

07/07/2023

# Final Report: CBI.ATTRACT

## Team: Random Bits



**Università  
degli Studi  
di Ferrara**



**UNIMORE**  
UNIVERSITÀ DEGLI STUDI DI  
MODENA E REGGIO EMILIA



*Idea<sup>s</sup>*



**ALMACUBE**

EMPOWERING INNOVATORS

Alma Mater Studiorum - Università di Bologna

## Abstract:

The intent of this report is to present the process of development and the proof of concept of a solution, obtained during the CBI Attract project. The final product is called Maia, a software solution which leverages the truly random and infinite bit stream capabilities of Random Power to apply differential privacy to medical datasets in order to obfuscate and preclude the identification of individuals within the dataset while permitting the use of the aggregated data for research and analysis purposes.

This report aims to present all the steps which were followed during the challenge-based innovation approach to arrive at the final solution.

## Index:

- Context
  - *CBI ATTRACT*
  - *Our Team*
  - *Our Technology*
  - *Our Partners*
- Discovery Phase
  - *What is our technology, and where does it come from?*
  - *Lexicon*
  - *Understanding the technology*
  - *Strengths and uniqueness*
  - *The Divergence Map*
  - *Collision week at CERN*
  - *First Milestone*
- Design Phase
  - *Convergence*
  - *Second Milestone*
- Development Phase
  - *Maia*
  - *Ideation of the prototype*
  - *Realization*
  - *Student project EXPO*
  - *Third milestone: The end of the project*

## Context

### CBI ATTRACT

CBI ATTRACT represents one division of the second phase of the ATTRACT EU project—an innovative endeavor that unites Europe's fundamental research and industrial communities to spearhead the advancement of cutting-edge detection and imaging technologies.

During the initial phase, the project successfully identified and provided financial support to 170 groundbreaking technological concepts within the realm of detection and imaging technologies across Europe. Furthermore, it facilitated a pilot program called "Young Innovators and Entrepreneurs," which engaged students from Aalto and Esade in applying design thinking methodologies to address societal challenges associated with various ATTRACT projects' technologies.

Building upon the foundation laid by the "Young Innovators and Entrepreneurs" initiative, ATTRACT introduced the ATTRACT Academy in the second phase. This network consists of ten university projects aimed at generating ideas for social innovation, drawing inspiration from the technologies developed in the research, development, and innovation projects of the first phase.

Within this network, CBI ATTRACT emerged as a collaboration between the University of Bologna, the University of Modena e Reggio Emilia, the University of Ferrara and Almacube. The project adopts a Challenged-Based Innovation (CBI) approach, wherein participating students apply design thinking principles to explore novel and socially impactful applications for the technologies developed in the R&D&I projects of phase one.

Five teams were assigned specific technologies that are typically utilized for particular purposes and were challenged to innovate new possibilities for their applications, all while considering the Sustainable Development Goals (SDGs) outlined in the United Nations' 2030 agenda.

## Our team:

Team Random bits is composed of six students with six different backgrounds. The team from the get-go has been aware of the advantages and opportunities that this heterogeneity brings to the project and how heterogeneity also supposes a challenge between communication, teamwork, and understanding of the technology and project objectives.

We would like to introduce ourselves in order for you to better comprehend the report and the work done by the group:



*Name:* Antonio Pelusi

*Age:* 23

*Contact:* antoniopelusi2000@gmail.com

*Major:* Computer Science

*What is your strength:* Sociability

*What is your weakness:* Adaptation

*Something about your daily life:* Prefer to work under a precise schedule & work better when under pressure.



*Name:* Valeria Rossi

*Age:* 26

*Contact:* valeria.rossi15@studio.unibo.it

*Major:* Theoretical Physics

*What is your strength:* Adaptability

*What is your weakness:* Outdated on social media and internet tools

*Something about your daily life:* I'm good at improvising



*Name:* Davide Pomante

*Age:* 25

*Contact:* davide.pomante@studio.unibo.it

*What is your strength:* Positivity

*What is your weakness:* Never angry

*Something about your daily life:* I do yoga and meditation



*Name:* Olivia Riccomi

*Age:* 24

*Contact:* oliviariccomi98@gmail.com

*Major:* Computer Engineering

*What is your strength:* Empathy

*What is your weakness:* easy distracted

*Something about your daily life:* I'm not morning person, I like photography and video making



*Name: Luis Felipe - Alvarez Vega*

*Age: 29*

*Contact: lfelipeav@gmail.com*

*Major: Legal Studies*

*What is your strength: Perseverant*

*What is your weakness: I'm too hard on myself if I make a mistake*

*Something about your daily life: I love videogames a lot*



*Name: Giovan Matteo Marcias*

*Age: 23*

*Contact: giovanmatteo.marcias@studio.unibo.it*

*Major: Digital Transformation Management*

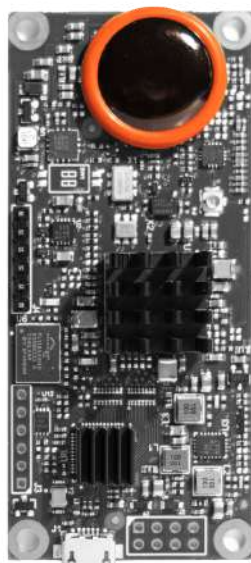
*What is your strength: Perseverant*

*What is your weakness: Sometimes I struggle to say no*

*Something about your daily life: Love practicing sports and being a morning person*

## Our technology:

The technology assigned to the group is Random Power, a Development of ASIC with a full system of random bit generator. The details of the functioning and current applications of the technology will be discussed in the following sections.



## Our Partners:

The main partner of the team is Random Power. Random Power is a Spin Off company of University of Insubria Como, Italy and AGH University of Science of Technology Kraków, Poland. As such, during the process of the programme, our team relied on actors from both the Como branch of Random power and the Kraków branch of Random Power.

From the Como branch, the main point of contact and source of support is *Professor Massimo Caccia*; Full professor at the Department of Science and Technology of Università degli Studi dell'Insubria, Massimo is the main inventor of the Random Power Technology.



From the Kraków branch, the main point of contact and source of support is *Professor Wojciech Kucewicz* ; Full professor in Microelectronics at the AGH-University of Science and Technology (Krakow, Poland), his main research focus is the design of integrated circuits for silicon detectors.



*Eleonora Musca*, Service Designer and Innovation Coach at Almacube; functioned as the team's main facilitator and guide in the challenged-based innovation process.



*Gaia Forghieri* PhD student in Quantum Physics at the Physics Department of the University of Modena; involved as Tech Mentor, she provided guidance on topics related to quantum physics to the team.



Additional support for the project comes from:

- Almacube the incubator, accelerator and innovation hub of the University of Bologna and Confindustria Emilia Area Centro.
- The universities of Bologna, Modena, and Ferrara.
- Ideasquare the innovation space at CERN.



