

# Phenotyping Genetic Diseases Through Artificial Intelligence Use of Large Datasets of Government-stored Facial Photographs: Concept, Legal Issues, and Challenges in the European Union

[Peter Kováč](#)<sup>\*</sup>, Bražinová Alexandra, [Varga Ivan](#), Mrázik Lukáš, Aláč Michal, Smatana Martin, [Andrej Thurzo](#)<sup>\*</sup>

Posted Date: 14 April 2023

doi: 10.20944/preprints202304.0344.v1

Keywords: Deep phenotyping; AI; 2D and 3D facial scans; Genetic diseases; Early treatment; Big data; 2D and 3D facial scans; Facial traits; Healthcare; Citizen privacy; Ethical concerns; EU legal framework; Forensic medicine; Orwellian ramifications



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Phenotyping Genetic Diseases through Artificial Intelligence Use of Large Datasets of Government-stored Facial Photographs: Concept, Legal Issues, and Challenges in the European Union

Peter Kováč<sup>1,2,3,\*</sup>, Alexandra Bražinová<sup>4</sup>, Ivan Varga<sup>5</sup>, Lukáš Mrázik<sup>3</sup>, Michal Aláč<sup>6</sup>, Martin Smatana<sup>7</sup> and Andrej Thurzo<sup>2,8,\*</sup>

<sup>1</sup> Department of Criminal Law and Criminology, Faculty of Law, Trnava University, Trnava, Slovakia

<sup>2</sup> forensic.sk Inštitút forenzných medicínskych expertíz s.r.o, Bratislava, Slovakia

<sup>3</sup> Kinstellar, Bratislava, Slovakia

<sup>4</sup> Institute of Epidemiology, Faculty of Medicine, Comenius University, Bratislava, Slovakia

<sup>5</sup> Institute of Histology and Embryology, Faculty of Medicine, Comenius University, Bratislava, Slovakia

<sup>6</sup> Department of Administrative Law, Faculty of Law, Trnava University, Trnava, Slovakia

<sup>7</sup> Faculty of Nursing and Health Professional Studies, Slovak Medical University, Bratislava, Slovakia

<sup>8</sup> Department of Orthodontics, regenerative and forensic dentistry, Faculty of Medicine, Comenius University, Bratislava, Slovakia

\* Correspondence: pkovac@gmail.com (P.K.); thurzo3@uniba.sk (A.T.)

**Abstract:** One in 12 babies is born with a rare genetic disease. Sadly, most cases are undetected until later age, missing time for early treatment and opportunity to prevent complications. Humanity has entered a new era where Big Data collected by governments, including 2D and 3D facial scans, are available. Many rare genetic diseases can be identified by artificial intelligence (AI) analysis of the facial photo. Phenotyping AI utilizations facilitate comprehensive and accurate genetic evaluations. AI processing of this Big Data to identify rare genetic diseases could bring unimaginable benefits to healthcare, although this would be a questionable step in terms of citizen privacy and could lead to future "Orwellian" ramifications with government abuse. Going forward, a balance must be found between protecting the privacy of citizens and the enticing use of AI for their health risks and cost savings through prevention. The unimaginable potential of AI early diagnostics from facial photos also raises various ethical and legal concerns. This paper presents concept, protentional methods and legal and other limitations within EU legal framework in contrast with potential benefits. This paper is focused on AI utilization to early diagnostic of rare genetic diseases. Shift of paradigm in the screening for rare genetic diseases in population with AI face analysis is expected to have a significant impact. The potential of AI algorithms utilizations similar to face2gene app in general population or systematically on Big governmental datasets recording facial traits changes in time can have significant impact on public health but at the same time give raise to profound concern as violation of one's privacy.

**Keywords:** deep phenotyping; AI; 2D and 3D facial scans; genetic diseases; early treatment; big data; 2D and 3D facial scans; facial traits; healthcare; citizen privacy; ethical concerns; EU legal framework; forensic medicine; orwellian ramifications

## 1. Introduction

Researchers often focus on what they can or cannot achieve, often forgetting to consider whether they should. The ethical principles of autonomy, good conduct, non-maleficence and justice apply to science [1].

Unique facial features are recognized symptoms of rare genetic diseases. With the boom of AI algorithms capable of determining phenotypic characteristics of genetic diseases based on facial photos (e.g., Face2gene app), a new era has begun. Rare genetic disorders can now be identified by analysing images of people's faces using a smartphone app powered by artificial intelligence - which has now been shown to be much more accurate than physician's evaluation.

