Persónugreining í gagnagrunni á heilbrigðissviði

Einar Árnason

Ágrip

Inngangur: Spurningin um persónugreinanleika er grundvallaratriði í allri umræðu um frumvarpið og lögin um gagnagrunn á heilbrigðissviði. Ef gögnin eru persónugreinanleg gilda þjóðréttarlegar skuldbindingar um að afla skuli fyrirfram samþykkis sjúklinga fyrir notkun heilsufarsupplýsinga í öðrum tilgangi en þeirra var aflað. Lögin ganga út frá því að dulkóðun í eina átt geri gögnin ópersónugreinanleg og því sé ekki þörf að afla fyrirfram samþykkis sjúklinga.

Niðurstöður: Með því að rekja sögu hugtaksins um persónugreinanleika í umræðunni sést að breytingar voru gerðar á skilgreiningum um persónugreinanleika. Fyrst var miðað við tilmæli ráðherranefndar Evrópuráðsins en síðar var tekin upp orðrétt skilgreining úr tilskipun Evrópusambandsins sem nú er þjóðréttarlega skuldbindandi fyrir Ísland. Breytingin var gerð til að bregðast við umsögn tölvunefndar sem kollvarpaði hugmyndafræðinni sem lagt hafði verið upp með varðandi persónugreiningu. Upplýsingar eru persónuuplýsingar ef til er lykill og engu máli skiptir hver gætir lykilsins. Dulkóðun í eina átt var þá sett fram sem aðferð til að gera að engu tilvist lykils. Þrátt fyrir það viðurkenna talsmenn gagnagrunnsins að til sé lykill.

Greint er frá því hvernig hægt er að smíða lykla að grunninum. Þar sem gagnagrunnurinn er langsum (longitudinal) og langtímasöfnun og -samtenging upplýsinga um hvern einstakling hlýtur dulkóðunaraðferðin að vera stöðug í tíma. Hver sá sem hefur aðferðina í höndunum getur fyrirhafnalítið búið til uppflettitöflu yfir nöfn eða kennitölur og fastanúmerin

Fyrirspurnir, bréfaskipti: Einar Árnason líffræðingur, Líffræðistofnun Háskólans, Grensásvegi 12, 108 Reykjavík. Sími: 525 4613; netfang: einar@lif.hi.is

Lykilorð: *persónugreinanleiki, gagnagrunnur á heilbrigðissviði, greiningarlykill, ættartré, samhengi upplýsinga.*

ENGLISH SUMMARY

Árnason E

Personal identifiability in the Health Sector Database

Læknablaðið 2001; 87: 807-16

Introduction: Personal identifiability is a fundamental question in the debate about the Bill and Act on the Health Sector Database (HSD). If the data are personally identifiable, Iceland's international committments dictate that a priori consent be obtained from patients for the use of their health records data. The HSD Act presumes that one way encryption renders the data non-personally identifiable and that therefore an a priori consent is not required.
Results: The history of the concept of personal identifiability during the debate on the HSD reveals changes made to the concept. In the first instance a reference was made to Recommendation R(97)5 of the Council of Europe Committee of Ministers which was changed by adopting a direct translation of the definition of personal data from the Directive 95/46/EC of the European Parliament and of the Council. These changes were made in response to the Data Protection Commission's opinion on the HSD Bill submitted to the Minister of Health that overturned the ideology previously used regarding indentifiability of persons. Information is identifiable if there exists a key and it makes no difference who holds the key. One way encryption was then adopted as a method that was supposed to mean that a key does not exist. Nevertheless, the database proponents now admit that a key exists.
The making of keys for opening up the database is discussed. The database is a longitudinal collection and linkage of records on each individual and therefore the method of encryption must remain stable. Therefore, anyone with access to the method can easily make a lookup-table containing side by side the names and the personal numbers produced by the encryption. Although it may be hard to go from a personal number directly back to a name, given the table it always is possible to look up what personal number belongs to a certain person or what person stands behind a certain personal number. This is a key. If the method of encryption was lost or access to it was not available it would nevertheless be possible to make a key. The intention is to encrypt the genealogy of the entire nation using the same encryption method used for the HSD. The genealogy of the nation with names is also generally available. The patterns of family trees become unique when one family is connected to another through marriage and childbirth. A comparison of the encrypted genealogy containing personal numbers with the same genealogy containing names is therefore a method for making a key.
Finally a key can be made from the context of general information. Even if the names were irreversibly removed there will be enough available bits of general information connected to a personal number to allow re-identification of the person in a large number of instances. This amounts to making a key.
Conclusions: The information in the Health Sector Database is personal information. Therefore it is both right and reasonable to obtain an a priori consent of patients for the transfer of their health data to the database as Iceland's international obligations stipulate. Anything less is unreasonable.

Key words: *personal identification, health sector database, keys, genealogy, context.*

Correspondence: Einar Árnason. E-mail: einar@lif.hi.is

sem dulkóðunin gefur. Þótt ekki sé raunhæft að finna beint af dulkóða hverjum hann tilheyrir er ætíð hægt að fletta upp í töflunni hvaða fastanúmer tilheyrir ákveðnum manni eða hvaða maður stendur að baki ákveðnu fastanúmeri. Þetta er lykill.

Ef dulkóðunaraðferðin færi forgörðum eða ekki væri aðgangur að henni væri samt hægt að búa til lykil að grunninum. Áætlað er að dulkóða ættartré þjóðarinnar með sömu aðferð og notuð er fyrir heilsufarsgrunninn. Ættartré þjóðarinnar með nöfnum er einnig þekkt sem almenn þekking í landinu. Munstur ættartrjáa verða einstök þegar ein ætt tengist við aðra með giftingum og barneignum. Samanburður á ættarmunstrum milli ættargrunns með fastanúmerum og sama ættargrunns með nöfnum er því leið til að smíða lykil.

Þá er hægt að smíða lykil af samhengi almennra upplýsinga. Jafnvel þótt búið sé að aftengja nöfn og kennitölur fylgja fastanúmeri nægar almennar upplýsingar til að unnt sé að endurþekkja einstaklinginn í velflestum tilfellum. Það jafngildir lykli.

Ályktanir: Upplýsingar í gagnagrunni á heilbrigðissviði teljast vera persónuupplýsingar. Því er réttmætt og sanngjarnt að afla verði fyrirfram samþykkis sjúklinga fyrir flutningi heilsufarsupplýsinga þeirra í grunninn eins og þjóðréttarlegar skuldbindingar landsins kveða á um. Allt annað er ósanngjarnt.

Inngangur

Lögin um gagnagrunn á heilbrigðissviði og fyrirætlun um smíði hans valda enn deilum. Málaferli eru hafin til þess að kanna hvort lögin standist stjórnarskrá og þjóðréttarlegar skuldbindingar. Gera má ráð fyrir að þær deilur haldi áfram þar til lögunum verður breytt eða þau felld úr gildi. Ástæðan er sú að ekki fær staðist sú grundvallarforsenda laganna, að um sé að ræða gögn *ópersónugreinanlegra* einstaklinga. Upplýsingarnar eru persónugreinanlegar og því gilda þjóðréttarlegar skuldbindingar Íslands um að afla fyrirfram samþykkis sjúklinganna fyrir notkun þeirra í öðrum tilgangi en þeirra var aflað. Í flestum tilfellum veitir læknir upplýsingunum móttöku eða aflar þeirra í þágu sjúklingins undir siða- og lagaskyldu um trúnað sem sjúklingurinn einn getur aflétt.

Hér er rakin saga hugtaksins um persónugreiningu í gagnagrunnsmálinu og leidd að því rök að gölluð röksemdafærsla um persónugreiningu og lykil búi að baki löggjöfinni. Þá er velt upp þeirri spurningu hvað sé lykill og greint frá leiðum til að smíða lykla til að ljúka upp grunninum með uppflettitöflu, samanburði ættartrjáa, eða af samhengi upplýsinga.

