DATA FOR DEVELOPMENT:

Building innovative solutions in energy, agriculture, and road safety at Innovative Solutions4Moldova Hackathon

2024

Data for Development:

Building innovative solutions in energy, agriculture, and road safety at Innovative Solutions4Moldova Hackathon

Authors:

United Nations Development Programme in Moldova

(Iulia Tvigun, Eugen Cepoi, Adam Kadduri)

UNDP is the leading United Nations organization fighting to end the injustice of poverty, inequality, and climate change. Working with our broad network of experts and partners in 170 countries, we help nations build integrated, lasting solutions for people and the planet.

Introduction

The COVID-19 pandemic, the war in Ukraine, and other trickling crises have highlighted people's and government's reliance on digital tools and the ever-increasing importance of data in our lives. A well-functioning national data ecosystem, tightly linked to crucial government decision-making processes, is critically important in ensuring that policies are human-centred and human-driven and in improving the quality of governance.

The Government of the Republic of Moldova set key priorities regarding data management throughout its National Statistical System Development Strategy 2023-2030 and National Digital Transformation Strategy 2023-2030, emphasizing the importance of improving statistical and administrative data collection and transparency.

While navigating complex interventions in governance, energy and climate, and inclusive economic development, UNDP Moldova identified several areas of development that require specific technical assistance and could greatly benefit from improved data collection and analysis mechanisms and tools.

Building on UNDP's collaboration with the Ministry of Labour and Social Protection in establishing the Energy Vulnerability Reduction Fund (EVRF) and platform, which aggregates various data on people's energy consumption, household composition and income levels, energy sources and appliances, it became natural to ask how we can use the data collected by the platform, and such tools as artificial intelligence and machine learning, to improve the accuracy of the information entered by the people, improve the classification mechanism of energy vulnerability, and predict energy consumption, and therefore, levels of compensation for those who are genuinely energy vulnerable.

Similarly, from UNDP's collaboration with the Ministry of Internal Affairs of the Republic of Moldova, Orange Moldova and Monday Lift on designing a Refugee Mobility Dashboard, which provided a situational analysis of refugees' location in the country, it became clear that a significant amount of data on human mobility, and therefore safety, was being harnessed by the Ministry and its subordinate agencies. This led to exploring how machine learning and IoT technologies can help us build real-time accident prediction and prevention systems to reduce car accidents and fatalities, thereby improving road safety.

Lastly, the compound refugee and energy crises increased food prices and curtailed economic growth and inflation, and the Republic of Moldova's vulnerability to droughts, etc., could not have avoided leaving its mark on the food security of the country. The Government must develop its capability to map, monitor, and anticipate shortages of supplies, and it must further increase its capacities to plan and target subsidies transparently to most in-need agricultural products affected by the abovementioned factors and climate change. These objectives also align with the Government of Moldova's objectives of integrating into the European Union, which put in place similar information systems and requires every country within the EU and those countries that wish to join the EU to do the same. The question remains: How can we improve land management and farming practices and increase crop yield and sustainability using data, satellite imagery, and artificial intelligence?

The grasp over the sectors through UNDP-led digitalization initiatives generated vast amounts of raw datasets, representing an untapped source of evidence and actionable insights capable of driving sustainable development and improving the well-being of its citizens across these critical sectors.

Some specific datasets targeting the challenges described above were harnessed through the power of collective intelligence and innovation – and delivered to the public through the Innovative Solutions4Moldova Hackathon, a three-day event that took place on 10-12 November 2023 in Chisinau, the Republic of Moldova.

UNDP Moldova organized the event with financial support from the German Federal Ministry for Economic Cooperation and Development (BMZ) in partnership with the Government of the Republic of Moldova,

Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Technovator, Premier Energy, Waze, and Orange Systems, which also hosted the event.

The hackathon brought together about 40 software developers, IT engineers, designers, data scientists and analysts, domain experts, business analysts, IoT professionals, blockchain specialists and innovators, as well as professionals in the field of artificial intelligence, data, and the design of innovative technological solutions. The purpose of the hackathon was to harness the power of technology and collaboration and bring together professionals to develop digital solutions that contribute to the sustainable development goals of the Republic of Moldova in three critical sectors – energy, agriculture, and road safety. All participants had access to relevant data sets to analyze and develop innovative digital solutions to address the issues and challenges UNDP and its partners put forward in this hackathon.