## DISCOVERY PHASE

What is our technology, and where does it come from?

As mentioned previously, the technology assigned to the group is Random Power, a Development of ASIC with a full system of random bit generator.

The random power chip generates random bit streams by processing the time series of self-amplified endogenous stochastic pulses in a dedicated Silicon structure. The principle, based on the quantum properties of matter, has been validated by applying the National Institute of Standard and Technology (NIST) protocols.

Random Power springs from a long-standing experience in the development of Silicon Detectors for sub-atomic particles, notably for experiments at CERN, the European Laboratory for Particle Physics. The idea of the Random Power chip was developed by Professor Massimo Caccia in an ingenious effort to turn an issue, the endogenous pulses created by the silicon chip, into an opportunity for the creation of a true randomness generator based on this intrinsic and unpredictable quality.

Understanding the working and function of the random power chip required the team to learn about some quantum physics, engineering, and computing concepts that were associated with it.

This was the first task the team undertook by developing a lexicon of associated concepts, divided into Linguistic concepts, Physics concepts, Hardware concepts, and Computer Science concepts.

## Lexicon:

### Linguistic Concepts:

#### Unpredictability:

This is usually perceived with a sense of uneasiness and discomfort. However, when it comes to securing our data, information, posts, pictures and whatever flows to (or from) the internet about us, protection relies on **unpredictability**. The impossibility for an eavesdropper or a hacker to break the walls protecting what sits on the net about us depends on secret keys, protocols, and ciphering algorithms, namely the **art & science of cryptography**. Keys, essential to locking & unlocking the doors, must be randomly generated. And true randomness implies unpredictability.

#### Paradigm-shifting:

A paradigm shift, a concept in the philosophy of science introduced and brought into the common lexicon by the American physicist and philosopher Thomas Kuhn, is a fundamental change in the basic concepts and experimental practices of a scientific discipline.

**Endless Stream:** is a sequence of elements made available over time without an end.

### Physics Concepts:

**Quantum [noun]:** the minimal amount of any physical entity involved in an interaction

**Quantum [adj.]:** What is beyond classical physics, what is very small and abides by a different set of laws, where intrinsic indeterminism arises

**Quantum properties of matter:** The main properties of quantum mechanics are the quantification of energy (quanta), the wave-particle duality, the uncertainty principle, and the correspondence principle. Connection to Big Idea about energy: Electronic transition in atoms corresponds to quantized energy.

**Semiconductors:** a material, such as silicon or germanium, that has electrical properties in between those of conductors and isolants. Such properties can be influenced by "doping" the material (putting a

different chemical in the lattice structure) and are at the heart of all modern technologies with devices such as transistors and diodes.

### Hardware Concepts:

**ASIC:** Application-Specific Integrated Circuit is an integrated circuit (IC) chip customized for a particular use

**Silicon device:** A silicon complementary metal-oxide-semiconductor field-effect transistor (CMOS FET) ring oscillator.

### Computer Science Concepts:

**Cybersecurity:** A mix of software and hardware systems devoted to protecting computer systems and networks from attacks from malicious actors that may result in unauthorized information disclosure, theft, or damage to hardware, software, or data, as well as from the disruption or misdirection of the services they provide.

**Robustness:** the ability of a computer system to cope with an error during execution and cope with erroneous input.

**Random bits feeding cyber-security systems:** cybersecurity systems are based on generated random bits that allow a higher level of security (the systems are in some ways more protected since the random bits are unknown)

**Bits:** The bit is the most basic unit of information in computing and digital communications. The bit also means binary storage in computing or a binary symbol in digital communications into which a bit, as a unit of information, may be stored or encoded. The name is a portmanteau of a binary digit.

## Understanding the technology:

After spending some time familiarizing itself with these concepts, a general understanding of the technology was developed.

The technology relies on a silicon photomultiplier, a small component that utilizes an electrical current to exponentially multiply photon spikes created naturally from the properties of the silicon chip.

Once the photon spikes have been multiplied, these responses are converted to “zeroes and ones” through a time-to-digital converter.

In layman’s terms, the technology is a chip that produces an endless stream of zeroes and ones that is characterized by a lack of predictability and patterns.

## Strengths and uniqueness:

The strengths of our technology are:

- True Random Number Generator (TRNG) → The device is a true random generator, which means it produces an endless truly random bitstream that is characterized by the lack of patterns, predictability and dependence from the previously generated string.
- Endogenous → Every single bit is generated from the chip. It does not need any additional computational power. This makes it easier to set-up.
- Robustness → It can be stable against temperatures and voltage variations. It can manage millions of electrons in a few tens of nanoseconds, with no risk of misregistering an event
- It is a very compact device with a really high bit extraction efficiency → random bit streams at 1 MHz for every mm<sup>2</sup> area device.
- Furthermore, it is an on-chip implementation → prototype dimensions are 8 cm x 3.5 cm x 0.4 cm
- There is no requirement of post-processing to cancel correlations → This is still a problem in other devices for streams of gigabit length

## The Divergence Map:

As it has been stated previously, the goal of the CBI programme is to find an innovative use for the assigned technology that will have a positive impact on society.

The main conceptual tool to conceive ideas related to the possible uses of the technology is the divergence map. In the center of the map, there is the assigned technology Random Power. Then, the map is made of circles:

- The inner circle represents the *functions*. Those represent literally what the Random Power chip does.
- The middle circle represents the *applications*. Those are the fields in which Random Power would have an impact.
- The outer circle represents the *needs* that our technology would solve, connected to each application.

Attached to every need, there are also the various Sustainable Development Goals established in the 2030 Agenda by the United Nations, in which the single need would have an impact.

## Functions:

What Random Power actually does is:

- Generating random and unique streams of bits
- Generating a random stream of bits
- Generating a unique stream of bits
- Generating an endless stream of bits
- Speeding up algorithms

These functions could be then utilized for specific applications and then paired with societal needs, each of them related to the UN's 2030 agenda Sustainable Development goals.

## Applications:

**Encryption:** Streams of true random bits play a crucial role in modern cryptography. They are used to generate strong cryptographic keys, which are essential for securing sensitive information. Random bits are unpredictable and provide a foundation for various cryptographic algorithms, such as symmetric and asymmetric encryption, digital signatures, and key exchange protocols. By using true randomness, cryptographic systems can resist attacks based on statistical analysis or pattern recognition, enhancing their resilience against adversaries.

**Hash Functions:** using the random bits in hash functions, a mathematical function that takes an input (or "message") and produces a fixed-size string of characters, known as the hash value or hash code. The primary purpose of a hash function is to convert data of arbitrary size into a compact and unique representation, commonly referred to as the hash digest. Hash functions are extensively used in cryptography and data integrity verification. They are designed to be fast and

efficient, producing a unique hash value for each unique input, and any slight modification in the input will result in a significantly different hash value.