The general public trying out the smart mobile app on themselves and close relatives, paediatricians using it as a practical tool for early detection of rare genetic diseases, and governmental health agencies considering large-scale epidemiological studies in societies with legal and moral constraints - these were the beginnings of deep phenotyping. DeepGestalt technology, and its app Face2Gene, had prominent impact on the diagnosis and management of genetic diseases (FDNA Inc., Boston, MA, USA; [www.face2gene.com](http://www.face2gene.com)) [2,3]. Possibly DeepGestalt was the first digital health technology that can identify rare genetic disorders in people based on facial features alone.

Since 2018 this new AI-powered smartphone app called Face2Gene can help physicians narrow down rare genetic disorders to ensure the best treatment plan for patients. Face2Gene is the first app to identify rare genetic disorders patients from just photos.

The Face2Gene app creates a list of ten genetic mutations suspected in a patient that doctors cannot readily identify. This allows them to create personalized treatment plans based on this information. The app was developed by U.S. based biotech company Facial Dysmorphology Novel Analysis (FDNA). The company believes the technology can save patients long waits and repeated clinic visits before they finally receive a diagnosis - giving them a better chance of early treatment. FDNA has published the results of a "milestone study" showing that facial analysis was able to identify the ten most common such conditions with 91% accuracy on 502 images within the group of topmost frequent rare diseases. FDNA, which offers facial recognition tools to clinical geneticists, developed Face2Gene using its own AI software called DeepGestalt [4,5].

However, the algorithm is only as good as its training data set — and there's a risk, especially where rare disorders affect only small numbers of people. There are several instances when the unbalanced data set composition unintentionally created biased AI [6]. A 2021 study on the performance of 2 prediction models for death by suicide after mental health visits accurately predicted suicide risk for visits of White, Hispanic, and Asian patients, but performance was poor for visits of Black and American Indian/Alaskan Native patients and patients without race/ethnicity reported [7,8]. Training data sets containing mostly Caucasian faces remains a concern. A 2017 study of children with an intellectual disability found that whereas Face2Gene's recognition rate for Down syndrome was 80% among white Belgian children, it was just 37% for black Congolese children [9]. The program's accuracy has improved slightly as more medical professionals upload patient photos to the app. There are now more than 200,000 images in the database [4].

Eventually, FDNA hopes to develop this technology further to help other companies filter, prioritize and interpret genetic variants of unknown significance in DNA analysis. But to train its models, FDNA needs data.

That is why the Face2Gene app is currently available free of charge to medical professionals, many of whom use the system as a sort of second opinion for diagnosing rare genetic disorders.

Possibly soon paediatricians and geneticists will have an app like this and will use it just like a tool more common than stethoscope. The AI approach even today helps doctors to narrow down the possibilities and save the cost of more expensive multi-gene panel testing.

## 2. Concept and Methods

The concept of this paper is to present the possibility to implement a comprehensive population screening with an option to adopt preventive measures that ensure the privacy of citizens. We see two scenarios for common use.

The first option is represented by technological solution that removes all identification of individuals from the AI algorithm output but provides epidemiological information on the prevalence of identified specific genetic diseases in specific age and sex groups. This scenario protects individual privacy while providing public health benefits to the society.

To study prevalence of genetic diseases at the population level, the individual identification is not necessary. To achieve the objective, the target photo database without any link to identity data needs to be screened by the face AI analysis. However, geographical characteristics of prevalence is desired in order to provide important public health information. Thus, the regions with locally increased prevalence of particular genetic disease can be identified. Personal data protection is a

challenge, given the low prevalence of most genetic diseases. Low level of data granularity will be necessary to maintain and guarantee the anonymity of individual cases. Prevalence on municipality or county level may be not sufficiently anonymous as only single case of genetic disease may exist within its population. Administrative region-level prevalence with population above tens or hundreds of thousands of inhabitants may provide for sufficient level of anonymity for most genetic diseases. Nomenclature of territorial units for statistics ("NUTS") used by Eurostat divides territory of the EU and the UK in ninety-two regions at NUTS 1 (major socio-economic regions), 242 regions at NUTS 2 (basic regions for the application of regional policies) and 1166 regions at NUTS 3 level (small regions for specific diagnoses) [10]. NUTS 2 regions meet the above requests.

