

UNIVERZITA KARLOVA

Právnická fakulta

Anna Vozanková

Genetické databáze a ochrana genetických údajů

Diplomová práce

Vedoucí diplomové práce: JUDr. Mgr. Martin Šolc, Ph.D.

Katedra občanského práva

Datum vypracování práce (uzavření rukopisu) : 28. 3. 2023

Prohlašuji, že jsem předkládanou diplomovou práci vypracovala samostatně, že všechny použité zdroje byly řádně uvedeny a že práce nebyla využita k získání jiného nebo stejného titulu.

Dále prohlašuji, že vlastní text této práce včetně poznámek pod čarou má 159 772 znaků včetně mezer.

Anna Vozanková

V Praze dne 28. 3. 2023

Tímto bych chtěla poděkovat vedoucímu této diplomové práce panu JUDr. Mgr. Martinu Šolcovi, Ph.D. za jeho cenné rady, připomínky a vstřícný přístup. Ráda bych tímto poděkovala též Mgr. Martinu Samkovi, se kterým jsem měla možnost značnou část této práce konzultovat. Velké díky patří rovněž mé rodině a přátelům za podporu po celou dobu mého studia.

OBSAH

Úvod.....	1
1. Ochrana osobnosti.....	4
1.1 Právo na soukromí.....	4
1.2 Ochrana osobních údajů.....	5
1.3 Zvláštní kategorie osobních údajů.....	7
2. DNA a sekvenování DNA.....	8
2.1 Cesta k DNA.....	8
2.2 Pojmy.....	10
2.3 Sekvenování DNA.....	12
2.4 Přehled některých metod sekvenování DNA.....	14
2.5 Aktuální projekty zaměřené na shromažďování genetických dat.....	15
3. Využití sekvenování DNA.....	17
3.1 Big data a data mining.....	20
4. Genetické databáze a genome browsery.....	22
4.1 Agregované genetické databáze.....	24
4.2 Genome browser – příklad.....	26
4.3 Genetické databáze a genome browsery v právním slova smyslu.....	28
4.3.1 Bezúplatná zákonná licence.....	30
4.4 Genome browsery – software.....	32
4.5 Genome browsers – Cloud Computing.....	34
4.5.1 Genome browsery – DBaaS.....	36
4.5.2 Genome browsery – SaaS.....	37
4.6 Free Software a Open Source.....	37
5. Genetické údaje.....	39
5.1 Právní základ zpracování genetických údajů – GDPR.....	42
5.2 Anonymizace a pseudonymizace údajů.....	45
5.3 Posouzení vlivu na ochranu osobních údajů (DPIA).....	48
5.4 DPIA – genome browsery a genetické databáze.....	50
5.5 Souhlas se zpracováním údajů.....	51
5.5.1 Udělení souhlasu pro konkrétní účel zpracování.....	52
5.6 Omezení zpracování genetických údajů genetickými databázemi.....	55
6. European Health Data Space.....	58
6.1 Sekundární využití elektronických zdravotních dat.....	59

Závěr	63
Seznam použitých zdrojů.....	66
Abstrakt.....	78
Abstract	79

Úvod

Tématem diplomové práce je fungování genetických databází a ochrana genetických údajů v kontextu současného stavu práva, technologií a vědy. Autorka zvolila téma diplomové práce s ohledem na svůj dlouhodobý zájem o genetiku, ochranu osobnosti a ochranu osobních údajů, přičemž při psaní diplomové práce vycházela autorka částečně ze své obhájené práce SVOČ. Zároveň je genetika a výzkum s ní spojený téma, které stále častěji rezonuje globální společností a má potenciál významně ovlivnit budoucí generace.

Soudobý výzkum a věda jsou často založeny na analýze a vyhodnocování velkého množství dat. V mnoha odvětvích se jedná o data, která jsou vysoce osobní. Zároveň je však pro dosažení co možná nejlepších výsledků výzkumů nutný také velmi široký přístup k nim. Je však nutné zajistit vhodný způsob jejich zpřístupnění, ukládání a spravování, k čemuž jsou stále častěji využívány moderní technologie. V medicínském a farmaceutickém odvětví se jedná především o genetické údaje, s jejichž zpracováním a užíváním je spojeno nejen mnoho potenciálních výhod, ale rovněž i nespočet rizik jejich zneužití.

Genetická vyšetření se stávají díky technologickému rozvoji stále dostupnějšími, což zároveň znamená stále narůstající množství zpracovávaných genetických údajů. Genetické údaje každého jednotlivce jsou unikátní a jsou tak již ze své povahy jedinečným identifikátorem konkrétní osoby. Vzhledem k charakteru genetických údajů, tj. jejich množství a objemu, ve kterém jsou zpracovávány a zejména k rozsahu informací, které jsou v nich zanesené a jejich prostřednictvím dostupné, je při nakládání s nimi nutné postupovat s vysokou mírou důslednosti a zodpovědnosti. S novými moderními technologiemi a metodami jejich zpracování se stále zvětšuje množina informací, které se lze dozvědět, byť jen z malého úseku lidské DNA. Také z tohoto důvodu je nutné zaručit dostatečné zabezpečení genetických údajů a tím tedy i co největší ochranu soukromí subjektů údajů.

Analýza a výzkum genetických údajů jsou slibným nástrojem pro vývoj léčivých přípravků a léčebných postupů pro léčbu mnoha závažných, nejen genetických, onemocnění. Aby byl takový výzkum efektivní, je třeba velké množství genetických dat od jednotlivých subjektů údajů co možná nejrůznějšího původu, přičemž tyto údaje musí být dále přístupné širokému okruhu vědců, lékařů a dalších odborníků. To vše za minimalizace rizik s tím spojených, jejich pravidelného vyhodnocování a za využití patřičných postupů při nakládání s genetickými údaji.

Proto je vhodné stanovit určité limity jejich zpracování a minimální standard jejich ochrany, o což se snaží právní úprava. Je však otázkou, jak efektivní jsou v tomto ohledu aktuální zákonem dané záruky a povinnosti a zdali se vůbec na určité způsoby zpracování genetických údajů uplatní. V této souvislosti je též třeba posoudit závažnost možných rizik z tohoto vyplývajících oproti možnému prospěchu pramenícímu z jejich výzkumu pro společnost.

Diplomová práce se zabývá především problematikou současného způsobu fungování genetických databází a zpracování genetických údajů, a to v souvislosti s aktuální právní úpravou a řešením jejich nedostatků *de lege ferenda*.

Práce je systematicky rozdělena do šesti kapitol a tematicky souvisejících podkapitol. První kapitola je zaměřena na vývoj práva v oblasti ochrany osobnosti a s tím souvisejícího práva na soukromí a ochranu osobních údajů včetně definic některých základních pojmů. Pro účely porozumění charakteru genetických údajů se autorka věnuje v druhé a třetí kapitole samotnému sekvenování DNA a jeho praktickému využití. Ve čtvrté kapitole je podrobně rozebráno fungování genetických databází a genomebrowserů, a to včetně způsobu jejich ochrany a zpřístupňování. Práce se v této souvislosti blíže zaměřuje na fungování a právní ochranu databází, software a jeho licencování. Stěžejní kapitola číslo pět pojednává o genetických údajích, základu jejich zpracování, udělování souhlasu s jejich zpracováním subjekty údajů, možnosti jejich anonymizace a vlivu současné právní úpravy v této souvislosti. Autorka v závěru této kapitoly posuzuje možné dopady důsledné aplikace právní úpravy na průběh výzkumu v oblasti genetiky. Obsahem poslední, šesté, kapitoly je možný dopad iniciativy Evropské unie European Health Data Space na zpracování, zejména sdílení a využití genetických údajů v rámci subjektů členských států Evropské unie.

Při zpracování diplomové práce autorka užívala převážně metody analytické, metody syntézy, metody deskriptivní a metody analogie. Diplomová práce obsahuje též kazuistickou analýzu předmětné problematiky a jejích jednotlivých prvků. Jelikož problematika genetických databází a zpracování genetických údajů je ve společnosti poměrně novým tématem, pracuje autorka zejména s odbornými články zahraničních autorů, právními předpisy a legislativními návrhy, internetovými zdroji, knižními publikacemi a místy též s judikaturou Soudního dvora Evropské unie.

Hlavním cílem této práce je přiblížit současný způsob zpracování genetických údajů a upozornit na problematiku spojenou především s jejich sdílením. Autorka práce má za to,

že téma vztahu genetiky a práva není v České republice dostatečně diskutováno, přestože význam genetiky stále rapidně roste. Druhým cílem diplomové práce je poukázat na nutnost zaručení vyšší ochrany genetických dat a přispět tak do předpokládané budoucí debaty ohledně regulace genetických údajů.

1. Ochrana osobnosti

1.1 Právo na soukromí

Pojem *právo na soukromí* byl použit S. D. Warren a L. D. Brandeisem v článku *The Right to Privacy* z roku 1890.¹ Tento článek je považován za jednu ze základních a klíčových prací v oblasti ochrany soukromí, jelikož v něm bylo poprvé právo na soukromí formulováno jakožto právo být ponechán o samotě.² Autoři v článku uvádí, že každý jedinec má subjektivní právo na ochranu svého soukromí a že toto právo by mělo být chráněno zákonem. Právo na soukromí bylo později označeno jako jedno ze základních lidských práv.

Po druhé světové válce byl kladen stále větší důraz na ochranu člověka jakožto jednotlivce před zásahy do jeho tělesné i duševní integrity. V roce 1950 byla přijata Evropská úmluva o ochraně lidských práv a svobod (EÚLP), která v čl. 8 zakotvila právo každého na respektování soukromého a rodinného života, obydlí a korespondence. Zákaz svévolného zasahování do soukromého života, rodiny, domova nebo korespondence, cti a pověsti a též právo na zákonnou ochranu před takovými zásahy nebo útoky je výslovně zakotven též v čl. 17 Mezinárodního paktu o občanských a politických právech z roku 1966.³

Potřeba ochrany soukromí jednotlivce je odražena také v mnoha aktech Evropské unie. Obdobně jako EÚLP zakládá čl. 7 Listiny základních práv Evropské unie (LZPEU) právo na respektování soukromého a rodinného života, obydlí a komunikace, přičemž v čl. 3 je explicitně stanoveno právo na nedotknutelnost lidské osobnosti a v čl. 8 právo na ochranu osobních údajů.

V českém právním řádu je právo na soukromí zakotveno primárně v Listině základních práv a svobod (LZPS)⁴. V čl. 7 odst. 1 LZPS je zaručena nedotknutelnost osoby a jejího

¹ WARREN, S. D., BRANDEIS, L. D. *The Right to Privacy*. In *Harvard Law Review*, 1890, roč. 4, č. 5, s. 195. Harvard Law Review Association 1890. [online]. [cit. 2023-03-17]. Dostupné z <https://doi.org/10.2307/1321160>

² V originálním znění článku bylo právo na soukromí definováno jako *the right of the individual to be alone*.

³ MATES, P. (ed.), VALOUŠEK M., FIALOVÁ E., LECHNER, T., HÁLOVÁ, M., SIVÁK, J., SOVOVÁ, O., BRUNA, E., BRUNOVÁ, M. *Ochrana osobnosti, soukromí a osobních údajů*. Nakladatelství Leges, s. r. o., podle stavu k 14. 8. 2019. ISBN: 978-80-7502-346-9.

⁴ Ústavní zákon č. 2/1993 Sb. ze dne 16. 12. 1992, ve znění ústavního zákona č. 162/1998 Sb., o vyhlášení Listiny základních práv a svobod (LZPS).

soukromí a v čl. 10 odst. 2 potom právo na ochranu před neoprávněným zasahováním do soukromého a rodinného života.

Aby bylo možné docílit efektivní ochrany soukromí jednotlivců, je právo na soukromí odraženo v soukromém právu, kde jsou stanoveny hmotněprávní podmínky jeho uplatňování. Zákon č. 89/2012 Sb., občanský zákoník, je mimo jiné přímo postaven na zásadě, že *každý má právo na ochranu svého života a zdraví, jakož i svobody, cti, důstojnosti a soukromí*.⁵ Občanský zákoník právo na soukromí zařazuje do tzv. osobnostních práv, která jsou upravena v oddílu s názvem „Osobnost člověka“⁶. Občanský zákoník zaručuje ochranu osobnosti člověka včetně všech jeho přirozených práv, přičemž *ochrany požívají zejména život a důstojnost člověka, jeho zdraví a právo žít v příznivém životním prostředí, jeho vážnost, čest, soukromí a jeho projevy osobní povahy*⁷ a výslovně přiznává člověku, do jehož osobnosti bylo zasaženo, právo domáhat se upuštění od takového neoprávněného zásahu či odstranění jeho následku.⁸

S rostoucím množstvím zpracovávaných a dostupných informací o každém jedinci a v souvislosti s rozvojem nových technologií, se zvyšuje riziko neoprávněných zásahů do soukromí jednotlivců. Zásahy do osobnosti člověka jsou v současné době stále častěji spojovány s neoprávněným zpracováním⁹ osobních údajů.

1.2 Ochrana osobních údajů

Osobní údaje jsou definovány jakožto *veškeré informace o identifikované nebo identifikovatelné fyzické osobě*. Jedná se o jakékoli skutečnosti vypovídající o určité osobě či jejím znaku, nejběžnějšími příklady jsou například jméno a příjmení, datum narození, identifikační čísla či kontaktní údaje, jde však též o informace o prvcích fyzické, fyziologické,

⁵ § 3 odst. 2 písm. a) zák. č. 89/2012 Sb., občanský zákoník ve znění pozdějších předpisů (OZ).

⁶ Osobnost člověka jakožto předmět soukromých práv zahrnuje vše, čím se člověk projevuje navenek ve vztahu ke svému okolí, a to jak po stránce fyzické, tak po stránce duchovní a duševní. (LAVICKÝ, P. a kol. Občanský zákoník I. Obecná část (§1-654). Komentář. 1. vydání. Praha: C. H. Beck, 2014, 394 s., ISBN 978-80-7400-529-9.)

⁷ § 81 OZ.

⁸ § 82 odst. 1 OZ.

⁹ Čl. 4 odst. 2 Nařízení Evropského parlamentu a Rady (EU) č. 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (GDPR).

genetické, psychické, ekonomické, kulturní nebo společenské identity této osoby.¹⁰ Osobní údaje se týkají z velké části samotné osobnosti člověka z čehož vyplývá též nutnost vysoké úrovně jejich ochrany.

Neoprávněným zpracováním osobních údajů může být zasaženo do osobnostních práv člověka, a to především do jeho soukromí, cti či důstojnosti. S určitými typy osobních údajů s ohledem na jejich povahu hrozí větší riziko spojené s jejich zneužitím. Je nutné, aby bylo s osobními údaji nakládáno vždy takovým způsobem, který bude míru rizika dostatečně zohledňovat. Osobám musí být též garantováno právo na ochranu před neoprávněnými zásahy. Významnost práva na ochranu osobních údajů je zřejmá i ze skutečnosti, že jeho základ je přímo v listinách týkajících se základních práv a svobod.

LZPEU stanoví základní požadavky na zpracování veškerých osobních údajů bez výjimky. Osobní údaje musí být zpracovány *korektně, k přesně stanoveným účelům a na základě souhlasu dotčené osoby nebo na základě jiného oprávněného důvodu stanoveného zákonem. Každý má právo na přístup k údajům, které o něm byly shromážděny, a má právo na jejich opravu.*¹¹ Čl. 10 odst. 3 LZPS stanoví, že *každý má právo na ochranu před neoprávněným shromažďováním, zveřejňováním nebo jiným zneužíváním údajů o své osobě.*

Ochrana osobních údajů je potom upravena nařízením Evropského parlamentu a Rady EU 2016/679 o ochraně osobních údajů (GDPR), přičemž v českém právním řádu se dále uplatní zákon č. 110/2019 Sb., o zpracování osobních údajů.

Cílem regulace zpracování osobních údajů je především umožnit subjektům údajů¹² ovlivnit a kontrolovat způsob nakládání s jejich osobními údaji a možnost tak v určité míře zabránit jejich nezákonnému zpracování. Výše uvedenými právními předpisy jsou též stanoveny základní podmínky a náležitosti zpracování osobních údajů jakýmkoli osobami, se snahou reflektovat rizika spojená se zpracováním různých kategorií osobních údajů.

¹⁰ Čl. 4 odst. 1 GDPR.

¹¹ Čl. 8 odst. 2 Listiny základních práv Evropské unie (LZPEU).

¹² Čl. 4 odst. 1 GDPR.

1.3 Zvláštní kategorie osobních údajů

Osobní údaje, které jsou považovány za obzvláště rizikové a citlivé, řadí GDPR do zvláštní kategorie osobních údajů, tj. takových osobních údajů, *kteřé vypovídají o rasovém či etnickém původu, politických názorech, náboženském vyznání či filozofickém přesvědčení nebo členství v odborech a zpracování genetických údajů, biometrických údajů za účelem jedinečné identifikace fyzické osoby a údajů o zdravotním stavu či o sexuálním životě nebo sexuální orientaci fyzické osoby.*¹³ Osobní údaje spadající do této kategorie požívají zvláštní ochrany, jelikož jsou z hlediska základních práv a svobod svou povahou obzvláště citlivé, a to z právě důvodu, že při jejich zpracování můžou hrozit závažná rizika zásahů do těchto základních práv a svobod.¹⁴

Při neoprávněném zpracování zvláštní kategorie osobních údajů, tj. citlivých osobních údajů, hrozí subjektům údajů závažná rizika. Jedná se zejména o různé druhy diskriminace subjektu údajů, zásah do soukromí či ovlivnění rozhodování týkající se např. poskytování zdravotních služeb, včetně veškerých negativních důsledků v těchto souvislostech.

Díky dynamickému vývoji moderních technologií dochází současně k významnému rozvoji v odvětví genetiky. Základním prvkem výzkumu v této oblasti jsou genetická data, která jsou zpracovávána ve stále větším a masivnějším množství. Právě za pomoci nových technologií, jako jsou různé interaktivní nástroje, umělá inteligence, objemná úložiště či způsoby zpracování a standardizace těchto údajů je pro vědce a výzkumníky snadnější s daty pracovat, efektivně je při výzkumu využívat a následně ověřovat výsledky svých studií. S dostupnějšími genetickými vyšetřeními a jejich komercializací jsou genetické údaje a jejich zpracování stále přístupnější, a to jak pacientům, zdravotníkům a odborníkům v rámci poskytování zdravotních služeb, tak široké veřejnosti.

Rostoucí poptávka po genetických datech je přirozeně odražena v rozšiřující se nabídce. To znamená mimo jiné neustále se zvětšující rozsah zpracovávaných genetických údajů, což v praxi často znamená kvantitu dat na úkor kvality jejich zabezpečení.

¹³ Čl. 9 odst. 1 GDPR.

¹⁴ Zejm. rec. 51 až 54 GDPR.

2. DNA a sekvenování DNA

2.1 Cesta k DNA

S počátkem moderní genetiky je spojován přírodovědec Gregor Johann Mendel, který ve svém díle *Versuche über Pflanzen-Hybriden* z roku 1866 (česky „Pokusy s rostlinnými hybridy“) definoval zákony dědičnosti.¹⁵ V roce 1869 Johannes F. Miescher zjistil, že v jádrech bílých krvinek je přítomna směs látek, která obsahuje nukleové kyseliny, včetně dnes již známé DNA.

Ve 20. letech 20. století přišel vědec Phoebus Aaron Theodore Levene na skutečnost, že nukleové kyseliny mohou být dále rozloženy na jednotlivé nukleotidy, které jsou složeny z fosfátové skupiny, pětiuhlíkatého cukru a jedné ze čtyř bází – adeninu (A), guaninu (G), cytozinu (C) a thyminu (T). Následně probíhaly další výzkumy, které měly za cíl upřesnit závěry, které vyplývaly z předchozích zjištění. Rok 1944 s sebou přinesl výzkum Oswalda Averyho, Colina MacLeoda a Maclyna McCartyho, a to v návaznosti na tzv. „Griffithův experiment“¹⁶ Fredericka Griffitha. Tento výzkum ukázal, že tzv. *transforming factor*¹⁷ – substance, která může přejít z buňky do jiné a s sebou též přinést určité znaky, které druhá buňka následně začne též vykazovat – původně predikovaný právě F. Griffithem, nejsou předpokládané proteiny, ale právě kyselina deoxyribonukleová neboli DNA.¹⁸

V roce 1946 zkoumal DNA různých organismů biochemik Erwin Chargaff, který přišel na skutečnost, že obsah jednotlivých dusíkatých bází ve vztahu k jednotlivým druhům organismů je různý, avšak v rámci jednoho konkrétního druhu zůstává stejný. Též vyšlo najevo, že množství bází A odpovídá množství bází T, a stejně tak báze G bázím C.¹⁹ Toto bylo však s jistotou potvrzeno až roku 1952 elegantním experimentem Alfreda Hersheyho a Marthy

¹⁵ MENDEL, G. J. *Versuche über Pflanzen-hybriden*. Experiments on Plant Hybrids (1866).

¹⁶ GRIFFITH, F. The Significance of Pneumococcal Types. 1928. [online] Journal of Hygiene. Cambridge University Press. 27 (2): 113–159. [cit. 2022-03-10]. Dostupné z: doi: 10.1017/s0022172400031879.

¹⁷NIH. National Human Genome Research Institute. 1944: DNA is "Transforming Principle". [online]. [cit. 2022-03-10] Dostupné z: <https://www.genome.gov/25520250/online-education-kit-1944-dna-is-transforming-principle>

¹⁸STEINMAN, R. M., MOBERG, C. L. Dostupné z: A triple tribute to the experiment that transformed biology. Discovering the Genetic Role of DNA. [online]. Journal of Experimental Medicine. 1994. 179 (2): 379–384. [cit. 2022-03-10]. Dostupné z: <https://doi.org/10.1084/jem.179.2.379>

¹⁹ The Guardian. WRIGHT, P. Obituary: Erwin Chargaff. 2002. [online]. [cit. 2022-04-30]. Dostupné z: <https://www.theguardian.com/news/2002/jul/02/guardianobituaries.obituaries>

Cowles Chase, spočívajícího ve zkoumání přenosu genetických informací DNA prostřednictvím modifikovaných virů do bakterií. Do jedné skupiny bakterií byly přeneseny viry obsahující radioaktivní síru nacházející se v bílkovinách a do druhé skupiny byly vloženy viry obsahující radioaktivní fosfor, který je zase přítomen pouze v DNA. Zjistilo se, že bakterie „nakažené“ virem obsahujícím fosfor následně za určitých podmínek svítí, kdežto druhá skupina bakterií nikoliv. Tím bylo dokázáno, že právě nukleová kyselina, konkrétně DNA, je chemikálií odpovědnou za přenos genetických informací.

V dubnu roku 1953 byl publikován krátký, ale výstižný článek ve vědeckém časopisu Nature Jamese D. Watsona a Francise H. C. Cricka, který byl uveden následujícím: *We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest*, který je považován za první odhad prostorové struktury DNA²⁰ - nyní všem známé dvoušroubovice. V roce 1980 obdržel vědec Frederik Sanger Nobelovu cenu za svůj výzkum z roku 1977 týkající se DNA konkrétního viru a počtu párů bází, které obsahuje. Roku 1981 byla vytvořena metoda sekvenace DNA a přečtena DNA lidské mitochondrie. V průběhu dalších let došlo k dalším průlomovým bodům v tomto odvětví. Příkladem může být dnes hojně využívaný princip polymerázové řetězové reakce (PCR) Karym B. Mullisem (r. 1983)²¹ či přečtení prvního kompletního genomu volně žijícího organismu (bakterie *Haemophilus influenzae*) týmem Craiga Ventera v roce 1995.²² Roku 2003 byl přečten první lidský genom, a to v rámci The Human Genome Project (r. 1990 – 2003) – třináctiletého mezinárodního vědeckého výzkumu, jehož cílem bylo určení párů bází, ze kterých se skládá lidská DNA, mapování genů lidského genomu z určitých hledisek a obecně porozumění lidským genům²³ – publikovaného

²⁰ WATSON, J., CRICK, F. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. [online]. Nature 171, 737–738 (1953). [cit. 2022-03-11]. Dostupné z: <https://doi.org/10.1038/171737a0>

²¹ KOSSAKOVSKI, F. The eccentric scientist behind the ‘gold standard’ COVID-19 test. 2021. [online]. National Geographic. [cit. 2022-03-11]. Dostupné z: <https://www.nationalgeographic.com/science/article/the-eccentric-scientist-behind-the-gold-standard-covid-19-pcr-test>

²² FLEISHMANN, R.D., ADAMS, M. D., WHITE, O., et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. [online]. Science 269, 496-512. July 28, 1995. [cit. 2022-03-11]. Dostupné z: <https://doi.org/10.1126/science.7542800>

²³ NIH. National Human Genome Research Institute. The Human Genome Project. [online]. [cit. 2022-03-11]. Dostupné z: <https://www.genome.gov/human-genome-project>

organizací Human Genome Organization. Teprve nedávno, dne 31. 3. 2022, byl oznámen úspěch mezinárodní skupiny vědců, kterým se podařilo sekvenovat kompletní lidský genom.²⁴

2.2 Pojmy

Pro lepší orientaci v předmětné problematice považuji za vhodné krátce vymezit některé pojmy.

DNA je kyselina, která nese informace o způsobu fungování jednotlivých buněk a je schopna v sobě tyto informace uchovávat. Skládá se z molekuly cukru (deoxyribózy), molekuly fosfátu a jedné z dusíkatých bází (adeninu, cytosinu, guaninu a thyminu). Díky těmto bázím a jejich konkrétnímu pořadí je možné správné fungování určitých buněk. Geny jsou ve dvoušroubovici DNA zaneseny ve formě kódů dusíkatých bází a tyto sekvence jsou přepsány do RNA²⁵ dále kontrolující a regulující výrobu bílkovin.²⁶

Lidský genom je soubor veškerých genetických informací zapsaných v lidské buňce, resp. v DNA. Tento pojem zastřešuje jak tzv. kódující úseky DNA, které jsou přepisované v proteiny, tak úseky nekódující, které plní jiné ochranné či regulační funkce. Každá z cca trilionu buněk lidského organismu obsahuje kompletní kopii lidského genomu.²⁷

Geny jsou části lidského genomu, které obsahují základní informace pro život organismu. Jedná se o úseky DNA kódující buď protein za pomoci jeho aminokyselin, nebo některý z druhů RNA,²⁸ jakožto prostředníka informací mezi DNA a proteiny.²⁹ Dle odhadů se počet

²⁴ PARK, A. The Human Genome Is Finally Fully Sequenced. 2022. [online]. Time. [cit. 2022-04-15]. Dostupné z: <https://time.com/6163452/human-genome-fully-sequenced/>

²⁵ Ribonukleová kyselina se skládá z nukleotidů tvořených vždy jednou ze čtyř základních bází (adeninu, uracilu, guaninu a cytosinu), cukrem ribózou a fosfátovou skupinou. (...) RNA molekuly slouží k nejrůznějším úkolům: od strukturních molekul, které tvoří základ buněčných organel, přes RNA molekuly, které kontrolují a regulují výrobu bílkovin, až po RNA molekuly, které katalyzují nejrůznější chemické reakce. (...) Asi nejznámější funkcí RNA je přenos informace mezi DNA a proteinem. (Ustav molekulární genetiky Akademie věd ČR, v. v. i., Akademie věd České republiky. STANĚK, D. RNA – temná hmota v našich buňkách. [online]. [cit. 2022-03-11]. Dostupné z: https://www.img.cas.cz/files/2012/10/RNA-temna_hmota_v_nasich_bukach.pdf)

²⁶ MUKHERJEE, S. The gene: an intimate history. New York: Scribner. 2017. ISBN 978-1-4767-3352-4. St. 500.

²⁷ *Ibid.* st. 499.