#### **Random noise/ Data Obfuscation:**

A true random stream of bits can play a significant role in data obfuscation by introducing an element of unpredictability and confusion. By incorporating random bits into the process of data obfuscation, sensitive information can be masked or disguised in a way that makes it challenging for unauthorized individuals to decipher or understand. The random stream can be used to generate random keys or masks, which can be applied to the original data, making it difficult to reverse-engineer or extract meaningful information without the corresponding keys or masks.

**Gaming:** Algorithms that leverage the true randomness of the technology could be implemented to make video games more random, particularly in the case of virtual card games or games that rely on randomness.

**Unique identifiers and Unique ID generators:** the different strings of bits or the native capabilities of the random power chip could be employed in creating an endless amount of unique identifiers for different purposes.

**Algorithm speed up:** since some algorithms require using the CPU to create pseudorandomness, an external random power chip providing the randomness instead could speed up the algorithm and reduce CPU usage.

#### **Machine Learning and Deep Learning:**

An endless, random stream of bits can bring several benefits to deep learning and machine learning. It can serve as a valuable source of randomness for various tasks, such as initializing model weights, shuffling training data, and introducing noise during training. Randomness plays a crucial role in preventing models from overfitting to specific patterns in the data and promotes generalization. By incorporating random bits, models can explore a broader range of possibilities, improving their ability to handle unseen or unpredictable inputs. Furthermore, random streams can be used in reinforcement learning to introduce stochasticity in decision-making processes, enabling agents to explore different actions and learn more robust and adaptive policies. In summary, an endless random stream of bits empowers deep learning and machine learning systems to enhance

their learning capabilities, improve generalization, and make more informed and diverse decisions.

### **Simulation of Complex Systems:**

An endless random stream of bits can be immensely valuable for simulations, especially when studying complex systems. By leveraging true randomness, researchers can inject unpredictable elements into simulations, enabling the exploration of a vast array of scenarios and system behaviours. The random stream can be utilized to generate random inputs, such as initial conditions or environmental factors, which can significantly impact the dynamics of the simulated system. This randomness allows for the modelling of various uncertainties and stochastic processes that exist in real-world systems. By incorporating an endless random stream of bits, simulations can capture the inherent complexity and variability of complex systems, providing insights into their behaviour, robustness, and sensitivity to different factors. Ultimately, this facilitates a deeper understanding of complex phenomena and aids in making informed decisions and predictions in fields like physics, biology, economics, and more.

### **Needs:**

Each of the previously identified applications could then be related to a need associated with one or more of the Sustainable development goals. Just a suggestion: The path that goes from the understanding of the technology to a specific application in which Random Power can give a substantial impact, it passes through the recognition of the needs, and the issues related to that, that we think our technology can satisfy. Being able to recognize and emphasize specific needs represents a key point in this project

**Encryption - Cryptocurrency, planned obsolescence, satellite security and super secure car keys:**

*Cryptocurrency:* The identified needs related to encryption could be reducing the environmental impact of cryptocurrency transactions by creating more efficient cryptographic keys and key exchange protocols.

*Fighting planned obsolescence:* Randomness can be utilized in various ways, such as generating unique identification codes, cryptographic keys, or randomization techniques for software updates. These

measures can enhance product longevity, increase security, and empower users to have more control over their devices.

Both these needs are related to *SDG 13 Climate Action* as they could aid in reducing energetic use and reduce waste; and *SDG 12 responsible consumption and production* by ensuring optimal use of resources.

*Satellite Security:* For satellite security, an endless random stream of bits can be used to generate cryptographic keys that are unique and virtually impossible to guess. Random bits serve as a crucial component in key generation algorithms, ensuring that each key is unpredictable and resistant to brute-force attacks. These keys can be used for secure communication, authentication, and encryption of data transmitted to and from satellites. By incorporating randomness, satellite systems can protect against unauthorized access, eavesdropping, and tampering, enhancing overall security and confidentiality.

*Super secure Car Keys:* In the context of car keys, an endless random stream of bits can play a role in the generation of secure and unpredictable key codes. Random bits can be utilized to create unique and random key codes for car key fobs or transponders. These codes can then be used for secure authentication and communication between the car and the key. The randomness introduced in the key codes makes it extremely difficult for attackers to replicate or guess valid key combinations, significantly enhancing the security of vehicle access and ignition systems.

Both of these needs are related to *SDG 16 Peace, Justice and Strong Institutions* as they could reduce violence through security.

**Hash functions - Improving privacy, digital voting and Genetic Research :**

*Improving Privacy:* Hash functions can leverage an endless random stream of bits to enhance privacy through the concept of salting. Salting is the process of adding a unique random value (the salt) to the input data before hashing. By incorporating an endless random stream of bits as a source of salt, hash functions can generate a different salt value for each input, ensuring that even identical inputs produce different hash outputs. This technique is particularly valuable in scenarios like password hashing, where maintaining privacy and preventing pre-computed attacks is crucial. This need is related to *SDG 16 Peace, Justice and Strong Institutions* as it could facilitate secure communications in a variety of ways.



*Digital Voting:* One key application is in the creation of unique identifiers for voters. By using an endless random stream of bits, hash functions can generate random and irreversible hash values or tokens that can serve as anonymized identifiers for each voter. These identifiers can be associated with individuals' voting records, ensuring privacy while maintaining the integrity of the voting process. This need is related to SDG 16 Peace, Justice and Strong Institutions since it could facilitate direct participation in democratic decisions.

*Genetic Research:* With an endless, random stream of bits, hash functions can generate unique and irreversible hash values for individual genetic profiles. These hash values can be used to represent genetic information while preserving privacy by dissociating it from personally identifiable information. Researchers can then work with the hash values instead of the original genetic data, allowing them to perform analyses and share insights without directly exposing sensitive information. This need is related to SDG 3 Good Health and Wellbeing, as it could facilitate health research while respecting privacy.

**Random noise - Ensuring obfuscation of sensitive data, and anonymity of the safety of witnesses**

*Obfuscation of sensitive data:* the random noise created by an endless stream of random bits enhances the security of sensitive data through encryption. It helps generate strong cryptographic keys, masks the original data, enables secure communication protocols, enhances the security of cryptographic algorithms, and protects against attacks by introducing unpredictability and complexity.

*The anonymity of safety of witnesses:* it can also ensure the anonymity and safety of witnesses by adding randomness to their data, making it difficult to identify them. This technique, known as noise infusion or differential privacy, protects sensitive information and prevents adversaries from linking the data back to specific individuals, enhancing witness protection.

This need is related to SDG 17 (Partnership for the Goals) because it promotes partnerships, collaboration, the rule of law, access to justice, and policy influence. These collective efforts contribute to creating a safer and more inclusive society that values the rights and protection of witnesses involved in legal proceedings.