The logistics of a hackathon

All participants were invited to come individually or in already-formed teams. For those who came individually, UNDP Moldova and Technovator helped to find other potential team members and form a team. The ten teams that were eventually formed could select from one of the 12 challenges across three subjects (energy, agriculture, and road safety) that UNDP Moldova and its partner organizations provided to work on.

A pool of 17 mentors (in data, artificial intelligence, user experience and user interface, product management, agriculture technology, experts in agriculture, road safety and energy) and three trackers were available for consultations throughout the hackathon. Every team was invited to schedule at least a 30-minute consultation a day with any of the proposed mentors. The mentors were helping teams understand and address the challenges, refine their ideas, test the prototypes, and prepare their final pitches.

Overall, three trackers, each responsible for one subject, were in constant contact with the teams, offering various types of support, from logistics and resources to subject matter expertise, to ensure that all teams were on track with developing their proposed solutions.

Depending on the selected challenge, all teams were given access to five datasets with over 300 thousand data records for all subjects (energy, agriculture, and road safety). The teams were also invited to use other open-source data sources and sets, such as roads data from OpenStreetMap and State Road Administration (harta.asd.md), and data about climate from clima.gov.md.

40 participants

10 teams

12 challenges

5 datasets

17 mentors

3 trackers

9 jury members

9 solutions

The objective of each team was to build functional prototypes and present their solutions in a pitching session to a jury on the final day. The pitched prototypes were evaluated by an independent jury of nine people (representatives of five ministries and several state agencies, civil society organizations and the tech community) – who selected one solution in each of the three abovementioned subjects. Each team presented their ideas on stage (using slides), showing their data, problem description, competitive advantages, and approach. The jury evaluated each solution, and the judging criteria were divided into four groups: 1) social impact (20 pts), 2) UX/UI (20 pts), 3) applicability (25 pts) and 4) minimum viable product (35 pts).

The Development Challenges

ENERGY

The rationale

In the energy sector, Moldova has experienced an energy crisis since February 2022 due to the war in Ukraine and Moldova's significant dependence on energy imports from Russia. The immediate effect of the energy crisis was further exacerbated by a rapid increase in the rate of inflation, driven by a quadrupling of electricity prices on the back of the gas supply shock. Combined with some pre-existing socio-economic challenges, this has seen the level of energy poverty rise to over 60 percent in Moldova. Although many initiatives have since been launched and implemented, such as the Energy Vulnerability Reduction Fund (EVRF) – a governmental compensation programme, the energy sector remains a top priority for the Government and solutions are needed in this sector to understand both the energy vulnerability of citizens and what solutions for renewable energy sources are available to try, pilot and test in Moldova. Specifically, there is a need to understand the process of the EVRF vulnerability and its classification hinges on the accuracy and truthfulness of the data entered by consumers that are part of the EVRF compensation scheme. Misrepresentations, whether unintentional (due to human errors) or deliberate, can lead to

misclassifications, affecting the distribution of benefits and potentially compromising the goals of the EVRF.

Data for governance could play a crucial role in reducing energy vulnerability in Moldova. By collecting and analyzing data on energy consumption, energy efficiency, and renewable energy sources, policymakers could better understand the country's energy landscape and identify areas where interventions are most needed. This could help prioritize investments in energy efficiency measures and renewable energy projects and inform public awareness campaigns to promote behaviour change and energy conservation.

However, significant data collection issues need to be addressed in Moldova. The country needs comprehensive and up-to-date energy data, and collecting and analyzing data is challenging due to limited resources and technical capacity. Improving data collection and analysis systems will require significant investment and technical assistance. Still, it is a crucial step towards reducing energy vulnerability and promoting sustainable development in Moldova.

Developing an AI-powered system to continuously validate and verify real-time data entries from household energy consumers on the compensatii.gov.md platform. The goal is to ensure that the provided information, such as income levels, family size, heating system type, and more, is consistent, accurate, and free from anomalies or potential manipulations.

The Challenges

1.

AI-assisted data verification for accurate registration

The process of energy vulnerability classification hinges on the accuracy and truthfulness of the data entered by consumers.