The outcome of this scenario is the population level prevalence data, without the possibility to identify any individual with particular genetic disease. Mapping the true population prevalence of particular genetic diseases is of high value as it allows for the assessment of validity and accuracy of routine screening.

The second option is the population screening for genetic disorders with the aim to adopt preventive measures for those at risk. In this scenario, the identification would be necessary at certain stage of the process.

In this scenario, data pseudonymisation would be necessary to ensure the privacy of screened population. The Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ("GDPR") governs the processing of personal data in the European Union ("EU"). GDPR Article 4 (5) defines pseudonymisation as processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person. To pseudonymise the target database the identity linked to the photo shall be replaced by some sort of code (e.g., alphanumeric). The identity and the assigned code can be kept in a separate database with secure access, which allows the identification of those at risk if preventive action would be possible and necessary.

The database of mandatory national ID schemes (NID) seems to be the ideal candidate for the population wide testing. But there are legal limitation and consideration for the population screening on these databases. The European Commission already proposed a new regulation to create European Health Data Space ('EHDS'), which would allow researchers, innovators, policy-makers and regulators at EU and Member State level to access relevant electronic health data to promote better diagnosis, treatment and well-being of natural persons, and lead to better and well-informed policies [11].

#### *General legal limitations and considerations*

The presented concept, while already technically possible [with available technologies and means] cannot be implemented without considering applicable legislation.

The access to the NID databases is strictly regulated. Such data are as accurate as possible and updated on a regular basis. Only fraction of data stored in NID would be required for the prevalence study. Test data set shall follow the GDPR data minimization principle and be limited to necessary information only. The minimal dataset should consist of the following:

- a) photos and NUTS 2 location only for the prevalence study
- b) photos and the assigned pseudonymous code for studies with option to identify those at risk and facilitate preventive actions, with separate database allowing identification when necessary.

To create databases required for both prevalence study and identification of those at risk would require amending the NID legislation.

From the privacy legislation perspective (e.g., GDPR) facial photo is personal data as it relates to identified or identifiable natural person. Considering the ultimate aim of proposed study– to identify generic disease, the photo shall be considered data concerning health as it reveals information about

health status. Facial images may be also deemed biometrical data and within the context of the intended aim they shall be considered also genetic data. Due to the sensitive nature of health related, biometric, and genetic data, data processing is subject to more stringent rules to the processing of personal data in general. We need to point out GDPR has also extraterritorial reach if personal data concerns EU data subject.

From GDPR perspective, processing personal data shall be lawful and based on the appropriate Article 6 legal basis including consent, contract to which data subject is party, legal obligation, protection of the vital interest, performance of a task conducted in the public interest or in the exercise of official authority. Health related, biometric and genetic data are deemed special category of personal data subject to GDPR Article 9, allowing to process such data only under certain conditions. Our assessment of data processing legal basis suitability is presented in the Table 1.

**Table 1.** Assessment of GDPR data processing legal title suitability.

<b>GDPR data processing legal basis</b>	<b>Prevalence study</b>	<b>Individual risk study</b>
<b>Article 6 (1) a)</b> <b>Data subject consent</b>	Unsuitable due to lacking possibility to identify data subject	Suitable with opt-in study design
<b>Article 6 (1) b)</b> <b>Performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract</b>	Unsuitable due to the absence of contract	Unsuitable due to the absence of contract
<b>Article 6 (1) c)</b> <b>For compliance with a legal obligation</b>	Suitable after adoption appropriate legislation	Only marginally suitable after adoption appropriate legislation
<b>Article 6 (1) d)</b> <b>To protect the vital interests of the data subject</b>	Unsuitable, need to protect one's life shall exist to exploit this legal basis	Generally unsuitable, need to protect one's life shall exist to exploit this legal basis
<b>Article 6 (1) e)</b> <b>Performance of a task carried out in the public interest or in the exercise of official authority</b>	Suitable after adoption appropriate legislation	Unsuitable due to the disproportional incursions in one's privacy
<b>Article 6 (1) f)</b> <b>Legitimate interests pursued by the controller or by a third party</b>	Unsuitable - interests and fundamental rights and freedoms of the data subject are prevailing	Unsuitable - interests and fundamental rights and freedoms of the data subject are prevailing
<b>Article 9 (2) a)</b> <b>Data subject has given explicit consent</b>	Unsuitable due lacking possibility to identify data subject	Suitable with opt-in study design
<b>Article 9 (2) b)</b> <b>Processing is necessary for carrying out the obligations and exercising specific</b>	Unsuitable – study aim outside of the scope	Unsuitable – study aim outside of the scope