**Unique identifiers - Improving efficiency in healthcare, creating identity, records in blockchain, privacy, verifying identity, and expanding the voting base**

*Improving efficiency in healthcare:* a unique identifier improves efficiency in healthcare by accurately identifying patients, enabling data integration and interoperability, facilitating research and analytics, streamlining clinical trials and patient recruitment, and supporting secure health information exchange.

Improving efficiency in healthcare is linked to SDG 10 (Reduce Inequality) by ensuring equitable access to healthcare services and SDG 3 (Good Health and Well-being) by enhancing affordability, quality, and accessibility of care for all individuals and communities.

*Creating or verifying digital identities:* a unique identifier ensures uniqueness, protects privacy, enables scalability, supports data integration, facilitates identity verification, and allows for identity anonymization.

Creating an identity is linked to SDG 10 (Reduce Inequality) by providing equal access to resources and opportunities, and to SDG 16 (Peace, Justice, and Strong Institutions) by promoting transparency, accountability, and the rule of law.

*Creating records in blockchain:* a unique identifier ensures immutability, verifying data integrity, facilitating consensus, enabling smart contract execution, supporting auditing and traceability, and promoting interoperability and data sharing, and enhances the overall reliability and trustworthiness of the records within the blockchain. This is linked to SDG 16 (Peace, Justice, and Strong Institutions) by promoting transparency, accountability, fraud prevention, efficient dispute resolution, secure digital identity, data integrity, and trust in public institutions.

*Privacy for social media users:* a unique identifier enables pseudonymity, user segmentation, anonymity in interactions, privacy-preserving analytics, minimizing data linkage, limiting data sharing with third parties, and overall enhancing privacy for social media users.

*Verifying identity:* a unique identifier provides uniqueness, secure verification, privacy protection, scalability, anonymity, and cross-platform interoperability for verifying identities.

Verifying identities is linked to SDG 10 (Reduce Inequality) by ensuring equal access to services, resources, and opportunities for all individuals. It helps prevent discrimination and barriers faced by marginalized populations, promoting inclusivity and reducing inequalities based on identity or social status; and to SDG 16 by

ensuring accountability, enhancing transparency, and contributing to safer and more equitable societies.

### Deep learning, Machine learning and Simulation of complex systems - predicting financial performance, and studying climate change events

An endless random stream of bits can greatly benefit machine learning and deep learning in predicting financial performance and studying climate change events.

By incorporating random bits, models can learn to handle unpredictable market conditions and simulate various scenarios, making them more robust and capable of accurate predictions. In financial forecasting, random bits can introduce noise and diversify training data, enabling models to better handle volatility and uncertainty. This need is related to SDG 8 Decent work and economic growth, as it could assist in understanding how current approaches to the economy of specific regions or territories sustain per capita economic growth.

In climate studies, random streams can be used to generate synthetic climate variables, allowing models to simulate different climate scenarios and assess the impacts of climate change. The randomness provided by an endless stream of bits enhances the ability of machine learning and deep learning models to capture the complexities of these domains, leading to more accurate predictions and valuable insights. This need is related to SDG 13 climate action, as a better understanding of environmental events would strengthen adaptive capacity to climate-related hazards and natural disasters.

## Collision week at CERN

Staying closely with the experts that work at CERN, such as *John Wood*, *Pablo Garcia Tello*, and *Markus Nordberg*, gave the team a lot of awareness about how new ideas are born, which is the relation between values and technology, and how to find or even to dig deeper into which specific fields we could implement our ideas. One important lesson is about the increased velocity that characterized the evolution of almost every phenomenon nowadays. In our present, and for sure more in the near future, we will be faced more and more with the exponential impact

of events, often catastrophic events, that affect and influence our life. Being aware of the increasing growth of a phenomenon is important because it allows you to understand its impact, make informed decisions, identify opportunities and anticipate challenges.

In connection with this project, that means that the specific innovation that we're looking for will face challenges and won't suppose a "panacea" for the many issues the world faces, but must be a new tool with which we can better face challenges that the future will present to us.

With regard to the innovation opportunities we had identified; the divergence map further changed and evolved during the collision week at CERN, first with the feedback provided by other participants in the CBI attract project and staff of CERN and CERN Ideasquare. Followed by changes proposed by the team that originated from discussions and exercises like the 6 thinking hats methodology developed by *Edward de Bono*.

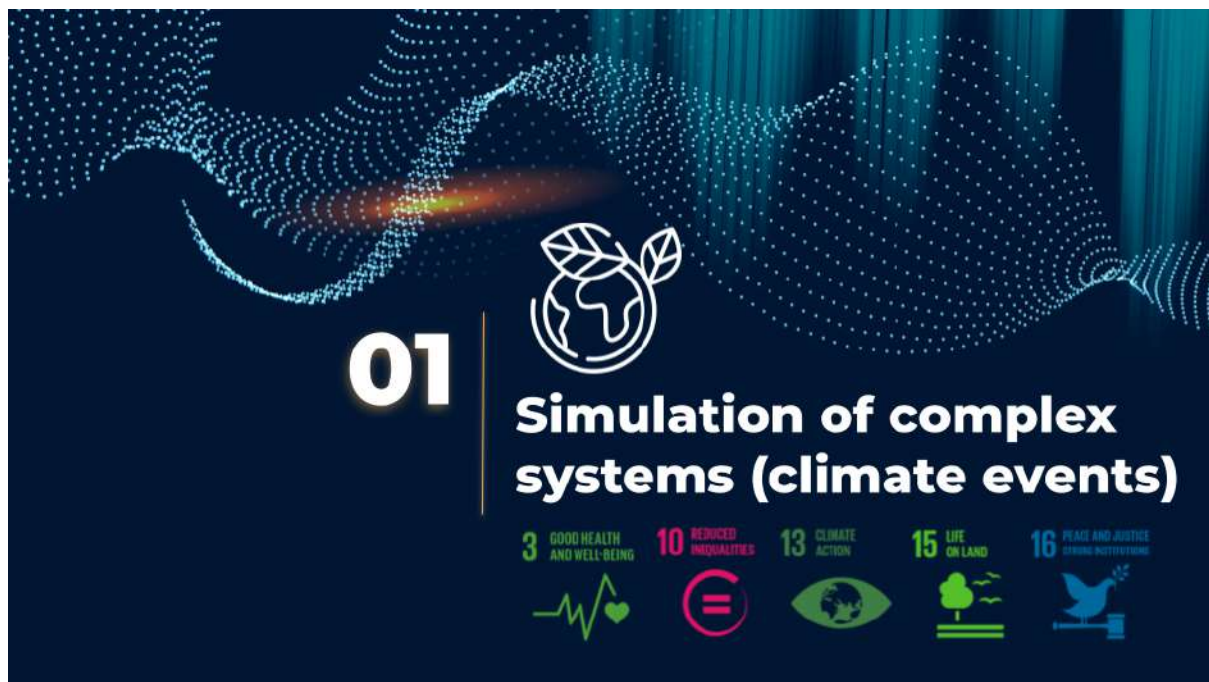


## First Milestone:

During the first milestone, the team was challenged to perform an initial conversion into 5 opportunities that were the most promising.