Misrepresentations, whether unintentional (due to human errors) or deliberate, can lead to misclassifications, affecting the distribution of benefits and potentially compromising the goals of the Energy Vulnerability Reduction Fund. An AI system capable of verifying data entries in real time can significantly reduce such discrepancies.

2.

Dynamic vulnerability category classifier using machine learning

The Energy Vulnerability Reduction Fund's success is accurately categorising consumers based on their energy vulnerability.

With changing economic conditions, seasonal adjustments, and evolving government regulations, it's crucial that the classification model remains agile. A dynamic ML model can intelligently adjust to these changes, ensuring that consumers continue to receive the appropriate benefits and support.

3.

AI/ML-driven anomaly detection for Energy Vulnerability Category Assignment

Assigning households to the correct energy vulnerability category is crucial for the appropriate allocation of resources and benefits.

Data anomalies can lead to households being assigned to the wrong category, either depriving them of necessary support or allocating resources that aren't needed. An efficient AI-driven system can help identify and rectify such anomalies before finalising category assignments.

4.

Predictive analysis of energy consumption patterns

Energy consumption patterns provide essential insights into households' vulnerability status.

Accurate prediction of these patterns helps better resource allocation, timely interventions, and strategic planning for energy supply and subsidies. By harnessing ML for predictive analysis, the Energy Vulnerability Reduction Fund can ensure proactive support to households, reducing the risk of energy crises or hardships.

5.

Using geospatial and infrastructure data, evaluate the ideal location for new power plants, wind farms, or solar installations.

With increasing global emphasis on sustainable energy, there's a dire need to strategically place renewable energy installations to maximize efficiency while minimizing environmental and infrastructural impact.

The optimal selection of sites can result in significant savings, increased energy production, and a reduced carbon footprint. By harnessing the capabilities of ML and geospatial analysis, decision-makers can make data-driven, precise choices for future energy projects.

The Datasets

In partnership with the Ministry of Labour and Social Protection, UNDP Moldova provided the participants with the data collected from energy consumers via the Energy Vulnerability Reduction Fund digital platform (known as SIVE and accessible to the public at compensatii.gov.md). Key fields included the applicants' year of birth, centralized heating consumption volumes during the cold period, electricity consumption volumes during the cold period, and gas consumption volumes during the cold period. It also details the type of heating available in the household, the total number of residents, applicants' sex, household street, primary heating source, and average monthly income for the last six months. This data is used by the mechanism employed by the Ministry of Labour and Social Protection of the Republic of Moldova to assess and provide necessary compensations for heating and energy costs during the colder months. This dataset was useful in Challenges 1 through 4.

For Challenge 5, related to green/renewable energy, the participants were given access to various open data sources to help evaluate the best locations for new power plants, wind farms, or solar installations. Climate data sourced from clima.gov.md provided historical data on wind patterns, precipitation levels, and temperature variations, which are essential for determining the suitability of renewable energy projects. Additionally, road data, accessed from OpenStreetMap and the State Road Administration (harta.asd.md), was crucial for ensuring that the chosen sites were accessible for construction, maintenance, and transportation of materials and equipment. These resources were vital for analysing and selecting optimal sites, considering factors like accessibility and climatic conditions.

The Solutions

Team Wise Energy

Team Wise Energy (winner) developed an ML model that predicts the most suitable locations for wind farms and solar installations based on historical weather data provided for the city of Chisinau. The team used two detailed datasets – one for solar energy (which included such information as the angle of rays, amount of UV, wind direction, temperature, length of day, etc.) and one for wind energy (which included information about various directions of wind). They built their solution using Arima, XGBoost, Python, Java, and ColabResearch Google. The team also concluded that to improve the accuracy of their prediction model, they needed more data, such as network nodes of different voltage types, access paths, licences issued and licences deployed, etc.

Team
Power Explorers

Team Power Explorers developed a model (and a dashboard) that predicts, based on 5-month historical consumer data, how much energy a consumer will need for the next month to help energy providers and consumers manage their energy consumption and, therefore, predict the amount of compensation the Ministry of Labour and Social Protection/the Government of the Republic of Moldova will have to pay. The jury noted that this solution could also work as an anti-fraud mechanism for cash payment for households that use wood/solid fuel to heat their homes because it can tell if a household is genuinely using wood as their (primary) energy source.