<b>rights of the controller or of the data subject in the field of employment and social security and social protection law</b>		
<b>Article 9 (2) c)</b>		
<b>Vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent</b>	Unsuitable, need to protect one's life shall exist to exploit this legal basis	Generally unsuitable, need to protect one's life shall exist to exploit this legal basis
<b>Article 9 (2) d)</b>		
<b>Legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim related to the member of such entities</b>	Unsuitable, whole study population would require to be members of foundation, association etc.	Unsuitable, whole study population would require to be members of foundation, association etc.
<b>Article 9 (2) e)</b>		
<b>Personal data are manifestly made public by the data subject</b>	Unsuitable – with publicly available data the aim of the study cannot be reached	Unsuitable – with publicly available data the aim of the study cannot be reached
<b>Article 9 (2) f)</b>		
<b>Establishment, exercise or defence of legal claims</b>	Unsuitable, no such legal claims existing	Unsuitable, no such legal claims existing
<b>Article 9 (2) g)</b>		
<b>Substantial public interest</b>	(Possibly) suitable, further study with respect to the study aim and design will be required	Unsuitable – the study design aims to identify the individual at risk with possible follow up
<b>Article 9 (2) h)</b>		
<b>Preventive or occupational medicine, assessment of the working capacity of the employee, medical diagnosis, provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and</b>	Suitable after adoption of the appropriate legislation to support medical diagnosis	Suitable after adoption of the appropriate legislation to support medical diagnosis

<b>subject to the conditions and safeguards</b>		
<b>Article 9 (2) i</b>		
<b>Public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices</b>	Possibly suitable, further study with respect to the study aim and design will be required	Unsuitable, the aim of the study is to protect individual health
<b>Article 9 (2) j</b>		
<b>Archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with GDPR Article 89(1)</b>	Possibly suitable in the case of scientific research, further study with respect to the study aim and design will be required Unsuitable with regards to archiving, historical or statistical purposes.	Unsuitable with regards to archiving, historical or statistical purposes as the aim of the study is to protect individual health. Also unsuitable for scientific research – considering ethical point of view, individual subject shall consent to participation in research.

Table 1 demonstrates that it is legally feasible to perform both population prevalence screening study and individual risk study. For the prevalence study, adoption of the appropriate legislation will create sufficient legal basis for the personal data processing in compliance with the GDPR. For individual risk study, the adoption of specific legislation shall be accompanied with additional steps and measures. Least problematic would-be opt-in study design where a subject or his/her legal representative would consent to the participation in the study.

Having an appropriate legal basis will, however, not be enough to comply with EU privacy regulation. Such regulation will in fact be likely the most important legal limitation for the intended studies, although other must be considered as well.

The GDPR contains a set of fundamental principle. One of the principles is transparency, which requires that appropriate information is provided to the individual in order to ensure their understanding of who and why processes personal data. With an appropriate information campaign and especially in an opt-in study, this principle could be satisfied. Another principle is “purpose limitation”, which requires that secondary uses of data are in line with the purpose for which such data was initially collected. In the case at hand, this may present a possibly conflict as even a population study could be considered incompatible with the initial purpose. On the other hand, processing in the public interest or scientific research benefit from certain exemptions, which may in turn inform the choice of appropriate legal bases. We consider the “data minimisation”, “accuracy” and “integrity and confidentiality” principles as not imposing a significant hurdle, although they will have to be considered carefully in detail at the point of designing the study(ies).

Any study (either prevalence, individual risk or any other study in scope of this article), will have to be subject to a detail assessment of the risks. In particular, a data protection impact assessment will have to be conducted which will evaluate the necessity, proportionality of the processing in relation to its purpose and the rights of the data subjects, including the measures addressing security and mitigating the identified risks.