²⁸ ŠÍPEK, A. RNA – Ribonukleová kyselina. 2008. [online]. Genetika - Biologie. [cit. 2010-02-11]. Dostupné z: <http://www.genetika-biologie.cz/ribonukleova-kyselina>

²⁹ KOHOUTOVÁ, M. Lékařská biologie a genetiky (II. díl). 1. vydání., Praha: Nakladatelství Karolinum, 2013. St. 202. ISBN 9788024637907.

lidských genů pohybuje kolem 25 000, přičemž každý člověk má dvě kopie každého genu, každou zděděnou po jednom svém rodiči.³⁰ Geny jsou ve dvoušroubovici DNA zaneseny ve formě kódů v nukleotidech,³¹ přičemž tyto sekvence jsou přepsány do RNA³² dále kontrolující a regulující výrobu bílkovin. Geny se mohou vyskytovat ve více formách, tzv. *alelách*, kdy u každé z nich lze nalézt nepatrný rozdíl v sekvenci nukleotidů DNA, která determinuje podstatu genu z molekulárně genetického pohledu.³³

Chromozomy jsou útvary v buňkách, ve kterých jsou uspořádány řetězce DNA a proteiny nesoucí genetické informace.³⁴

Polymorfismus je běžně se vyskytující genetická změna úseku DNA (frekvence jeho výskytu přesahovat v populaci 1 %).³⁵ Může spočívat buď v odlišnosti bází, tedy jejich pořadí, počtu nebo typů.³⁶

Genetické choroby či *geneticky podmíněné choroby* jsou taková onemocnění, u nichž se předpokládá, že pramení z určité genetické informace či konkrétní mutace genu a jsou jimi podmíněny. Každé onemocnění, tj. jeho průběh, původ a projevy, je u každého jedince určitým způsobem jedinečně modifikováno v reakci na unikátní genetické informace dané osoby.³⁷ Jejich původ může být různý, mohou být způsobeny téměř výhradně určitou variantou alely genu, resp. poškozením jednoho genu, tyto označujeme jako *monogenní nemoci*. Další

³⁰ Počet genů se pohybuje mezi 20 000 – 25 000, přičemž každý člověk má dvě kopie každého genu, každou zděděnou po jednom svém rodiči. (NIH. National Library of Medicine. Medlineplus Genetics. MedlinePlus.What is a gene? [online]. [cit. 2022-03-23]. Dostupné z: <https://medlineplus.gov/genetics/understanding/basics/gene/>

³¹ TEPLÁ, M.. Struktura DNA a RNA. [online]. Přírodovědecká fakulta Univerzity Karlovy v Praze. [cit. 2022-04-06]. Dostupné z: <http://www.studiumbiochemie.cz/na.html>

³² Ústav molekulární genetiky Akademie věd ČR. Op. cit. 25.

³³ NIH. National Human Genome Research Institute. Allele. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.genome.gov/genetics-glossary/Allele>

³⁴ MUKHERJEE, S. Op. cit. 26, st. 499.

³⁵ Pozn. Pokud je výskyt méně častý, tj. přítomný u méně než 1 % populace, jedná se o tzv. *mutaci*.

³⁶ Některé polymorfismy jsou zcela neškodné či mohou mít dokonce ochranný charakter. Potom mají za následek pouze rozdílnosti jednotlivých lidí, avšak některé mohou být rizikové a mít povahu spouštěčů genetických onemocnění či dokonce vést k vyšší náchylnosti k infekčním chorobám. (FERWERDA, B., MCCALL, M. B. B., ALONSO, S. (...) NETEA, M. G. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. 2007. [online]. [cit. 2022-03-23]. Dostupné z: <https://doi.org/10.1073/pnas.0704828104>)

³⁷ Doba genova. Genetické Choroby. [online]. [cit. 2022-03-23]. Dostupné z: <https://doba-genova.cz/geneticke-ochorenia/>

kategorii tvoří *chromozomální poruchy*, které spočívají v nižším či vyšším počtu genů v chromozomech či jejich částech, tedy změnou chromozomální struktury. Můžeme též hovořit o *komplexních (multigenních) nemocech*, které jsou odrazem určitých kombinací alel několika genů a vlivem dalších epigenetických i vnějších faktorů. V neposlední řadě je vhodné uvést kategorii onemocnění vznikajících na základě poruch *genomického imprintingu*, což je proces, který způsobí, že aktivita jedné z forem genu, tedy maternální či paternální kopie, které mají z počátku „stejnou váhu projevu“, bude utlumená a druhá plně funkční.³⁸

Genomika je věda, resp. obor, zabývající se studiem genomu různých organismů, jejímž cílem je především porozumět jejich struktuře, fungování a obecně informacím obsaženým v DNA, jejich vlastnostem a též komunikací mezi nimi.³⁹ Tento obor se zabývá především zkoumáním komplexních geneticky podmíněných onemocnění.⁴⁰

2.3 Sekvenování DNA

Sekvenování DNA neboli *sekvenace* či *sekvencování* je metoda, s jejíž pomocí je možné určit pořadí nukleotidů, resp. dusíkatých bází, v konkrétním úseku DNA. Z vyobrazeného úseku DNA jsou potom čitelné genetické informace obsažené v konkrétním úseku DNA. Z takto získaných údajů je následně možné dovodit, jaké úseky DNA obsahují geny samotné a jaké úseky nesou pouze regulativní instrukce, tj. např. aktivují či deaktivují geny.⁴¹ Z údajů mohou být též zřejmé genetické změny či anomálie, které mohou mít za následek různé choroby.⁴² Sekvenování DNA je v dnešní době jednou z nejvíce využívaných metod analýzy biologických materiálů, z tohoto důvodu se touto problematikou zabývá mnoho institucí a vývojářů, kteří se snaží o její co největší zefektivnění.

Základní metodou sekvenování DNA je klasická Sangerova metoda známá již ze sedmdesátých let dvacátého století, jejímž původcem je výše zmíněný F. Sanger. Nyní jsou

³⁸ KING, R. C., STANSFIELD, W. D., MULLIGAN, P. K.. A Dictionary of Genetics, Seventh Edition. 2007. [online]. Oxford University Press. [cit. 2022-03-23]. Dostupné z doi: 10.1093/acref/9780195307610.001.0001.

³⁹ NIH National Human Genome Research Institute. A Brief Guide to Genomics. [online]. [cit. 2022-03-23]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics>

⁴⁰ Státní zdravotní ústav. SOUČEK, P. Genetika a zdravotní stav populace. [online]. [cit. 2023-03-23]. Dostupné z: <http://www.szu.cz/tema/genetika>

⁴¹ NIH. National Human Genome Research Institute. DNA Sequencing Fact Sheet. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Fact-Sheet>

⁴² *Ibid.*

však více využívané tzv. metody druhé generace či metody nové generace (NGS). Tyto metody spočívají ve využívání bioinformatických⁴³ metod ke zpracování většího množství dat, které jsou potom porovnávány s tzv. referenčním genomem.⁴⁴ Tento genom může být použit pro mapování DNA a jakožto „záchytný bod“ při porovnávání genomů sekvenovaných od různých jedinců. Úseky DNA, které jsou již sekvenované, jsou označovány jakožto *reads*. Aby byl určitý úsek DNA považován za dostatečně přečtený, je třeba provést tuto operaci opakovaně, minimálně dvakrát. Pojem pro takto přečtený úsek DNA je *coverage*. Pokud se jedná o sekvenování celého genomu, je žádoucí, aby *coverage* obsahovala minimálně od desítek do tisíců *reads*, přičemž tato *coverage* nemusí být, a většinou není, rovnoměrná.⁴⁵ Jedná se o proces zahrnující zpracování velkého množství dat a potřebu četného ověřování výsledků.

Sekvenování genomu může probíhat jakožto *de-novo sekvenování*, které se provádí u organismů, jejichž genom zatím sekvenován nebyl, není známý a staví se tedy „od nuly“. V takových případech je nutná *coverage* zpravidla 20 až 100násobná. Při *resequencingu*, opakovaném sekvenování, tedy sekvenování genomů dalších jedinců určitého druhu, jejichž genomická sekvence je již známá, je třeba *coverage* v rozmezí 10 až 50násobném.

Sekvenovány jsou též kratší úseky genů například při zkoumání frekvence genetických anomálií v populaci, spojené se čtením PCR. Někdy je třeba určit sekvenci jen několika určitých genů či jejich částí, v takových případech je užíváno tzv. *targeted sequencing*, tedy cílené sekvenování, kdy jsou čteny pouze určité oblasti. Cílené sekvenování je hojně užíváno v klinické praxi, zejména figuruje v onkologické diagnostice,⁴⁶ a to hned z několika důvodů. Jelikož jsou sekvenovány pouze konkrétní úseky DNA, tyto části genu je možné přečíst velmi přesně a důkladně (s *coverage* až více než 1000násobnou), s vysokou citlivostí, a přitom

⁴³ Bioinformatika je obor, ve kterém se setkává biologie a informatika při studiu biologických dat jako jsou DNA, RNA a proteiny. (NIH. National Human Genome Research Institute. Bioinformatics. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.genome.gov/genetics-glossary/Bioinformatics>

⁴⁴ GONZAGA-JAUREGUI, C., LUPSKI, J. R., GIBBS, R. A. Human genome sequencing in health and disease. 2012. [online]. Annual Review of Medicine. 63:35-61. [cit. 2022-04-06]. Dostupné z: <https://doi.org/10.1146/annurev-med-051010-162644>

⁴⁵ KREJČÍ A., MÜLLER P., VOJTĚŠEK B. Bioinformatika a sekvenování nové generace. [online]. Klin Onkol 2015; 28 (Suppl 2): 2S93. [cit. 2022-04-06]. Dostupné z: <http://dx.doi.org/10.14735/amko20152S91>

⁴⁶ KOUBKOVÁ, L., VOJTĚŠEK, B., VYZULA, R. Sekvenování nové generace a možnosti jeho využití v onkologické praxi. [online]. Klin Onkol 2014; 27 (Suppl 1): S61–S68. [cit. 2022-04-06]. Dostupné z: doi: 10.14735/amko20141S61.

odpadá poměrně vysoká finanční zátěž. Díky těmto přednostem je možné detekovat genetické mutace.⁴⁷

2.4 Přehled některých metod sekvenování DNA

Klasickou a dlouho nejpoužívanější metodou je Sangerova metoda,⁴⁸ kterou se stanovuje uspořádání (resp. pořadí) nukleotidů v molekule DNA.⁴⁹ Využívá proces replikace DNA, tedy princip, který je využíván např. při PCR testování.⁵⁰ Tato metoda, jejímž přesným procesem se není nutné pro účely této práce blíže zabývat, je v dnešní době využívána zejména v rámci projektů menšího měřítka za účelem získání sekvence delšího úseku DNA.⁵¹

Po přelomu druhého tisíciletí nastal rozmach nových metod sekvenování DNA, sekvenování nové generace – NGS.⁵² Od té doby dochází ke stálému vývoji metod čtení DNA, dnes označovaným jako sekvenování třetí generace.⁵³ NGS s sebou přináší výhody větší efektivity a přesnosti, menší časové i finanční náročnosti a možnosti využití sekvenování v mnohonásobně větším měřítku.⁵⁴ Pro ilustraci, sekvenování celého lidského genomu na mezinárodní úrovni trvalo cca 13 let a na výzkumu se podílelo mnoho subjektů.⁵⁵ Společnosti Celera Genomics to samé později trvalo pouhé 2 roky,⁵⁶ přičemž byly užívány stovky zařízení a technologií umožňující jejich automatický provoz. Dnes jsou laboratoře

⁴⁷KOUBKOVÁ, L. Op. cit. 46.

⁴⁸ SANGER, F., COULSON, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. [online]. Journal of Molecular Biology, 94(3), 441–448. [cit. 2022-03-24]. Dostupné z: [https://doi.org/10.1016/0022-2836\(75\)90213-2](https://doi.org/10.1016/0022-2836(75)90213-2)

⁴⁹ VFU Brno, Fakulta veterinární hygieny a ekologie, Ústav biologie a chorob volně žijících zvířat. Sekvenování DNA, Enzymová Sangerova metoda. [online]. [cit. 2022-03-24]. Dostupné z: https://cit.vfu.cz/opvk2011/?title=popis_metod-sekvenovani&lang=cz

⁵⁰ Merck. Sanger Sequencing Steps & Method. [online]. [cit. 2022-03-24]. Dostupné z: <https://www.sigmaaldrich.com/CZ/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing>

⁵¹ Illumina. NGS vs. Sanger Sequencing. [online]. [cit. 2022-03-24]. Dostupné z: <https://www.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html>

⁵² *Ibid.*

⁵³ SLATKO, E. B., GARDNER, F. A., AUSUBEL, M. F. Overview of Next Generation Sequencing Technologies. 2018. [online]. [cit. 2022-03-24]. Dostupné z: doi:10.1002/cpmb.59.

⁵⁴ PATEL, N., FERNS, R. B., NASTOULI, E., KOZLAKIDIS, Z., KELLAM, P., MORRIS, S. Cost analysis of standard Sanger sequencing versus next generation sequencing in the ICONIC study. The Lancet (British edition) [online]. Elsevier, 2016, 388, S86-S86 [cit. 2022-04-05]. ISSN 0140-6736. Dostupné z: doi:10.1016/S0140-6736(16)32322-4.

⁵⁵ NIH. Op. cit. 23.

⁵⁶ Nature. The Human Genome Project. Nature News. Barranco, C. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.nature.com/articles/d42859-020-00101-9>

schopny tento proces zvládnout za pomoci jediného přístroje a jediného člověka v rozmezí jednoho dne či dokonce během několika hodin.⁵⁷ Při využívání NGS je největší důraz kladen právě na vyhodnocování tímto způsobem získaných dat velkého objemu,⁵⁸ ze kterých lze vyčíst velké množství údajů různé povahy, jako například pravděpodobnost rozvoje určitých onemocnění či jejich dědičnost.

Z lidské DNA jsou různými způsoby získávána rozmanitá genetická data různého rozsahu, objemu a obsahu a za různými účely. Aby však byla využitelná, zejména v oblasti výzkumu, je třeba jejich zpracování vhodným způsobem a umožnění jejich dalšího zpřístupnění.

2.5 Aktuální projekty zaměřené na shromažďování genetických dat

Genetická data jsou většinou zpracovávána za užití různých počítačových systémů, software, tzv. cloudů⁵⁹ apod. Mnohé společnosti jako např. Google⁶⁰ či Microsoft⁶¹ vyvíjejí úložiště a programy pro zpracování a nakládání s genetickými daty. Zpracováním genetických dat se zabývají také nadnárodní instituce, mezinárodní organizace a další zainteresované subjekty.

Zde je vhodné poukázat na iniciativu Evropské unie (EU) „1+ Million Genomes“ (1+MG), do které je zapojeno 24 členských států EU, Spojené království Velké Británie a Severního Irska (UK) a Norsko. Tento projekt si klade za cíl získat genetická data z více než jednoho milionu jednotlivců po celé Evropě, včetně pacientů s různými onemocněními a zdravých jedinců a vytvořit rozsáhlou veřejnou databázi genomů v Evropě, která bude sloužit k podpoře výzkumu lidské genetiky a usnadnění vývoje personalizované medicíny.⁶²

⁵⁷ Stanford Medicine. Fastest DNA sequencing technique helps undiagnosed patients find answers in mere hours. News Center. [online]. [cit. 2022-03-25]. Dostupné z: <https://med.stanford.edu/news/all-news/2022/01/dna-sequencing-technique.html>

⁵⁸ Generi Biotech. NOVOTNÁ, M. Princip NGS Metody. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.generi-biotech.com/cs/princip-ngs-metody/>

⁵⁹ Cloudflare. What is the cloud? Cloud definition. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.cloudflare.com/en-gb/learning/cloud/what-is-the-cloud/>

⁶⁰ Google Cloud Whitepaper. Handling genomic data in the cloud. 2019. [online]. [cit. 2022-03-25]. Dostupné z: <https://cloud.google.com/files/genomics-data-wp.pdf>

⁶¹ Microsoft. Microsoft Genomics. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.microsoft.com/en-us/genomics/>

⁶² European Commission. European '1+ Million Genomes' Initiative. [online]. [cit. 2022-03-26]. Dostupné z: <https://digital-strategy.europa.eu/en/policies/1-million-genomes>

Evropská Unie přijala v roce 2020 strategii nazvanou The European Health Data Strategy, která obnáší mimo jiné i vytvoření společného evropského datového prostoru, tzv. European Health Data Space (EHDS). Zaměřuje se na zjednodušení sdílení zdravotních údajů, do kterých spadají i genetická data.⁶³ Cílem strategie je vytvořit jednotný digitální prostor pro zdravotnická data v Evropě a dále podporovat výzkum preventivních postupů, léčebných metod, léčivých přípravků a zdravotnických prostředků na evropské úrovni. Tato iniciativa by měla umožnit efektivní sdílení a využití zdravotnických dat mezi členskými státy EU a zlepšit koordinaci a interoperabilitu systémů zdravotních služeb. Součástí má však být i zaručení kontroly občanů EU nad jejich osobními údaji a jejich soukromí.⁶⁴ EHDS je věnována poslední kapitola této práce.

Dalším zajímavým projektem v této oblasti je „100,000 Genomes Project“⁶⁵ zastřešovaný společností Genomics England v rámci Ministerstva zdravotnictví Spojeného království. Do tohoto projektu je zapojeno deset renomovaných farmaceutických společností včetně GlaxoSmithKline, AstraZeneca, Roche či GSK, které zkoumají sdílená genetická data. V rámci této iniciativy je cíleno na užití odvětvových odborných znalostí farmaceutických firem při zpracování a analýze genetických dat v souvislosti s výzkumem nových léčivých přípravků a inovativních léčiv, či znovuzavedení nebo jiné užití již známých léčivých přípravků, to vše s co největším prospěchem pro pacienty.⁶⁶

⁶³ Assessment of the EU Member States' rules on health data in the light of GDPR, Specific Contract No SC 2019 70 02 in the context of the Single Framework Contract Chafea/2018/Health/03. [online]. European Commission. [2022-03-26]. Dostupné z: https://ec.europa.eu/health/system/files/2021-02/ms_rules_health_data_en_0.pdf

⁶⁴ European Commission. Towards European Health Data Space. [online]. [cit. 2022-03-26]. Dostupné z: https://www.ema.europa.eu/en/documents/presentation/presentation-towards-european-health-data-space-mdorazil-dg-sante_en.pdf

⁶⁵ Genomics England. 100,000 Genomes Project. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.genomicsengland.co.uk/initiatives/100000-genomes-project>

⁶⁶ Pinset Masons, Out-Law News. Drugs companies collaborate to unlock potential of 100,000 Genomes project. 2015. [online]. [cit. 2022-03-26] Dostupné z: <https://www.pinsentmasons.com/out-law/news/drugs-companies-collaborate-to-unlock-potential-of-100000-genomes-project>

3. Využití sekvenování DNA

Využití sekvenování DNA má své místo v mnoha oblastech, od různých vědeckých oborů, přes lékařství, forenzní vědy⁶⁷ až po zemědělství⁶⁸. Pro každý obor je podstatný jiný způsob sekvenování DNA a od toho se dále odvíjejí užívané metody. Například pro fylogenetiku je výzkum založen na sekvenování DNA kódující RNA,⁶⁹ pro kriminalistiku mají naopak největší význam opakující se oblasti DNA zvané mikrosatelity, díky kterým lze od sebe odlišovat konkrétní jedince.⁷⁰ Ve forenzních vědách jsou potom genetické informace využívány zejména k identifikaci pachatelů či obětí trestné činnosti, k tvorbě forenzních databází či identifikaci nezvěstných osob.⁷¹ S každou metodou a specifiky konkrétních oborů jsou též spojeny různé právní problematiky jako je např. zákonný základ zpracování genetických údajů či způsoby jejich využití, ukládání a zpřístupňování. To, co mají tyto metody společné, jsou samotné genetické údaje.

Obor, se kterým je sekvenování DNA spojeno nejužší, je bez pochyby medicína, s ní spojený výzkum a farmacie. Sekvenování usnadňuje a urychluje diagnostiku nemocí, možné zjištění či predikci míry pravděpodobnosti či náchylnosti k určitým chorobám (např. kardiovaskulárním onemocněním⁷² nebo rakovině⁷³). Sekvenování dále usnadňuje a posouvá vývoj léčivých přípravků zejména v souvislosti s některými onemocněními. Slibnou léčbou je například tzv. genová terapie, která je využitelná při léčbě nemocí jako rakovina, cystická fibróza, srdeční choroby, některé typy diabetu, AIDS, hemofilie nebo alespoň na posílení organismu při odolávání těmto nemocem.⁷⁴ Při léčení konkrétních jedinců

⁶⁷ Illumina. Forensic genomics. NGS solutions for forensic applications. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.illumina.com/areas-of-interest/forensic-genomics.html>

⁶⁸ Illumina. Agrigenomics. Applying agricultural genomics to improve the food supply. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.illumina.com/areas-of-interest/agrigenomics.html>

⁶⁹ HILLIS, D.M., DIXON, M.T. Ribosomal DNA: molecular evolution and phylogenetic inference. [online]. Q Rev Biol. 1991 Dec;66(4):411-53. [cit. 2022-03-27]. Dostupné z: doi: 10.1086/417338.

⁷⁰ *Ibid.*

⁷¹ MAZURA, I. Genetický polymorfismus jako nástroj identifikace osob v kriminalistické a soudně-lékařské praxi. [online]. [cit. 2022-03-26]. Dostupné z: <https://soudni.lf1.cuni.cz/file/5717/Forenzn%C3%AD%20genetika.pdf>

⁷² O'DONNELL, C. J., NABEL, E. G. Genomics of Cardiovascular Disease. 2011. [online]. New England Journal of Medicine. [cit. 2022-04-05]. Dostupné z: doi: 10.1056/NEJMra1105239.

⁷³ NOGRADY, B. How cancer genomics is transforming diagnosis and treatment. Nature (London) [online]. England: Nature Publishing Group, 2020, 579(7800), S10-S11 [cit. 2022-04-05]. ISSN 0028-0836. dostupné z: doi:10.1038/d41586-020-00845-4.

⁷⁴ Mayo Clinic. Mayo Foundation for Medical Education and Research. Gene therapy. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.mayoclinic.org/tests-procedures/gene-therapy/about/pac-20384619>

je vhodné přihlížet v neposlední řadě k jejich genotypům⁷⁵ a vybírat tak vhodné léčebné postupy, což je díky případné znalosti genetických údajů daného pacienta možné.

Jednou z nejnovějších technologií, která je spjata se znalostí sekvenování DNA, je léčba modifikací genomů. Příkladem může být technologie genetických nůžek neboli CRISPR-Cas9, za kterou byly v roce 2020 oceněny Nobelovou cenou za chemii vědkyně Jennifer Doudna a Emmanuelle Charpentier.⁷⁶ Za pomoci metody CRISPR je možné modifikovat či opravit genom jakéhokoli organismu, a to efektivně, poměrně levně a rychle.⁷⁷ Vědci se domnívají, že v budoucnu bude možné díky metodě CRISPR cílit také na ničení karcinogenních nádorů⁷⁸ apod.⁷⁹ V září roku 2021 udělil Úřad pro kontrolu potravin a léčiv Spojených států amerických (FDA) souhlas ke klinickému testování genové terapie na bázi technologie CRISPR v souvislosti s léčbou HIV.⁸⁰

Díky stále většímu množství sekvenovaných genomů, jejich zpracování a diagnostice přichází vědci a lékaři na konkrétní geny či úseky, které mohou mít vliv na výskyt určitých onemocnění. Nacházeny a určovány jsou též další konkrétní genetické mutace, které jsou původci genetických onemocnění. V roce 1983 byla s využitím znalostí genetického polymorfismu zmapována první genetická choroba – dědičná Huntigtonova choroba, tj. porucha, jejímž důsledkem je např. ztráta kognitivních funkcí či nástup demence.⁸¹ Příkladem dalších geneticky podmíněných onemocnění jsou např. achondroplazie – disproporcionální trpaslictví⁸², cystická fibróza – onemocnění postihující dýchací a trávicí

⁷⁵ Soubor genetických informací organismu, který determinuje jeho fyzické, chemické, biologické a intelektuální charakteristiky. (MUKHERJEE, S. Op. cit. 26, st. 500.)

⁷⁶ The nobel Prize. The Nobel Prize in Chemistry 2020. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.nobelprize.org/prizes/chemistry/2020/summary/>

⁷⁷ Doudna Lab. Editing in Cells. [online]. [cit. 2022-03-27]. Dostupné z: https://doudnalab.org/research_areas/gene-fixing/

⁷⁸ ISAACSON, W. The Code Breaker: Jennifer Doudna, Gene Editing, and the Future of the Human Race. Simon & Schuster. 2021. ISBN-13. 978-6267012802. Kapitola 32.

⁷⁹ DOUDNA, J. A., STERNBERG, S. H. A crack in creation: Gene editing and the unthinkable power to control evolution. Vintage, 2018. ISBN-10: 9781784702762 Zejm. kapitola 6.

⁸⁰ Clinical Trials Arena. FDA approves first trial investigating CRISPR gene editing as HIV cure. 2021. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.clinicaltrialsarena.com/news/crispr-gene-editing-hiv-cure/>

⁸¹ NIH. National Human Genome Research Institute. 1983: First Disease Gene Mapped. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.genome.gov/25520322/online-education-kit-1983-first-disease-gene-mapped>

⁸² NIH. National Library of Medicine. Medline Plus. Achondroplasia. [online]. [cit. 2023-03-19]. Dostupné z: <https://medlineplus.gov/genetics/condition/achondroplasia/>

soustavu⁸³, hemoglobinopatie – srpkovitá anémie⁸⁴, hemofilie – porucha srážlivosti krve⁸⁵, chromozomové vady jako jsou např. Downův syndrom, Turnerův syndrom, Edwardsův syndrom⁸⁶, multigenové vady, tj. např. diabetes (1. a 2. typu) či poruchy genomického imprintingu, jehož příkladem je Praderův-Williho syndrom⁸⁷, který vede k neustálému pocitu hladu s následkem chronického přejídání a s tím spojenou obezitou.

Sekvenování genomu je v současné době ve zdravotnictví využíváno především k identifikaci pravděpodobnosti výskytu chorob u konkrétních jedinců, dědičnosti genetických poruch a onemocnění a reakce organismu na léčiva, což pomáhá farmaceutickému průmyslu zaměřit svůj výzkum na vývoj cílených terapií.⁸⁸ S dostupností stále většího množství vzorků lidských genomů v databázích je též možné provádět simulace, díky kterým lze určit předpokládané účinky léčivých přípravků při jejich užití konkrétním jedincem s poměrně vysokou pravděpodobností, což je nyní označováno za obor farmakogenomiky.⁸⁹ Se zvyšujícím se množstvím dat jsou vědci schopni snadněji najít vzory a odchylky v určitých sekvencích a zjistit jejich význam – příkladem je hledání konkrétní varianty lidského genomu odpovědné za schizofrenii, která nebyla v rámci 3 500 vzorků nalezena vůbec, tušení o ní vědci získali až po analýze 10 000 vzorků a teprve v rámci 35 000 vzorků bylo konečně možné tuto

⁸³ Mayo Clinic. Mayo Foundation for Medical Education and Research. Cystic Fibrosis. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.mayoclinic.org/diseases-conditions/cystic-fibrosis/symptoms-causes/syc-20353700>

⁸⁴ ISAACSON, W. The Code Breaker: Jennifer Doudna, Gene Editing, and the Future of the Human Race. Simon & Schuster. 2021. ISBN-13. 978-6267012802. Kapitola 32.

⁸⁵ Centers for Disease Control and Prevention. What is Hemophilia? [online]. [cit. 2023-03-19]. Dostupné z: <https://www.cdc.gov/ncbddd/hemophilia/facts.html>

⁸⁶ Encyclopedia Britannica. Britannica, T. Editors of Encyclopaedia. Chromosomal disorder., 2017. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.britannica.com/science/chromosomal-disorder>.

⁸⁷ NIH. National Library of Medicine. Medline Plus. Prader-Willi Syndrome. [online]. [cit. 2023-03-19]. Dostupné z: <https://medlineplus.gov/genetics/condition/prader-willi-syndrome/>

⁸⁸ Proclinical. How genome sequencing is aiding drug research and development. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.proclinical.com/blogs/2015-7/genome-sequencing-is-aiding-drug-research-and-development>

⁸⁹ KATSILA, T., PATRINOS, G. P. Whole genome sequencing in pharmacogenomics. Frontiers in pharmacology. [online]. Switzerland: Frontiers Media S.A, 2015, 6, 61-61 [cit. 2022-04-05]. ISSN 1663-9812. Dostupné z: doi:10.3389/fphar.2015.00061.

variantu určit s dostatečnou statistickou pravděpodobností.⁹⁰ Obdobné postupy a analýzy jsou vyvíjeny též na poli bioinformatiky.⁹¹

3.1 Big data a data mining

S ohledem na výše uvedené není překvapivé, že k dalšímu pokroku je nutný velký objem genetických dat, jejich zpracování a přístup k nim. Počet vzorků sekvenovaných genomů se neustále zvyšuje, což koreluje se zvyšováním dostupnosti genetického sekvenování široké veřejnosti, ať už pacientům či pouhým zájemcům o vyšetření.⁹²

Předpokládá se, že během dalšího desetiletí bude vygenerováno 2 až 40 miliard gigabytů genetických dat, která budou vědci a laboratoři neustále analyzována za pomoci různých výpočetních, statistických a vizualizačních metod a zpracovávají s cílem vytrždit podstatná data pro jejich výzkum a určit tak různé predikce pravděpodobnosti se základem v těchto datech.⁹³ Tyto údaje pochází od milionů lidí z celého světa a jsou často sdíleny napříč vědeckou komunitou pro výzkumné a analytické účely. S tím však také přichází nemálo právních, technických, bezpečnostních a dalších úskalí.