These opportunities were considered by the team to be the most interesting and the ones with the most probability to be implemented and impactful.

They are the following:



As was mentioned previously, the idea at the core of this opportunity is injecting unpredictable elements into simulations, enabling the exploration of a vast array of scenarios and system behaviors. The random stream can be utilized to generate random inputs, such as initial conditions or environmental factors, which can significantly impact the dynamics of the simulated system. This randomness allows for the modeling of various uncertainties and stochastic processes that exist in real-world systems.

One sector that would benefit greatly from better simulations is climate modeling, as more accurate predictions could help better inform scientists, policymakers and government.



The idea behind this opportunity is to use the random bitstream created by Random Power to generate parameters such as terrain features, weather patterns, vegetation distribution, or animal behaviors, adding diversity and realism to the virtual environment. This randomness helps create complex and dynamic systems that exhibit emergent behaviors, mirroring the organic and unpredictable nature of the real world. By incorporating an endless truly random stream of bits, virtual environments can capture the richness and complexity of natural systems, fostering more immersive and realistic experiences for users.

Natural complexity could be particularly useful in simulations used for training, for example for rescue operations, or in virtual therapeutic simulations.



As was mentioned before for this opportunity it is hoped to create a virtual identifier that ensures uniqueness, protects privacy, enables scalability, supports data integration, facilitates identity verification, and allows for identity anonymization.

One central use case for this opportunity would be for academic researchers that need to make sure their publications are only attributed to them.



Incorporating a true randomness generator can significantly benefit randomized algorithms, including those used in machine learning. True randomness ensures that the generated random numbers or bits are unpredictable and unbiased, which is crucial for the quality and effectiveness of randomized algorithms. In machine learning, true randomness can impact training processes in several ways. Firstly, it can enhance the diversity and exploration of the solution space during optimization tasks, allowing the algorithm to discover a wider range of potential solutions and avoid getting trapped in local optima. This can lead to improved generalization and better overall performance of machine learning models. Secondly, true randomness can play a role in data augmentation techniques by introducing variations and perturbations to the training data, facilitating better model robustness and preventing overfitting.





As mentioned previously under data obfuscation, the random noise created by an endless stream of random bits enhances the security of sensitive data through encryption. It helps generate strong cryptographic keys, masks the original data, enables secure communication protocols, enhances the security of cryptographic algorithms, and protects against attacks by introducing unpredictability and complexity.

The main applications for this could be communications and obfuscation for analysis that is respectful of the data subject.

Choosing the five opportunities meant making an evaluation of each of them about several key factors. This meant looking for evidence and interpreting this information to understand meanings, themes and patterns from the data that we're looking for. This analysis concerns several investigations in different fields such as: market analysis (understand and identify the target market and the target audience), get a customer evaluation (to evaluate the demand and check the degree of desirability), risk and mitigation (identify potential challenge), financial viability (how to assess the profitability and sustainability), determine values to assess the feasibility and scalability of the idea. What the team did was a qualitative analysis aimed to gain insight into the experiences, the perspectives and behaviors which are connected with the proposed applications.

All of this was necessary to understand the degree of feasibility in regard to each opportunity. From the knowledge that was taken analyzing Random Power, and its uniqueness, the focus switched to looking for the concrete possibilities that were hidden inside the five specific contexts that were chosen, and in which it was believed that the applications of the technology could provide serious and concrete improvements.

It was necessary to evaluate the proposed idea and build a network with experts, professors, or ideally stakeholders who were inside the specific fields that we were analyzing and could provide advice or suggestions.

From the technology to the benefits that a good application could generate, passing through the specific context of said application and the needs of its stakeholders, and taking care of the positive impact that must follow align with the sustainable development goals which are the cornerstone of the project.

As mentioned previously, these final five applications were considered by the team as the most feasible, interesting and with the most probability for impact, but they changed slightly after the visits to Como and Kraków as will be expanded upon in the following phase.

## DESIGN PHASE

### Convergence

After the first milestone, the challenge posed to the team is to converge once more in order to narrow down the chosen opportunities to two, based on criteria of feasibility and desirability.

At that point, the key effort was validation of the potential that each opportunity was believed to have and of the ease of eventual implementation.

In the Discovery phase the focus was thinking about every possible field of application in which the Random Power technology could be applied, in this phase the focus was finding evidence with which to support each idea.

To determine how the opportunities fit into both these categories, the team set out to conduct research and contact experts in the fields related to the opportunities to learn more about the different implications that each of them would have.

This required diving deeply into the opportunities and learning as much as possible from the technology itself and from the field of applications, the team needed to increase its awareness about what the tech was able to achieve (technical studying) and about how the issues identified were relevant for society(socio-economical studying).

In the scenarios we had identified, which were virtual possibilities or realities the team looked to enable with the technology, it was crucial to understand the weight in terms of impact between the recognition of the specific needs of the users and the possibility to solve (to enable) them with the function of the technology.

The team thus needed to narrow down on the selected opportunities. To do that, it was fundamental to meet our partners in Kraków and in Como.



During the visit to Como, the team connected with researchers working on the more creative side of the Random Power team. This visit led to the addition of opportunities for Random System network coding and Differential Privacy to the list of opportunities.

These two new opportunities, along with the simulation of complex systems, were evaluated as having the most potential for positive impact.

### 1. Simulation of complex systems

We wanted to increase the speed and precision of the simulation of complex systems, Particularly in Montecarlo simulations since they currently rely on pseudorandomness.

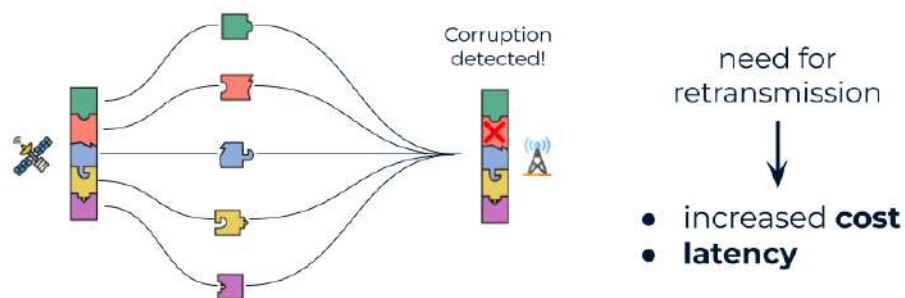
While using a pseudo-random number generator (PRNG) a bit stream could be replicated, implementing Random Power that would no longer be possible because every bit stream is unique and independent of the others. This allows to have more accurate results and a faster convergence time. However, the implications of PRNG inexplicably affect the performances of various machine learning models. This creates an issue for reproducibility in the kind of simulations that require to be reproduced.