The specific evaluation may present even unexpected risks, as the application of GDPR has reached in some cases extreme interpretations which may adversely impact data processing which is

at focus of this article. Judgment of the Court of Justice of the European Union in Case C-184/20 OT v Vyriausioji tarnybinės etikos komisija clarified that GDPR Article 9 which outlines sensitive personal data shall be interpreted rather broadly as even data that are liable to disclose indirectly the sexual orientation of a natural person are to be considered as sensitive personal data, for the purpose of those provisions. Following from this judgment, it seems that even any initial processing of photographs with the intention to derive indirectly information relating to genes or genetic diseases should be from the outset considered as sensitive data.

It remains unclear whether this interpretation would result in any processing of the same information to be considered as sensitive personal data. Such extreme interpretation might render use of social media or even CCTV in public spaces impossible and would likely result in re-categorization of a wide range of “normal” personal data (such as family photos), into sensitive genetic data. For the purposes of the processing discussed within this article, it is sufficient to say that the underlying data, even if simple photographs, would be deemed sensitive genetic data, from the initial point of collection.

AI used for both population prevalence study and individual risk study is a software designed and intended for diagnosis, prevention, prediction or prognosis of genetic disease. As such, it is covered by the definition of medical device under Article 2 (1) of the Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices (MDR). The study AI therefore will have to comply with the MDR and be certified as medical device of class I or eventually of class IIa.

We must note AI will become a subject to EU –wide regulation in the near future. The first proposal for an AI Act was published by the European Commission already in April 2021. [12] The Council of the European Union approved a compromise version of the proposed Artificial Intelligence Regulation (AI Act) on December 6, 2022, [13,14] European parliament will vote on the AI Act in 2023. Processing of health-related data may become subject and be facilitated by the proposed EDHS regulation [11].

We intentionally limited legal limitation and consideration only to those listed above as going into further details is beyond the scope of this paper [15].

#### *Privacy safeguards in the study design*

It is obvious the intended studies cannot be performed on the NID data itself but on databases specifically created and derived from it. Considering the sensitivity of the study databases, security and privacy by design approach shall be the leading principle. To minimize the security concerns, two separate data sets shall be created – one for population study and the other for individual risk study. Both data sets shall be anonymized/pseudonymized in as much as the reaching the aim of the study would permit.

To minimise the risk of database compromising or leakage, appropriate security measures shall be implemented. To use cloud storage or outsourced IT solutions would significantly increase the risk of security breach. Furthermore, any data transfer outside of the country of origin will be highly undesirable, and outside of the European Economic Area would also likely be problematic due to the international data transfer requirements set out under the GDPR. Dedicated IT infrastructure operated by the EU member states or its dedicated agencies as public health authority or national e-health operator would likely be preferred to commercially available solutions. Also, the tailored AI solution shall be operated on such dedicated infrastructure, as sending the sensitive personal data for processing via internet to the AI right holder constitutes additional risk. Court of Justice of the European Union (CJEU) addressed such risks e.g., in the Schrems II ruling (Data Protection Commissioner v. Facebook Ireland and Maximillian Schrems [JUDGMENT OF THE COURT (Grand Chamber) 16 July 2020, Case C-311/18), ECLI:EU:C:2020:559].

The legislation allowing the intended studies shall create an environment for fair and transparent processing the test data sets. To achieve the highest attainable standard of data and cyber security, minimal security standards and required safeguards shall be integral part of such legislation.



### 3. Discussion

Mandatory national ID schemes collect large amount of personal data into a centralized database. Following September 11, 2001, biometric identification has been increasingly included in these databases [16]. While national ID is issued at certain age, passport is required even for new-borns for international travel. As a result, national ID databases include 2D/3D photos combined with set of biometric and personal data covering majority of nation's population. Increasing global usage of social networks allows for scrapping the personal data and photos even for non-state players that have sufficient resources and determination [17].

Current advances in generative AI are transforming all relevant fields including medical education [18] and have significant impact on various forensic applications including human remains identification [19]. Omnipresent mobile devices can recognize 3D faces with increasing precision [20]. Various forensic applications can be significantly improved with AI application from morphometric analysis [21] or AI behavioral analyses and predictions in prevention of violent or murderous behavior [22] to age estimation in minors [23].

Genome-wide association studies (GWAS) have investigated the association between normal facial variation and millions of single nucleotide polymorphisms (SNPs). Over 50 loci associated with facial traits have already been identified [24].

Described AI technologies could facilitate automated population screening for certain genetic traits or disorders without the need to obtain any biological sample by nation-states or event private companies.

