or Democrat Party (85% accuracy) [5]. What confirmed how robust and reliable the modelling was becoming, was the predictive evaluation of the personality profile. The model was becoming consistently more accurate with less information: in his findings the (Kosinski et al) [6] model was now able profile an individual in ten facebook likes, to a greater accuracy than a co-worker, Seventy likes would profile the individual better than a friend, One hundred and fifty likes better than an individual's parents.

Images and videos that are created using deep learning and contain a real person acting or saying things they didn't do or say are called Deepfakes. If used for entertainment purposes, Deepfakes are fun, but people are creating Deepfakes for fake news and knowledge and, worse, Deepfake porn. In 2019, an application called DeepNude was

launched where you'll upload any image of girls , and it might generate a real-like nude image from it. It is quite disturbing that anyone could exploit woman images available online by creating nude images from DeepNude. After too much controversy, it was shut down. Still, it's just a matter of your time that somebody can again misuse Deepfake technologies by taking your publicly available videos or photos.

In recent years, China has received severe criticism thanks to mass surveillance of its people without their consent. They use over 200 million surveillance cameras and facial recognition to keep a constant watch on their people. China also mines their behavioral data captured on the cameras.To make it worse, China implemented a social credit system to rate the trustworthiness of its citizens and give them ratings accordingly based on their surveillance. People with high credit get more benefits and low credits loose benefits. But the more severe part is that each one this is often being determined by AI-based surveillance without people's knowledge and consent.

III. PROPOSED METHODOLOGY

Every time you attend the web for searches, browsing websites, or once you use mobile apps, you are doing not even realize that you simply are making a gift of your data either explicitly or without your knowledge. And most of the time you permit these companies to gather and process your data legally since you'd have clicked "I agree" button of terms and conditions of using their services. Apart from your information that you explicitly submit to your websites like Name, Age, Emails, Contacts, Videos, or Photo uploads, you also allow them to collect your browsing behavior, clicks, likes and dislikes. Reputed companies like Google and Facebook use this data to enhance their services and don't sell this to anyone. Still, there are instances where third party companies have scrapped sensitive user data by loopholes or data breaches. In fact, the only intention of the many companies is to gather user data by luring them into using their online services, and that they sell this data for vast amounts of money to third parties. The situation has worsened with the surge in malicious mobile apps whose primary purpose is to gather even that data from the phone that it didn't seek permission. These are primarily data collection apps disguised as a game or entertainment app. In today's world, smartphones contain very sensitive data like personal images, videos, GPS location, call history, messages, etc., and that we don't even know that our data is getting stolen by these mobile apps. Every now and then, such malware apps are removed from Play Store and Apple Store but not before they have already been download millions of times. We're seeing a social shift in the long term effects of privacy as billions more in venture investing targets our personal data for resale in a multitude of ways, people are starting to more deeply question their growing lack of data privacy and control. The challenges of protection of personal data in the context of Big Data Analytics (BDA) mainly connect to concepts such as profiling and prediction supported large datasets of private data. A secondary result of big data analytics is that combinations of non-personal data can still lead to the identification of persons and/or other sensitive .A dilemma put forward by data science is that data protection and data-driven innovation is diverging,and have even opposite premises.. We need to look for new ways to guarantee the protection of personal data while retaining the potential benefits of big data analytics

The Healthcare Industry already has many protections in situ, with HIPAA, and

frequently reminds us of how they're protecting us. Meanwhile our wearables are collecting a lot of health-related data on us. Who owns that data? And now that Google has bought Fitbit, what's that getting to mean for privacy? Add the many folks that have given away their DNA to seek out ancestors and therefore the Google Nightingale project to the privacy issue and it's clear HIPAA's getting to need an update[7].In Facial recognition what feels like a prequel to Minority Report, people's physical safety and movements are at risk. Citizens are taking measures to protect themselves from detection, trying to avoid arrests at protests or simply not wanting to share their whereabouts in public settings like an airport.As reported by CNN on varied defensive measures people have taken to protect their privacy - from rudimentary scarves and goggles to incredibly lifelike paper masks used to anonymize protesters. Big Think reported on designers using LED-equipped visors and transparent masks to guard identity. San Francisco became the primary city to ban the utilization of face recognition software by the town .We now know that voice-activated devices are listening all the time. Are our phones eavesdropping on us, too, as reported on Marketplace. It's inconclusive, but the investigation suggests there could also be more ways we divulge information than we realize. We also got to mention all the Ring and Nest doorbell systems. capturing video not just from your front entrance but all round the neighborhood, as well. Ring is actively collaborating with local enforcement - a practice that's raising privacy questions at the local level. We are being swiped right, All the clicks we've left behind are being used to rate us. And that can sometimes work against us. A recent piece from The NY Times outlined the industry-for-hire that makes a score for every one among us and sells it to businesses. Right, the list goes on and this is often just the start . Data privacy concerns reach voting and what data protection means to democracy. Facebook, Twitter, YouTube, TikTock, Google all have integrated with brands to hyper target us down to the tap, touch, and like. The truth is that there is only so much regular citizens can do without laws and policies that empower citizens to retake some personal data power.