1. Saga hugtaksins um persónugreiningu í umræðu um gagnagrunn

Með því að rekja nokkur helstu atriði úr sögu hugtaksins um persónugreiningu úr umræðunni um gagnagrunninn er leitt fram i) að hugmyndafræðin sem upphaflega var gengið út frá byggðist á röngum forsendum, ii) að tölvunefnd kollvarpaði hugmyndafræðinni og iii) að þrátt fyrir það virðist sú hugmyndafræði enn vera við lýði. Hér er meðal annars spurt hvort lykill sé til og í hvers höndum hann sé og hvaða merkingu skuli leggja í hugtakið *persónugreiningu*.

1.1. Skilgreiningar laganna

Í gagnagrunnslögunum eru meðal annars eftirfarandi skilgreiningar, í 3. grein (1):

„2. *Persónuupplýsingar:* Allar upplýsingar um persónugreindan eða persónugreinanlegan einstakling. Maður telst persónugreinanlegur ef unnt er að persónugreina hann, beint eða óbeint, svo sem með tilvísun í kennitölu eða einn eða fleiri þætti sem sérkenna hann í líkamlegu, lífeðlisfræðilegu, andlegu, efnalegu, menningarlegu eða félagslegu tilliti.

3. *Ópersónugreinanlegar upplýsingar:* Upplýsingar um einstakling sem ekki er persónugreinanlegur samkvæmt skilgreiningu 2. tölul.

4. *Dulkóðun:* Umbreyting orða eða talna í óskiljanlega runu af táknum.

5. *Dulkóðun í eina átt:* Umbreyting orða eða talna í óskiljanlega runu af táknum sem ekki er hægt að rekja til baka með greiningarlykli."

Persónugreinanleiki er samkvæmt þessu mjög víðtækur og ópersónugreinanleiki, andhverfa hans, að sama skapi takmarkaður. *Dulkóðun í eina átt* er skilgreind sem aðgerð sem á að gera að engu þann möguleika að persónugreina einstakling með greiningarlykli. Af skilgreiningunum er ljóst að ein og sér *dulkóðun* er ekki nægjanleg. Lykilatriði virðist vera að það sé *í eina átt*. Einstefna er samkvæmt þessu eitthvert tækniundur sem á að gera að engu tilvist lykils.

Í umræðunni var því einnig haldið fram að *ef* „sú tæknilega forsenda frumvarpsins að „dulkóðun í eina átt" feli í sér raunverulega og endanlega aftengingu persónuauðkenna" þá standist frumvarpið og lögin kröfur þjóðarréttar (2). Þó er viðurkenndur sá möguleiki að ekki verði talinn sá eðlismunur á dulkóðun með greiningarlykli og „dulkóðun í eina átt" að síðarnefnda tilvikið verði talið fela í sér endanlega og afdráttarlausa aftengingu persónuauðkenna (2).

1.2. Frumdrög að frumvarpi í júlí 1997

Kári Stefánsson lét Lögmenn á Skólavörðustíg 12 gera *Frumdrög að frumvarpi til laga um gagnagrunna á heilbrigðissviði*, dagsett 14. júlí 1997, sem hann sendi síðan heilbrigðisráðuneytinu 3. september sama ár (3). Ætlun höfunda var að frumvarpið yrði að lögum haustið 1997 og lögin tækju gildi 1. janúar 1998. Í 2. grein draganna er þessi skilgreining:

„*3. Persónuupplýsingar:* Upplýsingar er varða einkamálefni, heilsuhagi, fjárhagsmálefni eða önnur málefni nafngreinds eða nafngreinanlegs

einstaklings sem sanngjarnt er og eðlilegt að leynt fari. Einstaklingur skal eigi teljast nafngreinanlegur ef verja þyrfti verulegum tíma og mannafla til að nafngreining hans gæti átt sér stað. Þegar einstaklingur er ekki nafngreinanlegur skal litið svo á að upplýsingarnar séu ekki persónuupplýsingar."

1.3. Frumvarp og drög vorið og sumarið 1998

Þegar frumvarp um gagnagrunna (4) var lagt fram á 122. löggjafarþingi vorið 1998 var notast við þessa skilgreiningu væntanlegs rekstrarleyfishafa. Þó með þeirri viðbót að jafnvel þótt til sé lykill skuli einstaklingur ekki talinn persónugreinanlegur ef sá aðili sem hefur upplýsingar undir höndum hefur ekki aðgang að lyklinum:

„*4. Persónuupplýsingar:* Upplýsingar er varða einkamálefni, þar með talda heilsuhagi, fjárhagsmálefni eða önnur málefni persónugreinds eða persónugreinanlegs einstaklings sem sanngjarnt er og eðlilegt að leynt fari. Einstaklingur skal eigi teljast persónugreinanlegur ef verja þyrfti verulegum tíma og mannafla til að persónugreining hans gæti átt sér stað. Sama gildir ef persónugreining getur einungis átt sér stað með notkun greiningarlykils sem sá aðili er hefur upplýsingar undir höndum hefur ekki aðgang að. Þegar einstaklingur er ekki persónugreinanlegur skal litið svo á að upplýsingar sem hann varða séu ekki persónuupplýsingar í skilningi laga þessara."

Sú hugmyndafræði að einstaklingur teljist ekki persónugreinanlegur ef „*verja þyrfti verulegum tíma og mannafla til að persónugreining hans gæti átt sér stað*" er tekin úr tilmælum ráðherranefndar Evrópuráðsins nr. R(97)5 frá 13. febrúar 1997 (5) um verndun heilsufarsupplýsinga (Recommendation No R(97)5 of the Committee of Ministers to Member States on the Protection of Medical Data). Jafnframt segja frumvarpshöfundar varðandi notkun lykils að byggt sé á „vinnuferli sem tölvunefnd, sem starfar samkvæmt lögum um skráningu og meðferð persónuupplýsinga, nr. 121/1989, hefur nýlega mótað reglur um á sviði vísindarannsókna á heilbrigðissviði, en í skilmálum tölvunefndar er kveðið á um að rannsóknargögn skuli kóðuð með dulmálslykli áður en þau eru afhent rannsóknaraðila og varðveiti sérstakir tilsjónarmenn tölvunefndar síðan dulmálslykilinn." Þar með er ýjað að því að skilgreining frumvarpsins um persónuupplýsingar sé í samræmi við tölvulögin (nr. 121/1989; (6)) og jafnframt að ákvæði frumvarpsins varðandi notkun lykils fari eftir skilmálum tölvunefndar.

Frumvarpið um gagnagrunna var dregið til baka, endurskoðað af starfshópi heilbrigðisráðuneytisins og sent til umsagnar, meðal annars til vísindasiðanefndar og tölvunefndar í júlí 1998. Þau frumvarpsdrög innnihéldu sömu skilgreiningu um „verulegan tíma og mannafla" og var í frumvarpinu sem lá fyrir 122. þinginu (4) sem byggði á tilmælum Evrópuráðsins nr. R(97)5. Tölvunefnd hafði ýmislegt við þetta að athuga.

1.4. Umsögn tölvunefndar í september 1998

Tölvunefnd kollvarpaði allri hugmyndafræði frumvarpsins um persónugreiningu og þeirri aðferðafræði að miða við tilmæli Evrópuráðsins með umsögn um drög að frumvarpinu (dagsett 31. júlí 1998) til heilbrigðis- og tryggingaráðherra dagsett 4. september 1998 (7). Bréf tölvunefndar er undirritað af Þorgeiri Örlygssyni formanni nefndarinnar, en á þeim tíma var hann helsti sérfræðingur landsins um persónugreiningu og persónuvernd og ætla má að hann hafi verið helsti höfundur umsagnarinnar.

Tölvunefnd tekur fram í upphafi máls síns að „á árinu 1995 var samþykkt tilskipum ESB (8, innskot EÁ) um vernd einstaklinga að því er varðar meðferð persónuupplýsinga og frjálsan flutning slíkra upplýsinga, (þ.e. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data).". Jafnframt segir tölvunefnd að umrædd tilskipun verði felld undir EES samninginn, en „af því leiðir, að efnisákvæði tilskipunar ESB þarf að leiða í lög hér á landi.". „Í því felst, að almenn löggjöf um meðferð persónuupplýsinga þarf að vera í samræmi við efnisákvæði tilskipunarinnar, og gildir hið sama einnig um sérlöggjöf á þessu réttarsviði."