The most significant barrier preventing this to (ab)use AI is the privacy legislation such as European Union's General Data Protection Regulation (GDPR) [25,26]. From the GDPR perspective consent, legal obligation and public interest may be considered as viable options for the selection of appropriate legal basis. Therefore, the adoption of legislation allowing facial based AI population screening for certain health/genetic traits could be imaginable.

Unfortunately, the privacy legislation does not apply to personal data processing for the purpose of national security or by law enforcement. Widespread usage of CCTV allows to obtain still images or video of the suspected perpetrators. Having the national ID scheme and the investigatory powers of the law enforcement agencies allows linking and analysing information stored in different state and private operated databases to create unprecedented surveillance society.

Imagine the following scenario: a crime has been recorded using a CCTV, there are several suspects or witnesses. By using AI on the video or still images a genetic/health trait has been identified among the persons present at the scene of crime. Now the nationally available databases can be exploited to search for all persons having the same genetic/health traits – national ID database, national electronic medical records database etc. Facial kinship verification using AI is also potentially immensely powerful tool, allowing to identify possible relatives of the persons of interest [27].

Several commercial companies offering direct-to-consumer (DTC) genetic testing. Using a testing kit, sample of saliva or hair is collected, sent for DNA extraction, and analyse. The customer is provided with test result via email or website [28]. The DTC companies store the genetic data frequently linked to a person who requested the test. Law enforcement already started to utilize DTC data to track suspects or their relatives [29]. With an appropriate court order, the law enforcement agencies can compel the DTC companies to produce customers' data matching the health/genetic traits identified by AI. Legal instruments for international police cooperation can make these DTC data available for foreign governments and their respective agencies.

The available AI applications in genetics open new possibilities. Population screening using facial AI recognition of genetic and health traits would be cheap and effective tool identify individuals at risk with subsequent targeted prevention or health care provision to save the limited resources of health care system. It can also create a society oppressed and controlled beyond George Orwell's darkest fantasies of 1984. With no doubt, AI application in genetics can create extreme risk to privacy and anonymity. AI supported genetic application can be used to discriminate minorities based on race, ethnicity, or health traits. Ethical implication of AI based genetic application are yet

only partially understood and shall be carefully analysed to allow for effective legal regulation. Any such study shall be governed by ethical principles of autonomy, beneficence, non-maleficence, and justice.

#### 4. Conclusions

Population wide screening for genetic disorders without the necessity to obtain biological sample is already possible with available technologies. This conclusion applies to both epidemiological and individual risk study. The development and usage of such advanced technologies faces significant legal limitations, especially in the heavily regulated EU landscape. AI for proposed studies will have to be certified as medical device and eventually will be also subject for regulation by upcoming AI Act. The EU privacy legislation introduces additional significant barriers (or safeguards, depending on the point of view), but with properly drafted specific legislation, such studies could be eventually performed.

AI applied to big datasets derived from NID databases may generate unprecedented epidemiological data on genetic diseases and its development in the course of time. The comparison of data obtained from present sources and data which may be obtained with proposed studies may result in discovery of discrepancies. In such case, further analysis may be performed. With NID data on kinship, genetic diseases may be tracked over the generations which could bring valuable information from public health point of view.

Mere possibility of the studies not necessarily mean such studies shall be implemented without further discussion. While we presented significant potentially beneficial use cases, taking into consideration ethical aspects will be necessary in the future. One of such ethical dilemmas would be discussion whether the person who is an unwitting carrier of certain genetic disorders would like to get information in his/her carrier status. With the properly provided advance information and informed consent duly obtained such dilemma seems to be easy resolvable. But there may be several arguments to the contrary.

Unfortunately, the law enforcement and intelligence agencies are in some cases exempt from privacy protecting legislation and may get access to and process almost any type of personal data arguing by the national security. Therefore, the dystopia of Orwell's 1984 world could significantly underestimate of the advanced surveillance modalities readily available in 21<sup>st</sup> century.

**Author Contributions:** Conceptualization, supervision, investigation, writing – original draft preparation, writing—review and editing: PK and AT. Writing – original draft preparation: AB, IV, MS, LM and MA. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Slovak Grant Agency for Science KEGA Thurzo, grant number 054UK-4/2023 and APVV-22-0381.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** MA is the Director General of Slovak Information Service. The other authors declare no conflict of interest.