Při sekvenování DNA je neustále získáváno velké množství dat různého druhu a rozsahu, která jsou následně zpracovávána, ukládána a zpřístupňována, o čemž bude pojednáno níže. Tato data je možné podřadit pod tzv. *big data*⁹⁴, tedy rozmanitá komplexní data, přicházející ve velkém rozsahu a objemu z více autonomních zdrojů, přičemž je zde snaha o jejich rychlé vzájemné propojení a systematické zpracování. Pro práci s big data je třeba zajistit datovou integraci, jejich správu a následnou analýzu. Aby bylo možné využít výhody datového souboru v širší big data, je vhodné standardizovat přístup k jejich získávání a zpracování za použití

⁹⁰ UNITED STATES, PODESTA, J. Big data: seizing opportunities, preserving values. 2014. [online]. [cit. 2022-04-05]. Dostupné z: <http://purl.fdlp.gov/GPO/gpo64868>.

⁹¹ CHIMUKANGARA, B., WILKINSON, E., FISH, M., RAMSURAN, V. Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. *Frontiers in genetics* [online]. Frontiers Media S.A, 2020, 11, 544162-544162 [cit. 2022-04-05]. ISSN 1664-8021. Dostupné z: doi:10.3389/fgene.2020.544162.

⁹² PADMAGIRISON, S. Trustworthy AI provides business insights for the global leader in genomics. 2021. [online]. IBM. [cit. 2022-04-05]. Dostupné z: <https://www.ibm.com/blogs/journey-to-ai/2021/10/trustworthy-ai-provides-business-insights-for-the-global-leader-in-genomics/>

⁹³ NIH. National Human Genome Research Institute. Genomic data science. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/Genomic-Data-Science>

⁹⁴ Oracle. Co jsou big data? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.oracle.com/cz/big-data/what-is-big-data/#three>

vymezených technologií a sladit tyto postupy se způsobem ukládání, tj. v dnešní době většinou za užití cloudu a cloud computingu.⁹⁵

Big data je též pojem pro označování techniky sběru, spravování a zpracování dat velkého rozsahu, bez ohledu na jejich formu a strukturu, která je většinou automatizovaná. Aby bylo možné využívat potenciál přístupu k big data, je třeba tzv. *data miningu*, metody sloužící k extrakci a procesování důležitých informací z objemných datových souborů. V rámci data miningu jsou zpracovávány rozsáhlé soubory dat nalézáním určitého vzoru ve vztazích mezi nimi, a to jak automatizovaně za pomoci různých algoritmů a programů, tak též manuálně.⁹⁶ Jelikož výzkum a práce s genetickými daty zahrnuje porovnávání, zkoumání a hodnocení obrovského množství různých vzorků a údajů, vznikají stále nové specifické databáze a programy⁹⁷ obsahující různé nástroje, které ve svém celku cílí na umožnění a usnadnění takovéto činnosti.⁹⁸

⁹⁵ Viz kap. 4.5.

⁹⁶ WU, X., ZHU, X., WU, G. -Q., DING, W. Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*. Vol. 26, no. 1, pp. 97-107, [online]. [cit. 2023-03-16]. Dostupné z: doi: 10.1109/TKDE.2013.109.

⁹⁷ Např. JINXIANG, C., LI, F., WANG, M., LI, J., MARQUEZ-LAGO, T. T., LEIER, A., REVOTE, J., LI, S., LIU, Q., SONG, J. BigFiRSt: A Software Program Using Big Data Technique for Mining Simple Sequence Repeats From Large-Scale Sequencing Data [online]. 18. 1. 2022, 1-16. [cit. 2022-03-25]. Dostupné z: <https://doi.org/10.3389/fdata.2021.727216>

⁹⁸ Viz kap. 4.

4. Genetické databáze a genome browsery

Pro rozvoj výzkumu v této oblasti je žádoucí, aby data získána sekvenováním DNA byla ukládána takovým způsobem, který umožňuje rychlý, efektivní, jednoduchý a volný přístup, a který by měl zároveň zajišťovat dostatečné zabezpečení zpracovávaných údajů ve vztahu k ochraně soukromí osob. Za těmito účely jsou využívány genetické databáze a programy *sui generis* označované jako genome browsery. Genome browsery umožňují uživatelům vyhledávat, procházet a vizualizovat data celých genomů včetně jejich anotací za použití různých zdrojů, procházet strukturu genů, proteinů, jejich variace, provedené srovnávací analýzy apod.⁹⁹ Vědci jsou díky těmto programům schopni porovnávat konkrétní části lidských genomů napříč populací, porovnávat je s vlastními daty, získávat mnoho genetických údajů pro svůj výzkum ve spojení s hledáním konkrétních genetických variabilit nebo zobrazovat vyjádření nukleotidových polymorfismů. To vše s přístupem k informacím o jednotlivých dusíkatých bázích s ohledem na určité úseky DNA, jejich vlivu na evoluci atp.¹⁰⁰

Většina těchto genetických databází jsou tzv. agregované databáze. Tyto databáze obsahují genetická data z různých zdrojů, jako jsou např. klinické studie, genetické testy a jiné veřejné genetické databáze. Jejich přednost spočívá v tom, že obsahují data z mnoha zdrojů, které následně uceluje a získané informace spojuje, čímž ulehčuje samotné vyhodnocování a porovnávání předmětných dat vědci a výzkumnými institucemi.¹⁰¹

Příkladem genome browseru je Ensembl¹⁰², projekt Evropského Institutu Bioinformatiky (EBI), Evropské Laboratoře Molekulární Biologie (EMBL) a Wellcome Trust Sanger Institutu ve Spojeném království. Dalším příkladem je UCSC Genome Browser¹⁰³ Kalifornské

⁹⁹ Yourgenome. How are sequenced genomes stored and shared? 2016. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.yourgenome.org/facts/how-are-sequenced-genomes-stored-and-shared>

¹⁰⁰ HARTE, R.A., KAROLCHIK, D., KUHN, R.M., KENT, W.J., HAUSSLER, D. Databases and Genome Browsers. M.R. Speicher et al. (eds.), Vogel and Motulsky's Human Genetics, Problems and Approaches. Springer, Berlin, Heidelberg. 2010. [online]. [cit. 2023-03-16]. Dostupné z: doi: 10.1007/978-3-540-37654-5_44.

¹⁰¹ LEK, M., KARCZEWSKI, K. J., MINIKEL, E. V., et al. Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. 2016. [online]. Nature, 536(7616), 285–291. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1038/nature19057>

¹⁰² Ensembl. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ensembl.org/index.html>

¹⁰³ UCSC Genomics Institute. UCSC Genome Browser. [online]. [cit. 2023-03-19]. Dostupné z: <https://genome.ucsc.edu>

Univerzity v Santa Cruz, nebo Genome Data Viewer¹⁰⁴ zaštiťovaný Národním centrem pro biotechnologické informace (NCBI), jež je součástí Národní lékařské knihovny USA v areálu americké agentury National Institutes of Health (NIH).

Soubory genetických dat jsou prostřednictvím genome browserů a genetických databází většinou volně a bezplatně přístupné široké veřejnosti, přičemž data by neměla umožnit zpětné dohledání subjektů údajů. K efektivnímu využití takto přístupných dat a genome browserů je však nezbytné disponovat znalostmi genetiky a molekulární biologie, z tohoto důvodu jsou využívány především odborníky, tedy zejm. vědci pracujícími s genetickým materiálem, bioinformatiky a klinickými a akademickými pracovníky. Data obsažená v genome browsech pramení většinou ze spolupráce s různými výzkumnými projekty a databázemi jako jsou např. International Nucleotide Sequence Database Collaboration (INSDC)¹⁰⁵, Single Nucleotide Polymorphism database (dbSNP)¹⁰⁶, the Encyclopedia of DNA Elements (ENCODE)¹⁰⁷, a tzv. 1000 Genomes Project¹⁰⁸.

Ze zástupců genetických databází je vhodné zmínit např. americkou The Genome Sequence DataBase (GSDB) pod správou National Center for Genome Resources (NCGR), veřejně dostupnou databázi sekvencí nukleotidů a s nimi souvisejících biologických a bibliografických informací. Dle dostupných zdrojů jsou v této databázi přístupná data ve vlastnictví GenBank,¹⁰⁹ spadající pod NCBI. GenBank je taktéž veřejná databáze obsahující předmětná data týkající se sekvenovaného DNA, která je součástí GSDB Collaboration složené

¹⁰⁴ NIH. National Library of Medicine. Genome Data Viewer. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ncbi.nlm.nih.gov/genome/gdv/>

¹⁰⁵ International Nucleotide Sequence Database Collaboration. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.insdc.org>

¹⁰⁶ SHERRY, S. T., WARD, M. H., KHOLODOV, M., et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* [online]. [cit. 2022-04-05]. 2001;29(1):308-311. Dostupné z: doi: 10.1093/nar/29.1.308.

¹⁰⁷ Stanford University. Encode. [online], [cit. 2023-03-19]. Dostupné z: <https://www.encodeproject.org>

¹⁰⁸ The International Genome Sample Resource. 1000 Genomes Project Summary. 1000 Genomes. [online]. [cit. 2022-03-26]. Dostupný z: <https://www.internationalgenome.org/1000-genomes-summary>

¹⁰⁹ NIH. National Library of Medicine. GenBank Overview. [online]. [cit. 2022-04-05]. Dostupné z: <https://www.ncbi.nlm.nih.gov/genbank/>

dále z japonské DNA DataBank of Japan (DDBJ)¹¹⁰, the European Nucleotide Archive (ENA)^{111, 112}

4.1 Agregované genetické databáze

Agregace dat je proces, při kterém jsou získávána data ze zdrojů a následně slučována a zpracována tak, aby vznikla určitá ucelená a zjednodušená informace. Agregace dat může být prováděna pomocí různých metod a technologií, včetně statistických modelů, datových sítí a softwarových aplikací. Na úrovni jednotlivých záznamů se data agregují tak, aby byla vytvořena statistika nebo průměr z hodnot jednotlivých záznamů. Na vyšší úrovni se potom data agregují tak, aby byla získána ucelená informace o celkových trendech nebo vzorcích ve zpracovávaných datech.¹¹³

Agregované genetické databáze jsou v posledních letech stále důležitějším nástrojem pro genetiky, biologické vědce a další odborníky zabývající se výzkumem genetických poruch a onemocnění. Tyto databáze umožňují shromažďování, sdílení a vyhodnocování velkého množství genetických dat ze širokého spektra zdrojů. Jednou z výhod agregovaných databází je skutečnost, že zpracovávaná data jsou následně dostupná v ucelené a standardizované podobě tak, aby byla data z různých zdrojů srovnatelná.¹¹⁴

Jednou z největších a nejvýznamnějších agregovaných genetických databází je dbGaP (Database of Genotypes and Phenotypes), která byla vytvořena v rámci NIH v USA. Tato databáze obsahuje genotypová data z více než 250 studií, které zkoumají genetické faktory různých onemocnění a zdravotních stavů.¹¹⁵

¹¹⁰ MASHIMA, J., KODAMA, Y., FUJISAWA, T., KATAYAMA, T., OKUDA, Y., KAMINUMA, E., OGASAWARA, O., OKUBO, K., NAKAMURA, Y., TAKAGI, T. DNA Data Bank of Japan. *Nucleic acids research*, 45(D1), D25–D31. 2017. [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/nar/gkw1001>.

¹¹¹ European Nucleotide Archive. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk/ena/browser/home>

¹¹² HARGER, C., CHEN, G., FARMER, A., et al. The genome sequence DataBase. *Nucleic Acids Res.* [online]-2000;28(1):31-32. [cit. 2022-04-05]. Dostupné z: doi:10.1093/nar/28.1.31.

¹¹³ RAMAN, A. What is Data Aggregation? : A Comprehensive Guide 101. 4. 2. 2021. [online]. [cit. 2023-03-12]. Dostupné z: <https://hevodata.com/learn/data-aggregation/>

¹¹⁴ LAPPALAINEN, T., SCOTT, A. J., BRANDT, M., HALL, I. M. Genomic Analysis in the Age of Human Genome Sequencing. *Cell*, 177(1), 70–84. 2019. [online]. [cit. 2023-03-19] Dostupné z: <https://doi.org/10.1016/j.cell.2019.02.032>

¹¹⁵ NIH. National Library of Medicine. Database of Genotypes and Phenotypes (dbGaP). 2018. [online]. [cit. 2023-03-10]. Dostupné z: <https://www.ncbi.nlm.nih.gov/gap/>

Další významnou agregovanou genetickou databází je ExAC (Exome Aggregation Consortium), která umožňuje vědcům a lékařům analyzovat genetické varianty a jejich výskyt v populaci a také poskytuje informace o zdravotním stavu jedinců, jejichž genetická data jsou zahrnuta.¹¹⁶

Mezi další významné agregované genetické databáze patří též GnomAD (Genome Aggregation Database) a ClinVar (Clinical Variation). GnomAD obsahuje genotypová data více než 140 000 jedinců a umožňuje analýzu genetických variant v populaci.¹¹⁷ ClinVar je databáze, která shromažďuje informace o vztahu mezi genetickými variantami a různými zdravotními stavy, což pomáhá v diagnostice a léčbě různých onemocnění. ClinVar je provozován NCBI a cílem této databáze, resp. genome browseru, je shromažďovat, ukládat a sdílet informace o klinické významnosti genetických polymorfismů v souvislosti s nemocemi, včetně vzácných onemocnění^{118, 119}.

Výhodou agregovaných genetických databází je, že umožňují vědcům a lékařům přístup k velkému množství již určitým způsobem zpracovaných genetických dat, což může pomoci při identifikaci nových genetických variant a poruch. Agregované genetické databáze jsou používány pro mnoho účelů, včetně výzkumu lidské genetiky, genomiky a bioinformatiky a umožňují například identifikaci nových genetických variant a asociaci těchto variant s různými fenotypy, včetně onemocnění a fyziologických vlastností. Dále jsou tyto databáze hojně využívány při studiu, při analýze výsledků různých genomových studií či vyhodnocování vztahů mezi genetickými variantami a nemocemi.¹²⁰

¹¹⁶ LEK, M. Op. cit. 101.

¹¹⁷ GUDMUNDSSON, S., SINGER-BERK, M., WATTS, N.A., PHU, W., GOODRICH, J.K., SOLOMONSON, M.; Genome Aggregation Database Consortium; REHM, H.L., MACARTHUR, D.G., O'DONNELL-LURIA, A. Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat. Srpen 2022. [online]. [cit. 2023-03-11]. Dostupné z: doi: 10.1002/humu.24309.

¹¹⁸ Dle akčního programu Společenství pro vzácná onemocnění přijatého od roku 1999 do roku 2003 jsou vzácná onemocnění definována jakožto taková závažná onemocnění, která postihují méně než 5 z 10 000 osob. Rada Evropské unie ve svém doporučení ze dne 8. 6. 2009, o akci v oblasti vzácných onemocnění, 2009/C 151/02 uvádí, že vzácná onemocnění jsou *velmi složitá onemocnění s nízkou prevalencí, která ohrožují život nebo jsou chronicky invalidizující*. To pro představu znamená, že v Evropské unii trpí odhadem 29 milionů lidí nebo bude trpět 5 000 až 8 000 různými vzácnými onemocněními. (Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů o vzácných onemocněních: Výzvy pro Evropu. Brusel dne 11. 11. 2008. [online]. [cit. 2023-03-19]. Dostupné z: <https://eur-lex.europa.eu/legal-content/CS/TXT/HTML/?uri=CELEX:52008DC0679&from=CS>)

¹¹⁹ LANDRUM, M. J., CHITIPIRALLA, S., BROWN, G. R., (...). ClinVar: improvements to accessing data. Nucleic Acids Res. 2020. [online]. 48(D1):D835-D844. [cit. 2023-03-11]. Dostupné z: doi: 10.1093/nar/gkz972.

¹²⁰ LAPPALAINEN, T. Op. cit. 114.

Genetická data dostupná prostřednictvím agregovaných genetických databází jsou díky způsobu jejich zpracování standardizovaná, zprůměrovaná a navzájem provázaná. To znamená, že jsou ve většině případů v takovém formátu, že nevypovídají o konkrétním genomu jednotlivce, ale spíše o údajích spojených s daným úsekem DNA.

4.2 Genome browser – příklad

Pro bližší ilustraci fungování genome browserů uvádím základní principy a metody fungování jednoho z výše uvedených browserů – *Ensembl*.

Ensembl je databázový projekt, který poskytuje rámec pro uspořádání a organizaci informací ve vztahu k sekvencím genomů. Jedná se o komplexní elektronický zdroj obsahující automatické anotace lidského genomu (definici počátku a konce určitého úseku DNA, informace o konkrétním úseku a jeho funkcích), s ověřenými predikcemi genových lokací, které byly integrovány s externími zdroji dat a které jsou pravidelně aktualizovány. Jedná se o data v agregované podobě. Tato data jsou následně ukládána do databáze (MySQL)¹²¹ za účelem následné analýzy a zobrazení,¹²² a odtamtud jsou též přístupná.¹²³ Data dostupná v Ensembl a s ním spojených databází jsou též přístupná prostřednictvím dalších webových stránek, programů, open source software (k pojmu viz níže) a dalších specifických nástrojů často poskytovaných třetími stranami (BLAST, BLAT, BioMart a Variant Effect Predictor).¹²⁴ Ensembl je k dispozici buď jako interaktivní webová stránka, nebo jako prostý databázový soubor, tzv. *flat file*.

Jedná se o projekt tzv. open source software¹²⁵, software s otevřeným zdrojovým kódem či otevřeného software (OSS)¹²⁶, a free software¹²⁷, neboli software s důrazem na svobodné

¹²¹ Oracle. MySQL 8.0 Reference Manual. [online]. [cit. 2023-03-19]. Dostupné z: https://docs.oracle.com/cd/E17952_01/mysql-8.0-en/

¹²² Ensembl. The Ensembl public MySQL Servers. [online]. [cit. 2023-03-19]. Dostupné z: <https://m.ensembl.org/info/data/mysql.html>

¹²³ STABENAU, A., MCVICKER, G., MELSOPP, C., PROCTOR, G., CLAMP, M., & BIRNEY, E. The Ensembl core software libraries. *Genome research*, 14(5), 929–933. 2004. [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1101/gr.1857204>

¹²⁴ Ensembl. Ensembl Variation - Data access. [online]. [cit. 2023-03-19]. Dostupné z: https://m.ensembl.org/info/genome/variation/tools/data_access.html

¹²⁵ Open Source Initiative. [online]. [cit. 2023-03-19-]. Dostupné z: <https://opensource.org>

¹²⁶ JANSÁ, L. OTEVŘEL, P., ŠTEVKO, M. Softwarové právo, 3. aktualizované rozšířené vydání, r. 2018. Computer Press. St. 290. ISBN: 978-80-251-4914-0.

¹²⁷ Free Software Foundation. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.fsf.org>

užívání ze strany uživatelů,¹²⁸ se zaměřením na vývoj přenosného systému schopného pojmout genetická data a s tím související požadavky od sekvenční analýzy po ukládání dat a jejich vizualizaci.¹²⁹ To znamená, že všechna data zpracovaná v rámci projektu Ensembl jsou přístupná ke stažení a jsou vzdáleně přístupná pomocí databázového serveru.¹³⁰

Ensembl spadá pod EMBL's European Bioinformatics Institute (EMBL-EBI)¹³¹, který za účelem podpory vědeckého výzkumu odkazuje na FAIR Data Principles¹³², principy vyhledatelných (*findable*), dostupných (*accessible*), interoperabilních (*interoperable*) a opětovně využitelných (*reusable*) dat. FAIR Data Principles spočívají ve snadném vyhledávání dat fyzickými uživateli i stroji. Tato data jsou podložena kvalitními metadaty a jsou jednoznačně identifikovatelná, ukládána na vhodném úložišti (např. cloudu) a otevřeně dostupná, přičemž jsou jasně dány podmínky a možnosti jejich využití. Data by měla být ve standardizovaném formátu a jejich popis by měl obsahovat unifikované výrazy. Tato data by měla být dostupná pod co možná nejméně omezující licenci, přičemž uživatelé jsou informováni o vzniku dat, jejich obsahu a o možnostech nakládání s nimi.

Ensembl zvolil při licencování přístupných dat cestu tzv. Creative Commons licence (CC).¹³³ EMBL-EBI uvádí, že cílem je rámcově převzít licenci tohoto typu na všechny využívané zdroje dat v rámci Ensembl za účelem standardizace užívaných licencí a co nejmenšího omezení užívání dat, přičemž samozřejmě v některých případech připouští nezbytný souběh více licencí.¹³⁴ S ohledem na výše uvedené jsou užívány konkrétně CC0 licence, což znamená, že oprávnění vlastníci různých práv copyright či práv z ochrany databáze se pro tento účel ochrany svých práv vzdají, a to dokonce včetně práv autorských (což

¹²⁸ JANSÁ, L., et al. Op. cit. 126, st. 289.

¹²⁹ HUBBARD, T., BARKER, D., BIRNEY, E., et al. The Ensembl genome database project. *Nucleic Acids Res.* 2002;30(1):38-41. [online]. [cit. 2022-04-06]. Dostupné z: doi:10.1093/nar/30.1.38.

¹³⁰ HERRERO, J., MUFFATO, M., BEAL, K., FITZGERALD, S. et al. Ensembl comparative genomics resources, Database, Volume 2016, 2016, baw053 [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/database/baw053>

¹³¹ EMBL's European Bioinformatics Institute. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk>

¹³² GoFair. FAIR Principles. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.go-fair.org/fair-principles/>

¹³³ Creative Commons. [online]. [cit. 2023-03-19]. Dostupné z: <https://creativecommons.org>

¹³⁴ EMBL's European Bioinformatics Institute. Licensing of EMBL-EBI data resources. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk/licencing>

však např. u osobnostních autorských práv dle české právní úpravy nelze)¹³⁵. Je tedy umožněno tato data bez dalšího volně užívat.

Ve vztahu k software Ensembl využívá licenci Apache 2.0 licence¹³⁶, schválenou ze strany Open Source Initiative (OSI), díky které může uživatel svobodně užívat software k různým účelům – distribuci, jeho upravování, včetně redistribuce upravené verze software atp.¹³⁷ Apache licence je bezúplatná permissivní licence umožňující šíření pozměňovaného software pod stejnou licencí.¹³⁸

4.3 Genetické databáze a genome browsery v právním slova smyslu

Česká právní úprava odrážející směrnici Evropského parlamentu a Rady (EU) o právní ochraně databází¹³⁹ obsažená v ustanovení § 88 zákona č. 121/2000 Sb., o právu autorském (AutZ)¹⁴⁰ rozumí databází *soubor nezávislých děl, údajů nebo jiných prvků, systematicky nebo metodicky uspořádaných a individuálně přístupných elektronickými nebo jinými prostředky, bez ohledu na formu jejich vyjádření*. Databází je tedy určitý uspořádaný soubor dat, samostatně přístupných příslušnými prostředky ve vztahu k formě jejich zpracování, ať už v elektronické či jiné podobě. Genetická data bývají přístupná jak jednotlivě, tak v uskupeních, čímž je splněna podmínka kladená na uspořádání a podmínka individuálního přístupu k nim. Moderní genetické databáze mají většinou elektronickou formu výslovně zmiňovanou AutZ, přípustné jsou však i formy jiné. Do českého, resp. evropského, právního pojetí databáze se tedy řadí i soubory obsahující hmotný genetický materiál, biologické vzorky, často doprovázené též daty v elektronické podobě, označované jako tzv. biobanky, speciálně regulované např. technickou normou Biotechnologie – Biobanky – Obecné požadavky na biobanky.¹⁴¹

¹³⁵ § 11 odst. 4 zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších předpisů (AutZ).

¹³⁶ The Apache Software Foundation. Licenses. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.apache.org/licenses/>

¹³⁷ Ensembl. Software Licence. [online]. [cit. 2023-03-19]. Dostupné z: https://www.ensembl.org/info/about/legal/code_licence.html

¹³⁸ JANSÁ, L., et. al. Op. cit. 126, st. 293.

¹³⁹ Směrnice Evropského parlamentu a rady 96/9/ES ze dne 11. 3. 1996 o právní ochraně databází.

¹⁴⁰ Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších předpisů (AutZ).

¹⁴¹ ČSN EN ISO 20387 (010107) Biotechnologie – Biobanky – Obecné požadavky na biobanky. Česká verze evropské normy EN ISO 20387:2020.

Databáze jsou chráněny dvojnásobným způsobem – jako souborné autorské dílo, anebo jako výsledek činnosti pořizovatele databáze. V prvním případě je databáze chápána jako vlastní autorský duševní výtvar, který musí splňovat podmínku jedinečnosti, srov. definice díla dle AutZ jakožto *jedinečného výsledku tvůrčí činnosti autora, vyjádřeného v jakékoli objektivně vnímatelné podobě včetně podoby elektronické, trvale nebo dočasně, bez ohledu na jeho rozsah, účel nebo význam*.¹⁴² Na takovou databázi by se aplikovala úprava ochrany autorského díla. Z druhého pohledu je databáze považována za předmět zvláštní ochrany pořizovatele databáze dle speciální úpravy práva pořizovatele databáze dle AutZ,¹⁴³ přičemž pro vznik ochrany v tomto případě odpadá kritérium jedinečnosti. Toto zvláštní právo pořizovatele databáze má trvání 15 let od pořízení databáze, přičemž při zpřístupnění databáze před uplynutím této doby zaniká toto právo za 15 let od prvního takového zpřístupnění. Je možné, aby jedna tatáž databáze požívala za určitých podmínek ochrany dle obou výše zmíněných přístupů.¹⁴⁴

Genome browsery naplní vzhledem ke způsobu jejich fungování definici databáze dle AutZ a souvisejících právních předpisů. Genome browsery většinou zpracovávají jak svá vlastní data, tak data z již existujících genetických databází a umožňují jejich další zpracování za užití interaktivních nástrojů. Na rozdíl od databází jsou data obsažená v genome browserech uchovávána a přístupná zpravidla pouze v elektronické podobě. Genome browsery také nabízejí více doplňkových funkcí a často lepší uživatelské rozhraní.

Aby určitý genome browser naplnil definici databáze je nutné, aby součástí tohoto genome browseru byla právě ona systematicky uspořádaná samostatná data, tedy data sekvenovaných genomů, ať už získána vlastní přímou cestou, či souhrnem a příslušným uspořádáním a zpracováním těchto dat získaných z jiných databází. Jinými slovy, je žádoucí, aby v jejich rámci docházelo k užívání a zpřístupnění „vlastních dat“ a získaných informací, a to nikoli pouze na bázi přímého využívání databází třetích osob. Zpracováním vlastních dat naplní tyto genome browsery i požadavek podstatného vkladu uvedený v ustanovení § 88a AutZ,¹⁴⁵ zdůrazněného např. v rozhodnutí Soudního dvora EU (SDEU) ve věci Fixtures

¹⁴² § 2 odst. 1 věta 1. AutZ.

¹⁴³ § 88 a násl. AutZ.

¹⁴⁴ TELEC I., TŮMA P. Autorský zákon, Komentář, 2. vydání, Praha: C.H.Beck 2019. ISBN 978-80-7400-748-4. St. 33.

¹⁴⁵ § 88a AutZ.

Marketing Ltd proti Oy Veikkaus Ab.¹⁴⁶ V tomto rozhodnutí SDEU uvádí, že *pojem „vklad spojený s pořízením, ověřením nebo předvedením obsahu databáze“ (...) je třeba chápat tak, že se jedná o vklad určený ke zřízení dané databáze a označuje tedy prostředky vynaložené na získání existujících prvků a jejich shromáždění v uvedené databázi, ale nezahrnuje prostředky použité k samotnému vytvoření prvků tvořících obsah databáze.* Rozhodnutí SDEU ve věci British Horseracing¹⁴⁷ konkretizuje pojem podstatný vklad jakožto *kvalitativně nebo kvantitativně podstatný vklad* ve vztahu k prostředkům vynaloženým na vytvoření databáze.