## 2. Random Linear Network Coding (RLNC)

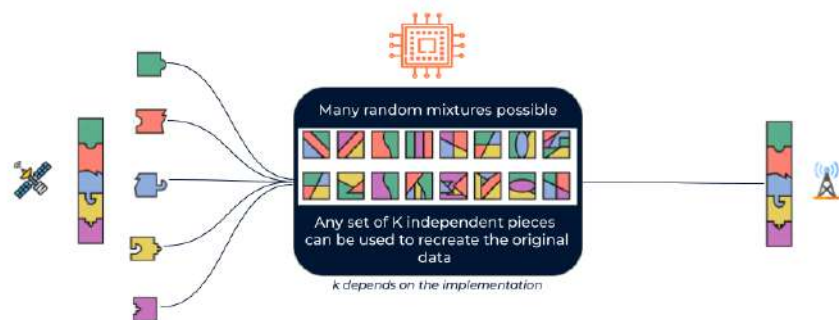
Random linear network coding is a technique used in computer networks to optimize data transmission and improve network efficiency. It involves combining and encoding multiple data packets as linear combinations before transmitting them across the network. Unlike traditional routing techniques that forward packets as is, random linear network coding introduces randomness and linear algebra into the process.

An endless random stream of bits can greatly enhance random linear network coding by providing a continuous source of random coefficients for the encoding process. When combining multiple data packets, random coefficients are crucial to creating unique linear combinations. By leveraging an endless random stream of bits, nodes in the network can continuously access new random coefficients, ensuring a diverse set of encoded packets with minimal repetition. This enhances the efficiency of data transmission by reducing packet collisions and maximizing network capacity. Moreover, the continuous availability of random coefficients contributes to the robustness of random linear network coding, allowing for error correction and packet recovery in the face of packet loss or network disturbances. Overall, the endless random stream of bits enables more effective and resilient data transmission through random linear network coding, making it a valuable resource in optimizing network performance.

### What happens without RLNC?



## What happens with RLNC?









### 3. Differential Privacy

Differential privacy is a technique used to protect the privacy of individuals when analysing aggregated sensitive data. It provides a mathematical framework for ensuring that the statistical analysis of a dataset does not reveal specific information about any individual within the dataset. The goal of differential privacy is to strike a balance between preserving the utility of the data for analysis and preventing the disclosure of personal or sensitive information.

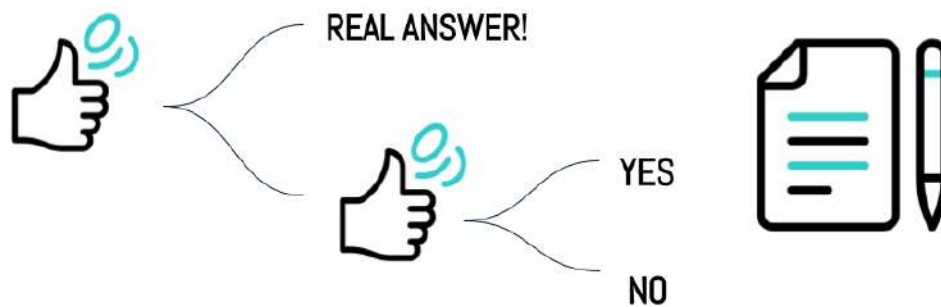
In simple terms, differential privacy adds a layer of noise or randomness to the data before releasing it or performing computations on it. This noise makes it difficult to determine the specific contribution of any individual data point, thus protecting their privacy. The amount of noise added is carefully controlled to balance privacy and data utility.

Let's focus on a simple explicit example: if you had a YES/NO questionnaire I would toss a coin upon registering your answer: if I get a head I'll register your real actual answer, if I get a tail instead I'll flip the coin again. If I get a head I'll register a YES, if I get a tail I'll register a NO. In this way the overall statistical trends of the datasets are preserved, while my personal data are covered by plausible deniability.

## What happens without DP?

|   | NAME       | AGE | DATA  |   |
|---|------------|-----|---|---|
|  | ██████████ | ██  |  | <p><b>95%</b><br/>identified!</p> <p><b>NETFLIX</b></p> |
|  | ██████████ | ██  |  |   |
|  | ██████████ | ██  |  |   |
|  | ██████████ | ██  |  |   |

## What happens with DP?



We want to implement Random Power using differential privacy to create a secure dataset of medical data where sensitive data of patients can be stored and utilized for analysis and medical research leading to better medical treatments and diagnosis.



## Second milestone

As part of the second milestone the team had to choose only one technology of the aforementioned technologies.

The choice was difficult as all the opportunities had their strengths and weaknesses. The team finally decided to choose differential privacy because it was the opportunity where there was more concrete evidence that supported the feasibility of the idea.

Even if the connection between privacy and the power of randomness was something that was already in the minds of the creators of the Random Power technology.

The team decided to generate something new by taking care of a concrete problem that exists in the health sector. That's how Maia, the group's final solution was created.



## DEVELOPMENT PHASE

### Maia



Maia is a software that applies differential privacy to datasets in order to allow an analysis of the aggregated data contained in a data set while concealing the individual information about any of the people that are within that dataset.

The team chose particularly health datasets as this is one sector where there is a huge need for both solutions, analysis and privacy.

As the Data saves lives initiative puts it “The challenge is to deliver this within the same healthcare budgets. New opportunities are arising for treatment through genetics, and better decision-making is possible using algorithms and artificial intelligence. Learning more from health data can support these changes and help achieve better care in a cost-efficient manner”.

The central idea is that by making privacy exponentially more robust, people would be encouraged to share their health data, which will contribute to research and eventually machine learning diagnoses or treatments based on this aggregated data.

Thus, the team was posed to create a prototype and proof of concept for this opportunity.