#### References

1. Shade, J.; Coon, H.; Docherty, A.R. Ethical Implications of Using Biobanks and Population Databases for Genetic Suicide Research. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **2019**, *180*, 601–608, doi:10.1002/ajmg.b.32718.

2. Latorre-Pellicer, A.; Ascaso, Á.; Trujillano, L.; Gil-Salvador, M.; Arnedo, M.; Lucia-Campos, C.; Antoñanzas-Pérez, R.; Marcos-Alcalde, I.; Parenti, I.; Bueno-Lozano, G.; et al. Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes. *Int J Mol Sci* **2020**, *21*, 1042, doi:10.3390/ijms21031042.
3. Home - Face2Gene Available online: <https://www.face2gene.com/> (accessed on 11 April 2023).
4. Gurovich, Y.; Hanani, Y.; Bar, O.; Nadav, G.; Fleischer, N.; Gelbman, D.; Basel-Salmon, L.; Krawitz, P.M.; Kamphausen, S.B.; Zenker, M.; et al. Identifying Facial Phenotypes of Genetic Disorders Using Deep Learning. *Nat Med* **2019**, *25*, 60–64, doi:10.1038/s41591-018-0279-0.
5. Latorre-Pellicer, A.; Ascaso, Á.; Trujillano, L.; Gil-Salvador, M.; Arnedo, M.; Lucia-Campos, C.; Antoñanzas-Pérez, R.; Marcos-Alcalde, I.; Parenti, I.; Bueno-Lozano, G.; et al. Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes. *International Journal of Molecular Sciences* **2020**, *Vol. 21, Page 1042* **2020**, *21*, 1042, doi:10.3390/IJMS21031042.
6. Thurzo, A.; Kosnáčová, H.S.; Kurilová, V.; Kosmeř, S.; Beňuš, R.; Moravanský, N.; Kováč, P.; Kuracinová, K.M.; Palkovič, M.; Varga, I. Use of Advanced Artificial Intelligence in Forensic Medicine, Forensic Anthropology and Clinical Anatomy. *Healthcare* **2021**, *9*, 1545, doi:10.3390/healthcare9111545.
7. Coley, R.Y.; Johnson, E.; Simon, G.E.; Cruz, M.; Shortreed, S.M. Racial/Ethnic Disparities in the Performance of Prediction Models for Death by Suicide After Mental Health Visits. *JAMA Psychiatry* **2021**, *78*, 726, doi:10.1001/jamapsychiatry.2021.0493.
8. Olczak, K.; Pawlicka, H.; Szymański, W. Root and Canal Morphology of the Maxillary Second Premolars as Indicated by Cone Beam Computed Tomography. *Australian Endodontic Journal* **2022**, doi:10.1111/AEJ.12624.
9. Lumaka, A.; Cosemans, N.; Lulebo Mampasi, A.; Mubungu, G.; Mvuama, N.; Lubala, T.; Mbuyi-Musanayzi, S.; Breckpot, J.; Holvoet, M.; de Ravel, T.; et al. Facial Dysmorphism Is Influenced by Ethnic Background of the Patient and of the Evaluator. *Clin Genet* **2017**, *92*, 166–171, doi:10.1111/cge.12948.
10. Eurostat NUTS - Nomenclature of Territorial Units for Statistics Available online: <https://ec.europa.eu/eurostat/web/nuts/background> (accessed on 26 March 2023).
11. European Commission Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52022PC0197&from=EN> (accessed on 26 March 2023).
12. Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts Available online: <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf> (accessed on 13 April 2023).
13. Transport, Telecommunications and Energy Council (Telecommunications) - Consilium Available online: <https://www.consilium.europa.eu/en/meetings/tte/2022/12/06/> (accessed on 13 April 2023).
14. Artificial Intelligence Act: Council Calls for Promoting Safe AI That Respects Fundamental Rights - Consilium Available online: <https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/> (accessed on 13 April 2023).
15. WHITE PAPER On Artificial Intelligence - A European Approach to Excellence and Trust.
16. Wilson, D. Australian Biometrics and Global Surveillance. *Int Crim Justice Rev* **2007**, *17*, 207–219, doi:10.1177/1057567707306650.