Jelikož zpracování, zejm. správa a zpřístupnění, genetických dat stran genome browserů obnáší mnoho dalších úprav a přidávání anotací, není pochyb, že jejich pořizovatelé vynakládají podstatný vklad v rámci kontroly obsahu genome browseru v zájmu zajištění jeho spolehlivosti, využitelnosti a funkčnosti a v rámci dalšího zpracování extrahovaných dat. Zvláštní práva pořizovatele databáze tedy ve vztahu k určitým genome browserům vzniknout mohou.

4.3.1 Bezúplatná zákonná licence

Zákon stanoví bezúplatnou zákonnou licenci pro ty uživatele databází, kteří je užívají, tj. databáze vytěžují či zužitkovávají,¹⁴⁸ pro svou osobní potřebu či pro účely vědecké nebo vyučovací, za podmínky uvedení pramenu a v rozsahu odůvodněném sledovaným nevýdělečným účelem.¹⁴⁹

V této souvislosti je vhodné poukázat na případnou problematiku jejich užívání farmaceutickými společnostmi či spíše s nimi spolupracujícími výzkumnými institucemi. Znamenalo by to, že takovéto společnosti mohou genetické databáze využívat sice pro svou osobní potřebu a pro vědecké účely při výzkumu léčivých přípravků, avšak následně by bylo třeba vyhodnotit oprávněnost jejich užívání při komercializaci takovýchto léčivých přípravků. Farmaceutické společnosti totiž téměř v každém případě sledují určitý výdělečný účel, přitom jsou též považovány za subjekty, které výzkum a vývoj léčivých přípravků pro léčbu genetických a dalších onemocnění stále efektivněji prohlubují. V případě vytěžování databází farmaceutickými společnostmi lze však mít spíše za to, že tyto společnosti jsou obecně

¹⁴⁶ Rozsudek SDEU ze dne 9. 11. 2004 ve věci Fixtures Marketing Ltd v Oy Veikkaus Ab., C-46/02.

¹⁴⁷ Rozsudek SDEU ze dne 9. 11. 2004 ve věci The British Horseracing Board Ltd a další, C-203/02.

¹⁴⁸ § 90 odst. 1 až 3 AutZ.

¹⁴⁹ § 92 písm. a) b) AutZ.

považovány za subjekty, které výdělečný účel sledují. Zákonná licence k vytěžování databází se v případě farmaceutických společností tedy s největší pravděpodobností neuplatní. Otázkou však zůstává oprávněnost vytěžování genetických databází odlišnými subjekty se zaměřením na vývoj léčivých přípravků, např. různými výzkumnými ústavy, a to v případě, kdy tyto výzkumné instituce své závěry a poznatky zkomercializují a získají na tomto základě majetkový prospěch.

Otázku oprávnění k vytěžování¹⁵⁰ dat z databází nově upravuje unijní směrnice 2019/790 (Směrnice 2019/790)¹⁵¹, a to ve vztahu k vysokým školám, výzkumným organizacím a institucím kulturního dědictví. Směrnice výslovně adresuje problematiku široké škály a různorodosti subjektů, které se cíleně věnují vědeckému výzkumu a poukazuje na vývoj moderních technologií zejména v oblasti výzkumu a inovací,¹⁵² zpracování velkého množství dat a možnost provádění *automatizované výpočetní analýzy informací v digitální podobě*.¹⁵³ Směrnice 2019/790 má za cíl *dosažení spravedlivé rovnováhy mezi právy a zájmy (...) nositelů práv na jedné straně a uživatelů na straně druhé*,¹⁵⁴ minimalizaci právní nejistoty ohledně vytěžování dat v souvislosti s vědeckým výzkumem, a zároveň podněcovat inovace i v soukromém sektoru.¹⁵⁵

Pojmem *vědecký výzkum* se pro účely této směrnice rozumí výzkum přírodních i humanitárních věd a za výzkumné organizace jsou kromě institucí vysokoškolského vzdělávání, považovány též výzkumné ústavy či nemocnice provádějící výzkum. Tyto výzkumné organizace zpravidla působí *buď na neziskové bázi, nebo v souvislosti s posláním ve veřejném zájmu uznávaným daným státem*, přičemž právní formy a jejich struktury se mohou napříč členskými státy lišit. V této souvislosti je přípustné, aby bylo toto uznání státu odražené např. formou financování určitého výzkumu z veřejných prostředků. Pod pojem *výzkumné organizace* by na druhou stranu neměly být řazeny organizace, které jsou

¹⁵⁰ „Vytěžováním textů a dat“ se pro účely Směrnice 2019/790 dle čl. 2 odst. 2 rozumí *jakákoli automatizovaná technika analýzy, jejímž cílem je analyzovat text a data v digitální podobě za účelem získání informací a která zahrnuje mimo jiné vzory, tendence a souvztáhnosti*.

¹⁵¹ Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu a o změně směrnic 96/9/ES a 2001/29/ES (Směrnice 2019/790).

¹⁵² Bod 5 odůvodnění Směrnice 2019/790.

¹⁵³ Bod 8 odůvodnění Směrnice 2019/790.

¹⁵⁴ Bod 6 odůvodnění Směrnice 2019/790.

¹⁵⁵ Bod 18 odůvodnění Směrnice 2019/790.

pod rozhodujícím vlivem obchodních podniků.¹⁵⁶ Směrnice 2019/790 dále zavádí výjimku ze zákazu dočasného nebo trvalého rozmnožování celé databáze nebo její části,¹⁵⁷ a z práva pořizovatele databáze zabránit vytěžování a/nebo zužitkování celého obsahu databáze nebo jeho kvalitativně nebo kvantitativně podstatné části,¹⁵⁸ stanovených směrnicí Evropského parlamentu a Rady 96/9/ES, a to *s ohledem na rozmnoženiny a extrakce zhotovené výzkumnými organizacemi (...) za účelem vytěžování textů a dat z děl nebo jiných předmětů ochrany, k nimž mají zákonný přístup, pro účely vědeckého výzkumu.*¹⁵⁹ Současně musí být splněny další podmínky ohledně jejich ukládání, zabezpečení a uchovávání.¹⁶⁰

Česká právní úprava databází, obsažená zejména v AutZ, bude muset na Směrnici 2019/790 reagovat, a to mimo jiné modifikací bezúplatné zákonné licence pro užití databáze pro účely vědeckého výzkumu¹⁶¹ a reflektovat do právní úpravy výjimky a omezení stanovené touto směrnicí. V důsledku takovéto úpravy by měla být situace ohledně užívání a vytěžování dat pro výzkumné účely přehlednější. Řešení otázky užívání genetických databází farmaceutickými společnostmi či spolupracujícími výzkumnými ústavy se zaměřením na výzkum v oblasti genetiky však může být stále poměrně nejisté, především v souvislosti s jejich zařazením mezi výzkumné organizace ve smyslu Směrnice 2019/790 a od toho se odvíjejícím oprávněním. Bude záležet zejména na teleologickém výkladu klíčových pojmů a Směrnice 2019/790 jako takové.

Dle účinné právní úpravy může být tedy volné užívání genetických databází bez dalšího, tj. například bez speciální licenční úpravy umožňující volný přístup k databázi, problematické.

4.4 Genome browsers – software

Software neboli *programové vybavení počítačů*, je složen z několika různých prvků (datových souborů, databází, manuálů).¹⁶² Software zahrnuje tedy jednoduše vše, co je

¹⁵⁶ Bod 12 odůvodnění Směrnice 2019/790.

¹⁵⁷ Čl. 5 písm. a) Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. 3. 1996 o právní ochraně databází (Směrnice 96/9/ES).

¹⁵⁸ Čl. 7 odst. 1 Směrnice 96/9/ES.

¹⁵⁹ Čl. 3 odst. 1 Směrnice 2019/790.

¹⁶⁰ Čl. 3 a čl. 4 Směrnice 2019/790.

¹⁶¹ § 92 písm. b) AutZ ve spojení s čl. 9 písm. b) Směrnice 96/9/ES.

¹⁶² TELEC I., et. al. Op. cit. 144, st. 29.

obsaženo v hardwaru, ale není hardwarem samotným.¹⁶³ Pojem *software* ani pojem *počítačový program* není AutZ ani judikaturou zatím přesně definován. Dle Směrnice 2009/24/ES je počítačovým programem program v jakékoliv formě, včetně těch, které jsou součástí technického vybavení (*hardware*). Tento výraz zahrnuje rovněž přípravné koncepční práce vedoucí k vytvoření počítačového programu za podmínky, že povaha těchto prací v pozdější etapě umožní vytvoření počítačového programu. Tato směrnice dále pro vyloučení pochybností stanoví, že předmětem ochrany je pouze vyjádření počítačového programu a že myšlenky a zásady, na kterých jsou založeny jednotlivé prvky programu včetně myšlenek a zásad, na kterých jsou založena rozhraní, nejsou chráněny autorským právem.¹⁶⁴ K užití software je možné udělit licenci a vykonávat k němu majetková práva, včetně jejich převodu. Zařazení komplexních genome browserů jakožto software však ve smyslu chápání tohoto pojmu z autorskoprávního pohledu není zcela vhodné. Při dispozici se software v souvislosti s licencováním, jde většinou o šíření rozmnoženiny počítačového programu jakožto součásti licence samotné.¹⁶⁵ V případě užívání genome browserů je však určitý software poskytován jako služba, jelikož hlavním smyslem je zpřístupnění výstupu z tohoto software, resp. získání konkrétních dat z určitého datového souboru, nikoli rozmnoženiny software jako takového. Proto bude na úpravu genome browserů spadat spíše úprava software jakožto služby, o čemž bude pojednáno v podkapitole 4.6.

V této souvislosti je vhodné ve zkratce pojednat o specifické ochraně počítačového programu. Na počítačový program se za určitých podmínek vztahuje ochrana dána AutZ¹⁶⁶, Směrnicí 2009/24/ES¹⁶⁷, Smlouvou Světové organizace duševního vlastnictví (WIPO) o právu autorském¹⁶⁸ a Dohodou o obchodních aspektech práv duševního vlastnictví (TRIPS)¹⁶⁹, a to ochrana počítačového programu jakožto díla literárního. Těmito podmínkami jsou

¹⁶³ JANSÁ, L., et. al. Op. cit. 126, st. 31.

¹⁶⁴ Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. 4. 2009 o právní ochraně počítačových programů (Směrnice 2009/24/ES).

¹⁶⁵ Rozsudek SDEU ze dne 3. 7. 2012. Ve věci UsedSoft GmbH v. Oracle International Corp, C-128/11.

¹⁶⁶ § 2 AutZ.

¹⁶⁷ Čl. 1 Směrnice 96/9/ES.

¹⁶⁸ Čl. 4 Smlouvy Světové organizace duševního vlastnictví (WIPO) o právu autorském z roku 1996.

¹⁶⁹ Čl. 10 Dohody o obchodních aspektech práv duševního vlastnictví (TRIPS).

původnost¹⁷⁰ a objektivní vnímatelnost.¹⁷¹ Při posuzování míry původnosti bude rozhodující zejm. charakter činnosti programátora při vytváření počítačového programu. V praxi jsou kritérii především *struktura a uspořádání programu (tj. organizace dat, posloupnost příkazů, instrukcí a volba algoritmů vyjádřená ve zdrojovém kódu), a způsob komunikace s uživatelem (tzv. look and feel jako způsob, jímž se program prezentuje uživateli a jak s ním komunikuje, resp. jak je jím užíván) vyjádřený zobrazením na obrazovce a kritérii práce uživatele s programem,*¹⁷² přičemž se bere v úvahu skutečnost, že programátor nemusí nutně vyvíjet přímo vlastní tvůrčí činnost, jelikož je při tvorbě běžné užívání různých softwarových nástrojů a vývoj funkcionalit software. Požadavek individuality zde odpadá, funkce vytvářených software mohou být stejné, jejich zdrojové kódy mohou být velmi podobné, stále však musí jít o program původní. Objektivní vnímatelnost je potom zkoumána z pohledu vyjádření zdrojového kódu, a to ať už elektronického či písemného, trvalého či dočasného.¹⁷³

Genome browsery často fungují jako software, jejichž činnost je úzce spjata s databázemi. V takovém případě bývá databáze buď součástí dílčí funkce software, výsledkem jeho činnosti nebo jinou jeho součástí. Pokud genome browser užívá ke svému fungování externě poskytované databáze jiných pořizovatelů databáze, náleží práva pořizovatele databáze právě těmto původním pořizovatelům. V takovém případě se potom na tyto genome browsery uplatní souběh ochrany jak software, tak databáze, včetně jejich specifik. Za určitých podmínek je možná ochrana genome browseru prostřednictvím zvláštního práva pořizovatele databáze i autorskoprávní ochrana díla a software, jakožto díla literárního.¹⁷⁴

4.5 Genome browsers – Cloud Computing

Cloud Computing je pojem pro poskytování IT služeb prostřednictvím virtuální infrastruktury poskytovatele (*cloudu*) za užití internetu a internetových prohlížečů.¹⁷⁵ Cloudem se dle jiné definice rozumí soubor zdrojů velkého rozsahu, které lze snadno užívat a virtualizovat, tyto cloudy se dají konfigurovat a jsou lehce přizpůsobitelné změnám v jejich

¹⁷⁰ § 2 odst. 2 věta první AutZ.

¹⁷¹ JANSÁ, L., et. al. Op. cit. 126, st. 35-36.

¹⁷² TELEČ I., et. al., st. 26.

¹⁷³ JANSÁ, L., et. al. Op. cit. 126, st. 35-36.

¹⁷⁴ *Ibid.*, st. 38.

¹⁷⁵ SVOBODA, J. Cloud computing. Systémová integrace. [online]. 2009, 16.2: 66-87. [cit. 2022-03-24]. Dostupné z: <http://old.cssi.cz/system/files/all/Osvoboda.pdf>

využívání a souvisejících služeb a tím zajistit co nejefektivnější využívání zdrojů.¹⁷⁶ Cloud Computing s sebou přináší různé výhody jako je např. snížení nákladů na provoz fyzických data center, vyšší bezpečnost dat díky profesionálním bezpečnostním a zálohovacím systémům, vysokou specializaci poskytovatele úložiště, úpravu odpovědnosti za škodu a sankce, rychlejší a snadnější zpracování množství dat atp.¹⁷⁷

Cloud Computing můžeme dělit do tří základních kategorií, a to *Infrastructure as a Service (IaaS)*, *Software as a Service (SaaS)* a *Platform as a Service (PaaS)*.¹⁷⁸

IaaS jsou takové služby, kdy poskytovatel cloudu zajišťuje určitou infrastrukturu jako např. úložiště, server či síťové zdroje, business intelligence a vysokovýkonné výpočetní prostředí a zpřístupňuje je zákazníkům za pomoci internetového připojení. Do této kategorie dále spadá též *Database as a Software (DBaaS)*¹⁷⁹ – software, který umožňuje uživatelům nastavovat, užívat a provozovat databáze pomocí jednoduchých nástrojů bez potřeby zařizování vlastního hardwaru či dokonce bez potřeby instalace speciálního software, a *Data Storage as a Service (DSaaS)*. Jako IaaS funguje např. Microsoft Azure či Google Compute Engine (GCE). SaaS, poskytování software jako služby, spočívá ve zpřístupnění infrastruktury a software zákazníkům, třetím osobám či objednateli dle jeho zadání, a to dokonce bez omezení určitou maximální kapacitou využívaného serveru. Mezi SaaS se řadí např. Google Workspace či Dropbox. Poslední kategorií jsou PaaS, jejichž podstatou je vývoj tvorby a vývoje vlastních aplikací objednatelů, a to prostřednictvím poskytování vývojových platform, vývojářských nástrojů, cloudových aplikací, operačních systémů a dalších služeb. Zástupci PaaS jsou např. Microsoft Azure, Google App Engine a Apache Stratos.¹⁸⁰

Cloudy mohou být soukromé (private) – poskytované a užívané jednou osobou, touto osobou také často vytvořené, tj. např. interní cloud či servery, provozovány v rámci jedné

¹⁷⁶ BARTOŠEK, V., PETRUCHA, R. Cloud Computing: moderní směr poskytování IT služeb, Systémová integrace. [online]. 2010. [cit. 2022-03-25]. Dostupné z: <http://3z894f6.257.cz/cloud-computing-moderni-smer-poskytovani-it-sluzeb> St. 79. a JANSKA L. OTEVŘEL P., ŠTEVKA M., Softwarové právo, 3. aktualizované rozšířené vydání. 2018. St. 338.

¹⁷⁷ JANSKA, L., et. al. Op. cit. 126, st. 342 a st. 338-339.

¹⁷⁸ Microsoft Azure. Co je IaaS? Infrastruktura jako služba. [online]. [cit. 2022-03-25]. Dostupné z: <https://azure.microsoft.com/cs-cz/overview/what-is-iaas/#overview>

¹⁷⁹ IBM. What is Database-as-a-Service (DBaaS)? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.ibm.com/cloud/learn/dbaas>

¹⁸⁰ IBM. SaaS vs PaaS vs IaaS: Whats the Difference & How to Choose. [online]. [cit. 2022-03-25]. Dostupné z: <https://blogs.bmc.com/saas-vs-paas-vs-iaas-whats-the-difference-and-how-to-choose/?print-posts=pdf>

konkrétní společnosti, nebo veřejné (public) cloudy – vyvinuté, vlastněné, zpřístupňované a provozované poskytovatelem, který tyto cloudy dále poskytuje třetím osobám, resp. zákazníkům. Pokud je cloud užíván a sdílen v rámci určité skupiny osob s obdobnými zájmy a potřebami, hovoříme o tzv. komunitním (community) cloudu. Hybridní cloudy, jak už název sám napovídá, jsou cloudy využívající více výše uvedených přístupů, kdy jsou např. různí poskytovatelé provozují odlišné části cloudové infrastruktury, které jsou však řízeny a integrovány jedním software.¹⁸¹

S ohledem na způsob fungování genome browserů lze konstatovat, že pro genome browsery je koncepce Cloud Computingové služby, zejm. veřejné, komunitní a hybridní, ve formě SaaS či DBaaS velmi vhodná. Vzhledem k principům fungování genome browserů je zřejmé, že tyto modely pokrývají téměř veškeré žádoucí aspekty jejich datových úložišť, provozu a zpřístupňování dat. Tyto dva modely budou níže více rozebrány.

4.5.1 Genome browsery – DBaaS

Genome browsery umožňují svým uživatelům snadný, jednoduchý a vzdálený přístup ke genetickým datům z databází. Jejich cílem je zaměřením jak na zpřístupnění těchto dat jako takových, tak na poskytování dalších dílčích služeb, jako je např. výše zmíněné vyhledávání určitých sekvencí či vizualizace genomů. Aby uživatelé mohli efektivně genome browsery užívat, je třeba zajistit obsluhu, vedení a práci s informacemi v rámci databáze a monitorovat prováděné operace. Genome browsery vykonávají tuto kontrolu nad databázemi za pomoci tzv. *application programming interface* (API)¹⁸², které umožňují jednoduše a bezpečně výměnu dat a komunikaci mezi jednotlivými prvky. Databáze je následně uživatelům zpřístupněna pomocí určitého webového nástroje a je jim umožněno s databází pracovat ve výše uvedeném smyslu. Téměř všechny administrativní úkony s tím spojené jsou obstarávané poskytovatelem, tedy správcem browseru, což pro uživatele znamená minimální zatížení a velký prostor pro snadnou a efektivní práci s předmětnými daty.¹⁸³ Charakter fungování a provozování genome browserů naplňuje uváděné prvky DBaaS, je tedy možné uzavřít, že na některé genome browsery by se mohl vztahovat tento režim.

¹⁸¹ JANSÁ, L., et. al. Op. cit. 126, st. 340-341.

¹⁸² IBM. What is an API (application programming interface)? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.ibm.com/cloud/learn/api>

¹⁸³ The App Solutions. CLOUD Service Models Explained: PaaS vs. SaaS vs. IaaS vs. DbaaS. [online]. [cit. 2023-03-18]. Dostupné z: <https://theappsolutions.com/blog/cloud/cloud-service-models/>

4.5.2 Genome browsersy – SaaS

Genome browsersy fungují nejčastěji jako SaaS, což může být považováno možná za nejvhodnější řešení. Jako příklad lze uvést právě Ensembl či UCSC Genome Browser.¹⁸⁴ Genome browsersy často obsahují soubor komplexních prvků, nástrojů a aplikací, jejichž pomocí je možné pracovat s určitými genetickými daty, přičemž tato data jsou odebírána především z databází třetích osob. Tyto genome browsersy jsou zaměřeny zejm. na zajištění stálého aktualizovaného a uceleného prostředí pro jednotlivé uživatele, kteří v jejich rámci mohou tvořit např. vlastní portfolia dat v rámci poskytovaného úložiště, a to s využitím genome browsersy nabízených nástrojů, včetně big data a data miningu, popř. s využitím umělé inteligence (AI), a to vše ve webovém rozhraní. Uživatelům je tedy umožněn přístup k datům odkudkoli s připojením internetu, bez potřeby instalace programu do zařízení. Na rozdíl od DBaaS charakteristickým prvkem takových genome browserů není jen samotná databáze a práce s ní, ale možnost uživatele pracovat v rámci takového genome browseru se samotnými daty, případnou vlastní činností atp., což nejlépe odpovídá modelu SaaS.

4.6 Free Software a Open Source

Aby mohly genetické databáze a genome browsersy co nejefektivněji sloužit svému účelu, tj. systematickému zpřístupňování genetických dat uživatelům, resp. vědcům a odborníkům, za účelem podpory co nejširšího vědeckého pokroku, je nutné uživatelům umožnit volný, resp. co nejméně restriktivní, přístup k datům, nástrojům a databázím. Genome browsersy v této souvislosti často odkazují na výše zmíněné FAIR Principles.¹⁸⁵ Jak již bylo výše uvedeno v souvislosti s browserem Ensembl, nejvhodnějším nabízejícím se řešením je pro genome browsersy jejich sdílení jakožto Free software a Open Source software (FOSS). Free software, spadající pod projekt GNU¹⁸⁶, dává jeho uživatelům svobodu užívat tento software za jakýmkoli účelem, možnost přizpůsobovat funkční atributy software svým potřebám, svobodu tento software dále šířit včetně jeho upravených kopií.¹⁸⁷ OSS cílí na další šíření ve volném režimu, otevřený zdrojový kód, jakýkoli účel jeho užívání, nemožnost jeho transformace do komerčního software atp. Při kombinaci obou přístupů je potom dokonce

¹⁸⁴TRAJANOSKI Z. Computational Medicine, Tools and Challenges, 2012. [online]. Springer Vienna. [cit. 2023-03-18]. St. 30. ISBN: 978-3-7091-0947-2. Dostupné z doi: <https://doi.org/10.1007/978-3-7091-0947-2>

¹⁸⁵ Viz podkapitola 4.2.

¹⁸⁶ GNU Operating System. What is GNU? [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org>

¹⁸⁷ JANSÁ, L., et. al. Op. cit. 126, st. 289.

možné se vyhnout nutnosti užití stejných licencí při jejich distribuci.¹⁸⁸ Se šířením free software a OSS jsou slučitelné určité typy licencí, které jsou Free Software Foundation,¹⁸⁹ dohlížející na free software, podporovány. Mezi tyto licence řadíme např. GNU General Public License¹⁹⁰, Mozilla Public License¹⁹¹, ISC License¹⁹², či permissivní MIT licence užívaná UCSC Genome Browser¹⁹³, nebo Apache licence.¹⁹⁴

Jak je z koncepce genome browserů a genetických databází, jakožto neplacených programů s minimálně omezeným přístupem, zřejmě, největším přínosem jejich existence a užívání je snadná dostupnost genetických dat, která jsou následně využívána především odborníky při vývoji nových léčivých přípravků a postupů, či vědci při jejich výzkumu. Podstatou jsou ona samostatná data, bez kterých by provoz těchto systémů neměl smysl.

¹⁸⁸ JANSÁ, L., et. al. Op. cit. 126, st. 290.

¹⁸⁹ *The Free Software Foundation (FSF), neboli nadace pro svobodný software, je nezisková organizace s celosvětovým posláním podporovat svobodu uživatelů počítačů.* (Free Software Foundation. The Free Software Foundation, [Online]. [cit. 2023-03-18]. Dostupné z: <https://www.fsf.org/>)

¹⁹⁰ GNU General public License, Version 3, 29 June 2007. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org/licenses/gpl-3.0.html>

¹⁹¹ Mozilla. Mozilla Public License, Version 2. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.mozilla.org/en-US/MPL/2.0/>

¹⁹² Internet Systems Consortium. ISC License. [online]. [cit. 2023-03-18] Dostupné z: <https://www.isc.org/licenses/>

¹⁹³ UCSC Genomics Institute. UCSC Licence. [online]. [cit. 2023-03-18]. Dostupné z: <https://genome.ucsc.edu/license/>

¹⁹⁴ GNU Operating System. Various Licenses and Comments about Them. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org/licenses/license-list.html#GPLCompatibleLicenses>

5. Genetické údaje

Genetickými údaji jsou dle Nařízení Evropského parlamentu a Rady Evropské Unie č. 2016/679 ze dne 27. 4. 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (GDPR), *osobní údaje*¹⁹⁵ *týkající se zděděných nebo získaných genetických znaků určité fyzické osoby, které vyplývají z analýzy biologického vzorku dotčené fyzické osoby, zejména chromozomů nebo kyseliny deoxyribonukleové (DNA) či ribonukleové (RNA), anebo z analýzy jiného prvku, která umožňuje získat rovnocenné informace.*¹⁹⁶ Genetické údaje jsou tedy osobní údaje vypovídající o genetických znacích určitého subjektu údajů, které jsou předmětem ochrany. Autorka má za to, že pojem *genetická data* může být v praxi chápán jako pojem širší než pojem genetické údaje. Genetická data mohou zahrnovat samotné genetické údaje, ve smyslu definice dle GDPR, tak např. veškeré další doplňující informace, specifikace, bližší informace týkající se metod zpracování daného vzorku DNA či obecné poznatky o genetických predispozicích, které nemusí být nutně osobními údaji. Pro účely této práce lze však považovat pojmy genetické údaje a genetická data za synonyma.

Genetické údaje jsou potom řazeny do zvláštní kategorie osobních údajů, kterým je přiznána vyšší míra ochrany. Není pochyb o tom, že na data získaná sekvenováním DNA by se měl vztahovat zvláštní režim nakládání s nimi, jejich ochrany a zpracování. Zde se však otevírá otázka možnosti anonymizace genetických údajů, a tedy jejich vyloučení z působnosti právní úpravy ochrany osobních údajů ve smyslu GDPR,¹⁹⁷ o čemž bude pojednáno zejm. v podkapitole 5.2.

Genetická data, která jsou předmětem zpracování genome browsery a obsahem genetických databází, jsou tedy ve své podstatě citlivé údaje. Může se tak jevit poněkud překvapivým, že jsou ve výše uvedené formě takto jednoduše a volně přístupná, přičemž je zde poměrně velké riziko jejich zneužití. Genetické údaje každého jednotlivce jsou unikátní a aby se dostaly do dispozice databáze či genome browseru, a následně k jejich koncovým uživatelům, je třeba zakotvení úpravy celého procesu. Tato úprava by se měla vztahovat zejména na získávání těchto dat a informovaný souhlas subjektů, bezpečné nakládání s daty ze strany správců a zpracovatelů, včetně ukládání, zabezpečování, dalšího sdílení a obecně

¹⁹⁵ Čl. 4 odst. 1 GDPR.

¹⁹⁶ Rec. 34 GDPR.

¹⁹⁷ Rec. 26 GDPR.

jejich užití, přičemž po celou dobu musí být dbáno na jejich ochranu, ochranu soukromí subjektů dat a dalšího s tím souvisejícího. Jelikož je toto téma z pohledu právní úpravy velmi rozsáhlé, je nutné se v této práci zaměřit pouze na některé dílčí otázky, tedy možnost anonymizace genetických údajů, aplikace právní úpravy a plnění povinností v souvislosti s jejich zpracováním.