Tölvunefnd tekur fram að í frumvarpinu „virðist með öllu litið fram hjá" ofangreindri tilskipun ESB 95/46/EC sem segir „að upplýsingar um einstaklinga eru persónuupplýsingar, ef til er greiningarlykill að dulkóðuðum upplýsingum. … Gerir tilskipunin engan greinarmun eftir því, hvort verja þurfi verulegum tíma og mannafla til þess að persónugreining geti átt sér stað." Hugtakið um *verulegan tíma og mannafla* er reyndar hvergi að finna í tilskipun ESB heldur er það komið úr tilmælum ráðherranefndar Evrópuráðsins. Af þessu ætti að vera ljóst að það er tilskipun ESB en ekki tilmæli ráðherranefndar Evrópuráðsins sem verður að leggja til grundvallar lagasetningunni.

Þá segir tölvunefnd að „hæpið sé að halda því fram" að frumvarpið taki mið af tölvulögunum (nr. 121/1989) (6) því samkvæmt þeim lögum eru upplýsingar „jafnan persónuupplýsingar, ef til er greiningarlykill að dulkóðuðum upplýsingum. ... Tölvunefnd hefur jafnan byggt á því ... að dulkóðaðar upplýsingar um einkamálefni einstaklinga séu persónuupplýsingar í skilningi laganna, og nefndin telur í því sambandi engu máli skipta, hvort sá aðili, sem upplýsingarnar hefur undir höndum, hefur aðgang að greiningarlyklinum eða ekki."

Að lokum segir tölvunefnd: „þýðingamikið er, að skilgreining frumvarpsins á hugtakinu persónuupplýsingar orki ekki tvímælis.". Leggur tölvunefnd

áherslu á „að bæði ákvæði almennrar löggjafar um skráningarmálefni hér á landi (nú lög nr. 121/1989) og ákvæði sérlöggjafar um skráningarmálefni (til dæmis fyrirhuguð löggjöf um gagnagrunn á heilbrigðissviði) fullnægi skilyrðum og skilgreiningum tilskipunar ESB, eftir að hún er orðin þjóðréttarlega skuldbindandi fyrir Íslands hönd.“.

1.5. Viðbrögð og ný umsögn tölvunefndar

Þessi skýra afstaða tölvunefndar kollvarpaði í raun skilgreiningum frumvarpsins um persónugreiningu. Hér töluðu helstu sérfræðingar ríkisins um persónugreiningu og persónuvernd. Viðbrögð frumvarpshöfunda voru að nema á brott ákvæði sem byggðu á tilmælum ráðherranefndar Evrópuráðsins og ákvæði um lykil þótt hann væri í vörslu annars en rannsakenda. Þess í stað var tekin upp orðrétt þýðing úr tilskipun ESB (95/46/EC) sem tölvunefnd sagði þjóðréttarlega bindandi. Hún hljóðar svo á ensku:

> „For the purposes of this Directive
> (a) personal data shall mean any information relating to an identified or identifiable natural person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;“

Þegar gagnagrunnsfrumvarpið var lagt fram aftur á 123. löggjafarþinginu í október 1998 tók skilgreining persónuupplýsinga mið af tilskipuninni (8,9):

> *„Persónuupplýsingar:* Allar upplýsingar um persónugreindan eða persónugreinanlegan einstakling. Maður telst persónugreinanlegur ef unnt er að persónugreina hann, beint eða óbeint, svo sem með tilvísun í kennitölu eða einn eða fleiri þætti sem sérkenna hann í líkamlegu, lífeðlisfræðilegu, andlegu, efnalegu, menningarlegu eða félagslegu tilliti.“

Hér er *an identification number* þýtt sem *kennitala* sem er ónákvæmt því texti tilskipunarinnar nær ekki einvörðungu yfir það sem á Íslandi er kallað kennitala heldur hverskyns kennitákn eða fastanúmer einstaklinga (identification number eða personal number).

Enn á ný gerði tölvunefnd athugasemdir með umsögn til heilbrigðis- og trygginganefndar Alþingis, dagsett 26. október 1998 (10):

> „Í tilskipun Evrópusambandsins er hugtakið persónuupplýsingar víðfemt og tekur til allra upplýsinga, álita og umsagna sem beint eða óbeint má tengja tilteknum einstaklingi, þ.e. allra upplýsinga sem eru persónugreindar eða persónugreinanlegar. Af a-lið 2. gr. tilskipunarinnar leiðir að upplýsingar teljast persónugreinanlegar ef unnt er að persónugreina þær á grundvelli einhvers auðkennis, beint eða óbeint, með tilvísun í kennitölu eða annað auðkenni, með eða án greiningarlykils. Í 26. gr. formála tilskipunarinnar segir að meginreglur um vernd skuli gilda um allar persónugreindar eða persónugreinanlegar upplýsingar og að til þess að ákveða hvort upplýsingar séu persónugreinanlegar (rekjanlegar) skuli tekið mið af öllum aðferðum sem megi hugsa sér að ábyrgðaraðili eða annar aðili geti beitt til að bera kennsl á viðkomandi einstakling. Af því leiðir og að meginreglur um vernd skuli ekki gilda um upplýsingar sem hafa með öllu verið aftengdar einstaklingum og útilokað gert að rekja þær til einstakra manna.
>
> Í aðalatriðum eru til tvær leiðir til að tryggja persónuvernd í slíkum gagnagrunni. Annars vegar sú að „aftengja“ persónuupplýsingar persónuauðkennum og hins vegar sú að „dulkóða“ upplýsingarnar eins og það er gjarnan nefnt. Gagnagrunnsfrumvarpið miðar við að upplýsingar um einstaka menn verði dulkóðaðar fyrir flutning í gagnagrunninn. Er gert ráð fyrir því að upplýsingar í grunninum verði uppfærðar reglulega þegar nýjar upplýsingar bætast við. Til þess er nauðsynlegt að greina megi hvar eldri upplýsingar um sama mann sé að finna og því verða upplýsingar í grunninum ekki aftengdar heldur aðeins dulkóðaðar. Munur þessara tveggja aðferða, dulkóðunar og aftengingar, felst í aðalatriðum í því að þegar persónuupplýsingar eru dulkóðaðar fær viðkomandi einstaklingur nýtt og tilbúið skráningar- eða persónuauðkenni, en til er greiningalykill sem gerir það kleift að persónugreina upplýsingarnar. Þegar upplýsingar eru hins vegar aftengdar persónuauðkennum fær viðkomandi einstaklingur sem fyrr tilbúið skráningar- eða persónuauðkenni, en að því auðkenni er hins vegar enginn greiningarlykill. Í því tilviki teljast upplýsingar vera ópersónugreinanlegar, nema þær megi persónugreina með öðrum hætti, s.s. með tilvísun í tiltekna þætti sem sérkenna hinn skráða í líkamlegu, lífeðlisfræðilegu, andlegu, efnalegu, menningarlegu eða félagslegu tilliti, sbr. a-lið 2. gr. tilskipunarinnar.
>
> Með hliðsjón af öllu framanrituðu telur Tölvunefnd ekki að fái staðist sú fullyrðing að í grunninum verði ópersónugreinanlegar heilsufarsupplýsingar. Er því lagt til að því orði verði sleppt úr ákvæði 1. gr.“

Ekki var farið að þeirri tillögu tölvunefndar að sleppa orðinu um ópersónugreinanleika. Skilgreining tilskipunarinnar um persónuupplýsingar verður síðan að lögum (eins og að framan greinir). Þar sem ekki gengur að hafa greiningarlykil hver sem geymir hann, eins og tölvunefnd benti á, var tekin upp *dulkóðun í eina átt*. Í framhaldinu er því haldið fram að þar sem ekki er unnt að rekja sig beint til baka eftir dulkóðun nafna eða kennitalna í eina átt þá sé þar með ekki til greiningarlykill. Dulkóðun í eina átt er þannig það lykilatriði frumvarpsins sem ætlað er að tryggja að ákvæði tilskipunarinnar sé virt.