Team
Proton Energy

Team Proton Energy developed an Intelligent Anomaly Detection System for the Governmental on-bill compensation platform for energy vulnerability. The system automatically analyzes and discovers data patterns and automatically classifies fraudulent/abnormal/suspicious records. The improved UX/UI of the system/dashboard reduces by 30% the time spent on operational activities of the Ministry's staff. If automated further, it can be trained to correct statistical anomalies by itself. The team used fully open sources (SCIKIT-Learn), TensorFlow, and FASTAPI to train network statistical modules and wrap them all into the backend and REACT for the system's front end.

Lessons learned

About the datasets

The datasets provided for challenges 1-4 were small (only five months, i.e., November – March 2022-2023) and contained missing values. This means that the team had to do a great deal of pre-processing to use this information. However, even with these flaws, the AI model developed by the Proton Energy Team was able to find/cluster the data and test the solution accurately.

The datasets provided for challenge 5, which deals with identifying ideal locations for new power plants, wind farms, or solar installations, are noted. To improve the accuracy of their prediction model, they needed more data, such as network nodes of different voltage types, access paths, licences issued and licenses deployed, etc., which were missing from the provided datasets.

About the solutions

Digital solutions can be versatile. While Energy Explorers developed a solution to predict energy consumption in households and, therefore, predict the level of energy vulnerability and respective on-bill compensation, the jury noted that the developed model could also work as an anti-fraud mechanism for cash payment.

By predicting the energy consumption for any household and seeing a difference in the consumption pattern, the Ministry of Labour and Social Protection can determine if the household is genuinely using wood as their primary energy source to heat the homes and reject a fraudulent claim for such compensation.

Cash payment for solid fuel is a mechanism that has always been available to households that use solid fuel (wood) to heat their homes in Moldova to compensate for their expenses for procuring timber. The applicants had to file a paper-based claim at the Ministry of Labour and Social Protection to benefit from such a payment. However, with the start of the heating season 2023-2024, the Ministry of Labour and Social Protection digitalized this mechanism by adding such an option into the digital form available on the Energy Reduction Vulnerability System known as compensatii.gov.md.

AGRICULTURE

The rationale

In modern agriculture, there's an urgent need to precisely map and delineate agricultural fields or parcels to aid in better land management, efficient farming practices, and integrated data analysis. Traditionally, cadastral maps have been the go-to tool for delineating property and land ownership in Moldova. Still, the tool is somewhat lacklustre because a great deal of data was wrongly recorded when cadastral maps were first introduced, starting with the year 200. No updates or upgrades of the data were made for years.

A cadastral map is a detailed, large-scale map that delineates the boundaries and ownership of land parcels. It is an essential tool for land administration, urban planning, property taxation, and legal disputes regarding land ownership. Government authorities or land registry offices maintain them, and they can be used to provide an accurate and official record of land ownership and boundaries.

On the other hand, rapid advancement in satellite technology and high-resolution imagery is becoming increasingly accessible to public institutions and private businesses, including farms and farmers. It can help us gain insights about the land below in unprecedented detail and thereby can be a tool that can be easily used by farmers, landowners, and related authorities to ensure that their agricultural boundaries are accurate, up-to-date, and harmonized with cadastral information which could lead to a new era of precision agriculture.

Therefore, the challenge is to develop a solution that accurately delineates and maps agricultural field boundaries. The solution should accurately identify and map the boundaries of farm fields or parcels, compare and contrast these boundaries with traditional cadastral delineations, highlight any discrepancies and offer actionable insights for landowners, farmers, or agricultural agencies. This should be developed as a prototype or easy-to-use tool that farmers, landowners and related authorities can use to ensure their agricultural boundaries are accurate, up-to-date, and harmonized with cadastral information.

The Challenges

1.

Resolving discrepancies: Precise agricultural field mapping vs. traditional cadastral delineation

Traditionally, cadastral maps have been the go-to for delineating property and land ownership.

However, in modern agriculture, there's an urgent need to precisely map and delineate agricultural fields or parcels to aid in better land management, efficient farming practices, and integrated data analysis.