Rec. 51 GDPR stanoví, že *svou povahou obzvláště citlivé osobní údaje z hlediska práv a svobod, zasluhují zvláštní ochranu, jelikož by při jejich zpracování mohla vzniknout závažná rizika pro základní práva a svobody*.¹⁹⁸ Genetické údaje o jednotlivcích mnoho vypovídají – za pomoci genetických údajů mohou být identifikováni, může být určena pravděpodobnost výskytu nemocí v jejich organismu a dokonce mohou mít vypovídající hodnotu týkající se biologicky příbuzných rodinných příslušníků.¹⁹⁹ Třetí osoby s přístupem k těmto datům, mohou mít *de facto* informace o etnickém původu, fyzických předpokladech, dědičnosti a pravděpodobnosti rozvoje určitých onemocnění a aktuálním zdravotním stavu subjektu údajů.²⁰⁰ Navíc s vývojem technologie, bioinformatiky a výzkumu jsou data stále senzitivnější.²⁰¹ Vědci uvádějí, že jsou údajně již nyní za pomoci genetických dat schopni určit též např. sexuální orientaci jedince²⁰², jeho politické názory, inteligenci²⁰³ či sklony ke kriminalitě²⁰⁴.

¹⁹⁸ Rec. 51 GDPR.

¹⁹⁹ PIKE, E. R. Securing sequences: ensuring adequate protections for genetic samples in the age of big data. *Cardozo law review*. [online]. Cardozo School of Law, 2016, 37(6), 1977 [cit. 2022-04-05]. ISSN 0270-5192. Dostupné z: <http://cardozolawreview.com/wp-content/uploads/2018/08/PIKE.37.6.pdf>

²⁰⁰ HELLMAN, D. What makes genetic discrimination exceptional? [online]. *American journal of law & medicine*. United States: Sage Publications, 2003, 29(1), 77-116 [cit. 2022-04-05]. Dostupné z: doi: 10.1017/S0098858800002343.

²⁰¹ DANKAR, F. K., PTITSYN, A., DANKAR, K. S.. The development of large-scale de-identified biomedical databases in the age of genomics-principles and challenges. [online]. *Human genomics*. England: BioMed Central, 2018, 12(1), 19-19 [cit. 2022-04-05]. ISSN 1479-7364. Dostupné z: doi:10.1186/s40246-018-0147-5.

²⁰² NGUN, T. C., VILAIN, E. The biological basis of human sexual orientation: is there a role for epigenetics?. *Advances in genetics*. [online]. United States, 2014, 86, 167-184 [cit. 2022-04-05]. ISSN 0065-2660. Dostupné z: doi:10.1016/B978-0-12-800222-3.00008-5.

²⁰³ DEARY, I. J., JOHNSON, W., HOULIHAN, L. M. Genetic foundations of human intelligence. *Human genetics*. [online]. Berlin/Heidelberg: Springer-Verlag, 2009, 126(1), 215-232 [cit. 2022-03-28]. ISSN 0340-6717. Dostupné z: doi:10.1007/s00439-009-0655-4.

²⁰⁴ MORLEY, K. I., HALL, W. D. Is There a Genetic Susceptibility to Engage in Criminal Acts?. *Trends and issues in crime and criminal justice*. [online]. Woden: Australian Institute of Criminology, 2003, (263), 1 [cit. 2022-04-05]. ISSN 0817-8542. ISBN 0 642 53816 6. Dostupné z: https://www.aic.gov.au/sites/default/files/2020-05/tandi263_0.pdf

Zneužití genetických dat je velmi snadné a má vysoký potenciál zasáhnout do základních práv a svobod jejich subjektu. Může docházet např. ke genetické diskriminaci²⁰⁵ jedinců ze strany pojišťoven ve vztahu k podmínkám poskytování pojištění, zaměstnavatelů při uzavírání a ukončování pracovních vztahů či stanovení zaměstnaneckých podmínek, i k obecné diskriminaci ze strany společnosti. Některé hrozby, jako např. riziko diskriminace fyzických osob na základě informací o jejich zdravotním stavu, ve spojení se zneužitím genetických dat, jsou zmíněny přímo v GDPR a ve stanovisku Pracovní skupiny 29²⁰⁶ o technikách anonymizace dat.²⁰⁷ Zákaz diskriminace osoby z důvodu jejího genetického dědictví je též výslovně uveden v čl. 11 Úmluvy o biomedicíně.²⁰⁸ Důsledkem může být mimo jiné ztráta či porušení zákonné ochrany osobních údajů a soukromí jednotlivce při zpětném dohledání subjektu údajně anonymizovaných údajů.²⁰⁹

Problematice úpravy genetických dat se věnovala Pracovní skupina 29, která v roce 2004 přijala dokument nazvaný jako „Working Document on Genetic Data“²¹⁰, ve kterém byly tyto údaje mimo jiné blíže specifikovány, stejně jako důvody potřeby vyšší ochrany těchto údajů. Pro srovnání, tento dokument, i když v některých svých částech už neaktuální, se věnuje problematice genetických údajů v podstatně větším rozsahu než GDPR. V tomto dokumentu bylo poznamenáno, že z genetických údajů lze vyčíst informace jak o subjektu údajů, tak o jeho pokrevních příbuzných a dalších skupinách, kterých je součástí; genetické údaje nejsou ve většině případů subjektu údajů samotnému známy, jejich podoba nezávisí na jeho vůli a není

²⁰⁵ NIH. National Library of Medicine. MedlinePlus. What is genetic discrimination?: Medlineplus Genetics. MedlinePlus.[online]. [cit. 2022-03-26]. Dostupné z: <https://medlineplus.gov/genetics/understanding/testing/discrimination/>

²⁰⁶ Pracovní skupina 29 byla ustanovena na základě čl. 29 směrnice Evropského parlamentu a Rady 95/46/ES ze dne 24. října 1995 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů, jakožto nezávislý poradní orgán na ochranu dat a soukromí složený z vedoucích zástupců dozorových úřadů členských států Evropské unie. Od účinnosti GDPR se jedná o Evropský sbor pro ochranu osobních údajů (EPDB).

²⁰⁷ Rec. 71 a 75 GDPR.

²⁰⁸ Úmluva na ochranu lidských práv a důstojnosti lidské bytosti v souvislosti s aplikací biologie a medicíny: Úmluva o lidských právech a biomedicíně (1997), Rada Evropy (Úmluva o biomedicíně).

²⁰⁹ GEROSKI, A. S., ed. by: LIEDTKE, B., SAMPATH, N., SISTLA, M. Abuse of Our Genetic Data Is the Next Privacy Scandal. [online]. Spring 2019 Journal. [cit. 2022-04-11]. Dostupné z: <https://bjppj.berkeley.edu/2019/02/27/spring-2019-abuse-of-our-genetic-data-is-the-next-privacy-scandal/>

²¹⁰ Pracovní skupina 29. Working Document on Genetic Data ze dne 17. 5. 2004, 12178/03/EN WP 91. [online] [cit. 2022-03-28]. Dostupné z: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2004/wp91_en.pdf (Working Document).

možné je modifikovat (což je však zastaralá informace)²¹¹ či že rozsah informací možný vyčíst z genetických údajů se s novými technologiemi stále zvětšuje a bude možné tyto údaje využít pro různé účely.²¹²

GDPR v současném znění některé problematiky uvedené v tomto dokumentu sice reflektovalo, avšak nikoli ve vyčerpávajícím rozsahu.²¹³ Některé aspekty zpracování a správy genetických údajů se tedy řídí dle ne zcela vhodné právní úpravy, což má za následek např. nejednotné pojetí závazných povinností pro genome browsery týkající se ochrany zpracovávaných dat a s tím též možné mezery v jejich zabezpečování.

5.1 Právní základ zpracování genetických údajů – GDPR

Genetická data, jak bylo uvedeno výše, jsou citlivými osobními údaji a požívají ochranu zvláštní kategorie údajů. Dle GDPR platí obecný zákaz jejich zpracování,²¹⁴ který může být prolomen při splnění určitých podmínek.²¹⁵ Zpracování je definováno velmi široce a obsahuje téměř každou situaci, kdy se osoba dostane s daty do kontaktu, včetně jejich *shromáždění, zaznamenání, uspořádání, strukturování, uložení, přizpůsobení nebo pozměnění, vyhledání, nahlédnutí, použití, zpřístupnění přenosem, šíření nebo jakékoliv jiné zpřístupnění, seřazení či zkombinování, omezení, výmazu nebo zničení*.²¹⁶

Zpracování genetických údajů je tedy možné při udělení výslovného souhlasu *se zpracováním těchto osobních údajů pro jeden nebo více stanovených účelů subjektem údajů*,²¹⁷ pokud je jejich zpracování *nutné pro ochranu životně důležitých zájmů subjektu údajů nebo jiné fyzické osoby v případě, že subjekt údajů není fyzicky nebo právně způsobilý udělit souhlas*,²¹⁸ je *nezbytné pro účely preventivního nebo pracovního lékařství (...), lékařské diagnostiky, poskytování zdravotní nebo sociální péče či léčby nebo řízení systémů a služeb*

²¹¹ Příkladem může být např. technologie tzv. genetických nůžek neboli CRISPR-Cas9 (CRISPR). Za pomoci CRISPR je možné modifikovat, tedy též efektivně, poměrně levně a rychle, opravit genom jakéhokoli organismu. (Doudna Lab. Op. cit. 77)

²¹² Pracovní skupina 29. Op. cit. 210.

²¹³ Srov. např. rec. 34 GDPR.

²¹⁴ Čl. 9 odst. 1 GDPR.

²¹⁵ Čl. 9 odst. 2 GDPR.

²¹⁶ Čl. 4 odst. 2 GDPR.

²¹⁷ Čl. 9 odst. 2 písm. a) GDPR.

²¹⁸ Čl. 9 odst. 2 písm. c) GDPR.

zdravotní nebo sociální péče na (...) při splnění podmínek a záruk (...) ²¹⁹, je nezbytné z důvodů veřejného zájmu v oblasti veřejného zdraví, jako je ochrana před vážnými přeshraničními zdravotními hrozbami nebo zajištění přísných norem kvality a bezpečnosti zdravotní péče a léčivých přípravků nebo zdravotnických prostředků, ²²⁰ či je nezbytné pro účely archivace ve veřejném zájmu, pro účely vědeckého či historického výzkumu nebo pro statistické účely (...) které je přiměřené sledovanému cíli, dodržuje podstatu práva na ochranu údajů a poskytuje vhodné a konkrétní záruky ²²¹ pro ochranu základních práv a zájmů subjektu údajů ²²². Zpracování osobních údajů pro účely vědeckého výzkumu by přitom mělo být chápáno v širokém smyslu. ²²³ Ve vztahu ke genetickým databázím a genome browserům je možné uvažovat zejména o udělení souhlasu a o nezbytnosti zpracování těchto dat v souvislosti s vědeckým výzkumem.

Udělení takového souhlasu může být samo o sobě problematické, vzhledem k nutnosti stanovení určitého účelu jejich zpracování. Tato data jsou využívána s různými záměry a s vědeckým pokrokem je možné z genetických dat vyčíst stále více informací a spojovat je s výzkumy různého zaměření. Navíc pro laika je velmi obtížné si některé způsoby využití dat v tomto ohledu představit. Je sice možné stanovit účel zpracování velmi zešíroka, je však otázkou, zdali by byl takový souhlas dostatečným titulem pro zpracování těchto údajů, vzhledem k požadavku jeho konkrétnosti a informovanosti. Problematické souhlasu se zpracováním osobních údajů je věnována podkapitola 5.5.

Zpracování genetických dat by mohlo být tedy v tomto ohledu postaveno především na jejich přínosu pro vědecký výzkum, tedy na výjimkách z obecného zákazu zpracování citlivých údajů stanoveném v čl. 9 odst. 1 GDPR. Tyto výjimky jsou zakotveny v čl. 9 odst. 2 písm. j) ve spojení s čl. 89 odst. 1 GDPR, budou se však na zpracování vztahovat pouze za předpokladu, že bude zároveň naplněna alespoň jedna z podmínek zákonnosti zpracování vymezených v čl. 6 GDPR. Příkladem může být zpracování založené na oprávněném zájmu ²²⁴ (např. pro účely cílení reklamy na zákazníky dodavateli služeb), avšak pokud by správce

²¹⁹ Čl. 9. odst. 3 GDPR a Čl. 9 odst. 2 písm. h) GDPR.

²²⁰ Čl. 9 odst. 2 písm. i) GDPR.

²²¹ Srov. rec. 156 GDPR.

²²² Čl. 9 odst. 2 písm. j) GDPR.

²²³ Rec. 159 GDPR.

²²⁴ Čl. 6 odst. 1 písm. f) GDPR.

využíval k cílení i údaje o zdravotním stavu subjektů údajů, tak k tomu potřebuje výslovný souhlas²²⁵. Z výše uvedeného je zřejmé, že jednotlivé výjimky dle čl. 9 odst. 2 GDPR většinou naplní konkrétní odpovídající požadavky dle čl. 6 odst. 1 GDPR, nikoli však vždy.

Zpracování zvláštní kategorie osobních údajů ve veřejném zájmu či pro účely vědeckého výzkumu musí mít též základ v právu Evropské unie nebo členského státu, které je *přiměřené sledovanému cíli, dodržuje podstatu práva na ochranu údajů a poskytuje vhodné a konkrétní záruky pro ochranu základních práv a zájmů subjektu údajů*.²²⁶ Zde se otevírá otázka, zdali je v současné době českým právem či právem Evropské unie pro zpracování genetických údajů za výzkumnými účely tento zákonný základ dán. Zpracováním osobních údajů pro vědecké účely se částečně zabývá český zákon č. 110/2019 Sb., o zpracování osobních údajů (ZZOÚ), a to v ustanovení § 16. Toto ustanovení konkretizuje výše zmíněná náležitá opatření pro zajištění ochrany subjektů údajů a stanoví odchylky od obecné úpravy pro zpracování osobních údajů. Takovými opatřeními mohou být např. pseudonymizace, minimalizace, šifrování osobních údajů, jmenování pověřence, opatření k zajištění trvalé důvěrnosti, integrity, dostupnosti a odolnosti systémů a služeb zpracování apod.²²⁷

Je diskutabilní, zdali lze toto ustanovení samo o sobě považovat za zákonný základ zpracování zvláštní kategorie osobních údajů ve smyslu čl. 9 odst. 2 písm. j) GDPR. Lze si představit výklad, že ustanovení § 16 ZZOÚ jako takové naplňuje předpoklady dané výjimky a je ho tudíž možné považovat za legitimní právní základ pro zpracování citlivých údajů pro účely vědeckého výzkumu. Autorka se však spíše přiklání k názoru, že genetické údaje pro tyto účely pouze na základě předmětného ustanovení ZZOÚ zpracovávat nelze, a to především s ohledem na vysokou míru obecnosti tohoto ustanovení a nedostatečné zohlednění specifik charakteru genetických údajů a výzkumu s nimi spojeného. V této souvislosti autorka poukazuje kromě znění čl. 9 odst. 2 písm. j) GDPR též na rec. 52 a rec. 53 GDPR. Autorka má za to, že v současné době pro tento účel zpracování genetických údajů není českým právem ani právem Evropské unie zákonný základ bez dalšího dán.

Pro úplnost je třeba uvést, že prostřednictvím ustanovení § 16 ZZOÚ využil český zákonodárce možnosti stanovené v čl. 89 odst. 2 GDPR a vymezil pro účely zpracování

²²⁵ Čl. 9 odst. 2 písm. a) GDPR.

²²⁶ Čl. 9 odst. 2 písm. j) GDPR.

²²⁷ Čl. 16 odst. 1 zák. č. 110/2019 Sb., o zpracování osobních údajů (ZZOÚ) ve spojení s čl. 89 odst. 1 GDPR.

osobních údajů pro vědecké účely určité odchylky od obecné úpravy.²²⁸ Pokud by bylo zpracování genetických údajů založeno na předmětné výjimce, znamenalo by to tedy mimo jiné poměrně mnoho možností odklonu od obecné regulace zpracování osobních údajů též dle GDPR. Jedná se např. o možnost jejich dalšího zpracování, tj. bez plného „účelového omezení“ či možnost delší doby jejich ukládání,²²⁹ povinnost poskytování dalších informací subjektům údajů²³⁰ či neuplatnění úpravy práva na výmaz, pokud je pravděpodobné, že by jejich výmaz znemožnil nebo vážně ohrozil splnění cílů zpracování.²³¹ Zvýhodněný režim dle GDPR by se měl však uplatnit pouze v případě, že budou dodržovány specifické odvětvové i etické standardy a výzkum bude cílen na rozvoj kolektivních znalostí a společenského blaha.²³²

5.2 Anonymizace a pseudonymizace údajů

S ohledem na výše uvedené je zřejmé, že zpracování genetických údajů podléhá regulaci stanovené právními předpisy, zejm. GDPR a ZZOÚ. Oprávněné zpracování genetických údajů je tedy možné pouze na zákonném základě a při jejich zpracování musí být zaručena ochrana zájmů subjektu údajů. Za tímto účelem právní úprava stanoví správcům a zpracovatelům povinnost poskytovat vhodné záruky a přijímat opatření s cílem zajistit dodržování zásad zpracování osobních údajů.²³³

Pseudonymizace²³⁴ údajů může být opatřením ve výše uvedeném smyslu, jestliže její užití splní sledovaný účel, přičemž pokud lze údaje anonymizovat, musí být přijato takovéto účinnější opatření.²³⁵

²²⁸ VLACHOVÁ, B., MAISNER, M. Zákon o zpracování osobních údajů. Komentář. 1. vydání. Praha: C. H. Beck, 2019. St. 34-37. ISBN 978-80-7400-760-6.

²²⁹ Čl. 5 odst. 1 písm. b) a e) GDPR.

²³⁰ Čl. 14 odst. 5 písm. b) a rec. 62 GDPR.

²³¹ Čl. 17 odst. 3 písm. d) a rec. 65 GDPR.

²³² European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. [online]. 2020, s. 12, [cit. 2022-03-28]. Dostupné z: https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

²³³ Čl. 5 odst. 1 GDPR.

²³⁴ Dle čl. 4 odst. 5 GDPR se „pseudonymizací“ rozumí *zpracování osobních údajů tak, že již nemohou být přiřazeny konkrétnímu subjektu údajů bez použití dodatečných informací, pokud jsou tyto dodatečné informace uchovávány odděleně a vztahují se na ně technická a organizační opatření, aby bylo zajištěno, že nebudou přiřazeny identifikované či identifikovatelné fyzické osobě.*

²³⁵ Čl. 89 odst. 1 GDPR.

Přímou definici pojmu *anonymizace* GDPR neuvádí, vysvětluje však pojem *anonymní informace*, a to jakožto *informace, které se netýkají identifikované či identifikovatelné fyzické osoby, ani na osobní údaje anonymizované tak, že subjekt údajů není nebo již přestal být identifikovatelným*,²³⁶ přičemž při určování identifikovatelnosti fyzických osob, *by se mělo přihlídnout ke všem prostředkům, jako je například výběr vyčleněním, o nichž lze rozumně předpokládat, že je správce nebo jiná osoba použije pro přímou či nepřímou identifikaci dané fyzické osoby. Ke stanovení toho, zda lze rozumně předpokládat použití prostředků k identifikaci fyzické osoby, by měly být vzaty v úvahu všechny objektivní faktory, jako jsou náklady a čas, které si identifikace vyžádá, s přihlédnutím k technologii dostupné v době zpracování i k technologickému rozvoji*.²³⁷ Význam tohoto pojmu lze též nepřímo dovodit např. ze čl. 89 odst. 1 věty poslední GDPR, a to jako proces či zpracování, které *neumožňuje nebo které přestane umožňovat identifikaci subjektů údajů*.

Anonymizované údaje jsou potom takové údaje, které ani nepřímo nepomáhají v identifikaci určitého člověka a nejsou s ním tedy nijak spojitelná.²³⁸ Některé zdroje uvádí, že nelze tolerovat, aby po anonymizaci údajů zůstalo jakkoli velké riziko re-identifikace subjektů údajů, v opačném případě nelze označit takové údaje za anonymní.²³⁹ GDPR přitom stanoví, že ochrana a úprava osobních údajů dána tímto předpisem se na anonymní informace nevztahuje.

Data přístupná v genetických databázích a genome browserech jsou sice údajně anonymizována a jsou přijímána opatření k ochraně jejich subjektů, avšak lidský genom, který je v případě každého jedince unikátní, jak bylo zmíněno výše, je možná ten nejosobnější a nejunikátnější identifikátor již ze své podstaty.

Vzhledem k jedinečnosti těchto údajů při jejich zpracování, či pouze jejich částí, tedy např. pouze z určitého úseku lidského genomu obsahující jakoukoli vzácnou genetickou

²³⁶ Rec. 26 GDPR.

²³⁷ Rec. 26 GDPR a k identifikovatelnosti srov. Rozsudek SDEU ze dne 19. 10. 2016 ve věci Patrick Breyer vs. Bundesrepublik Deutschland, C-582/14.

²³⁸ European Commission. Co jsou to osobní údaje? [online]. [cit. 2022-03-28]. Dostupné z: https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_cs

²³⁹ FINCK, M., PALLAS, F.. They who must not be identified—distinguishing personal from non-personal data under the GDPR. International data privacy law [online]. Oxford: Oxford University Press, 2020, 10(1), 11-36 [cit. 2022-04-05]. ISSN 2044-3994. Dostupné z: doi:10.1093/idpl/ipz026.

informaci, existuje vždy riziko, že subjekt údajů bude reidentifikován.²⁴⁰ V roce 2019 bylo dokázáno, že dokonce jen za využití informací o jednonukleotidových polymorfismech (SNPs)²⁴¹ jedinců je možné konkrétního jedince identifikovat, a to přes veškerou snahu genome browserů a genetických databází a jimi přijatých opatření k ochraně dat a soukromí subjektů dat,²⁴² což bylo dále potvrzeno navazujícími experimenty.²⁴³ Dle stanoviska Pracovní skupiny 29 z roku 2014²⁴⁴ týkajícího se technik anonymizace dat je nutné vzít v úvahu, že i anonymizovaný soubor dat může představovat zbytkové riziko, ať už samostatně či nepřímo ve spojení s dalšími informacemi z jiných zdrojů. K tomuto lze poukázat na studii, ze které bylo zřejmé, že lze identifikovat konkrétní osobu pouze na základě veřejně dostupných genetických údajů a dalších zdrojů.²⁴⁵ Tato studie prokázala, že anonymizace výzkumných genetických dat by neměla být nadále považována za realistickou.²⁴⁶ Ve stanovisku je dále uvedeno, že anonymizace by neměla být pokládána za samostatně dostatečné opatření ochrany dat a potenciální hrozby jejich zneužití by měly být správci dat pravidelně analyzovány.²⁴⁷

Jak bylo uvedeno výše, v případě některých agregovaných genetických databází jsou data zpracována takovým způsobem, že je téměř zcela nemožné subjekty údajů reidentifikovat.

²⁴⁰ MALIN, B., SWEENEY, L. How (not) to protect genomic data privacy in a distributed network: using trail re-identification to evaluate and design anonymity protection systems. *Journal of biomedical informatics* [online]. US: Elsevier, 2004, 37(3), 179-192 [cit. 2022-04-06]. ISSN 1532-0464. Dostupné z: doi:10.1016/j.jbi.2004.04.005.

²⁴¹ NIH. National Library of Medicine. MedlinePlus. What are single nucleotide polymorphisms (snps)?: Medlineplus Genetics. [online]. [cit. 2022-03-29]. Dostupné z: <https://medlineplus.gov/genetics/understanding/genomicresearch/snp/>

²⁴² VON THENEN, N., AYDAY, E., CICEK, A. E.. Re-identification of individuals in genomic data-sharing beacons via allele inference. *Bioinformatics* [online]. England: Oxford University Press, 2019, 35(3), 365-371 [cit. 2022-04-06]. ISSN 1367-4803. Dostupné z: doi:10.1093/bioinformatics/bty643.

²⁴³ SHABANI, M., MARELLI, L. Re-identifiability of genomic data and the GD9,PR [online]. 24. 5. 2019, 1-5 [cit. 2022-03-25]. Dostupné z: doi:10.15252/embr.201948316.

HUMBERT, M., HUGUENIN, K., HUGONOT, J., AYDAY, E., HUBAUX, J. De-anonymizing Genomic Databases Using Phenotypic Traits. *Proceedings on Privacy Enhancing Technologies 2015*. [online]. 2015 (2):1–16. [cit. 2023-03-18]. Dostupné z: doi: 10.1515/popets-2015-0020.

²⁴⁴ Article 29 Data Protection Working Party. (2014) Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN WP216. [online]. [cit. 2023-03-19]. Dostupné z: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

²⁴⁵ GYMREK, M., MCGUIRE, A. L., GOLAN, D., HALPERIN, E., ERLICH, Y. Identifying Personal Genomes by Surname Inference. *Science*. [online]. United States: American Association for the Advancement of Science, 2013, 339(6117), 321-324 [cit. 2022-04-06]. ISSN 0036-8075. Dostupné z: doi:10.1126/science.1229566.

²⁴⁶ QUINN, P., QUINN, L. Big genetic data and its big data protection challenges. *The computer law and security report* [online]. Elsevier, 2018, 34(5), 1000-1018 [cit. 2022-04-06]. ISSN 0267-3649. Dostupné z: doi:10.1016/j.clsr.2018.05.028.

²⁴⁷ Article 29 Data Protection Working Party. Op. cit. 244.

Data v agregované podobě je možné v některých případech za anonymizované považovat, stejně jako určité úseky DNA, které jsou pro všechny jednotlivce společné. Ani v těchto případech však nelze zcela vyloučit riziko zpětné identifikace subjektů údajů, a to s ohledem na možné nepředvídatelné či vzácné nuance jejich genetických informací, které mohou být i z dat v agregované databázi zřejmé.

Z výše uvedeného lze usoudit, že pouhý proces anonymizace genetických dat nemůže být v tomto kontextu považován za efektivní. S ohledem na tyto skutečnosti by měly být vždy genetické údaje v takové formě, v jaké jsou na těchto platformách dostupné, považovány nanejvýše za údaje pseudonymizované.²⁴⁸ To by však znamenalo též nutnost aplikace příslušné právní úpravy stanovené zejm. GDPR a zaručení jejich příslušné ochrany.

5.3 Posouzení vlivu na ochranu osobních údajů (DPIA)

Jednou ze zvláštních povinností stanovenou GDPR v souvislosti se zpracováním osobních údajů je tzv. Data Protection Impact Assessment (DPIA), tedy posouzení vlivu na ochranu osobních údajů. Správce osobních údajů musí před jejich zpracováním provést posouzení vlivu zamýšlených operací zpracování na jejich ochranu, a to *pokud je pravděpodobné, že určitý druh zpracování, zejména při využití nových technologií, bude s přihlédnutím k povaze, rozsahu, kontextu a účelům zpracování, mít za následek vysoké riziko²⁴⁹ pro práva a svobody fyzických osob.*²⁵⁰ Toto je nutné zejména v případě, že se jedná o rozsáhlé zpracování citlivých údajů.²⁵¹ DPIA obsahuje systematický popis zamýšlených operací zpracování a jeho účely, popř. včetně uvedení oprávněných zájmů správce, dále hodnocení nutnosti a proporcionality zpracování v souvislosti s jeho účelem, posouzení rizik pro subjekty příslušných údajů a plánovaná opatření, bezpečnostní mechanismy v souvislosti s těmito riziky a zajištěním ochrany osobních údajů, včetně doložení souladu s GDPR, s ohledem na práva a oprávněné zájmy subjektů údajů a dalších dotčených osob. DPIA má být přezkoumáváno vždy minimálně v případech, že se změní rizika zpracování dotčených

²⁴⁸ ENISA. European Union Agency for Cyber Security. Deploying pseudonymisation techniques. The case of the Health Sector. 2022. [online] [cit. 2023-03-19]. Dostupné z: <https://www.enisa.europa.eu/publications/deploying-pseudonymisation-techniques>

²⁴⁹ Pokyny pro posouzení vlivu na ochranu údajů a stanovení, zda „je pravděpodobné, že zpracování údajů bude mít za následek vysoké riziko“ pro účely nařízení 2016/679, přijaté dne 4. 4. 2017 a dne 4. 10. 2017 v aktualizovaném znění, 17/CS, WP 248 rev.01 (Pokyny pro posouzení vlivu na ochranu údajů). [online]. [cit. 2023-03-26]. Dostupné z: https://www.uouu.cz/assets/File.ashx?id_org=200144&id_dokumenty=31892

²⁵⁰ Čl. 35 odst. 1 ve spojení s rec. 90 a 91 GDPR.