Þetta stenst ekki. Dulkóðun í eina átt merkir ekki að ekki sé til lykill. Dulkóðun í eina átt merkir einungis að erfitt eða reiknifrekt er að rekja sig beint til baka frá fastanúmeri að kennitölu eða nafni. Með því að taka þessa skilgreiningu inn er verið að koma aftur með hugmyndina að „verulegan tíma og mannafla“ þurfi til. Þeirri hugmyndafræði var tölvunefnd þegar búin að hafna enda er hún ekki hluti að tilskipun ESB. Þvert á móti ber að taka mið af öllum þeim aðferðum sem með sanngirni má hugsa sér líklegt að ábyrgðaraðili eða hvaða annar aðili sem er kynni að beita til að bera kennsl á einstaklinginn (8).

Embættismenn ríkisins, sem falið er að framfylgja þessum lögum, halda því sumir fram að *í skilningi laganna sé um ópersónugreinanleg* gögn að ræða. Sumir gagnrýnendur hafa kallað þetta flatjarðarkenninguna: ef lagatexti fullyrðir að jörðin sé flöt þá er hún flöt *í skilningi laganna*. Þar sem gagnagrunnslögin segi að dulkóðun í eina átt merki umbreytingu persónuauðkenna í fastanúmer sem ekki sé hægt að rekja til baka með greiningarlykli að þá sé ekki til lykill í skilningi laganna.

1.6. Viðurkennt að lykill er enn til

Bæði Kári Stefánsson og heilbrigðishópur gagnagrunnsdeildar Íslenskrar erfðagreiningar hafa nýlega staðfest að lykill er til. Í viðtali við Kára Stefánsson í *New Scientist* 15. júlí 2000 (11) segir hann varðandi samtengingu erfðaupplýsinga við heilsufarsuplýsingar:

„Once we have identified a family with one of these diseases, what we will do is to go to those people and ask them to give us blood so that we can isolate DNA. ... When we do this, we will ask for their permission to cross-reference their names with the help of the health-care database. But in order to do this, we will have to get their explicit, signed consent.

(NS:) Does this mean that you can identify individuals from the database?

No. The information in the database will be encrypted and the keys will be kept by the Data Protection Commission of Iceland.“

Kári Stefánsson viðurkennir því að til séu lyklar og segir að þeir verði í vörslu Persónuverndar. Lyklarnir eru sagðir vera í vörslu annars aðila en þess sem hefur gögnin. Þeirri hugmyndafræði hafði tölvunefnd þegar hafnað þegar hún sagði það „engu máli skipta, hvort sá aðili, sem upplýsingarnar hefur undir höndum, hefur aðgang að greiningarlyklinum eða ekki.“.

Í grein í Morgunblaðinu 27. febrúar 2001 (12) segir heilbrigðishópur ÍE að heilsufarsupplýsingar verði gerðar ópersónugreinanlegar:

„Mjög háþróaðar tæknilausnir hafa verið hannaðar og verða notaðar til að þrídulkóða kennitölu einstaklinga í eina átt. Hver dulkóðun er gerð eftir sérstökum dulkóðunarlykli sem stenst afar strangar tæknilegar öryggiskröfur. Til að afkóða kennitöluna og persónugreina þannig heilsufarsupplýsingar sem einnig eru kóðaðar og dulkóðaðar þyrfti að nota alla þrjá lyklana í réttri röð. Svo að slíkt geti ekki gerst er gert ráð fyrir að dulkóðunarlyklarnir verði í höndum þriggja mismunandi aðila (heilbrigðisstofnananna sjálfra, Persónuverndar og ÍE). Þessi sjálfvirka þrefalda dulkóðun brenglar kennitölur þannig að mögulegt verður að uppfæra gögn einstaklinga þegar þau koma í MGH (Miðlægan Gagnagrunn á Heilbrigðissviði), án þess þó að þau verði nokkurn tíma persónugreinanleg eftir að þau eru afrituð úr sjúkraskýrslum. Þrátt fyrir slíkar öryggisráðstafanir sem fullyrða má að séu einstakar í sögu íslenskra vísindarannsókna draga jafnvel aðilar, sem kunnugir eru vísindarannsóknum af þessu tagi, í efa að gögnin verði í raun ópersónugreinanleg.“

Hér viðurkennir einnig starfshópur ÍE að til eru lyklar að upplýsingunum og að unnt er að persónugreina einstaklinga með því að beita lyklunum. Að segja að „afar strangar tæknilegar öryggiskröfur“ séu gerðar merkir væntanlega að „verulegan tíma og mannafla“ þyrfti til að brjóta kóðann. Vel má vera að svo sé en þeirri hugmyndafræði hafnaði tölvunefnd og hún kemur málinu ekki lengur við, enda er sú hugmyndafræði ekki hluti af tilskipun ESB.

1.7. Persónugreinanleg gögn

Eftir að hafa rakið þessa sögu er niðurstaða mín sú að vissulega sé til lykill eða lyklar sem hægt er að nota til að persónugreina einstakling. Engu máli skiptir, hvort sá aðili, sem upplýsingarnar hefur undir höndum, hefur aðgang að greiningarlyklinum eða ekki og hvort greiningarlykillinn er í einum hluta eða fleiri. Ekki er hægt að fallast á að dulkóðun í eina átt geri upplýsingar ópersónugreinanlegar sem ekki er hægt að rekja til baka með greiningarlykli. Sú forsenda lokaútgáfu frumvarpsins og gagnagrunnslaganna að ekki sé hægt að ljúka upp grunninum með lykli stenst því ekki. Enda staðfesta nú Kári Stefánsson og sérfræðingar ÍE það að til er lykill.

Sú forsenda frumvarpsins og laganna að ekki sé hægt að rekja sig til baka með lykli er því ekki rétt. Gögnin eru því persónugreinanleg samkvæmt gagnagrunnslögunum (1) og samkvæmt persónuverndarlögunum (13) sem bæði byggja á tilskipun ESB (95/46/EC). Ísland er nú þjóðréttarlega skuldbundið skilyrðum og skilgreiningum tilskipunar ESB. Ef Ísland ætlar að uppfylla þær skuldbindingar ber að afla fyrirfram samþykkis fyrir flutning gagna í gagnagrunn á heilbrigðissviði. Margföld dulkóðun í eina átt breytir engu um það.

Í lögunum (1) og í frumvarpinu og greinargerð með því (9) er því haldið fram, eins og fram kemur í þessum skilgreiningum, að *dulkóðun í eina átt* sé aðferð til að gera upplýsingar um persónugreinanlegan

einstakling ópersónugreinanlegar. Því er haldið fram að einstaklingur sé ópersónugreinanlegur vegna þess að ekki er til lykill sem ljúki upp upplýsingum um hver þessi einstaklingur er. Þá er því jafnframt haldið fram með skilgreiningunum að ekki sé hægt að persónugreina einstakling í grunninum með tilvísun í neina þá þætti sem *sérkenna hann í líkamlegu, lífeðlisfræðilegu, andlegu, efnalegu, menningarlegu eða félagslegu tilliti.*

Þetta stenst ekki. Það eru að minnsta kosti þrjár leiðir sem nota má til að smíða lykla að grunninum: i) að smíða lykil með uppflettitöflu, ii) að smíða lykil með samanburði á munstrum ættartrjáa og iii) að smíða lykil af samhengi upplýsinga.