17. Rezende, I.N. Facial Recognition in Police Hands: Assessing the 'Clearview Case' from a European Perspective. *New Journal of European Criminal Law* **2020**, *11*, 375–389, doi:10.1177/2032284420948161.
18. Thurzo, A.; Strunga, M.; Urban, R.; Surovková, J.; Afrashtehfar, K.I. Impact of Artificial Intelligence on Dental Education: A Review and Guide for Curriculum Update. *Educ Sci (Basel)* **2023**, *13*, 150, doi:10.3390/educsci13020150.
19. Thurzo, A.; Jančovičová, V.; Hain, M.; Thurzo, M.; Novák, B.; Kosnáčová, H.; Lehotská, V.; Moravanský, N.; Varga, I. Human Remains Identification Using Micro-CT, Spectroscopic and A.I. Methods in Forensic Experimental Reconstruction of Dental Patterns After Concentrated Acid Significant Impact. *Molecules* **2022**, *27*, 4035, doi:10.3390/molecules27134035.
20. Thurzo, A.; Strunga, M.; Havlínová, R.; Reháková, K.; Urban, R.; Surovková, J.; Kurilová, V. Smartphone-Based Facial Scanning as a Viable Tool for Facially Driven Orthodontics? *Sensors* **2022**, *22*, 7752, doi:10.3390/s22207752.
21. Bianchi, I.; Oliva, G.; Vitale, G.; Bellugi, B.; Bertana, G.; Focardi, M.; Grassi, S.; Dalessandri, D.; Pinchi, V. A Semi-Automatic Method on a Small Italian Sample for Estimating Sex Based on the Shape of the Crown of the Maxillary Posterior Teeth. *Healthcare* **2023**, *Vol. 11*, Page 845 **2023**, *11*, 845, doi:10.3390/HEALTHCARE11060845.
22. Alqahtani, S.M.; Almutairi, D.S.; Binaqeel, E.A.; Almutairi, R.A.; Al-Qahtani, R.D.; Menezes, R.G.; Alqahtani, S.M.; Almutairi, D.S.; Binaqeel, E.A.; Almutairi, R.A.; et al. Honor Killings in the Eastern Mediterranean Region: A Narrative Review. *Healthcare* **2023**, *Vol. 11*, Page 74 **2022**, *11*, 74, doi:10.3390/HEALTHCARE11010074.
23. Cameriere, R.; Scendoni, R.; Ferrante, L.; Mirtella, D.; Oncini, L.; Cingolani, M. An Effective Model for Estimating Age in Unaccompanied Minors under the Italian Legal System. *Healthcare* **2023**, *Vol. 11*, Page 224 **2023**, *11*, 224, doi:10.3390/HEALTHCARE11020224.
24. Richmond, S.; Howe, L.J.; Lewis, S.; Stergiakouli, E.; Zhurov, A. Facial Genetics: A Brief Overview. *Front Genet* **2018**, *9*, doi:10.3389/fgene.2018.00462.
25. Gruschka, N.; Mavroeidis, V.; Vishi, K.; Jensen, M. Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data); IEEE, December 2018; pp. 5027–5033.
26. Goddard, M. The EU General Data Protection Regulation (GDPR): European Regulation That Has a Global Impact. *International Journal of Market Research* **2017**, *59*, 703–705, doi:10.2501/IJMR-2017-050.
27. Wu, X.; Feng, X.; Cao, X.; Xu, X.; Hu, D.; López, M.B.; Liu, L. Facial Kinship Verification: A Comprehensive Review and Outlook. *Int J Comput Vis* **2022**, *130*, 1494–1525, doi:10.1007/s11263-022-01605-9.
28. Kalokairinou, L.; Howard, H.C.; Slokenberga, S.; Fisher, E.; Flatscher-Thöni, M.; Hartlev, M.; van Hellemond, R.; Juškevičius, J.; Kapelenska-Pregowska, J.; Kováč, P.; et al. Legislation of Direct-to-Consumer Genetic Testing in Europe: A Fragmented Regulatory Landscape. *J Community Genet* **2018**, *9*, 117–132, doi:10.1007/s12687-017-0344-2.
29. Basch, C.H.; Hillyer, G.C.; Samuel, L.; Datuowei, E.; Cohn, B. Direct-to-Consumer Genetic Testing in the News: A Descriptive Analysis. *J Community Genet* **2022**, *14*, 63–69, doi:10.1007/s12687-022-00613-z.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.