²⁵¹ Čl. 35 odst. 3 písm. b).

osobních údajů.²⁵² Dle výsledků DPIA je následně správce údajů povinen podniknout případné další kroky.²⁵³ Při posuzování nutnosti provádění DPIA má správce postupovat též dle kritérií uvedených v Pokynech Pracovní skupiny 29 pro posouzení vlivu na ochranu údajů.²⁵⁴

Cílem DPIA je určení rizikovitosti zpracování osobních údajů pro jejich subjekty při přípravě či změně jejich zpracování,²⁵⁵ je jedním z hlavních nástrojů zajišťování souladu s GDPR při potenciálním vysoce rizikovým zpracováním údajů.²⁵⁶ Vzhledem k výše uvedenému je ve vztahu ke genetickým údajům a jejich zpracování genetickými databázemi a genome browsersy bez pochyby zřejmé, že při uplatnění GDPR a příslušné právní úpravy, bude v souladu s touto úpravou ve většině případů nutné provádět DPIA. Při zpracování osobních údajů genetickými databázemi a genome browsersy bude tedy nejčastěji docházet k naplnění hned několika kritérií stanovených výše uvedenými pokyny Pracovní skupiny 29. Zejména zpracování citlivých údajů a osobních údajů o zranitelných osobách, zpracování údajů ve velkém rozsahu vzhledem k počtu dotčených subjektů údajů, době zpracování a územnímu rozsahu, popř. zpracování propojených nebo kombinovaných souborů osobních údajů, které pocházejí z více různých zpracování, zpracování pro účely nad rámec původního účelu či zpracování údajů při použití nových technologických řešení a organizačních opatření. Ke vzniku povinnosti provádění DPIA je dle těchto pokynů nutné splnění minimálně dvou uvedených podmínek.²⁵⁷ Toto vyplývá mimo jiné přímo z demonstrativního výčtu příkladů uvedených v Pokynech pro posouzení vlivu na ochranu údajů, kde je výslovně uvedeno, že s vysokou pravděpodobností bude za určitých podmínek třeba provedení DPIA dokonce v případě ukládání *pseudonymizovaných osobních citlivých údajů týkajících se zranitelných subjektů údajů získaných v rámci výzkumných projektů nebo klinických hodnocení za účelem archivování*.²⁵⁸

²⁵² Čl. 35 odst. 7 a odst. 11 GDPR.

²⁵³ Čl. 36 GDPR.

²⁵⁴ Pokyny pro posouzení vlivu na ochranu údajů. Op. cit. 249.

²⁵⁵ Úřad pro ochranu osobních údajů. Návod k posouzení vlivu na ochranu osobních údajů u návrhů právních předpisů (DPIA). [online] [cit. 2022-03-19]. Dostupné z: <https://www.uoou.cz/navod-k-posouzeni-vlivu-na-ochranu-osobnich-udaju-u-navrhu-pravnich-predpisu-dpia/ds-5344>

²⁵⁶ Pokyny pro posouzení vlivu na ochranu údajů. Op. cit. 249.

²⁵⁷ *Ibid*

²⁵⁸ *Ibid*.

Správce osobních údajů je povinen zavést vhodná technická a organizační opatření, s ohledem na povahu, rozsah, kontext, účely a možná rizika, k zajištění zpracování těchto údajů v souladu s GDPR, což musí být též schopen doložit, přičemž je povinen tato opatření v případě potřeby revidovat a aktualizovat.²⁵⁹ V případě neprovedení DPIA,²⁶⁰ jeho nesprávného provedení²⁶¹ či jeho nepředložení dozorovému orgánu ke konzultaci²⁶², je možné správci uložit správní pokutu do výše 10 milionů EUR, nebo jedná-li se o podnik, do výše 2 % z celosvětového ročního obrátu.²⁶³

S ohledem na skutečnost současné nemožnosti anonymizace genetických údajů, a tedy nutnost aplikace právní úpravy ochrany osobních údajů, vyplývá provozovatelům genetických databází a genome browserů, jakožto správcům osobních údajů v nich obsažených, povinnost provádění DPIA, a to s ohledem na naplnění ustanovení v čl. 35 GDPR.

5.4 DPIA – genome browsery a genetické databáze

Provádění DPIA a zejména vyhodnocování možných rizik při zpracování genetických údajů může samo o sobě být velmi problematické. Byť jen rámcový přehled možných rizik a jejich následků při této činnosti a porušení konkrétních práv a svobod jak subjektů údajů, tak třetích osob, by byl velmi široký a téměř jistě neúplný. Rizika pro subjekty údajů musí být zvažována nejen z pohledu citlivosti genetických dat jako takových, jejich typu a způsobů jejich užití, ale též z hlediska informací, které lze z těchto údajů získat.²⁶⁴ Je možné si představit mnoho nepříznivých situací – od výše uvedeného neoprávněného zásahu do soukromí a integrity jednotlivých subjektů údajů či jeho příbuzných, přes potenciální genetické diskriminace určitých jedinců ve vztahu ke zdravotnímu či sociálnímu pojištění a přístupu ke zdravotní péči jako takové, až po různé důsledky pro konkrétní skupiny subjektu údajů atp. Též by bylo nutné uvést opatření a záruky přijaté za účelem minimalizace těchto rizik a zároveň prokázat, že jsou vhodná a dostačující, v souladu s GDPR a dalšími zásadami zpracování

²⁵⁹ Čl. 24 odst. 1 GDPR.

²⁶⁰ Srov. čl. 35 odst. 1, odst. 3 a odst. 4 GDPR.

²⁶¹ Srov. čl. 35 odst. 2 odst. 7 až odst. 9 GDPR.

²⁶² Srov. čl. 36 odst. 3 písm. e) GDPR.

²⁶³ Čl. 83 odst. 4 písm. a) GDPR.

²⁶⁴ STAUNTON, C., SLOKENBERGA, S., PARZIALE, A., MASCALZONI, D. Appropriate Safeguards and Article 89 of the GDPR: Considerations for Biobank, Databank and Genetic Research. *Frontiers in genetics* [online]. Switzerland: Frontiers Media S.A, 2022, 13, 719317-719317 [cit. 2022-04-06]. ISSN 1664-8021. Dostupné z: doi:10.3389/fgene.2022.719317.

osobních údajů.²⁶⁵ Pro genome browsery užívající ke zpracování genetických údajů další různé nástroje a programy by to také znamenalo povinnost provádění DPIA také pro každý takovýto nástroj. Za určitých podmínek by mohlo být teoreticky z hlediska efektivity přijatelné vypracování DPIA obecně v rámci komplexního projektu celého genome browseru, přičemž představitelné je to bez pochyby minimálně ve vztahu k určitému okruhu užívaných aplikací a nástrojů se stejným zaměřením a s podobnými riziky.²⁶⁶

Paul Quinn a Liam Quinn uvádí, že posouzení a vyhodnocení veškeré možné újmy v této souvislosti a přijatých opatření k jejich minimalizaci by byl *obrovský úkol vyžadující užití multidisciplinární perspektivy z pohledu etiky, práva, genetiky a sociologie*.²⁶⁷ Zpracování genetických údajů a jejich shromažďování může mít totiž negativní vliv na základní lidská práva a svobody, jako je například právo na soukromí, zachování důstojnosti, na rodinný život, ochranu osobních údajů, nediskriminaci, svobodu vědeckého bádání a umělecké tvorby apod.²⁶⁸ Je tedy otázkou, zda by byli správci genetických údajů schopni řádně takové posouzení provést a zda by byly jejich závěry dostatečně objektivní a relevantní, tedy zdali by provedení DPIA bylo v těchto případech schopno dosáhnout svého účelu a být efektivním podpůrným mechanismem ochrany osobních údajů.

5.5 Souhlas se zpracováním údajů

Další možností zpracování genetických údajů, tj. citlivých osobních údajů, je dle GDPR udělení již výše zmíněného souhlasu subjektu údajů.²⁶⁹ Aby mohly být osobní údaje zpracovávány na základě souhlasu, musí být takový souhlas udělen dobrovolně, pro konkrétně stanovený účel, subjekt údajů musí být předem jasně a srozumitelně informován o relevantních skutečnostech a o tom, že je oprávněn souhlas kdykoli odvolat.²⁷⁰ Souhlas také musí být udělen výslovně a jednoznačně.²⁷¹

²⁶⁵ STAUNTON, C. et. al. Op. cit. 264.

²⁶⁶ Global Alliance for Genomics and Health. GDPR Brief: The Data Protection Impact Assessment and Genomic Health Research. 2019. [online]. [cit. 2022-03-29]. Dostupné z: <https://www.ga4gh.org/news/gdpr-brief-the-data-protection-impact-assessment-and-genomic-health-research/>

²⁶⁷ QUINN, P., QUINN, L. Op. cit. 246.

²⁶⁸ STAUNTON, C. et. al. Op. cit. 264 a LZPS.

²⁶⁹ Čl. 9 odst. 2 písm. a) GDPR.

²⁷⁰ Čl. 4 odst. 11 GDPR.

²⁷¹ GDPR v kostce, Praktický průvodce povinnostmi pro podniky a spolky. C. H. BECK, 1. vydání, 2018. St. 18.

Tato práce je zaměřena na zpracování genetických údajů převážně pro účely vědeckého výzkumu a výzkumu v oblasti zdravotnictví. Udělování souhlasu pro účely poskytování zdravotních služeb konkrétní osobě není ani v případě genetických údajů nutně problematické, jelikož lze poměrně přesně definovat účel zpracování těchto údajů, nebo je jejich zpracování založeno na jiném základu dle čl. 9 odst. 2 GDPR.²⁷²

Aby byl souhlas platný a účinný, je nezbytné, aby splňoval veškeré výše uvedené náležitosti. V souvislosti se zpracováním genetických údajů může být však tato podmínka velmi obtížně splnitelná. Podstatou souhlas je totiž skutečnost, že subjekt údajů si je vědom důvodu a účelu zpracování, způsobu nakládání s jeho osobními údaji, a to tak, aby subjekt údajů plně pochopil veškeré okolnosti v této souvislosti. Aby mohl subjekt údajů udělit platný souhlas, musí mu být dostatečně jasně a srozumitelně vysvětleny veškeré informace o zpracování údajů, což může pro správce osobních údajů, zejm. při zpracování genetických údajů pořizovateli genetických databází a genome browserů, znamenat značnou administrativní či personální zátěž.

5.5.1 Udělení souhlasu pro konkrétní účel zpracování

Subjekt údajů musí pořizovateli genetické databáze či genome browseru udělit souhlas ke zpracování osobních údajů, který je vázán na konkrétní účel zpracování. Je otázkou, zdali je vůbec možné, aby byl pořizovatel genetické databáze, resp. genome browseru, schopen dostatečně určitě vymežit účel zpracování údajů, jelikož ani pořizovatel genetické databáze si není ve většině případech vědom veškerých účelů, pro které bude databáze v budoucnu využita či dokonce pro které účely již využívána je.

Aby mohl pořizovatel genetické databáze či genome browseru oprávněně zpracovávat osobní údaje na základě uděleného souhlasu, je nutné stanovit sice dostatečně konkrétní, ale zároveň také dostatečně obecný účel zpracování údajů.

Udělení souhlasu zkrátka pro účely vědeckého výzkumu je z pohledu současné praxe považováno za příliš široké a zcela zřejmě nedostatečně určité. Souhlas subjektu údajů ke zpracování osobních údajů obecně za účelem vědeckého výzkumu by byl tedy za předpokladu, že nebude daný výzkum blíže vymezen, neplatný. Vymezení předmětných výzkumů musí též splňovat požadavky na postačující konkretizaci, tj. například není možné

²⁷² Zejm. čl. 9 odst. 2 písm. h) či c) GDPR.

udělit souhlas ani čistě pro účely lékařského výzkumu, pokud by subjektu nebyly známy záměry, pro které budou údaje využity.

Část odborné veřejnosti má však za to, že i souhlas udělený ke zpracování osobních údajů za obecným účelem může být dostatečně informovaný. Mark Sheenan uvádí, že souhlas k obecně vymezenému účelu, tzv. *široký souhlas*, neznamená nutně skutečnost, že subjekt údajů nemá dostatečné povědomí o záměrech, pro které jsou jeho údaje zpracovávány. Subjekt údajů při udělování širokého souhlasu musí přijmout významné rozhodnutí týkající se skutečnosti, zdali přenechá další rozhodování o zpracování svých údajů na jiných osobách či institucích, které následně údaje zpracovávají a zpřístupňují je třetím osobám. K přijetí takového rozhodnutí je nutné, aby byl subjekt údajů dostatečně informován o relevantních skutečnostech v této souvislosti, tedy například o fungování rozhodování ohledně sdílení dat, etických standardech správce údajů, podmínkách, zamýšleném využití údajů, podmínkách, které budou muset být splněny pro povolení přístupu ke zpracovávaným datům apod.²⁷³

Široký souhlas by měl vždy též obsahovat základní náležitosti, a to a) popis jakýchkoli předvídatelných rizik pro subjekt údajů, b) popis výhod předpokládaných výzkumných záměrů pro subjekt údajů, které lze rozumně předvídat, c) rozsah, v jakém budou uchovávány osobní údaje vedoucí k identifikaci subjektu údajů a d) informaci o tom, že účast na výzkumu je dobrovolná a subjekt údajů může souhlas kdykoli odvolat.²⁷⁴ Takto udělený souhlas ke zpracování osobních údajů by byl tedy dostatečně informovaný a mohl by být co do účelu považován za přiměřeně konkrétní.²⁷⁵

Jedním ze způsobů možného přijetí širokého souhlasu se zpracováním genetických údajů by mohl být proces zahrnující fungování nezávislé etické komise. Subjekt údajů by udělil obecný souhlas se zpracováním svých osobních údajů pro výzkumné účely, přičemž o zpřístupnění dat by dále v každém jednotlivém případě rozhodovala etická komise. Subjekt údajů by tak mohl určitým způsobem delegovat své rozhodování na nezávislou komisi, která by následně vyhodnocovala konkrétní účely, pro které by potenciálně zpracovávané údaje

²⁷³ SHEEHAN, M. Can Broad Consent be Informed Consent? *Public Health Ethics*. 2011. [online] 4(3):226-235. [cit. 2023-03-16]. Dostupné z doi: 10.1093/phe/phr020.

²⁷⁴ MALOY, J. W., BASS, P. F., 3rd. Understanding Broad Consent. *The Ochsner journal*, 20(1), 81–86. 2020. [online]. [cit. 2023-03-21]. Dostupné z: <https://doi.org/10.31486/toj.19.0088>

²⁷⁵ GRADY, C., ECKSTEIN, L., BERKMAN, B., et. al. Broad Consent for Research With Biological Samples: Workshop Conclusions. 2015. [online]. *The American journal of bioethics: AJOB*, 15(9), 34–42. [cit. 2023-03-21]. Dostupné z: <https://doi.org/10.1080/15265161.2015.1062162>

zpřístupnila. Tento způsob respektuje vůli subjektu údajů a zároveň umožňuje poměrně účinný mechanismus zajišťující co nejbezpečnější a přiměřeně časově úsporné zpřístupňování údajů obsažených v genetických databázích a genome browserech.²⁷⁶

Autorka se přiklání k názoru, že souhlas splňující dané náležitosti by bylo možné považovat za dostatečně informovaný, přičemž široce vymezený účel by na základě teleologického výkladu za výše uvedených předpokladů splňoval podmínku konkrétního účelu dle čl. 6 odst. 1 písm. a) GDPR, čímž by byla dána zákonnost zpracování osobních údajů.²⁷⁷

Nelze mít za to, že široký souhlas lze obecně používat pro jakýkoli typ zpracování osobních údajů, tento přístup se však zdá být poměrně racionální a efektivní právě pro případ zpracování genetických údajů v genetických databázích a genome browserech, a to dále za předpokladu, že genetická databáze či genome browsery, nebudou bez dalšího přístupné široké veřejnosti a vždy bude v konkrétních případech rozhodováno o zpřístupnění databáze třetím subjektům.

Pokud bude souhlas udělen pro konkrétní specifikovaný výzkum, jehož náležitosti a záměry budou subjektu údajů prokazatelně známy a bude s nimi dostatečně srozuměn, přičemž budou splněny též další náležitosti souhlasu, bude tento souhlas platný. Je však nutné, aby byl takovýto souhlas udělen pro každý jednotlivý výzkumný projekt, pro jehož účely bude databáze či genome browser využíván. To by v praxi znamenalo, že by pořizovatel databáze, resp. genome browseru, musel všechny subjekty údajů opakovaně žádat o udělení dalších souhlasů pro každý další účel. Kromě zátěže s tím spojené na straně správce či zpracovatele je však toto řešení zatěžující i samotné subjekty údajů. Každý subjekt údajů bude muset správci údajů opakovaně udělovat souhlas se zpracováním, což by znamenalo poměrně zásadní časovou a administrativní náročnost. Kromě výše uvedených důsledků může opakované žádání o souhlasy ze strany pořizovatele databáze či genome browseru subjekty údajů odradit od jejich udělování. Je prokázáno, že čím více a častěji je po subjektech údajů vyžadováno udělení opětovného souhlasu, tím více mají subjekty údajů tendenci souhlasy neudělovat. To následně vede k poměrně nepředvídatelné fluktuaci dostupných dat v databázi, a především k nejistotě

²⁷⁶ STEINSBEKK, K. S., SOLBERG, B. Biobanks – When is Re-consent Necessary? [online]. Public Health Ethics, Vol. 4(3). 2011. St. 243 [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/phe/phr031>

²⁷⁷ Čl. 6 odst. 1 písm. a) ve spojení s čl. 9 odst. 2 písm. a) GDPR.

pořizovatele databáze a jejích uživatelů ve vztahu k rozsahu a možnosti využití zpracovávaných údajů.²⁷⁸

Pro účely zpracování genetických dat genetickými databázemi a genome browsery se s ohledem na výše uvedené jeví jako nejefektivnější řešení užití širokého souhlasu za splnění určitých podmínek zajišťujících bezpečné a účelné zpracovávání těchto údajů třetími subjekty. Pokud nebude možné postavit zpracování genetických údajů na jiném důvodu dle GDPR, nemusí být nutnost udělení souhlasu od subjektu údajů nepřiměřeně limitující, a to ani pro pořizovatele databáze, ani pro subjekty údajů. Je na místě podotknout, že nutností aktivního jednání subjektů údajů ve formě udělení souhlasu by byla zajištěna větší míra participace subjektů údajů na rozhodování o zpracování údajů a samotné zpracování by se mohlo jevit jako o to legitimnější.

5.6 Omezení zpracování genetických údajů genetickými databázemi

S ohledem na nemožnost anonymizace genetických údajů se na jejich zpracování uplatní příslušná právní úprava. To především znamená, že pořizovatelé databází a genome browserů budou muset prokazovat, na jakém základě genetické údaje zpracovávají a jaké záruky anebo opatření za účelem zabezpečení osobních údajů přijali.

Nutnost aplikace příslušné právní úpravy stanovené zejm. GDPR na genetické databáze, genome browsery a data v nich obsažená, by mohla vést k omezení jejich přístupnosti a dalším restriktivním opatřením. V takovém případě by se též mohl zpomalit a ztížit výzkum, k jehož podpoře primárně genome browsery a genetické databáze existují. GDPR uvádí, že *zpracování osobních údajů by mělo sloužit lidem, přičemž právo na ochranu osobních údajů není právem absolutním a musí být posuzováno v souvislosti se svou funkcí ve společnosti a v souladu se zásadou proporcionality musí být v rovnováze s dalšími základními právy.*²⁷⁹ Bude tedy nutné vyhodnotit a posoudit především závažnost a pravděpodobnost rizika potenciálního zneužití těchto dat v porovnání s hodnotou možného vědeckého přínosu a jeho využití v praxi. Pravděpodobnost a závažnost rizika ve vztahu k právům a svobodám subjektů údajů by se měly určovat dle *povahy, rozsahu, kontextu a účelů zpracování*, přičemž by k hodnocení rizik mělo docházet z objektivního hlediska.²⁸⁰

²⁷⁸ STEINSBEKK, K. S. Op. cit. 276, st. 236-250.

²⁷⁹ Rec. 4 GDPR.

²⁸⁰ Rec. 76 GDPR.

Pro zmírnění negativních dopadů na vědecký výzkum a klinické užití genome browserů a genetických databází mají správci možnost uchýlit se k přijetí kodexů chování,²⁸¹ ve kterých by se zavázali k řádnému uplatňování ustanovení GDPR vzhledem ke specifickým daného odvětví a k jejich konkrétním potřebám – např. s ohledem na zpracování genetických dat za účelem vědeckého výzkumu. K těmto kodexům je potom přihlíženo mimo jiné i při zpracovávání DPIA²⁸² a mohou znamenat usnadnění dodržování zejména unijní právní úpravy zpracování údajů.²⁸³ Přestože přijetí takovýchto kodexů, ve kterých bude poukazováno na riziko zpětné identifikace subjektů „anonymizovaných“ genetických dat, může pro správce znamenat usnadnění procesu zpracování a sdílení údajů, musí stále brát v této souvislosti v úvahu skutečnost přítomnosti zbytkového rizika. Toto riziko by se správci měli snažit za užití vhodných metod a technologií minimalizovat, a to vždy s ohledem na konkrétní potřeby, cíle a příslušné způsoby zpracování dotčených údajů.²⁸⁴

GDPR dává zpracovatelům a správcům také možnost získání osvědčení o ochraně údajů, pečeti či známk udělených určitými subjekty či dozorovými úřady²⁸⁵ za účelem prokázání existence vhodných záruk jimi poskytovaných v souvislosti s doložením souladu ochrany údajů s GDPR. Pro genome browsery by mohla tato certifikace znamenat větší důvěryhodnost a transparentnost, jelikož díky těmto osvědčením, mohou subjekty údajů snadno a rychle posoudit úroveň poskytované ochrany jejich údajů.²⁸⁶ Získáním těchto osvědčení se však nikterak nesnižuje odpovědnost provozovatelů genome browserů za soulad s GDPR.²⁸⁷

Přijímání vhodných opatření je v tomto odvětví zcela nezbytné, vzhledem k důsledkům, které by mohly nastat při neoprávněném zásahu do soukromí subjektů údajů či zneužití těchto údajů, tj. např. ztracení věrohodnosti správců údajů, stejně jako důvěry subjektů údajů.²⁸⁸

²⁸¹ Čl. 40 a 41 GDPR.

²⁸² Čl. 35 odst. 8 GDPR.

²⁸³ European Union Agency for Fundamental Rights and Council of Europe. Handbook on European data protection law. 2018. St. 181. [online]. [cit. 2022-04-20]. Dostupné z: https://www.echr.coe.int/documents/handbook_data_protection_eng.pdf

²⁸⁴ SHABANI, M. Op. cit. 243.

²⁸⁵ Čl. 43 GDPR.

²⁸⁶ European Union Agency for Fundamental Rights and Council of Europe. Op. cit. 283, st. 183.

²⁸⁷ Čl. 42 odst. 4 GDPR.

²⁸⁸ SHABANI, M. Op. cit. 243.

To by mohlo mít mimo jiné za následek pokles přílivu nových genetických vzorků, které jsou však pro klinickou praxi a výzkum v dnešní době esenciální.

6. European Health Data Space

Stran Evropské unie je zřejmá snaha o stanovení právního rámce bezpečného a zároveň efektivního zpracování genetických údajů. Úprava zpracování zdravotnických dat, jejich sdílení a využití veřejnými i soukromými subjekty členských států, je předmětem projektu European Health Data Space.

European Health Data Space (EHDS), evropský prostor pro zdravotní data, je nová iniciativa Evropské unie, která si klade za cíl zlepšit zdravotní péči a výzkum v Evropské unii podporováním sdílení zdravotnických dat mezi členskými státy. EHDS má vytvořit bezpečnou a interoperabilní platformu pro sdílení zdravotnických dat napříč Evropskou unií, což umožní zejm. zdravotnickým pracovníkům a vědcům na úrovni výzkumu snadnější přístup a využití těchto dat.²⁸⁹

V rámci EHDS má být zpracováváno velké množství elektronických dat, a to převážně osobních údajů vypovídajících o zdravotním stavu jedinců včetně genetických údajů, tedy údajů spadajících do zvláštní kategorie údajů.^{290, 291} Návrh nařízení o EHDS zohledňuje veškeré požadavky na zpracování osobních údajů stanovené GDPR a nařízením o ochraně fyzických osob v souvislosti se zpracováním osobních údajů orgány, institucemi a jinými subjekty Unie a o volném pohybu těchto údajů²⁹².

Lze předpokládat, že EHDS bude mít významný dopad na zdravotnictví a výzkum v Evropské unii, jelikož umožní zpracování velkého množství různých zdravotnických dat, která mohou být následně použita ke zlepšení výsledků léčby pacientů, vývoji nových léčebných postupů a léčivých přípravků. EHDS by měl minimalizovat stávající problémy týkající se fragmentace a ochrany údajů vypovídajících o zdravotním stavu subjektů údajů, protože zajistí, že tato data budou sbírána a sdílena v zabezpečeném režimu s poskytnutím vhodných záruk. EHDS má též potenciál podpořit vývoj více integrovaného a efektivního

²⁸⁹ Evropská komise. European Health Data Space. [online]. [cit. 2023-03-18]. Dostupné z: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

²⁹⁰ Čl. 9 odst. 1 GDPR.

²⁹¹ Čl. 2 odst. 2 písm. a) Návrhu nařízení Evropského parlamentu a Rady o evropském prostoru pro zdravotní data ze dne 3. 5. 2022 COM(2022) 197 final (Návrh o EHDS).

²⁹² Nařízení Evropského parlamentu a Rady 2018/1725 ze dne 23. 10. 2018 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů orgány, institucemi a jinými subjekty Unie a o volném pohybu těchto údajů a o zrušení nařízení (ES) č. 45/2001 a rozhodnutí č. 1247/2002/ES.

zdravotnického systému v EU, což by mohlo mít významné výhody pro pacienty a poskytovatele zdravotních služeb.

Cílem EHDS je tedy kromě zlepšení poskytování přeshraničních zdravotních služeb v rámci členských států Evropské unie, také poskytování občanům Evropské unie širší pravomoci nad spravováním svých osobních údajů vypovídajících o zdravotním stavu daného jedince. Subjekty údajů, resp. pacienti, budou mít ke svým údajům přístup v rámci elektronického a digitálního prostředí a budou moci ovlivnit způsob zpracování svých osobních údajů.²⁹³ Pacientům bude například umožněno kontrolovat a zvolit poskytovatele zdravotních služeb či jiné třetí osoby, se kterými budou dálkově sdílet údaje o svém zdravotním stavu. Pacienti budou moci též skrze dálkový přístup sdílet různé dokumenty, jako jsou například snímky, laboratorní výsledky, elektronické recepty, výsledky vyšetření apod., skrze svou standardizovanou zdravotnickou dokumentaci.

6.1 Sekundární využití elektronických zdravotních dat

EHDS se ve velkém rozsahu zaměřuje také na sekundární využití elektronických zdravotních dat²⁹⁴, tedy zpracování těchto údajů pro některé účely vymezené v návrhu o EHDS. Jedná se zejména o využití elektronických zdravotních dat pro potřeby výzkumu a inovací v oblasti zdravotnictví, činnosti ve veřejném zájmu v oblasti veřejného zdraví, přípravu oficiálních statistik, vývoj, testování a hodnocení zdravotnických prostředků, zdravotnických digitálních aplikací či systémů umělé inteligence či dokonce poskytování *individualizované zdravotní péče spočívající v posouzení, udržení nebo obnovení zdravotního stavu fyzických osob na základě zdravotních dat jiných fyzických osob*.²⁹⁵

Za účelem správy sekundárního využití zpracovávaných dat určí členské státy tzv. *subjekty pro přístup ke zdravotním datům*, jejichž úkolem bude sledovat uplatňování ustanovení návrhu o EHDS týkajících se sekundárního využití dat. Tyto subjekty by měly následně spolupracovat jak s orgány Evropské unie, patientskými a jinými zainteresovanými

²⁹³ Tímto prostředím by měla být platforma MyHealth@EU, definovaná v čl. 2 odst. 2 písm. t) Návrhu o EHDS jako *přeshraniční infrastruktura pro primární využití elektronických zdravotních dat tvořená kombinací vnitrostátních kontaktních míst pro digitální zdravotnictví a ústřední platformy pro digitální zdravotnictví*. Platforma MyHealth@EU je již zavedena v deseti členských státech EU, avšak pouze v minimálním rozsahu, který umožňuje velmi omezené služby.