2. Lykill smíðaður með uppflettitöflu

Að halda því fram að dulkóðun í eina átt merki að ekki sé hægt að rekja sig til baka með greiningarlykli gildir einungis í þröngum tæknilegum skilningi. Ef nafnið *Pétur Pálsson* (eða kennitala hans 010476-4878) er sett í gegnum dulkóðun til dæmis með *hakkafalli* (hash function) í eina átt (14,15) kæmi til dæmis út *fastanúmerið* 012578f77e5820f2c5bdfcd48ec273ce. Og ef nafnið *Pálína Pétursdóttir* (eða 020587-5988) er sett í gegnum sömu einstefnu dulkóðun kæmi út fastanúmerið 5ce18b1caca938a8b881613445377 23f. Ef við hefðum einungis í höndum fastanúmerin eða táknin 012578f77e5820f2c5bdfcd48ec273ce eða 5ce18b1caca938a8b881613445377 23f væri afar erfitt að ráða beint af táknunum að annað þeirra táknaði Pétur Pálsson og hitt Pálínu Pétursdóttir. Ef þetta væri allt og sumt væru einstaklingar ópersónugreinanlegir, þar sem ekki er unnt að fara beint til baka frá fastanúmeri að einstaklingi. Einstaklingar eru þó ópersónugreinanlegir einungis í þessum þrönga skilningi, *að fara beint til baka.*

Í rekstri gagnagrunnsins verður hins vegar til lykill sem gerir þetta að engu. Gagnagrunnurinn er langsum (longitudinal) og langtímasöfnun og -samtenging upplýsinga um sérhvern einstakling (16). Gagnagrunnurinn verður uppfærður reglulega og þegar nýjar upplýsingar verða til um einhvern einstakling (til dæmis vegna heimsóknar til læknis) verða þær upplýsingar að rata á réttan stað í gagnagrunninum og tengjast við aðrar upplýsingar um þennan sama einstakling. Sama gildir um uppfærslu ættfræði- og erfðafræðigagnagrunna. *Eitthvað* eða *einhver* hlýtur því að „*vita*“ hver einstaklingurinn er og hvar hann er í grunninum eða grunnunum þremur sem tengja má saman. Þetta *eitthvað* eða *einhver* er *aðferðin* sem notuð er við dulkóðunina. Þar sem um er að ræða endurtekna uppfærslu grunnsins verður aðferðin ætíð að vera sú sama, hún verður að vera stöðug í tíma. Aðferðin er því *lykill* þar sem ætíð er fyrirhafnarlítið hægt að búa til *uppflettitöflu* sem tengir nöfn eða kennitölur einstaklinga við dulkóðað fastanúmer og öfugt.

2.1. Dulkóðun, einkvæm vörpun nafna

Dulkóðun er ekkert annað en einkvæm vörpun nafna eða kennitalna yfir á annað form. Með einstefnudulkóðuninni fær einstaklingurinn nýtt og tilbúið skráningar- eða persónuauðkenni í stað kennitölu, svokallað fastanúmer. Í ýmsum gögnum um gagnagrunninn (15,17) er rætt um *hakkafall* sem aðferð fyrir slíka vörpun í eina átt. *Hakkafall H* er *aðferð* til að umbreyta *inntaki, m,* í úttaksstreng tákna sem hefur fasta lengd, það er að breyta því í hakkagildið *h* eða í táknum $H(m) \rightarrow h$.

Í dulkóðunarfræðum eru gerðar þær meginkröfur til hakkafalls að i) inntakið megi hafa hvaða lengd sem er, ii) úttakið hafi fasta lengd, iii) auðvelt sé að reikna $H(x)$ fyrir eitthvert gefið inntak x, iv) $H(x)$ sé í eina átt og v) $H(x)$ sé laust við árekstra (14).

Hakkafall H er sagt vera *í eina átt* ef erfitt er að finna andhverfu fallsins. Það að „erfitt er að finna andhverfu fallsins“ merkir að ef gefið er eitthvert hakkagildi h þá er það reikningslega séð illgerlegt (afar erfitt eða reiknifrekt) að finna eitthvert inntak x þannig að $H(x) \rightarrow h$. Ef gefið er eitthvert inntak x og það er reikningslega séð illgerlegt að finna eitthvert annað inntak y, sem ekki er jafnt og x, þannig að $H(x) = H(y)$ (að sama hakkið komi út) þá er hakkafallið H sagt vera laust við árekstra.

Notkun aðferða í þessa veru er í stuttu máli það sem virðist búa að baki því sem gagnagrunnslögin kalla *dulkóðun í eina átt* sem ekki er hægt að rekja til baka með greiningarlykli (15). Endurtekin dulkóðun í eina átt mundi taka fastanúmer sem fæst úr fyrri dulkóðun sem inntak fyrir þá næstu. Taka má MD5 (Message Digest) (14) sem dæmi um algrím sem nota mætti fyrir slíka einstefnu dulkóðun. MD5 tekur til dæmis inntak af hvaða lengd sem er og „meltir skilaboðin“ og býr til 128 bita „fingrafar“. Slík föll eru gjarnan notuð fyrir rafræna undirskrift skjala. Ég nota það hér til að útbúa sýnidæmi um gerð uppflettitöflu.

2.2. Uppflettitafla

Jafnvel þótt ekki sé hægt að leysa beint $H(x) \rightarrow h$ verður fallið eða föllin sem gagnagrunnurinn gerir ráð fyrir samt sem áður að vera stöðug í tíma því ella væri ekki unnt að uppfæra gagnagrunninn. Þess vegna getur hver sá sem hefur aðgang að fallinu (eða föllunum) ætíð búið til töflu sem inniheldur hlið við hlið *inntakið* og *úttakið* úr fallinu:

Uppflettitafla fyrir nöfn eða kennitölur og dulkóðuð nöfn eða kennitölur er tafla (tafla I) sem geymir hlið við hlið safn nafna (eða kennitalna) og dulkóðaðra fastanúmera. Fletta má upp í töflunni til að finna dulkóðað fastanúmer fyrir gefið nafn og öfugt. Slík uppflettitafla er greiningarlykill (18). Hvernig smíða má uppflettitöflu fyrir alla þjóðina var reyndar lýst í greinargerð með frumvarpinu (15): með því að keyra þjóðskrána í gegnum dulkóðunar-

ferlið og setja hlið við hlið nafnið eða kennitöluna sem fór inn og dulkóðað fastanúmer sem kemur út fyrir hvert nafn eða kennitölu. Slíka töflu má einnig smíða fyrir afmarkaðan hóp manna. Ef tekin yrði ákvörðun um að fara til baka í gagnagrunninum, til dæmis ef Alþingi heimilaði opnun grunnsins með lagabreytingu eða dómstóll dæmdi að opna skyldi grunninn vegna dómsmáls, þá tekur það einungis augnablik að búa til uppflettitöflu fyrir alla þjóðina eða fyrir þann hóp sem óskað væri eftir að finna. Einungis þyrfti að kalla til handhafa dulkóðunarfallsins (eða fallanna) og renna nöfnum eða kennitölum hópsins í gegn.

Ferlið við flutning upplýsinga er að heilbrigðisstofnun einstefnudulkóðar kennitölu yfir í fastanúmer og afritar síðan og tengir heilbrigðisupplýsingarnar úr sjúkraskránni við fastanúmerið í stað kennitölunnar. Þetta er síðan sent til dulkóðunarstofu Persónuverndar. Landlæknir einstefnudulkóðar kennitölur þeirra sem hafa sagt sig úr grunninum að fullu eða að hluta til og sendir þá skrá til dulkóðunarstofu. Dulkóðunarstofa skal sía burt upplýsingar um þá sem sagt hafa sig úr grunninum, endurdulkóða síðan fastanúmer og senda þau ásamt áföstum upplýsingum áfram til gagnagrunnsins. Það er því ljóst að það verða margir handhafar að fallinu fyrir einstefnudulkóðun. Hver sem er þeirra gæti búið til uppflettitöflu yfir þau fastanúmer sem fara til dulkóðunarstofu.