2.

CropSight: Satellite imagery for Crop Detection and Management

With the rapid advancement in satellite technology, high-resolution imagery is becoming increasingly accessible.

This presents an opportunity to gain unprecedented insights about the land below. For the agricultural sector, recognizing crops from space and subsequently managing them using this technology could lead to a new era of precision agriculture.

3.

SoilSense AI: Predict Soil Health Metrics for optimized farming

Design a solution that analyzes data from various sources (such as imagery, sensors, and historical records) to predict soil health metrics like pH, organic matter, and nutrient content and recommends actionable steps based on predictions for farmers to optimize crop yield and soil sustainability.

4.

Pestpredictor: AI-driven early detection and management for pest outbreaks

Develop a system that uses AI to analyze imagery and environmental data for early signs of pest activity, predicts potential pest outbreaks based on patterns, weather data, and historical trends, and offers real-time pest management and containment solutions.

5.

YieldMax: Machine learning forecasts for crop yields

Design an AI-powered solution that analyzes current season data, including weather patterns, soil metrics, and pest activity, to predict crop yields and offer recommendations for maximizing yield based on predictive analytics.

The Datasets

Three datasets were provided to the participants for the subject of technology in agriculture.

The satellite imagery from Sentinel-2 consisted of georeferenced satellite or drone images from the Sentinel-2 platform, available in GeoTiff format with 1, 2, and 3 bands. The photos could be used for various applications such as land use analysis, environmental monitoring, and agricultural assessment. The high-resolution geospatial data provided valuable insights for mapping and monitoring landscape changes.

The Cadastral Parcel data set from the Cadastre Department of the Agency of Public Services included detailed information about land parcels. It contains fields such as the area of the land, the unique cadastral code assigned to each property, land use type describing how the land is utilized (e.g., constructions, garden, development), and the type of property, indicating whether the property is public or private. This data is crucial for urban planning, property management, and legal purposes, providing a comprehensive view of land distribution and usage.

The “Images from Drones” data set focuses on agricultural applications, providing high-resolution aerial imagery captured by drones. These images are essential for monitoring crop health, assessing soil conditions, detecting pest infestations, and optimizing irrigation practices. The georeferenced imagery ensures precise mapping of fields and accurate assessment of agricultural conditions.

The solutions

Team WISE Agricultura

Team WISE Agricultura developed an AI-based algorithm called “Agro Data” that, based on weather information and cadastral maps, can forecast agricultural production and estimate production risks by crop for up to one year. They used two datasets – weather information and cadastral maps- and Google Earth, Geoportal, DJI Terra, meteo, and Arima to develop and train the algorithm. The solution can also work based on imagery of agricultural lands provided by drones and provide on-demand agricultural advice to Moldovan farmers.

Team DEAL

Team DEAL developed a digital solution called Smart Land Management that compares imagery of agricultural lands from the Public Institution Cadastre and imagery produced using drones. The solution shows land discrepancies and provides the opportunity to improve agricultural land management. It also includes information on the types and health of detected crops. The team used TS, Docker, and Meta to develop the algorithm. The solution also has the potential to provide customized and personalized agricultural advice to farmers if further development is needed.

Team
Strawberry Vision

Team Strawberry Vision (winner) worked on two challenges simultaneously – crop detection and management and prediction of soil health for optimized farming. The team built one platform that farmers and public officials could use. The platform's crop detection, management and prediction side utilizes satellite imagery to identify and manage different crops efficiently, offering insights for improved crop management. The soil health prediction algorithm aids farmers in making informed decisions for fertilization, irrigation and crop rotation based on predictive soil health metrics. When used by public officials, the platform will help them decide what agricultural lands need intervention.

Lessons learned

The participants found it challenging to understand the content and representation of cadastral data. They needed extensive explanation and guidance from data experts and mentors, including the imagery and data provided by Drone Assistance. The company representatives had to explain what data can be extracted from their maps, which leads us to the conclusion that a non-trained farmer could not operate with the data either.

Similarly, with the satellite imagery provided by Sentinel, the mentors had to explain to the participants how to correctly use the colour bands to highlight specific indicators in the context of the hackathon challenges – different crops. It is paramount that when working with satellite imagery in the future, hackathon organizers and mentors provide more explicit instructions on the use of satellite images so that participants can process them and generate images and data that represent the delimitation of land and crops.