## Ideation of the prototype

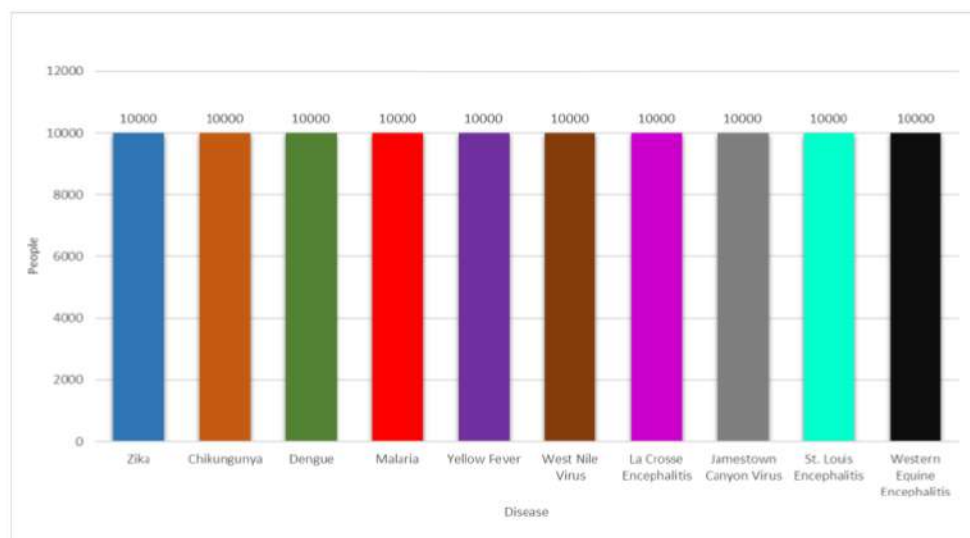
Since the team worked with a software-based technology, it developed a program through which its full potential could be exploited.

As a means to fulfill this, a mock dataset containing information about patients with a disease was built, to demonstrate the impact that DP would have on its data.

## Realization

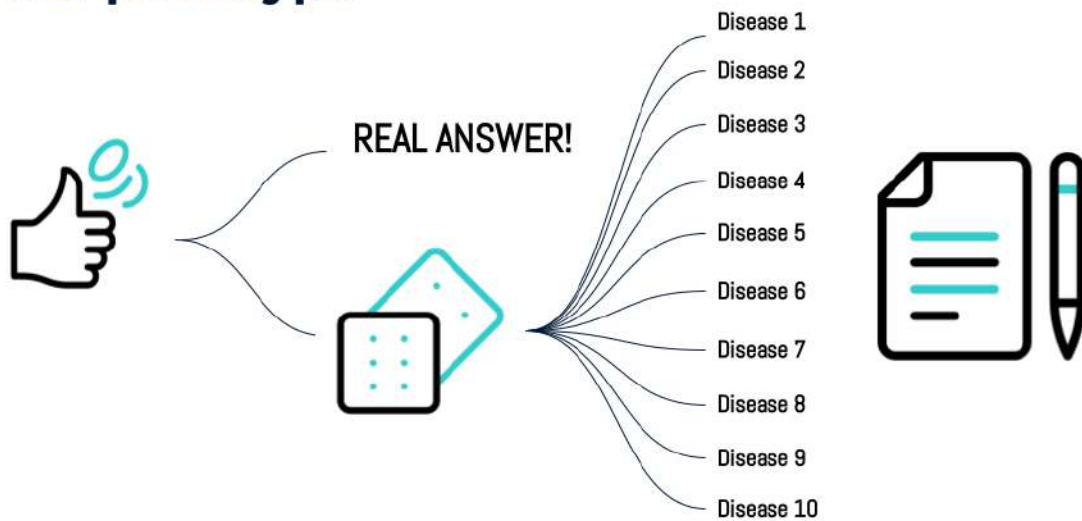
As a first step, an **artificial dataset** of 100,000 people was created. Each of them was assigned one out of 10 diseases transmitted by mosquitoes. This was done in a uniform distribution so that for each disease there would be 10,000 people with it.

## The raw dataset

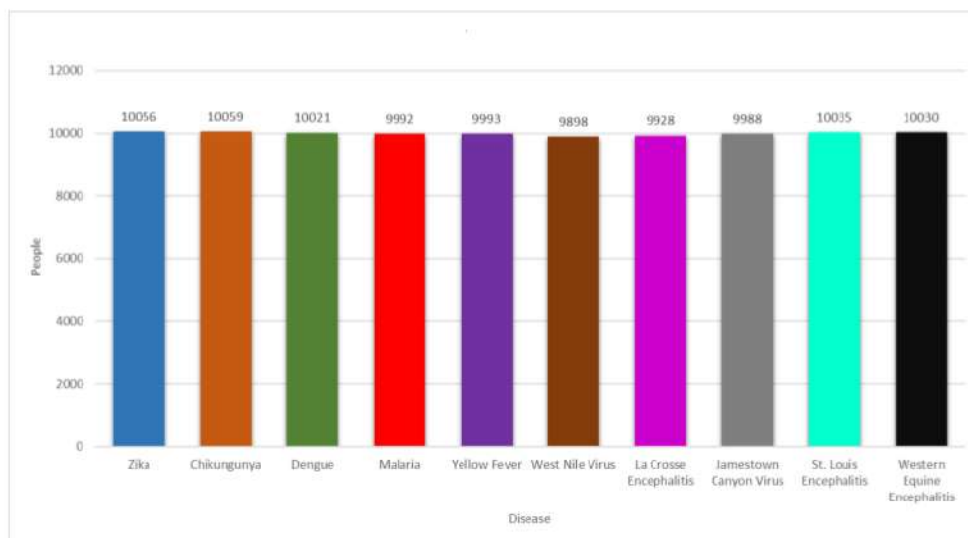


The DP algorithm flips a coin for each person, to choose whether to keep the original disease or to randomly assign them a new one with the toss of a 10-sided virtual dice. This procedure is embedded in the function `plot_with_DP`, which also produces a histogram of the new dataset.

## Our prototype



## ...and with differential privacy

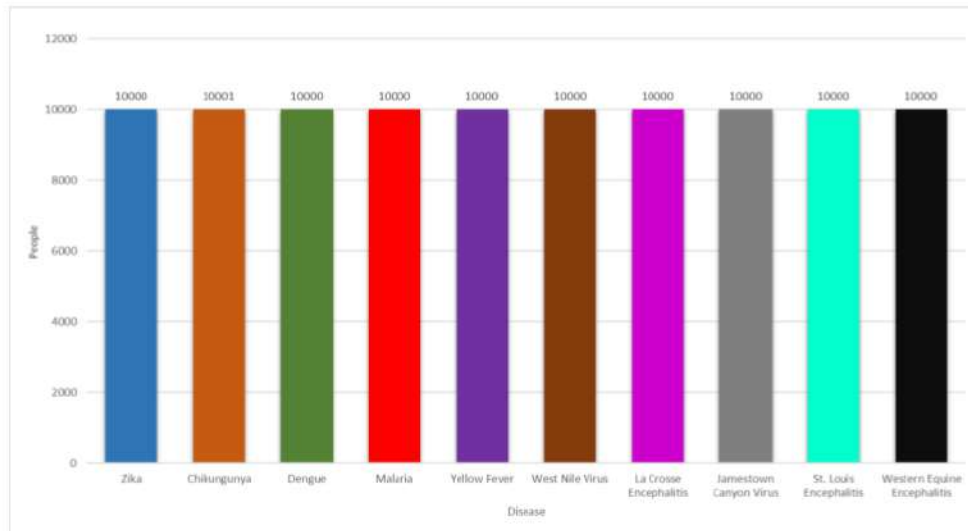


As seen above, there are no statistically significant deviations from the uniform distribution.