2.3. Persónugreining við undirbúning gagna

Til að flytja gögn í gagnagrunn á heilbrigðissviði þarf að opna þau og lesa og skrá á rafrænan hátt. Á þessu stigi eru gögnin að fullu persónugreinanleg. Þetta gildir um öll gögn sem fyrir eru og ætlun er að flytja í gagnagrunninn. Þetta gildir einnig um gögn allra þeirra 20.000 manna sem hafa hafnað því að taka þátt í gagnagrunninum með bréfi til landlæknis því þeir sem undirbúa gögnin til flutnings eiga ekki að vita hverjir hafa sagt sig úr gagnagrunninum. Þá gildir þetta einnig um allar upplýsingar um látið fólk, heilsufarsskýrslur þeirra verða opnaðar og lesnar og undirbúnar fyrir flutning í gagnagrunninn. Þessi skoðun allra heilsufarsupplýsinga er gerð í öðrum tilgangi en þeirra var aflað. Jafnframt verða í óþökk þeirra fleiri en 20.000 manna sem sagt hafa sig úr grunninum lesin gögn um þá í öðrum tilgangi en þeim var upphaflega safnað, þau búin til flutnings og send í átt til gagnagrunnsins. Ef mistök verða hjá dulkóðunarstofu gætu þau gögn einnig farið alla leið í grunninn þrátt fyrir að bann hafi verið lagt við því.

3. Persónugreining með samanburði á munstrum ættartrjáa

Gagnagrunnslögin heimila samtengingu upplýsinga úr gagnagrunni á heilbrigðissviði við gagnagrunn með ættfræðiupplýsingum. Samkvæmt öryggisskilmálum sem gerðir voru fyrir Persónuvernd (17) verður ættfræðigagnagrunnur rekstrarleyfishafa dulkóðaður á sama hátt og gagnagrunnur á heilbrigðissviði. Sama gildir um gagnagrunn með erfðafræðilegum upplýsingum sem einnig er fyrirhugað að tengja við hina grunnana. Grunnarnir þrír verða að hafa sömu fastanúmer (eða að minnsta kosti einkvæma svörun á milli) til þess að samtenging grunnanna sé möguleg.

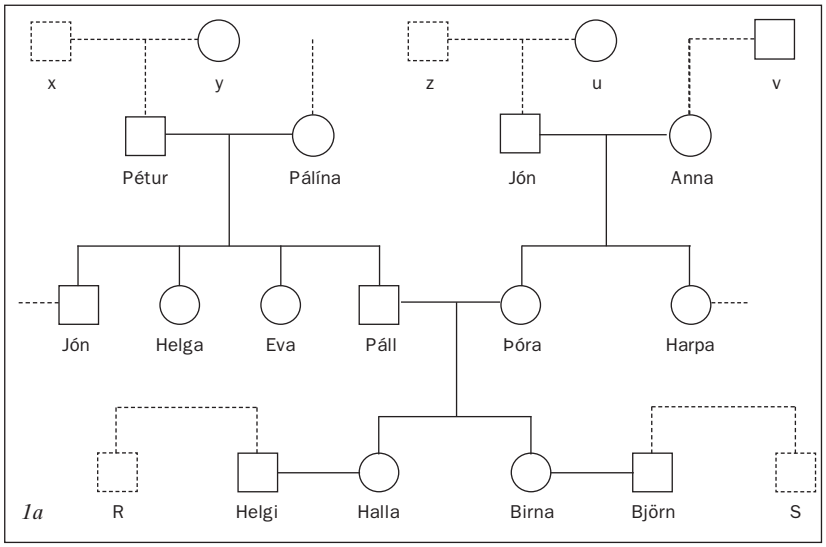
Tafla I. *Uppflettitafla fyrir nöfn einstaklinga og fastanúmer þeirra gerð með MD5 hakkafallinu (hash function).*

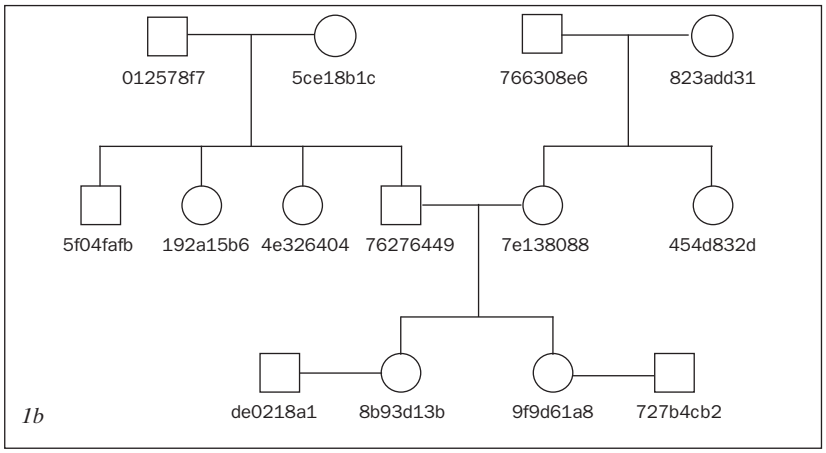
Stak	Inntak, x	Hakkafall	Úttak, h (hakk) eða fastanúmer
x_1	Pétur Pálsson	$H(x_1)$	012578f77e5820f2c5bdfcd48ec273ce
x_2	Pálína Pétursdóttir	$H(x_2)$	5ce18b1caca938a8b88161344537723f
x_3	Jón Jónsson	$H(x_3)$	766308e6bf587715483772c8f5b1c3c6
x_4	Anna Hallsdóttir	$H(x_4)$	823add31e1475737bb8d8351ca914c0f
x_5	Jón Pétursson	$H(x_5)$	5f04fafb74f2ecd7d66d2274d6fb78ad
x_6	Helga Pétursdóttir	$H(x_6)$	192a15b61372ea29a955027f7b7cfd59
x_7	Eva Pétursdóttir	$H(x_7)$	4e326404c5b0d75dcb76a3a7b634c320
x_8	Páll Pétursson	$H(x_8)$	762764495e433a7a16d2f27dd3a4b236
x_9	Þóra Jónsdóttir	$H(x_9)$	7e138088511b5d6a68e2860bfb7848c8
x_{10}	Harpa Jónsdóttir	$H(x_{10})$	454d832d78a28fbd3c3fce5904377934
x_{13}	Helgi Helgason	$H(x_{13})$	de0218a178a2e2bafc20279c57914626
x_{14}	Halla Bjarnadóttir	$H(x_{14})$	8b93d13bf2583ecd43681a93bcb091c2
x_{15}	Birna Bjarnadóttir	$H(x_{15})$	9f9d61a835cacf167bc97ff06bfde8ba
x_{12}	Björn Geirsson	$H(x_{12})$	727b4cb208be9eb929a9526e1200197b

Ættfræðigagnagrunnur er hins vegar einnig til hjá rekstraleyfishafa með nöfnum (og/eða kennitölum) einstaklinga. Þær upplýsingar eru einnig til annars staðar í þjóðfélaginu og fyrirætlanir eru um að veita almenningi netaðgang að þeim gagnagrunni. Þar sem sami gagnagrunnur er til bæði með nöfnum eða kennitölum og með fastanúmerum (dulkóða) getur sá sem hefur aðgang að báðum grunnunum, þeim með fastanúmerum og þeim ódulkóðaða, persónugreint einstaklinga í hinum dulkóðaða grunni með samanburði á ættarmunstrum.

Fræðilega séð er geysihár fjöldi mögulegra ættartrjáa á milli einstaklinga í einhverjum hópi (fjöldi möguleika er veldisfall af fjölda einstaklinga). Hið raunverulega ættartré þessa hóps einstaklinga er því líklega einstakt og frábrugðið að lögun frá ættartré annars hóps jafnmargra einstaklinga. Fjöldi barna og kynferði og tengingar einnar ættar við aðrar ættir með giftingu og barneignum eru einstakt munstur sem nota má til að þekkja fjölskyldur. Það eru um það bil 2.500 sex barna fjölskyldur og innan við 20.000 tveggja barna fjölskyldur í landinu. Aðrar algengar fjölskyldumyndir eru á milli þessa í fjölda. Tengingar fjölskyldna geta gert munstur þeirra einstök og þar með auðkennanleg. Jónsætt og Gunnuætt eru einstakar eins og aðrar ættir landsins og þekkjast hvort sem einstaklingarnir í ættartrénu eru tilgreindir með nöfnum eða dulkóðuðu fastanúmeri.