Another difficulty encountered in this challenge, notably Challenge 2, is the general insufficiency of data and, therefore, low accuracy of the solutions. The solution developed by the Strawberry Vision, although selected as the best, had a low accuracy of the model that detected the crops correctly due to a modest number of land images. The provided dataset must have had more images of the crops in different seasons/periods/years to have been sufficient for the model to be trained and provide accurate results. There are no institutions, at the moment, in the Republic of Moldova that can provide such data; however, the Agency for Agricultural Intervention and Payments (AIPA), subordinated to the Ministry of Agriculture and Food industry, could potentially collect data (take pictures and store metadata, drone imagery) of the agricultural lands they are subsidizing as a bi-product of their business as usual, and subsequently make it available to farmers, academia, the public, etc for research and innovation purposes.

Similarly, open-source data can be helpful, but due to its constant update and evolution, we could envisage a stalling of the accuracy of the built model as it will require government human resources who will be able to continually feed the model to keep the level of accuracy up.

Lastly, on a positive note, it must be noticed that while the data was not used to its full potential, the winning team proved to be creative and used the dataset intended for Challenge 5 (YieldMax) to solve Challenge 3 (SoilSense) and made the connection between soil condition and vegetation starvation.

ROAD SAFETY

The rationale

In recent years, Moldova has witnessed a noticeable fluctuation in car accidents. A brief overview of the car accidents data from 2014 to 2022 reveals periods of increase and decrease in numbers, while the numbers peaked in 2017 with 2,640 accidents. In 2022, there were 2,292 car accidents; 2022 represents an 11% increase from 2021, and by October 2023, the nation has already recorded 1,333 accidents. These alarming figures underline an urgent need for intervention.

By predicting accidents in real time, proactive measures can be taken to prevent accidents and the loss of lives, reducing the human and economic costs associated with road accidents. One solution can be to design a real-time accident prediction and prevention system using Machine Learning and IoT technologies. Creating a model that can predict the likelihood of accidents at specific road segments and intersections and develop solutions that can actively alert drivers, traffic management systems, and emergency services to prevent accidents could lead to fewer accidents and better road safety for drivers and pedestrians.

The challenges

The proposed challenges for approvement were the following:

1.

Real-time accident prediction and prevention

Accidents on roads lead to injuries, fatalities, and traffic congestion.

By predicting accidents in real time, we can take proactive measures to prevent them, reducing road accidents’ human and economic costs.

2.

AI-powered pedestrian and cyclist safety

Pedestrians and cyclists are at higher risk of accidents due to their vulnerability.

AI can help improve their safety by detecting potential dangers and providing alerts.

The Datasets

Regarding mobility and road safety, the organizers provided the participants with two datasets, which both could have served to design solutions for Challenges 1 and 2.

The “Road accidents” data set from the National Inspectorate of Public Security provides detailed information on road accidents. Key fields include the cause of the accident, type of accident, participant category, such as minors, cyclists, pedestrians, or drivers, and the date of the accident. It also includes geographic coordinates, weather conditions, the number of fatalities, the number of injured, the number of vehicles involved, road conditions, and the type of vehicle, such as bicycles or motorbikes. This data is essential for traffic incidents analysis, improving road safety measures, and formulating public safety policies.

The “Data on Roads” from the State Road Administration contained detailed information about road infrastructure. Key fields include whether there is a bridge on the road, the code defining the map layer, and the type of road, such as motorway, residential, or living street. It also includes the maximum speed limit and the road’s name, i.e., Str. Mihai Eminescu, whether the road is a one-way road, road index numbers, such as M1, R16, G130, and L123, and whether there is a tunnel on the road. This data is crucial for transportation planning, road maintenance, and navigation systems.