²⁹⁴ V čl. 2 odst. 2 písm. a) Návrhu o EHDS je pojem „elektronická zdravotní data“ definován jakožto *zdravotní data a genetická data definovaná v nařízení (EU) 2016/679, jakož i data týkající se determinant zdraví nebo data zpracovávaná v souvislosti s poskytováním zdravotnických služeb, zpracovávaná v elektronické podobě*.

²⁹⁵ Čl. 34 odst. 1 Návrhu o EHDS.

organizacemi, tak mezi sebou navzájem. Stejně tak by měly úzce spolupracovat s orgány pro ochranu osobních údajů. Při rozhodování v rámci své kompetence by měly být subjekty nestranné, a to zejm. při vydávání povolení k datům²⁹⁶.²⁹⁷

Kompetence subjektů pro přístup ke zdravotním datům budou spočívat především v zajišťování účinného sekundárního využití dat. Jedná se zejm. o rozhodování o žádostech o přístup k datům a vydávání příslušných povolení k datům, přijímání opatření nutných k zachování důvěrnosti práv duševního vlastnictví, spolupráce s držiteli dat²⁹⁸, spolupráce s dalšími subjekty při stanovování vhodných opatření a požadavků pro přístup k elektronickým zdravotním datům v bezpečném zpracovatelském prostředí apod.²⁹⁹ Subjekty pro přístup ke zdravotním datům budou mít též řadu informačních povinností směrem k dalším evropským institucím, orgánům Evropské unie, dozorovým úřadům členských států a v neposlední řadě též směrem k subjektům údajů, tedy pacientům, samotným.

EHDS tedy usnadní, mimo jiné, též sdílení a zpřístupňování genetických údajů, a to dokonce při splnění veškerých požadavků stanovených právními předpisy Evropské unie a členských států na bezpečné zpracování zvláštní kategorie údajů. Sekundární využití elektronických zdravotních dat bude možné pouze za splnění daných náležitostí a za podmínky vhodného zabezpečení předmětných údajů. Při sekundárním využití se uplatní zásada minimalizace dat a omezení účelu. Takto využívaná data budou sdílena primárně v anonymizovaném formátu, pokud tak bude možné dosáhnout účelu jejich zpracování oprávněným uživatelem dat. Data lze zpřístupnit též pseudonymizovaná, přičemž informace potřebné ke zpětné identifikaci údajů budou přístupné pouze subjektům pro přístup

²⁹⁶ Dle čl. 2 odst. 2 písm. aa) Návrhu o EHDS se „povolením k datům“ rozumí *správní rozhodnutí vydané subjektem pro přístup ke zdravotním datům nebo držitelem dat pro uživatele dat, aby mohl zpracovat elektronická zdravotní data, která jsou uvedena v povolení k datům, pro účely sekundárního využití, jež jsou uvedeny v povolení k datům, a to na základě podmínek stanovených v tomto nařízení.*

²⁹⁷ Bod 42 až 50 odůvodnění Návrhu o EHDS.

²⁹⁸ „Držitelem dat“ je dle čl. 2 odst. 2 písm. y) Návrhu o EHDS *fyzická nebo právnická osoba, která je subjektem nebo orgánem v odvětví zdravotnictví nebo péče nebo která provádí výzkum v souvislosti s těmito odvětvími, jakož i orgány, instituce a jiné subjekty Unie, které mají v souladu s tímto nařízením, platnými právními předpisy Unie nebo vnitrostátními právními předpisy provádějícími právo Unie právo nebo povinnost, nebo v případě neosobních dat prostřednictvím kontroly technického návrhu produktu a souvisejících služeb možnost zpřístupnit určitá data, včetně jejich registrace, poskytování, omezení přístupu nebo výměny.*

²⁹⁹ Čl. 37 Návrhu o EHDS.

ke zdravotním datům a uživatelé dat nebudou oprávněni subjekty údajů ani jinak reidentifikovat.³⁰⁰

Přestože bude stran veškerých zúčastněných institucí zapojených do EHDS vyvíjena snaha o maximální zabezpečení zpracovávaných údajů a zpracování dat v rámci EHDS bude zcela v souladu s příslušnou legislativou Evropské unie, riziko spojené s neoprávněným zpracováním dat nevymizí úplně. A to dokonce ani v případě anonymizace dotčených údajů. Návrh o EHDS dokonce v bodě 64 svého odůvodnění výslovně uvádí, že *některé kategorie elektronických zdravotních dat mohou zůstat zvláště citlivé, i když jsou v anonymizované podobě, a nejsou tedy osobní, jak již výslovně předpokládá akt o správě dat³⁰¹. I v situacích, kdy se používají nejmodernější techniky anonymizace, zůstává zbytkové riziko, že by mohla být nebo se stát dostupnou schopnost opětovné identifikace, a to nad rámec prostředků, jejichž použití je přiměřeně pravděpodobné.* Toto riziko se týká například údajů ve vztahu k pacientům trpícím vzácnými onemocněními či jiných typů zpracovávaných údajů v závislosti na úrovni podrobnosti dat a popisu charakteristik subjektů údajů, počtu dotčených osob nebo rozsahu zpracovávaných údajů o daném subjektu údajů. Riziko zpětné identifikace bez pochyby zůstává i při zpracování genetických údajů. Dle Návrhu o EHDS bude nutné v každém takovém případě vyhodnotit rizikovost zpětné identifikace subjektu údajů a zohlednit specifika zpracovávaných údajů, použitých technik anonymizace a veškerých dalších relevantních skutečností v této souvislosti. Na tomto základě bude nezbytné přijmout vhodná ochranná opatření a postupovat v souladu s příslušnými ustanoveními aktu o správě dat, které stanoví další povinnosti uživatele dat.³⁰²

Z výše uvedeného je zřejmé, že některé anonymizované údaje, které nejsou předmětem ochrany ani stran GDPR, nelze bez dalšího považovat za nerizikové a neomezeně je zpracovávat bez jakýchkoli dalších opatření. Vyšší mírou opatrnosti, přijímáním vhodných záruk a opatření a existencí dohledu a nad zpracováním takto rizikových údajů lze však existující rizika spojená s jejich neoprávněným užitím poměrně zásadně minimalizovat. Tento přístup nemusí nutně znamenat nepřiměřené omezení možnosti užití předmětných údajů pro účely výzkumu. V případě, že výzkumné instituce a další uživatelé dat budou schopni

³⁰⁰ Čl. 44 Návrhu o EHDS.

³⁰¹ Návrh nařízení Evropského parlamentu a Rady o evropské správě dat ze dne 25. 11. 2020 COM(2020) 767 final (Akt o správě dat).

³⁰² Zejm. čl. 5 odst. 13 Aktu o správě dat.

zaručit dostačující zabezpečení a odůvodní svůj účel, za kterým údaje zpracovávají, není předpokládáno jakékoli větší či neadekvátní omezení.

Vzhledem k tomu, že cílem EHDS je především zlepšení poskytování přeshraničních zdravotních služeb a podpora výzkumu v oblasti zdravotnictví, nebylo by v zájmu subjektů spolupracujících na EHDS nepřiměřeně omezit sdílení zpracovávaných elektronických zdravotních dat. EHDS má zrychlit a usnadnit inovace a průběh různých studií zaměřených na zdravotnictví, vývoj léčivých přípravků, zdravotnických prostředků, zdravotnických aplikací a umělé inteligence. Je však zároveň nutné zajistit subjektům údajů, jejichž údaje jsou pro tyto účely využívány, dostatečné zabezpečení těchto dat a zamezit tak neoprávněnému zásahu do jejich soukromí. Jelikož převážná většina elektronických zdravotních dat, včetně genetických údajů, spadá do zvláštní kategorie údajů, je nezbytně nutné přijmout veškerá možná opatření k minimalizaci rizik spojených se zpracováním těchto citlivých údajů, a to dokonce přesto, že budou tyto údaje určitým způsobem anonymizovány.

Závěr

Cílem diplomové práce bylo především popsat současné fungování genetických databází a genome browserů a způsob zpracování genetických údajů zejména pro účely výzkumu. Vzhledem k tomu, že genetické údaje každého jednotlivce jsou jeho unikátním identifikátorem, mohou být také snadno zneužity, čímž jsou subjekty údajů vystaveny riziku zásahu do jejich osobnosti a soukromí, jakožto základních práv. Práce se z toho důvodu ze začátku zaměřila na základy osobnostních práv a možnosti jejich uplatňování. Následně bylo pojednáno o samotném grosu genetiky, tedy o DNA, sekvencování DNA a jeho metodám. Díky genetické variabilitě každé fyzické osoby jsou genetické informace všech osob od sebe odlišné. Přestože jsou některé úseky DNA přítomny téměř v každém lidském organismu ve stejné podobě, není možné vyloučit, že i vcelku obecná genetická informace nebude obsahovat jakékoli nuance, na jejichž základě bude možné daného jedince identifikovat.

Práce dále zohlednila stále dynamičtější vývoj výzkumu a vědy v oblasti genetiky a též zvětšující se objem zpracovávaných genetických údajů. V této souvislosti bylo uvedeno, že genetická data lze označit za tzv. big data, tedy různorodé informace zpracovávané ve velkém objemu a velkou rychlostí. Dále byla práce zaměřena na ukládání a zpřístupňování genetických dat mnohými zpracovateli v různých režimech, zejm. jakožto databáze, software obsahující databáze či SaaS, a využití moderních technologií. Pro lepší porozumění problematice demonstrovala práce fungování genome browseru na konkrétním příkladu browseru Ensembl. Bylo poukázáno na skutečnost, že převládá snaha o co možná největší dostupnost genetických dat, nejméně omezující přístup k nim a co nejsnadnější další zpracování genetických údajů. Toto je nejčastěji umožňováno na základě zásad free software a open source a užíváním příslušných permisivních licencí, které práce blíže popisuje. V souvislosti s fungováním genetických databází bylo mimo jiné krátce pojednáno též o vlivu nové evropské legislativy na zákonnou licenci týkající se vytěžování genetických databází.

Dle pořizovatelů genetických databází jsou zpracovávaná data anonymizována a díky této skutečnosti se na jejich zpracování nevztahují právní předpisy týkající se ochrany osobních údajů. Z úvodních kapitol však vyplývá, že genetická data nelze z převážné většiny již ze své podstaty zcela anonymizovat a příslušná legislativa se na tyto případy uplatní. Platí, že vzhledem k charakteru genetických údajů, možnosti jejich využití, jejich vlastnostem, současné nemožnosti jejich anonymizace a jejich citlivosti, je tedy nutné zajistit jejich náležitou ochranu a přijmout vhodná opatření a zaručit soulad jejich zpracování s GDPR a další příslušnou právní úpravou. Jedním z těchto opatření je např. provedení DPIA, což může být

však v tomto kontextu poněkud problematické a neúčelné. Otázce způsobu provedení DPIA pořizovateli genetických databází se práce věnuje a navrhuje způsoby, jakými lze DPIA provést efektivně. Především se však v této části práce zaměřuje na základ zpracování genetických údajů jakožto zvláštní kategorie osobních údajů, a to zejména jakožto zpracování nezbytné pro účely vědeckého výzkumu a zpracování na základě souhlasu subjektu údajů.

V práci byla dále vyhodnocena potenciální rizika vyplývající ze zneužití genetických údajů, a možný přínos jejich širokého zpřístupňování dalším zpracovatelům z různých odvětví z hlediska proporcionality. S ohledem na skutečnost, že k jejich využití v plném rozsahu je třeba určitých odborných znalostí a na fakt, že s větším množstvím dostupných genetických údajů se též zvyšuje pravděpodobnost možnosti léčby různých závažných onemocnění, je dle názoru autorky přípustné jejich zpřístupňování veřejnosti s minimálním omezením, avšak za splnění určitých podmínek. Z genetických dat lze s vědeckým pokrokem vyčíst stále větší množství informací, a to nejen o primárních subjektech údajů, ale též jejich genetických příbuzných. Je tedy žádoucí zajistit minimalizaci rizik zásahů do základních práv širokého okruhu osob a jejich odpovídající komplexní ochranu.

Autorka považuje za vhodné, aby zpracovatelé a správci genetických databází a genome browserů, přijímali vhodná a transparentní opatření, záruky a vyvinuli veškerou snahu o jejich vysoké zabezpečení a předcházení rizik, o čemž budou subjekty údajů a své koncové uživatele dostatečně informovat. Zůstává však otázkou, zdali dokonce splnění všech povinností vyplývajících ze současné příslušné právní úpravy, je schopné zaručit vhodnou ochranu těchto dat. Správci a zpracovatelé genetických údajů by tedy měli kromě povinných postupů přistupovat též k užití dalších dobrovolných nástrojů, jako jsou etické kodexy či osvědčení o ochraně osobních údajů. Pro úplnost je též třeba dodat, že v případě některých agregovaných genetických databází je riziko zpětné identifikace subjektů údajů velmi nízké, jelikož dostupná data jsou zobecněná a často zprůměrovaná. V případě určitých genetických databází obsahujících genetické údaje týkající se vzácných onemocnění je však nebezpečí reidentifikace subjektů údajů vysoké a hrozí tak vážná rizika spojená se zneužitím osobních údajů. Vzhledem k neustálému rozvoji výzkumu, technologií a s tím také možností využití genetických údajů, je však nutná daná potenciální rizika a příslušná opatření pravidelně vyhodnocovat a aktualizovat.

Poslední kapitola práce vyhodnocuje potenciální dopady iniciativy Evropské unie – European Health Data Space, která má usnadnit poskytování přeshraničních zdravotních služeb a sdílení zdravotních osobních údajů pacientů mezi členskými státy a dalšími zainteresovanými

subjekty. Díky EHDS by se v budoucnu měl mimo jiné stanovit jasnější rámec zpracování genetických údajů pro účely výzkumu a zavést obligatorní přijetí relevantních opatření pro zamezení rizik spojených s neoprávněným zpracováním genetických dat.

Autorka dochází k názoru, že nutnost aplikace příslušných opatření stanovených právními předpisy při zpracování genetických údajů, neznamená nutně zásadní či disproporcionální omezení vědeckého výzkumu v oblasti genetiky. S ohledem na citlivost genetických údajů a vážnost rizik spojených s jejich zneužitím považuje autorka za vhodné, aby bylo zpracování genetických údajů speciálně regulováno. Regulace však musí zohledňovat velký význam a přínos genetických údajů pro vědce a další odborníky a nesmí znamenat nepřiměřené omezení probíhajícího a budoucího výzkumu. Autorka práce má za to, že genetické databáze a genetické údaje by měly být k dispozici pro různé výzkumné záměry, jejich užití by však mělo být též stran států, jiných nezávislých subjektů, institucí Evropské unie či mezinárodních organizací náležitě regulováno a kontrolováno.

Seznam použitých zdrojů

LITERATURA

- DOUDNA, J. A., STERNBERG, S. H. A crack in creation: Gene editing and the unthinkable power to control evolution. Vintage. 2018. ISBN-10: 9781784702762
- ISAACSON, W. The Code Breaker: Jennifer Doudna, Gene Editing, and the Future of the Human Race. Simon & Schuster. 2021. ISBN-13. 978-6267012802
- JANSA, L. OTEVŘEL, P., ŠTEVKO, M. Softwarové právo. 3. aktualizované rozšířené vydání. 2018. Computer Press. ISBN: 978-80-251-4914-0
- KOHOUTOVÁ, M. Lékařská biologie a genetika (II. díl). 2. vydání. Praha: Nakladatelství Karolinum, 2017. ISBN: 978-80-246-3790-7
- LAVICKÝ, P. a kol. Občanský zákoník I. Obecná část (§1-654). Komentář. 1. vydání. Praha: C. H. Beck, 2014, 394. ISBN 978-80-7400-529-9
- MATES, P. (ed.), VALOUŠEK M., FIALOVÁ E., LECHNER, T., HÁLOVÁ, M., SIVÁK, J., SOVOVÁ, O., BRUNA, E., BRUNOVÁ, M. Ochrana osobnosti, soukromí a osobních údajů. Nakladatelství Leges, s. r. o., podle stavu k 14. 8. 2019. ISBN: 978-80-7502-346-9
- MENDEL, G. J., MÜLLER-WILLE, S. (ed.), HALL, K. (ed.), DOSTÁL, O. (ed.), RADICK, G., RYCHLÍK, M. Versuche über Pflanzen-hybriden. Experiments on Plant Hybrids. New Translation with Commentary. Masarykova univerzita, 2020. ISBN: 978-80-210-9554-0
- MUKHERJEE, S. The gene: an intimate history. New York: Scribner. 2017. ISBN 978-1-4767-3352-4
- TELEC I., TŮMA P. Autorský zákon. Komentář. 2. vydání. Praha: C.H.Beck 2019. ISBN 978-80-7400-748-4
- VLACHOVÁ, B., MAISNER, M. Zákon o zpracování osobních údajů. Komentář. 1. vydání. Praha: C. H. Beck, 2019. ISBN 978-80-7400-760-6

ODBORNÉ ČLÁNKY

- CHIMUKANGARA, B., WILKINSON, E., FISH, M., RAMSURAN, V. Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. *Frontiers in genetics* [online]. Frontiers Media S.A, 2020, 11, 544162-544162 [cit. 2022-04-05]. ISSN 1664-8021. Dostupné z: doi:10.3389/fgene.2020.544162
- DANKAR, F. K., PTITSYN, A., DANKAR, K. S.. The development of large-scale de-identified biomedical databases in the age of genomics-principles and challenges. [online]. *Human genomics*. England: BioMed Central, 2018, 12(1), 19-19 [cit. 2022-04-05]. ISSN 1479-7364. Dostupné z: doi:10.1186/s40246-018-0147-5
- DEARY, I. J, JOHNSON, W., HOULIHAN, L. M. Genetic foundations of human intelligence. *Human genetics*. [online]. Berlin/Heidelberg: Springer-Verlag, 2009, 126(1), 215-232 [cit. 2022-03-28]. ISSN 0340-6717. Dostupné z doi:10.1007/s00439-009-0655-4

- FERWERDA, B., MCCALL, M. B. B., ALONSO, S. et. al. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. 2007. [online]. [cit. 2022-03-23]. Dostupné z: <https://doi.org/10.1073/pnas.0704828104>
- FINCK, M., PALLAS, F. They who must not be identified—distinguishing personal from non-personal data under the GDPR. *International data privacy law* [online]. Oxford: Oxford University Press, 2020, 10(1), 11-36 [cit. 2022-04-05]. ISSN 2044-3994. Dostupné z: doi:10.1093/idpl/ipz026
- FLEISHMANN, R.D., ADAMS, M. D., WHITE, O., et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. [online]. *Science* 269, 496-512. July 28, 1995. [cit. 2022-03-11]. Dostupné z: <https://doi.org/10.1126/science.7542800>
- GONZAGA-JAUREGUI, C., LUPSKI, J. R., GIBBS, R. A. Human genome sequencing in health and disease. 2012. [online]. *Annual Review of Medicine*. 63:35-61. [cit. 2022-04-06]. Dostupné z: <https://doi.org/10.1146/annurev-med-051010-162644>
- GRADY, C., ECKSTEIN, L., BERKMAN, B., et. al. Broad Consent for Research With Biological Samples: Workshop Conclusions. 2015. [online]. *The American journal of bioethics: AJOB*, 15(9), 34–42. [cit. 2023-03-21]. Dostupné z: <https://doi.org/10.1080/15265161.2015.1062162>
- GRIFFITH, F. The Significance of Pneumococcal Types. 1928. [online] *Journal of Hygiene*. Cambridge University Press. 27 (2): 113–159. [cit. 2022-03-10]. Dostupné z doi: 10.1017/s0022172400031879
- GUDMUNDSSON, S., SINGER-BERK, M., WATTS, N.A., PHU, W., GOODRICH, J.K., SOLOMONSON, M.; Genome Aggregation Database Consortium; REHM, H.L., MACARTHUR, D.G., O'DONNELL-LURIA, A. Variant interpretation using population databases: Lessons from gnomAD. *Hum Mutat*. 2022. [online]. [cit. 2023-03-11]. Dostupné z doi: 10.1002/humu.24309
- GYMREK, M., MCGUIRE, A. L, GOLAN, D., HALPERIN, E., ERLICH, Y. Identifying Personal Genomes by Surname Inference. *Science*. [online]. United States: American Association for the Advancement of Science, 2013, 339(6117), 321-324 [cit. 2022-04-06]. ISSN 0036-8075. Dostupné z doi:10.1126/science.1229566
- HARGER, C., CHEN, G., FARMER, A., et al. The genome sequence DataBase. *Nucleic Acids Res*. [online]- 2000;28(1):31-32. [cit. 2022-04-05]. Dostupné z: doi:10.1093/nar/28.1.31
- HARTE, R.A., KAROLCHIK, D., KUHN, R.M., KENT, W.J., HAUSSLER, D. Databases and Genome Browsers. M.R. Speicher et al. (eds.), *Vogel and Motulsky's Human Genetics, Problems and Approaches*. 2010. Springer, Berlin, Heidelberg. [online]. [cit. 2023-03-16]. Dostupné z doi: 10.1007/978-3-540-37654-5_44
- HELLMAN, D. What makes genetic discrimination exceptional? [online]. *American journal of law & medicine*. United States: Sage Publications, 2003, 29(1), 77-116 [cit. 2022-04-05]. Dostupné z doi: 10.1017/S0098858800002343
- HERRERO, J., MUFFATO, M., BEAL, K., FITZGERALD, S. et al. Ensembl comparative genomics resources, Database, Volume 2016, 2016, baw053 [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/database/baw053>

- HILLIS, D.M., DIXON, M.T. Ribosomal DNA: molecular evolution and phylogenetic inference. [online]. Q Rev Biol. 1991 Dec;66(4):411-53. [cit. 2022-03-27]. Dostupné z doi: 10.1086/417338
- HUBBARD, T., BARKER, D., BIRNEY, E., et al. The Ensembl genome database project. Nucleic Acids Res. 2002;30(1):38-41. [online]. [cit. 2022-04-06]. Dostupné z doi:10.1093/nar/30.1.38
- HUMBERT, M., HUGUENIN, K., HUGONOT, J., AYDAY, E., HUBAUX, J. De-anonymizing Genomic Databases Using Phenotypic Traits. Proceedings on Privacy KOUBKOVÁ, L., VOJTĚŠEK, B., VYZULA, R. Sekvenování nové generace a možnosti jeho využití v onkologické praxi. [online]. Klin Onkol 2014; 27 (Suppl 1): S61–S68. [cit. 2022-04-06]. Dostupné z: doi: 10.14735/amko20141S61. Enhancing Technologies 2015. [online]. 2015 (2):1–16. [cit. 2023-03-18]. Dostupné z doi: 10.1515/popets-2015-0020
- JINXIANG, C., LI, F., WANG, M., LI, J., MARQUEZ-LAGO, T. T., LEIER, A., REVOTE, J., LI, S., LIU, Q., SONG, J. BigFiRSt: A Software Program Using Big Data Technique for Mining Simple Sequence Repeats From Large-Scale Sequencing Data [online]. 18. 1. 2022, 1-16. [cit. 2022-03-25]. Dostupné z: <https://doi.org/10.3389/fdata.2021.727216>
- KATSILA, T., PATRINOS, G. P. Whole genome sequencing in pharmacogenomics. Frontiers in pharmacology. [online]. Switzerland: Frontiers Media S.A, 2015, 6, 61-61 [cit. 2022-04-05]. ISSN 1663-9812. Dostupné z doi:10.3389/fphar.2015.00061
- KING, R. C., STANSFIELD, W. D., MULLIGAN, P. K. A Dictionary of Genetics, Seventh Edition. 2007. [online]. Oxford University Press. [cit. 2022-03-23]. Dostupné z doi: 10.1093/acref/9780195307610.001.0001
- KOUBKOVÁ, L., VOJTĚŠEK, B., VYZULA, R. Sekvenování nové generace a možnosti jeho využití v onkologické praxi. [online]. Klin Onkol 2014; 27 (Suppl 1): S61–S68. [cit. 2022-04-06]. Dostupné z doi: 10.14735/amko20141S61
- KREJČÍ A., MÜLLER P., VOJTĚŠEK B. Bioinformatika a sekvenování nové generace. [online]. Klin Onkol 2015; 28 (Suppl 2): 2S93. [cit. 2022-04-06]. Dostupné z: <http://dx.doi.org/10.14735/amko20152S91>
- LANDRUM, M. J., CHITIPIRALLA, S., BROWN, G. R., (...). ClinVar: improvements to accessing data. Nucleic Acids Res. 2020. [online]. 48(D1):D835-D844. [cit. 2023-03-11]. Dostupné z doi: 10.1093/nar/gkz972
- LAPPALAINEN, T., SCOTT, A. J., BRANDT, M., HALL, I. M. Genomic Analysis in the Age of Human Genome Sequencing. Cell. 177(1), 70–84. 2019. [online]. [cit. 2023-03-19] Dostupné z: <https://doi.org/10.1016/j.cell.2019.02.032>
- LEK, M., KARCZEWSKI, K. J., MINIKEL, E. V., et al. Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. 2016. [online]. Nature, 536(7616), 285–291. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1038/nature19057>
- MALIN, B., SWEENEY, L. How (not) to protect genomic data privacy in a distributed network: using trail re-identification to evaluate and design anonymity protection systems. Journal of biomedical informatics [online]. US: Elsevier, 2004, 37(3), 179-192 [cit. 2022-04-06]. ISSN 1532-0464. Dostupné z doi:10.1016/j.jbi.2004.04.005
- MALOY, J. W., BASS, P. F., 3rd. Understanding Broad Consent. The Ochsner journal, 20(1), 81–86. 2020. [online]. [cit. 2023-03-21]. Dostupné z: <https://doi.org/10.31486/toj.19.0088>

- MASHIMA, J., KODAMA, Y., FUJISAWA, T., KATAYAMA, T. et. al. DNA Data Bank of Japan. *Nucleic acids research*, 45(D1), D25–D31. 2017. [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/nar/gkw1001>
- MORLEY, K. I., HALL, W. D. Is There a Genetic Susceptibility to Engage in Criminal Acts?. *Trends and issues in crime and criminal justice*. [online]. Woden: Australian Institute of Criminology, 2003, (263), 1 [cit. 2022-04-05]. ISSN 0817-8542. ISBN 0 642 53816 6. Dostupné z: https://www.aic.gov.au/sites/default/files/2020-05/tandi263_0.pdf
- NGUN, T. C., VILAIN, E. The biological basis of human sexual orientation: is there a role for epigenetics?. *Advances in genetics*. [online]. United States, 2014, 86, 167-184 [cit. 2022-04-05]. ISSN 0065-2660. Dostupné z doi:10.1016/B978-0-12-800222-3.00008-5
- NOGRADY, B. How cancer genomics is transforming diagnosis and treatment. *Nature (London)*. [online]. England: Nature Publishing Group, 2020, 579(7800), S10-S11 [cit. 2022-04-05]. Dostupné z doi:10.1038/d41586-020-00845-4
- O'DONNELL, C. J., NABEL, E. G. Genomics of Cardiovascular Disease. 2011. [online]. *New England Journal of Medicine*. [cit. 2022-04-05]. Dostupné z doi: 10.1056/NEJMra1105239
- PATEL, N., FERNS, R. B., NASTOULI, E., KOZLAKIDIS, Z., KELLAM, P., MORRIS, S. Cost analysis of standard Sanger sequencing versus next generation sequencing in the ICONIC study. *The Lancet (British edition)* [online]. Elsevier, 2016, 388, S86-S86 [cit. 2022-04-05]. Dostupné z doi:10.1016/S0140-6736(16)32322-4
- PIKE, E. R. Securing sequences: ensuring adequate protections for genetic samples in the age of big data. *Cardozo law review*. [online]. Cardozo School of Law, 2016, 37(6), 1977 [cit. 2022-04-05]. ISSN 0270-5192. Dostupné z: <http://cardozolawreview.com/wp-content/uploads/2018/08/PIKE.37.6.pdf>
- QUINN, P., QUINN, L. Big genetic data and its big data protection challenges. *The computer law and security report* [online]. Elsevier, 2018, 34(5), 1000-1018 [cit. 2022-04-06]. Dostupné z doi:10.1016/j.clsr.2018.05.028
- SANGER, F., COULSON, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. [online]. *Journal of Molecular Biology*, 94(3), 441–448. [cit. 2022-03-24]. Dostupné z: [https://doi.org/10.1016/0022-2836\(75\)90213-2](https://doi.org/10.1016/0022-2836(75)90213-2)
- SHABANI, M., MARELLI, L. Re-identifiability of genomic data and the GDPR [online]. 24. 5. 2019, 1-5 [cit. 2022-03-25]. Dostupné z doi:10.15252/embr.201948316
- SHEEHAN, M. Can Broad Consent be Informed Consent? *Public Health Ethics*. 2011. [online] 4(3):226-235. [cit. 2023-03-16]. Dostupné z doi 10.1093/phe/phr020
- SHERRY, S. T., WARD, M. H., KHOLODOV, M., et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* [online]. [cit. 2022-04-05]. 2001;29(1):308-311. Dostupné z doi: 10.1093/nar/29.1.308.
- SLATKO, E. B., GARDNER, F. A., AUSUBEL, M. F. Overview of Next Generation Sequencing Technologies. 2018. [online]. [cit. 2022-03-24]. Dostupné z: doi:10.1002/cpmb.59
- STABENAU, A., MCVICKER, G., MELSOPP, C., PROCTOR, G., CLAMP, M., & BIRNEY, E. The Ensembl core software libraries. *Genome research*, 14(5), 929–933. 2004. [online]. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1101/gr.1857204>