This approach is also useful to anonymize dynamical datasets.

For example, "Bob"s data will be appended to the dataset:

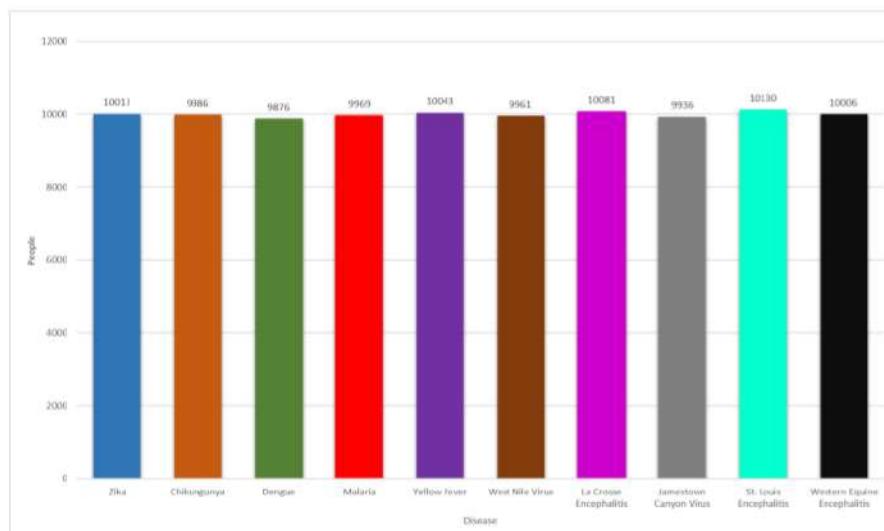
## For dynamical datasets



Bob is readily identifiable as having disease number 2 since it is the only histogram's bar having a value of 10001.

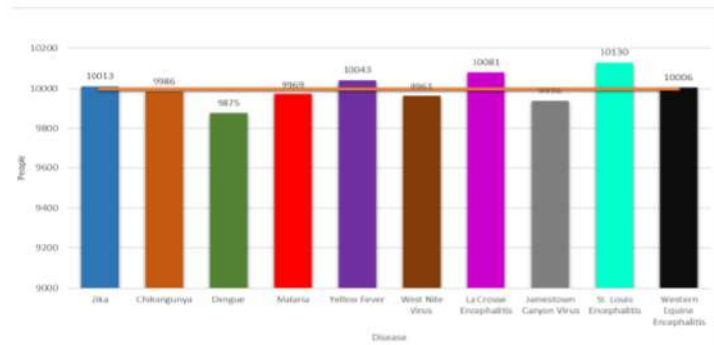
With the team's algorithm instead, there would be no information on Bob, because each individual person's information is concealed:

## ...people information are concealed

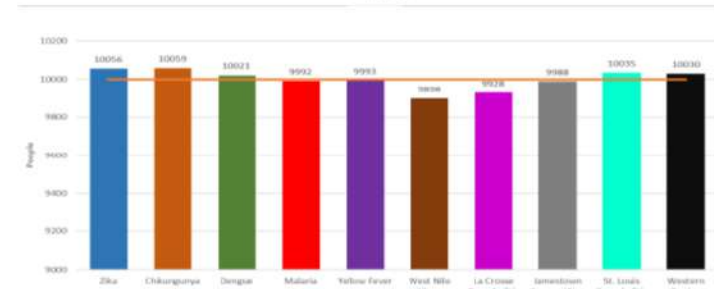


This initial algorithm was implemented using the pseudorandom number generator of Linux and the truly random bits of Random Power. The histogram generated with Random Power is "flatter", has less variation from the uniform distribution, thus promising improved accuracy in the results that can be derived from the dataset.

## PRNG



## Random Power



The code using Linux's PRNG is publicly accessible in the following link:

<https://github.com/antoniopelusi/simpleDP/blob/main/dp-decimal.ipynb>

While the Random Power version using truly random bits involves proprietary material from the company, and cannot be freely shared due to Intellectual Property reasons.

## Student Project EXPO

The team had the opportunity to showcase the solution and prototype at the ATTRACT Student EXPO 2023, which took place at the Almacube site in Bologna. This event brought together all the students who participated in the ten ATTRACT Academy projects across the EU.

As part of the expo, the team created visual elements to present the proof of concept for Maia, like the Maia Logo and the team poster.



# MAIA

UNLOCK THE POWER OF DATA

Nowadays we produce a huge amount of medical data that could be invaluable for research, and yet this potential lies untouched because of the sensitivity of such information.

## UNLOCK IT.

Our solution fueled by Random Power allows collection, access and analysis of aggregated health data.

MAIA opens up the future of secondary use of medical data for faster diagnosis, better therapies and more effective policies.



### RANDOM BITS

|                          |                 |
|--------------------------|-----------------|
| Giovan Matteo MARCIAS    | Olivia RICCIONI |
| Luis Felipe ALVAREZ VEGA | Antonio PELUSI  |
| Davide POMANTE           | Valeria ROSSI   |

All participants had the chance to meet each other, and share ideas, experiences, and knowledge. The EXPO started with a four-minute pitch made by all the different students, where each group explained its challenge and the solution they developed.

After the pitch, all students moved to the booth spaces where each team could better present or answer questions or doubts in a direct setting to the people who were interested in their project.



## Third milestone: The end of the project

At the end of the project, as always, everyone started to think about what we've done in these months during this journey called Cbi Attract. If our roots grow from understanding the Random Power technology our solution, Maia, even if it represents the concrete representation of what we've done, it involves so much more because inside there are a multitude of aspects that converge on it. What we have unlocked is not just the power of data (this is the Maia's motto) but a lot of energy and strength with which we have broken a lot of walls of our previous comfort zone creating a new state of consciousness about our possibilities.

We are glad for what we have done and all the experiences that we have lived during these months, but we are also good with ourselves and aware about all the challenges that we have matched and for the result that we have reached. It was not always easy but facing all the difficulties led us to grow as individuals, as a group and strengthen our ideas.

Another important aspect concerns how we as a single person with different backgrounds and knowledge made a collaboration which has allowed us to leverage a wide range of perspective, knowledge and skills resulting in innovative solutions that transcend traditional boundaries of research. Creating interdisciplinary groups, for sure, represents an innovation in the university research that provides a consistent qualitative improvement and an upgrade in our capacity to analyse the complexity of the phenomena regarding our world. The most important innovations that will characterize our future will be born from collaboration among groups of researchers.

We also want to thank the CEO of Random Power Massimo Caccia, who was an important support till the end and believed in a possible positive future for our idea.

Another thank you is for our coach Eleonora Musca who led the whole project and helped us many times trying to always look for things that can enhance even the difficult situations.