Auðveldara er að bera kennsl á tiltekna sex barna fjölskyldu en á tiltekna tveggja barna fjölskyldu því þær eru færri og möguleg fjölskyldumunstur eru fleiri. Algengustu fjölskyldur eru tveggja og þriggja barna fjölskyldur. Ef fæðingarröð er þekkt eru um það bil 5.000 fjölskyldur af hverri gerð tveggja barna





Mynd 1. *Samanburður ættartrjáa í ættargrunni a) með nöfnum og kennitölum og b) í ættargrunni með þrídulkóðuðum nöfnum og kennitölum. Mynd 1a sýnir ættartengsl nafngreindra einstaklinga úr ættargrunni sem inniheldur nöfn og/eða kennitölur. Einungis fornöfn eru gefin til að spara plass. Brotnar línur tákna tengingar við náin skyldmenni og þaðan yfir í ættartré allrar þjóðarinnar. Mynd 1b sýnir munstur ættartengsla sem fannst í sama ættargrunni sem inniheldur einungis dulkóðuð fastanúmer. Munstur ættartrésins með dulkóðum smellpassar við hluta af efra trénu. Einungis fyrstu átta stafir fastanúmers eru gefnir til að myndin sé greinilegri.*

fjölskyldna í landinu. Með tengingum við aðrar fjölskyldur verða þær fljótt einstakar og hægt að bera kennsl á þær.

Á myndum 1a og 1b er tekið dæmi um tvær ættir og tengingar þeirra. Mynd 1a er fengin úr ættargrunni sem inniheldur nöfn og kennitölur. Mynd 1b er fengin með upplýsingum úr gagnagrunni sem inniheldur ættartengsl og dulkóðuð fastanúmer: nöfn einstaklinga (eða kennitölur) hafa verið dulkóðuð í eina átt með aðferðum sem sagðar eru „mjög örugg" dulkóðun sem erfitt er að brjóta til baka. Þrátt fyrir það er persónugreining möguleg því fjölskyldumunstrin eru þau sömu og þau eru eina munstrið sem er einmitt svona í báðum grunnunum. Það má því lesa beint af myndinni hver er hver. Til dæmis, hver er 7e138088?

4. Persónugreining af samhengi upplýsinga

Allir kannast við leikinn *Hver er maðurinn* sem leikinn hefur verið í útvarpi og sjónvarpi, á árshátíðum stofnana, félaga og víðar. Einhver kemur fram í dulargervi og breytir rödd sinni. Keppendur fá að spyrja: er þetta karl (eða kona), leikur hann á hljóðfæri eða í knattspyrnu og svo framvegis. Maðurinn í dulargervinu svarar skrækri eða djúpri röddu, já eða nei, eftir atvikum. Að lokum finna þátttakendur út af samhengi upplýsinganna sem fram koma hver huldumaðurinn er og nefna hann.

Jafnvel þótt ekki kæmi til þekking á lykli með uppflettitöflu eða af samanburði ættartrjáa er samt hægt að þekkja einstaklinga í gagnagrunninum af sambærilegu samhengi upplýsinga og gerist í leiknum (19). Þegar persónuauðkenni, til dæmis nafn, kennitala eða gsm símanúmer, hafa verið afmáð með óafturkræfum hætti og skipt á þeim og einnota dulkóða er talað um aftengd (2,7) gögn (de-identified). Dulkóðanum geta fylgt lýðupplýsingar og heilsufarsupplýsingar. Eftir því sem slíkum upplýsingaþáttum er fjölgað þrengist hringurinn og að lokum verður samsetning slíkra upplýsingabita einstök. Með slíkri samsetningu er hægt að benda á einstaklinginn með fullri vissu eins og ef um fingraför væri að ræða. Á þennan hátt væri unnt að smíða lykil að dulkóða jafnvel þótt um aftengd gögn væri að ræða.

Talað er um að endurþekkja (re-identify) (19) einstaklinginn með slíkum upplýsingum og því samhengi sem setja má upplýsingarnar í. Þetta er miklu auðveldara hjá fámennri þjóð eins og Íslendingum en hjá fjölmennari þjóð. Tæknin hefur einnig breytt öllu í þessu sambandi. Með interneti á upplýsingaöld eru æ meiri almennar upplýsingar aðgengilegar hverjum sem er (19). Slíkar almennar upplýsingar er unnt að nota til að mynda einstaka samsetningu. Þar með er hægt að setja upplýsingar sem fylgja dulkóða í samhengi og leysa þannig gátuna hver er hver og hverjum tilheyra viðkvæmar persónuupplýsingar sem fylgja með dulkóðanum.

Sem dæmi má taka kennitölur einstaklinga sem hafa verið dulkóðaðar hvort sem er í einnota dulkóða eða í fastanúmer eins og gert verður í gagnagrunninum. Þeim fylgja almennar upplýsingar um kyn, fæðingardag og ár, hæð, búsetu, sem og misviðkvæmar heilsufarsupplýsingar svo sem uppskurð við botnlanga, krabbamein í maga eða brjóstum eða sykursýki (tafla II sem dæmi). Einnig mætti taka sem dæmi einstaklinga með sjúkdóma sem þykja ennþá viðkvæmari, svo sem geð- eða kynsjúkdóma.

Meðalfjöldi fæðinga á ári á Íslandi er rúmlega 4.200. Að meðaltali eru því 11-12 fæðingar á dag. Fáir dagar hafa fleiri en 20 fæðingar. Með upplýsingum um fæðingardag og ár er því búið að þrengja hringinn niður í 20 manns hið mesta (18). Með upplýsingum um kynferði helmingast hópurinn: að meðaltali fæðast sex stúlkur eða drengir og sjaldan fleiri en 10 stúlkur eða drengir á dag. Með því að bæta við hæð og búsetu eða augnlit er án vafa hægt að þekkja flesta ef ekki alla einstaklingana. Upplýsingarnar, sem eru sambærilegar við þær sem beðið er um fyrir vegabréf manna, nægja því til að bera kennsl á einstaklinginn (18) án þess að til komi nafn eða kennitala. Það er því hægt að greina hvaða einstaklingar eru haldnir þeim sjúkdómum sem fylgja með upplýsingunum í töflu II.

Einstaklingar haldnir ennþá „viðkvæmari“ sjúkdómum eru greinanlegir á sama hátt af slíkum dulkóðuðum lista. Karlmaður sem fæddist 2. febrúar 1979 er einn af (að meðaltali) sex karlmönnum sem fæddust þann dag á landinu. Hann er 176 cm og býr á Dalvík. Þetta hlýtur að vera Helgi. Hann er með sykursýki. Það þarf hvorki lykil, ættartré né persónuauðkenni að fylgja með til þess.

Tafla II. *Persónugreining af samhengi upplýsinga um þrídulkóðaða einstaklinga án notkunar lykils og ættarupplýsinga. Kyn, fæðingardagur og ár, hæð og búsetusveitarfélag eru almennar upplýsingar sem nægja til að persónugreina dulkóðaða einstaklinga. Viðkvæmar heilsufarsupplýsingar úr töflu sem þessari má því tengja nafngreindum einstaklingum. Einungis fyrstu átta stafir dulkóða eru teknir með til að minnka umfang.*

		Fæðingar-				Krabbamein í			
Dulkóði	Kyn	dagur	ár	Hæð	Búseta	maga	brjóstum	Sykursýki	Botnlangi
012578f7	1	2306	1922	177	Reykjavík	1	0	0	0
5ce18b1c	0	1101	1927	165	Reykjavík	0	0	0	0
766308e6	1	0312	1928	189	Akranes	0	0	0	0
823add31	0	0506	1930	178	Akranes	0	1	0	0
5f04fafb	1	1009	1942	182	Reykjavík	0	0	0	0
192a15b6	0	0404	1945	166	USA	0	0	0	1
4e326404	0	3101	1949	164	Höfn	0	0	0	1
76276449	1	1508	1951	176	Akureyri	0	0	0	0
7e138088	0	1705	1955	172	Akureyri	0	0	0	0
454d832d	0	1910	1958	170	Reykjavík	0	1	0	1
de0218a1	1	0202	1979	176	Dalvík	0	0	1	0
8b93d13b	0	1508	1980	170	Dalvík	0	0	0	1
9f9d61a8	0	2111	1981	177	Reykjavík	0	0	0	0
727b4cb2	1	1309	1980	192	Reykjavík	0	0	0	0

5. Lokaorð

Hér hafa verið tekin dæmi um það hvaða aðferðum unnt er að beita af sanngirni til að smíða lykla sem nota má til að persónugreina einstaklinga í gagnagrunni á heilbrigðissviði. Persónugreining er ekki fjarlægur, fræðilegur möguleiki (18) heldur er tiltölulega auðvelt að persónugreina einstaklinga. Persónugreining er möguleg með lykli og dulkóðun í eina átt breytir engu um það. Ef til dæmis Alþingi breytti lögunum síðar og segði heimilt að fara til baka eða ef dómstóll dæmdi í einhverju máli að opna skyldi gagnagrunninn þá er, tæknilega séð, hægt að gera það á augabragði. Forsenda gagnagrunnslaganna um ópersónugreinanleg gögn stenst því ekki.