- STAUNTON, C., SLOKENBERGA, S., PARZIALE, A., MASCALZONI, D. Appropriate Safeguards and Article 89 of the GDPR: Considerations for Biobank, Databank and Genetic Research. *Frontiers in genetics* [online]. Switzerland: Frontiers Media S.A, 2022, 13, 719317-719317 [cit. 2022-04-06]. ISSN 1664-8021. Dostupné z: [doi:10.3389/fgene.2022.719317](https://doi.org/10.3389/fgene.2022.719317)
- STEINMAN, R. M., MOBERG, C. L. Dostupné z: A triple tribute to the experiment that transformed biology. *Discovering the Genetic Role of DNA*. [online]. *Journal of Experimental Medicine*. 1994. 179 (2): 379–384. [cit. 2022-03-10]. Dostupné z: <https://doi.org/10.1084/jem.179.2.379>
- STEINSBEKK, K. S., SOLBERG, B. Biobanks – When is Re-consent Necessary? [online]. *Public Health Ethics*, Vol. 4(3). 2011. [cit. 2023-03-19]. Dostupné z: <https://doi.org/10.1093/phe/phr031>
- TRAJANOSKI Z. *Computational Medicine, Tools and Challenges*, 2012. [online]. Springer Vienna. [cit. 2023-03-18]. ISBN: 978-3-7091-0947-2. Dostupné z doi: <https://doi.org/10.1007/978-3-7091-0947-2>
- VON THENEN, N., AYDAY, E., CICEK, A. E..Re-identification of individuals in genomic data-sharing beacons via allele inference. *Bioinformatics* [online]. England: Oxford University Press, 2019, 35(3), 365-371 [cit. 2022-04-06]. Dostupné z: [doi:10.1093/bioinformatics/bty643](https://doi.org/10.1093/bioinformatics/bty643)
- WARREN, S. D., BRANDEIS, L. D. The Right to Privacy. In *Harvard Law Review*, 1890, roč. 4, č. 5, s. 195. Harvard Law Review Association 1890. [online]. [cit. 2023-03-17]. Dostupné z <https://doi.org/10.2307/1321160>
- WATSON, J., CRICK, F. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. [online]. *Nature* 171, 737–738 (1953). [cit. 2022-03-11]. Dostupné z: <https://doi.org/10.1038/171737a0>
- WU, X., ZHU, X., WU, G. -Q., DING, W. Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*. Vol. 26, no. 1, pp. 97-107, [online]. [cit. 2023-03-16]. Dostupné z: [doi: 10.1109/TKDE.2013.109](https://doi.org/10.1109/TKDE.2013.109)

INTERNETOVÉ ZDROJE

- Berkley Publib Policy Journal. GEROSKI, A. S., ed. by: LIEDTKE, B., SAMPATH, N., SISTLA, M. Abuse of Our Genetic Data Is the Next Privacy Scandal. [online]. Spring 2019 Journal. [cit. 2022-04-11]. Dostupné z: <https://bppj.berkeley.edu/2019/02/27/spring-2019-abuse-of-our-genetic-data-is-the-next-privacy-scandal/>
- Centers for Disease Control and Prevention. What is Hemophilia? [online]. [cit. 2023-03-19]. Dostupné z: <https://www.cdc.gov/ncbddd/hemophilia/facts.html>
- Clinical Trials Arena. FDA approves first trial investigating CRISPR gene editing as HIV cure. 2021. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.clinicaltrialsarena.com/news/crispr-gene-editing-hiv-cure/>
- Cloudflare. What is the cloud? Cloud definition. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.cloudflare.com/en-gb/learning/cloud/what-is-the-cloud/>
- Creative Commons. [online]. [cit. 2023-03-19]. Dostupné z: <https://creativecommons.org>

- ČSSI. BARTOŠEK, V., PETRUCHA, R. Cloud Computing: moderní směr poskytování IT služeb, Systémová integrace. [online]. 2010. [cit. 2022-03-25]. Dostupné z: <http://3z894f6.257.cz/cloud-computing-moderni-smer-poskytovani-it-sluzeb>.
- ČSSI. SVOBODA, J. Cloud computing. Systémová integrace. [online]. 2009, 16.2: 66-87. [cit. 2022-03-24]. Dostupné z: <http://old.cssi.cz/system/files/all/0svoboda.pdf>
- Doba genova. Genetické Choroby. [online]. [cit. 2022-03-23]. Dostupné z: <https://doba-genova.cz/geneticke-ochorenia/>
- Doudna Lab. Editing in cells. [online]. [cit. 2022-03-28]. Dostupné z: https://doudnalab.org/research_areas/gene-fixing/
- EMBL's European Bioinformatics Institute. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk>
- EMBL's European Bioinformatics Institute. Licensing of EMBL-EBI data resources. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk/licensing>
- Encyclopedia Britannica. Britannica, T. Editors of Encyclopaedia. Chromosomal disorder., 2017. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.britannica.com/science/chromosomal-disorder>.
- ENISA. European Union Agency for Cyber Security. Deploying pseudonymisation techniques. The case of the Health Sector. 2022. [online] [cit. 2023-03-19]. Dostupné z: <https://www.enisa.europa.eu/publications/deploying-pseudonymisation-techniques>
- Ensembl. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ensembl.org/index.html>
- Ensembl. Ensembl Variation - Data access. [online]. [cit. 2023-03-19]. Dostupné z: https://m.ensembl.org/info/genome/variation/tools/data_access.html
- Ensembl. Software Licence. [online]. [cit. 2023-03-19]. Dostupné z: https://www.ensembl.org/info/about/legal/code_licence.html
- Ensembl. The Ensembl public MySQL Servers. [online]. [cit. 2023-03-19]. Dostupné z: <https://m.ensembl.org/info/data/mysql.html>
- European Commission. Co jsou to osobní údaje? [online]. [cit. 2022-03-28]. Dostupné z: https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_cs
- European Commission. European '1+ Million Genomes' Initiative. [online]. [cit. 2022-03-26]. Dostupné z: <https://digital-strategy.europa.eu/en/policies/1-million-genomes>
- European Commission. European Health Data Space. [online]. [cit. 2023-03-18]. Dostupné z: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
- European Nucleotide Archive. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ebi.ac.uk/ena/browser/home>
- Free Software Foundation. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.fsf.org>
- Generi Biotech. NOVOTNÁ, M. Princip NGS Metody. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.generi-biotech.com/cs/princip-ngs-metody/>
- Genetika - Biologie. [ŠÍPEK, A. RNA – Ribonukleová kyselina. 2008. [online]. cit. 2010-02-11]. Dostupné z: <http://www.genetika-biologie.cz/ribonukleova-kyselina>

- Genomics England. 100,000 Genomes Project. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.genomicsengland.co.uk/initiatives/100000-genomes-project>
- Global Alliance for Genomics and Health. GDPR Brief: The Data Protection Impact Assessment and Genomic Health Research. 2019. [online]. [cit. 2022-03-29]. Dostupné z: <https://www.ga4gh.org/news/gdpr-brief-the-data-protection-impact-assessment-and-genomic-health-research/>
- GNU Operating System. GNU General public License, Version 3, 29 June 2007. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org/licenses/gpl-3.0.html>
- GNU Operating System. Various Licenses and Comments about Them. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org/licenses/license-list.html#GPLCompatibleLicenses>
- GNU Operating System. What is GNU? [online]. [cit. 2023-03-18]. Dostupné z: <https://www.gnu.org>
- GoFair. FAIR Principles. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.go-fair.org/fair-principles/>
- Google Cloud Whitepaper. Handling genomic data in the cloud. 2019. [online]. [cit. 2022-03-25]. Dostupné z: <https://cloud.google.com/files/genomics-data-wp.pdf>
- Hevo Data. RAMAN, A. What is Data Aggregation? : A Comprehensive Guide 101. 4. 2. 2021. [online]. [cit. 2023-03-12]. Dostupné z: <https://hevodata.com/learn/data-aggregation/>
- IBM. PADMAGIRISON, S. Trustworthy AI provides business insights for the global leader in genomics. 2021. [online]. [cit. 2022-04-05]. Dostupné z: <https://www.ibm.com/blogs/journey-to-ai/2021/10/trustworthy-ai-provides-business-insights-for-the-global-leader-in-genomics/>
- IBM. SaaS vs PaaS vs IaaS: Whats the Difference & How to Choose. [online]. [cit. 2022-03-25]. Dostupné z: <https://blogs.bmc.com/saas-vs-paas-vs-iaas-whats-the-difference-and-how-to-choose/?print-posts=pdf>
- IBM. What is an API (application programming interface)? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.ibm.com/cloud/learn/api>
- IBM. What is Database-as-a-Service (DBaaS)? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.ibm.com/cloud/learn/dbaas>
- Illumina. Agrigenomics. Applying agricultural genomics to improve the food supply. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.illumina.com/areas-of-interest/agrigenomics.html>
- Illumina. Forensic genomics. NGS solutions for forensic applications. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.illumina.com/areas-of-interest/forensic-genomics.html>
- Illumina. NGS vs. Sanger Sequencing. [online]. [cit. 2022-03-24]. Dostupné z: <https://www.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html>
- International Nucleotide Sequence Database Collaboration. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.insdc.org>
- Internet Systems Consortium. ISC License. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.isc.org/licenses/>

- Mayo Clinic. Mayo Foundation for Medical Education and Research. Gene therapy. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.mayoclinic.org/tests-procedures/gene-therapy/about/pac-20384619>
- Mayo Clinic. Mayo Foundation for Medical Education and Research. Cystic Fibrosis. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.mayoclinic.org/diseases-conditions/cystic-fibrosis/symptoms-causes/syc-20353700>
- Merck. Sanger Sequencing Steps & Method. [online]. [cit. 2022-03-24]. Dostupné z: <https://www.sigmaaldrich.com/CZ/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing>
- Microsoft Azure. Co je IaaS? Infrastruktura jako služba. [online]. [cit. 2022-03-25]. Dostupné z: <https://azure.microsoft.com/cs-cz/overview/what-is-iaas/#overview>
- Microsoft. Microsoft Genomics. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.microsoft.com/en-us/genomics/>
- Mozilla. Mozilla Public License, Version 2. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.mozilla.org/en-US/MPL/2.0/>
- National Geographic. KOSSAKOVSKI, F. The eccentric scientist behind the ‘gold standard’ COVID-19 test. 2021. [online]. [cit. 2022-03-11]. Dostupné z: <https://www.nationalgeographic.com/science/article/the-eccentric-scientist-behind-the-gold-standard-covid-19-pcr-test>
- Nature. The Human Genome Project. Nature News. Barranco, C. [online]. [cit. 2022-03-25]. Dostupné z: <https://www.nature.com/articles/d42859-020-00101-9>
- NIH. National Human Genome Research Institute. A Brief Guide to Genomics. [online]. [cit. 2022-03-23]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics>
- NIH. National Human Genome Research Institute. DNA Sequencing Fact Sheet. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Fact-Sheet>
- NIH. National Human Genome Research Institute. 1944: DNA is "Transforming Principle". [online]. [cit. 2022-03-10]. Dostupné z: <https://www.genome.gov/25520250/online-education-kit-1944-dna-is-transforming-principle>
- NIH. National Human Genome Research Institute. 1983: First Disease Gene Mapped. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.genome.gov/25520322/online-education-kit-1983-first-disease-gene-mapped>
- NIH. National Human Genome Research Institute. Allele. [online]. [cit. 2023-03-18]. Dostupné z: <https://www.genome.gov/genetics-glossary/Allele>
- NIH. National Human Genome Research Institute. Bioinformatics. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.genome.gov/genetics-glossary/Bioinformatics>
- NIH. National Human Genome Research Institute. Genomic data science. [online]. [cit. 2022-04-06]. Dostupné z: <https://www.genome.gov/about-genomics/fact-sheets/Genomic-Data-Science>
- NIH. National Human Genome Research Institute. The Human Genome Project. [online]. [cit. 2022-03-11]. Dostupné z: <https://www.genome.gov/human-genome-project>

- NIH. National Library of Medicine. GenBank Overview. [online]. [cit. 2022-04-05]. Dostupné z: <https://www.ncbi.nlm.nih.gov/genbank/>
- NIH. National Library of Medicine. Database of Genotypes and Phenotypes (dbGaP). 2018. [online]. [cit. 2023-03-10]. Dostupné z: <https://www.ncbi.nlm.nih.gov/gap/>
- NIH. National Library of Medicine. Genome Data Viewer. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.ncbi.nlm.nih.gov/genome/gdv/>
- NIH. National Library of Medicine. Medline Plus. Achondroplasia. [online]. [cit. 2023-03-19]. Dostupné z: <https://medlineplus.gov/genetics/condition/achondroplasia/>
- NIH. National Library of Medicine. Medline Plus. Prader-Willi Syndrome. [online]. [cit. 2023-03-19]. Dostupné z: <https://medlineplus.gov/genetics/condition/prader-willi-syndrome/>
- NIH. National Library of Medicine. Medlineplus Genetics. MedlinePlus.What is a gene? [online]. [cit. 2022-03-23]. Dostupné z: <https://medlineplus.gov/genetics/understanding/basics/gene/>
- NIH. National Library of Medicine. MedlinePlus. What are single nucleotide polymorphisms (snps)?: Medlineplus Genetics. [online]. [cit. 2022-03-29]. Dostupné z: <https://medlineplus.gov/genetics/understanding/genomicresearch/snp/>
- NIH. National Library of Medicine. MedlinePlus. What is genetic discrimination?: Medlineplus Genetics. MedlinePlus.[online]. [cit. 2022-03-26]. Dostupné z: <https://medlineplus.gov/genetics/understanding/testing/discrimination/>
- Open Source Initiative. [online]. [cit. 2023-03-19]. Dostupné z: <https://opensource.org>
- Oracle. Co jsou big data? [online]. [cit. 2022-03-25]. Dostupné z: <https://www.oracle.com/cz/big-data/what-is-big-data/#three>
- Pinset Masons, Out-Law News. Drugs companies collaborate to unlock potential of 100,000 Genomes project. 2015. [online]. [cit. 2022-03-26] Dostupné z: <https://www.pinsentmasons.com/out-law/news/drugs-companies-collaborate-to-unlock-potential-of-100000-genomes-project>
- Přírodovědecká fakulta Univerzity Karlovy v Praze. TEPLÁ, M. Struktura DNA a RNA. [online]. [cit. 2022-04-06]. Dostupné z: <http://www.studiumbiochemie.cz/na.html>
- Proclinical. How genome sequencing is aiding drug research and development. [online]. [cit 2022-03-26]. Dostupné z: <https://www.proclinical.com/blogs/2015-7/genome-sequencing-is-aiding-drug-research-and-development>
- Stanford Medicine. Fastest DNA sequencing technique helps undiagnosed patients find answers in mere hours. News Center. [online]. [cit. 2022-03-25]. Dostupné z: <https://med.stanford.edu/news/all-news/2022/01/dna-sequencing-technique.html>
- Stanford University. Encode. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.encodeproject.org>
- Státní zdravotní ústav. SOUČEK, P. Genetika a zdravotní stav populace. [online]. [cit. 2023-03-23]. Dostupné z: <http://www.szu.cz/tema/genetika>
- The Apache Software Foundation. Licenses. [online]. [cit. 2023-03-19]. Dostupné z: <https://www.apache.org/licenses/>

- The App Solutions. CLOUD Service Models Explained: PaaS vs. SaaS vs. IaaS vs. DbaaS. [online]. [cit. 2023-03-18]. Dostupné z: <https://theappsolutions.com/blog/cloud/cloud-service-models/>
- The Free Software Foundation. [Online]. [cit. 2023-03-18]. Dostupné z: <https://www.fsf.org/>
- The Guardian. WRIGHT, P. Obituary: Erwin Chargaff. 2002. [online]. [cit. 2022-04-30]. Dostupné z: <https://www.theguardian.com/news/2002/jul/02/guardianobituaries.obituaries>
- The International Genome Sample Resource. 1000 Genomes Project Summary. 1000 Genomes. [online]. [cit. 2022-03-26]. Dostupný z: <https://www.internationalgenome.org/1000-genomes-summary>
- The nobel Prize. The Nobel Prize in Chemistry 2020. [online]. [cit. 2022-03-27]. Dostupné z: <https://www.nobelprize.org/prizes/chemistry/2020/summary/>
- Time. PARK, A. The Human Genome Is Finally Fully Sequenced. 2022. [online]. [cit. 2022-04-15]. Dostupné z: <https://time.com/6163452/human-genome-fully-sequenced/>
- UCSC Genomics Institute. UCSC Genome Browser. [online]. [cit. 2023-03-19]. Dostupné z: <https://genome.ucsc.edu>
- UCSC Genomics Institute. UCSC Licence. [online]. [cit. 2023-03-18]. Dostupné z: <https://genome.ucsc.edu/license/>
- Úřad pro ochranu osobních údajů. Návod k posouzení vlivu na ochranu osobních údajů u návrhů právních předpisů (DPIA). [online] [cit. 2022-03-19]. Dostupné z: <https://www.uoou.cz/navod-k-posouzeni-vlivu-na-ochranu-osobnich-udaju-u-navrhu-pravnich-predpisu-dpia/ds-5344>
- VFU Brno, Fakulta veterinární hygieny a ekologie, Ústav biologie a chorob volně žijících zvířat. [Sekvenování DNA, Enzymová Sangerova metoda. [online]. cit. 2022-03-24]. Dostupné z: https://cit.vfu.cz/opvk2011/?title=popis_metod-sekvenovani&lang=cz
- Yourgenome. How are sequenced genomes stored and shared? 2016. [online]. [cit. 2022-03-26]. Dostupné z: <https://www.yourgenome.org/facts/how-are-sequenced-genomes-stored-and-shared>

PRÁVNÍ PŘEDPISY

- Nařízení Evropského parlamentu a Rady (EU) 2018/1725 ze dne 23. 10. 2018 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů orgány, institucemi a jinými subjekty Unie a o volném pohybu těchto údajů a o zrušení nařízení (ES) č. 45/2001 a rozhodnutí č. 1247/2002/ES
- Nařízení Evropského parlamentu a Rady (EU) č. 2016/679 ze dne 27. 4. 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (GDPR)
- Návrh nařízení Evropského parlamentu a Rady (EU) o evropském prostoru pro zdravotní data ze dne 3. 5. 2022 COM(2022) 197 final (Návrh o EHDS)
- Návrh nařízení Evropského parlamentu a Rady o evropské správě dat ze dne 25. 11. 2020 COM(2020) 767 final (Akt o správě dat)
- Směrnice Evropského parlamentu a Rady (EU) 2009/24/ES ze dne 23. 4. 2009 o právní ochraně počítačových programů (Směrnice 2009/24/ES)

Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu a o změně směrnic 96/9/ES a 2001/29/ES (Směrnice 2019/790)

Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu a o změně směrnic 96/9/ES a 2001/29/ES (Směrnice 2019/790)

Směrnice Evropského parlamentu a Rady (EU) 96/9/ES ze dne 11. 3. 1996 o právní ochraně databází (Směrnice 96/9/ES)

Ústavní zákon č. 2/1993 Sb. ze dne 16. 12. 1992, ve znění ústavního zákona č. 162/1998 Sb., o vyhlášení Listiny základních práv a svobod (LZPS)

Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších předpisů (AutZ)

Zákon č. 89/2012 Sb., občanský zákoník ve znění pozdějších předpisů (OZ)

MEZINÁRODNÍ SMLOUVY

Listina základních práv Evropské unie (LZPEU)

Evropská úmluva o ochraně lidských práv ve znění Protokolů č. 11, 14 a 15 (EÚLP)

Smlouva Světové organizace duševního vlastnictví (WIPO) o právu autorském z roku 1996

Úmluva na ochranu lidských práv a důstojnosti lidské bytosti v souvislosti s aplikací biologie a medicíny: Úmluva o lidských právech a biomedicíně (1997), Rada Evropy (Úmluva o biomedicíně)

JUDIKATURA

Rozsudek Soudního dvora Evropské unie ze dne 9. 11. 2004 ve věci Fixtures Marketing Ltd v. Oy Veikkaus Ab., C-46/02

Rozsudek Soudního dvora Evropské unie ze dne 9. 11. 2004 ve věci The British Horseracing Board Ltd a další, C-203/02

Rozsudek Soudního dvora Evropské unie ze dne 3. 7. 2012 ve věci UsedSoft GmbH v. Oracle International Corp, C-128/11

Rozsudek Soudního dvora Evropské unie ze dne 19. 10. 2016 ve věci C-582/14, Patrick Breyer v. Bundesrepublik Deutschland

OSTATNÍ

Article 29 Data Protection Working Party. (2014) Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN WP216. [online]. [cit. 2023-03-19]. Dostupné z: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Assessment of the EU Member States' rules on health data in the light of GDPR, Specific Contract No SC 2019 70 02 in the context of the Single Framework Contract Chafea/2018/Health/03.

- [online]. European Commission. [2022-03-26]. Dostupné z:
https://ec.europa.eu/health/system/files/2021-02/ms_rules_health-data_en_0.pdf
- ČSN EN ISO 20387 (010107) Biotechnologie – Biobanky – Obecné požadavky na biobanky.
Česká verze evropské normy EN ISO 20387:2020
- European Commission. Towards European Health Data Space. [online]. [cit. 2022-03-26].
Dostupné z: https://www.ema.europa.eu/en/documents/presentation/presentation-towards-european-health-data-space-mdorazil-dg-sante_en.pdf
- European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. [online]. 2020, s. 12, [cit. 2022-03-28]. Dostupné z:
https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf
- European Union Agency for Fundamental Rights and Council of Europe, 2018, Handbook on European data protection law. [online]. [cit. 2022-04-20]. Dostupné z:
https://www.echr.coe.int/documents/handbook_data_protection_eng.pdf
- GDPR v kostce, Praktický průvodce povinnostmi pro podniky a spolky. C. H. BECK, 1. vydání, 2018. ISBN: 978-80-7400-704-0
- MAZURA, I. Genetický polymorfismus jako nástroj identifikace osob v kriminalistické a soudně-lékařské praxi. [online]. [cit. 2022-03-26]. Dostupné z:
<https://soudni.lf1.cuni.cz/file/5717/Forezn%C3%AD%20genetika.pdf>
- Oracle. MySQL 8.0 Reference Manual. [online]. [cit. 2023-03-19]. Dostupné z:
https://docs.oracle.com/cd/E17952_01/mysql-8.0-en/
- Pokyny pro posouzení vlivu na ochranu údajů a stanovení, zda „je pravděpodobné, že zpracování údajů bude mít za následek vysoké riziko“ pro účely nařízení 2016/679, přijaté dne 4. 4. 2017 a dne 4. 10. 2017 v aktualizovaném znění, 17/CS, WP 248 rev.01 (Pokyny pro posouzení vlivu na ochranu údajů). [online]. [cit. 2023-03-26]. Dostupné z:
https://www.uoou.cz/assets/File.ashx?id_org=200144&id_dokumenty=31892
- Pracovní skupina 29. Working Document on Genetic Data ze dne 17. 5. 2004, 12178/03/EN WP 91. [online] [cit. 2022-03-28]. Dostupné z: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2004/wp91_en.pdf (Working Document)
- Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů o vzácných onemocněních: Výzvy pro Evropu. Brusel dne 11. 11. 2008. [online]. [cit. 2023-03-19]. Dostupné z: <https://eur-lex.europa.eu/legal-content/CS/TXT/HTML/?uri=CELEX:52008DC0679&from=CS>
- UNITED STATES, PODESTA, J. Big data: seizing opportunities, preserving values. 2014. [online]. [cit. 2022-04-05]. Dostupné z: <http://purl.fdlp.gov/GPO/gpo64868>
- Ústav molekulární genetiky Akademie věd ČR, v. v. i., Akademie věd České republiky. STANĚK, D. RNA – temná hmota v našich buňkách. [online]. [cit. 2022-03-11]. Dostupné z:
https://www.img.cas.cz/files/2012/10/RNA-temna_hmota_v_nasich_bukach.pdf

Genetické databáze a ochrana genetických údajů

Abstrakt

Tato diplomová práce se zabývá problematikou zpracování genetických údajů, které jsou získávány metodou sekvenování DNA. Získaná data o lidském genomu představují revoluční podklad pro vědecké výzkumy a otevírají cestu pro léčbu některých doposud nevyléčitelných nemocí jako např. rakovina, AIDS či Alzheimerova choroba.

Práce se zabývá především právní úpravou ochrany genetických údajů. Popisuje zpracování genetických údajů, včetně jejich využívání a šíření, a to zejména prostřednictvím pořizovatelů genetických databází a genome browserů pro účely vědeckého výzkumu. Z pohledu platné právní úpravy České republiky a Evropské unie práce analyzuje dosavadní stupeň ochrany genetických dat a v tomto kontextu poukazuje na zjevné nedostatky. Práce se však neomezuje pouze na popis právních předpisů, ale danou problematiku důkladně hodnotí *de lege ferenda*.

Cílem této diplomové práce je zhodnotit aktuální praxi zpracování genetických údajů, zejména jejich zpřístupňování, ukládání, možností anonymizace či způsobů jejich zneužití. Práce se věnuje otázce uplatnění a důsledkům současné právní regulace stanovené především Nařízením Evropského parlamentu a Rady (EU) č. 2016/679 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů (GDPR), nevynechává však ani zhodnocení možných dopadů navrhované budoucí právní úpravy. Zvláštní pozornost je přitom věnována rizikům v souvislosti se zneužitím genetických údajů zejména v kontextu ochrany osobnosti a možné reidentifikace subjektu údajů.

Poslední část práce je věnována zpracování zdravotních dat v souvislosti s návrhem Nařízení Evropského parlamentu a Rady (EU) o evropském prostoru pro zdravotní data, European Health Data Space a sekundárnímu využití těchto dat pro výzkumné účely.

Klíčová slova: genetická databáze, genetické údaje, ochrana osobních údajů

Genetic Databases and Protection of Genetic Data

Abstract

This thesis deals with the processing of genetic data obtained by DNA sequencing. The data obtained from the human genome represents a revolutionary basis for scientific research and open the way for the treatment of some so far incurable diseases such as cancer, AIDS, and Alzheimer's disease.

The thesis primarily focuses on the legal regulation of the protection of genetic data. The thesis describes the processing of genetic data, including its use and dissemination, mainly through providers of genetic databases and genome browsers for scientific research purposes. From the perspective of the current legal framework in the Czech Republic and the European Union, the thesis analyzes the existing level of protection of genetic data and points out evident shortcomings. However, the thesis is not limited to a description of legal regulations but thoroughly evaluates the issue *de lege ferenda*.

The aim of this thesis is to evaluate the current practice of processing genetic data, particularly its accessibility, storage, the possibility of its anonymization, and ways of its misuse. The thesis focuses on the issue of the implementation and consequences of the current legal regulation mainly established by the Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons regarding the processing of personal data (GDPR), but also assesses the possible impacts of the proposed future legal regulation. Special attention is given to the risks associated with the misuse of genetic data, especially in the context of the protection of personality and possible re-identification of the data subjects.

The last part of the thesis is devoted to the processing of healthcare data in connection with the Proposal of the European Parliament and Council Regulation (EU) on the European Health Data Space and the secondary use of the data for research purposes.

Key words: genetic database, genetic data, personal data protection