The EU's GDPR was a blunt first instrument .Just trying to turn things off by playing whack-a-mole won't work. we need new innovations focused on protections that are more conversation driven and transparent. Tech companies today are built by a number of the neatest people in business – they ought to be ready to work within the bounds of latest laws to repair this. Finding ways to claw back and respectfully manage that data will prove essential to all users. Until now, consumers have been willing to lend their data (or have unknowingly given it away) to urge convenience or information reciprocally[8] . Once they fully realize the results of this bargain they're going to be looking to government and business to safeguard data and hand control back to them, the customer. Business needs to start thinking now about how to counteract the fear and distrust flooding the marketplace. The should provide verifiable solutions, traceability and transparency. They should be willing balance upholding privacy concerns without annoying users with privacy notifications and too many restrictions.

In Europe, the General Data Protection Regulation (GDPR) that came into force in 2018 regulates the collection and use of personal data.[1] Data protection law does not refer explicitly to AI or Machine Learning but there is a significant focus on large-scale automated processing of personal data and automated decision-making. This means that where AI uses personal data it falls within the

scope of the regulation and GDPR principles apply. This can be through the use of personal data to train, test or deploy an AI system. Failure to comply with the GDPR may result in enormous penalties for the companies involved

Cape Privacy may be a company that states its focus as providing software which integrates into a company's existing data science and machine learning infrastructure, enabling all parties to figure together on projects and policies. That is true start up speak, with an initial specialise in the implementation of privacy policies at the event level. The company works to help programmers work to oversee privacy policies in the data and between applications. What its link to the legal and compliance groups is more rudimentary, there's clearly a vision showing intent to enhance collaboration because the solution grows.

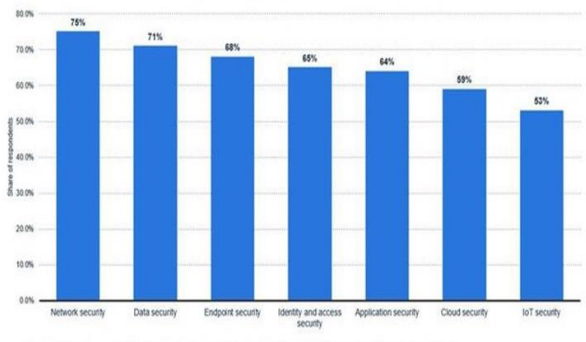

Figure 1. top AI use cases for cyber security in organization in selected countries as per past years

Now many countries have created their own data regulation policies to bring more transparency between these online platforms and the users. Most of those policies are centered around giving users more authority in what data they will share and be told about how the platform would process their data. A very documented example of this is often the GDPR law that came into existence a few of years back for the EU countries. It gives EU people more control of their personal data and the way it's processed by the businesses .Large companies like Google, Facebook, Amazon, Twitter, YouTube, Instagram, and LinkedIn literally own a majority of users' social data across the planet . Being, reputed giants, they ought to be extra careful to not leak any data to malicious people either intentionally or unintentionally.

The people of AI communities, especially the thought leaders, should raise their voices against the unethical use of AI on the private data of the users without their knowledge. And also they ought to educate the planet that this practice can cause such a disastrous social impact. Already many institutes are teaching AI ethics as a topic and also offering it as course.Finally, we should always remember that, despite government regulations, these are just policies and therefore the responsibility lies with us as individuals. We have to take care about what data we are uploading on social platforms and mobile apps and always inspect what permissions we are giving them to access and process our data, let us not merely "accept" anything blindly in "terms and conditions" that comes our way on these online platforms.

There are many concerns round the ethics in AI within the synthetic Intelligence Community thanks to the social biases and therefore the prejudices it can create. But processing personal data with AI without people's consent and further misusing it raises the concerns of AI ethics to subsequent level. AI emergence within the world remains in nascent age, and that we all need to intensify to form sure that our creating AI by misusing personal information should not become a daily occurrence within the future.

IV. CONCLUSION

Emerging artificial intelligence is affecting data privacy in very different way which should be restricting as it disobeys the basic privacy of an individual. some protection techniques are developing but it is still not clear they will be restrict this with changing artificial intelligence. Protecting privacy and securing digital data will continue to be a fundamental risk issue as AI becomes more mainstream in healthcare, raising numerous legal and ethical questions. Thus, it'll be obligatory healthcare leaders, AI developers, policymakers, data scientists, and other experts to spot vulnerabilities and consider innovative and proactive strategies to deal with them.

V. REFERENCES

[1]. Artificial Intelligence and the Privacy Paradox of Opportunity, Big Data and The Digital Universe Gary Smith Amity University Dubai Dubai, United Arab Emirates ,2018

[2]. Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation Bernd Carsten Stahl | De Montfort University David Wright | Trilateral Research,2019

[3]. J. Bossmann, "Top 9 Ethical Issues in Artificial Intelligence," World Economic Forum, 21 Oct. 2016;

[4]. W. Knight, "The Dark Secret at the Heart of AI," MIT Technology Review, 11 Apr. 2017.

[5]. C. Sandvig et al., "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms," Data Discrim. Convert. Crit. Concerns Product. Inq., 2014.

[6]. M. d'Aquin et al., "Towards an 'Ethics in Design' Methodology for AI Research Projects," AAAI/ACM Conf. Artificial Intelligence, Ethics, and Society, 2018.

[7]. S. Tan et al., "Detecting Bias in Black-Box Models Using Transparent Model Distillation," Oct. 2017.

[8]. J.A. Kroll et al., "Accountable Algorithms," Univ. Pa. Rev., 2016.

[9]. O.J. Erdélyi and J. Goldsmith, "Regulating Artificial Intelligence Proposal for a Global Solution," AAAI/ACM Conference on Artificial Intelligence, Ethics and Society, 2018.