The solutions

Team Driver Focus

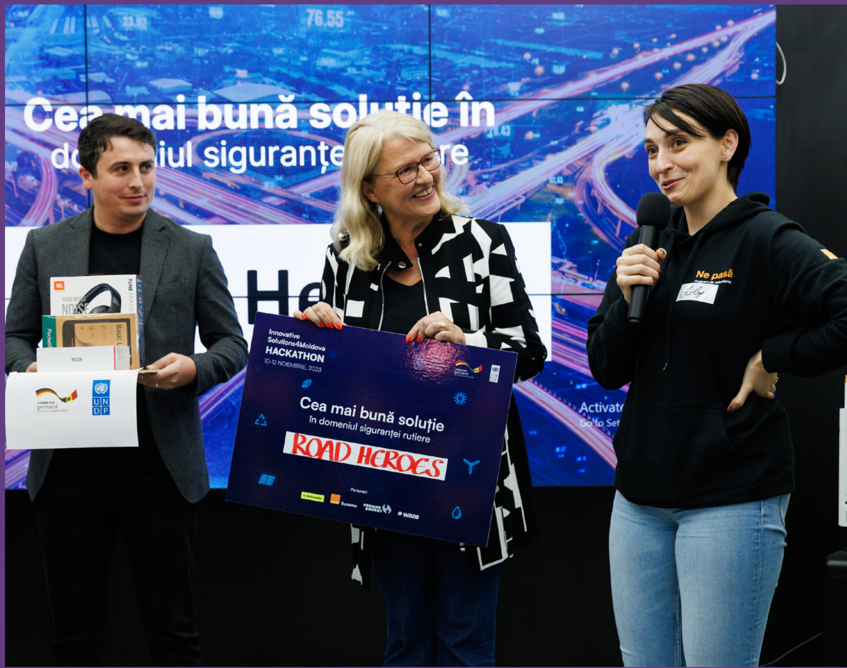
Team Driver Focus developed an AI-assisted mobile application called Safe Road, which enables all traffic participants (vehicle drivers, bike drivers, cyclists, pedestrians, etc.) to video-record and report in real time any traffic abnormalities and hazards, including infrastructure flaws. The AI automatically detects these flaws and confirms them with the drivers, pedestrians, etc. Public authorities and first responders then become alert and act to remove any such hazards.

Team Street Safers

Team Street Safers developed an algorithm to predict and prioritize traffic lights for pedestrians and bicyclists. Built using Invipo, the solution uses historical traffic accident data, accesses real-time traffic light data and electronic sights, and sends emergency notifications. The solution can also be scaled up to ease traffic for vehicles, but the focus is on pedestrians and alternative transportation.

Team Road Heroes

Team Road Heroes (winner) with the 'Safe Embrace' solution, an algorithm that analyzes historical road accident data, predicts the probability of road hazards and accidents and alerts drivers in real-time. The solution can be integrated with any mobile map application, such as Waze and Google Maps. The team found that most accidents resulting in serious injuries or death happened on weekdays, in the afternoon, on dry roads and in sunny weather, contrary to popular belief that accidents happen because of bad weather or roads. The team concluded that to scale the solution up for all types of mobility participants (pedestrians, cyclists) and to improve the accuracy of predictions, they would need additional data, such as traffic density, road infrastructure, type and condition of the vehicle, and profile (pedestrian, cyclist, car driver, etc.)



Lessons learned

The hackathon showed that the participants, once again, used mainly/only the data provided by the organizers (“Road accidents”, “Data on Roads”) rather than open data such as OpenStreetMap. The analysis was not as deep as expected, and the solutions were proposed based on the problem/business model rather than on the insights drawn from the data.

Nevertheless, the participants observed that more data on road design and infrastructure is needed to enable decision-makers and policymakers to understand better why road accidents happen. The winning team found that most accidents resulting in serious injuries or death occurred on weekdays, in the afternoon, on dry roads with sunny weather, which is contrary to popular belief that accidents happen because of bad weather or bad roads; the data shows that the answers as to why these accidents happen could lie in either road design or drivers’ behaviour. If this assumption is correct, we need to determine what new data type needs to be collected: data about driver experience, age, and the number of incidents in the past (data from insurance companies). Additionally, evaluating the road infrastructure with accident hotspots and conditions would be essential to understand the underlying causes comprehensively. This data could be vital as it would allow for identifying high-risk areas and patterns in driver behaviour that contribute to accidents. Other valuable data might include traffic density, vehicle types involved in accidents, road maintenance schedules, and even environmental factors like lighting conditions at different times of the day. More effective interventions can be designed to improve road safety and reduce accident rates by collecting and analysing this information.

The hackathon also helped determine that to improve the quality of accident prediction; the data must be collected for at least a year rather than five months, as provided at the hackathon. Additionally, as pointed out by the winning team, the data could reflect more diverse traffic participants’ profiles, such as users of alternative transportation means (scooters, etc.).

Overall Recommendations

The organization of the Innovative Solutions4Moldova Hackathon was an overall successful event that resulted in 12 innovative solutions, designed with artificial intelligence, machine learning, satellite, and drone imagery, etc. and many datasets and assistance provided by UNDP's Governmental partners – Ministry of Labour and Social Protection, Ministry of Internal Affairs, Cadastre Department of the Agency of Public Services, National Inspectorate of Public Security, State Road Administration, etc. and private partners Orange Systems, Dron Assistance, Premier Energy, Waze, Automobile Club from Moldova, Technovator.

Generally, the observations and recommendations can be summarized as 'we need more data' and 'we need better data'. While the participants could design solutions with the datasets provided, the data was not used to its full potential due to either the data not being 'historic'-enough – reflecting a more extended period or due to its quality (erroneous data), or sometimes being insufficient or missing altogether. There are several recommendations that we can draw from the hackathon organization. Still, the recommendations below relate to the datasets and how the data is collected.

The Agency for Agricultural Intervention and Payments (AIPA),

subordinated to the Ministry of Agriculture and Food Industry, **should collect data on the agricultural lands** they are subsidizing as a byproduct of their business as usual. When farmers are applying for subsidies, and AIPA is verifying the lands, it could take pictures either with the camera or drones and store metadata so it can subsequently make it available to farmers, academia, the public, etc., for research and innovation purposes, including hackathons, such as ours. This will enable a more accurate and effective detection of crops.

The National Inspectorate of Public Security (INSP) must collect more data on road design and infrastructure and more profiles of traffic participants,

including users of alternative transportation (such as scooters), to improve accident prediction models. The winning team in the Road Safety Challenge found that most accidents resulting in serious injuries or death happened on weekdays, in the afternoon, on dry roads with sunny weather, which is contrary to popular belief that accidents happen because of bad weather or faulty roads, the data shows that the answers as to why these accidents happen could lie in either road design or drivers' behaviour. If this assumption is correct, we must determine what new data type must be collected.

The quality of the datasets must be improved in terms of their age and accuracy.

The participants needed some help with the datasets provided for Energy Challenges 1-4 by the Ministry of Labour and Social Protection, harnessed through the Energy Vulnerability Reduction Fund, which contained missing values, which meant that the teams had to spend valuable hackathon time pre-processing. Similarly, several teams noted that the provided data for various challenges across all three subjects were young/small. For instance, the data provided for the Energy Challenge was, by design, reflecting the information on gas and electricity consumption for only five months, i.e., November – March 2022-2023, which is the cold season, for which the compensations are requested. Perhaps it would have been helpful to provide the participants with energy consumption data for the rest of the months of the year and/or provide energy consumption data for the cold season for several consecutive years. Other participants found similar issues with the data provided regarding the challenges to agriculture (see the first recommendation above).

Promote the use of and provide direct links to more open data at the beginning of the hackathon.

Generally, the participants mostly used the datasets provided by the hackathon organizers. However, all three challenge themes could have benefited from complementing the existing data with open data, such as Open Street for the Road Safety Challenges, for the Energy Challenge 5 (Using geospatial and infrastructure data, evaluate the ideal location for new power plants, wind farms, or solar installations), and Copernicus data (the European Union’s Space Programme) for the Agriculture Challenges.

Organizers and/or mentors must provide more precise instructions on using some types of data and more general assistance to make sense of the datasets.

Such hackathons as the Innovative Solutions4Moldova Hackathons warrant some understating of working with data and data from the participants. The better this understanding is, the better the solutions that hackathon participants develop. However, sometimes mentors must provide more guided, one-to-one assistance (with the teams) when understanding such data as the land and crop imagery harnessed by drones (such as the data supplied by Drone Assistance) or the satellite imagery provided by Sentinel. Alternatively, a separate section at the beginning of the hackathon can be dedicated to working and understating data.