Grunnregla laga (20) er að allir séu jafnir frammi fyrir lögunum. Hugmyndafræði fyrstu gerða frumvarpa um gagnagrunn byggðu á hugtaki fengnu úr tilmælum ráðherranefndar Evrópuráðsins að ópersónugreinanleiki miðist við það sem þarf verulegan tíma og mannafla til að leysa. Að mínu áliti getur þetta samt ekki verið grunnregla við setningu laga. Ef hugtök ráðherranefndarinnar væru notuð væru þeir sem hafa nægan tíma, mannafla og auð hafnir yfir lög en það brýtur í bága við frumregluna að allir séu jafnir fyrir lögum. Þá væru grunnatriði laganna einnig háð stöðu tækniþróunar, sem einnig er vafasamt.

Ég tel að að það sé grundvallarmunur á tilskipun Evrópusambandsins (95/46) og tilmælum Evrópuráðsins (R(97)5). Tilskipunin, sem er nú orðin þjóðréttarlega skuldbindandi fyrir Ísland, er lögleidd með persónuverndarlögunum og skilgreining gagnagrunnslaganna um persónuupplýsingar er komin úr henni. Munurinn á tilskipuninni og tilmælunum er sá að tilskipunin skilgreinir á afar víðan hátt einungis hvað átt er við með persónugreiningu en ræðir ekki sérstaklega hvað átt er við með ópersónugreiningu. Það gera tilmælin hins vegar. Þar með varpar tilskipunin sönnunarbyrðinni á hvern þann sem heldur því fram að hann sé að vinna með ópersónugreinanleg gögn.

Dulkóðun persónauðkenna í eina átt jafngildir ekki aftengingu upplýsinganna þar sem grunnurinn er langsum söfnun upplýsinga. Langsum tenging gagna merkir að dulkóðunaraðferðin má ekki breytast í tíma því ella væri ekki hægt að uppfæra gagnagrunninn. Dulkóðun sömu kennitölu mun því ætíð gefa sama fastanúmer. Hver sá sem getur sent kennitölur í gegnum dulkóðunarferlið og séð hvað kemur út getur því smíðað uppflettitöflu sem er lykill (16).

Jafnvel þótt ekki kæmi til möguleiki að fletta upp kennitölum og dulkóðuðum fastanúmerum er persónugreining samt sem áður einnig möguleg með ályktunum. Ættartré verða fljótt einstök þegar fjöldi manna í hóp eykst. Samanburður á munstrum ættartrjáa úr ættargrunni með dulkóðuðum fastanúmerum og ættargrunni með kennitölum eða nöfnum er því leið til að smíða lykil að grunninum. Þá er unnt að persónugreina einstaklinga af samhengi almennra upplýsinga (18).

Af framansögðu er ljóst að sanngjarnt er að ætla að til að persónugreina einstaklinga í gagnagrunni á heilbrigðissviði yrði beitt aðferðum í þá veru sem hér hefur verið lýst. Einstaklingar eru því persónugreinanlegir í gagnagrunni á heilbrigðissviði. Því er bæði réttmætt og sanngjarnt að aflað verði fyrirfram samþykkis sjúklinga fyrir notkun heilsufarsupplýsinga þeirra í öðrum tilgangi en þeirra var aflað eins og þjóðréttarlegar skuldbindingar Íslands standa til (7,8). Allt annað er ósanngjarnt.

Heimildir

1. Lög um gagnagrunn á heilbrigðissviði nr. 139/1998, desember 1998.
2. Björgvinsson DÞ, Arnardóttir OM, Matthíasson VM. Álitsgerð um ýmis lögfræðileg efni í frumvarpi til laga um gagnagrunn á heilbrigðissviði. Lagastofnun Háskóla Íslands. Íslensk erfðagreining sendi alþingismönnum. Alþingi. Erindi nr. Þ 123/9, komudagur 28.10.98, október 1998.
3. Stefánsson K. Frumdrög að frumvarpi til laga um gagnagrunna á heilbrigðissviði, 14 júlí 1997. URL http://www.mannvernd.is/login/IE_frumdrog_ggr_fax.doc.
4. Frumvarp til laga um gagnagrunna á heilbrigðissviði, Þskj. 1134, apríl 1998. (Lagt fyrir Alþingi á 122. löggjafarþingi 1997-1998.)
5. Recommendation No R(97)5 of the Committee of Ministers to Member States on the Protection of Medical Data, 13 February 1997.
6. Lög um skráningu og meðferð persónuupplýsinga nr. 121/1989, maí 1989.

7. Tölvunefnd. Umsögn tölvunefndar um drög að frumvarpi til laga um gagnagrunn á heilbrigðissviði. Beint til Ingibjargar Pálmadóttur heilbrigðisráðherra, 4. september 1998.
8. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 24 October 1995.
9. Frumvarp til laga um gagnagrunn á heilbrigðissviði. Þskj. 109, október 1998. (Lagt fyrir Alþingi á 123. löggjafarþingi 1998-1999.)
10. Tölvunefnd. Umsögn tölvunefndar um frumvarp til laga um gagnagrunn á heilbrigðissviði. Beint til heilbrigðis- og tryggingganefndar Alþingis, 26. október 1998.
11. Masood E. Gene Warrier. Opinion Interview with Kári Stefánsson. New Scientist magazine 15 July 2000; 167: 42.
12. Heilbrigðishópur gagnagrunnsdeildar Íslenskrar erfðagreiningar. Ópersónugreinanleg gagnasöfnun til ábyrgra vísindarannsókna. Morgunblaðið, 21. febrúar 2001: 46-7.
13. Lög um persónuvernd og meðferð persónuupplýsinga nr. 77/2000, 23. maí 2000.
14. RSA Laboratories. RSA Laboratories' Frequently Asked Questions About Today's Cryptography, Version 4.1. RSA Security Inc., 2000. URL http://www.rsa.com/rsalabs/faq/index.html.
15. Sigurðsson G, Björnsdóttir SH, Björnsson BÞ. Fylgiskjal VI. með frumvarpi til laga um gagnagrunn á heilbrigðissviði, Þskj. 109. Stiki ehf.: Minnisblað um feril heilsufarsupplýsinga frá heilbrigðisstofnun í miðlægan gagnagrunn, 29. september 1998.
16. Anderson R. Iceland´s medical database is insecure. BMJ 1999; 319: 59.
17. Admiral Management Services Limited. Security Target for an Icelandic Health Database. Admiral Management Services Limited, 5 January 2000. URL http://www.personuvernd.is/tolvunefnd.nsf/pages/gagnagrunnur.
18. Benediktsson O. Persónugreinanleiki í gagnagrunni á heilbrigðissviði, 13. september 2000. URL http://www.mannvernd.is/greinar/OBgreinanleikiMV.html.
19. Sweeney L. Re-identification of de-identified medical data. National Committee on Vital and Health Statistics Subcommittee on Privacy and Confidentiality, 1998. URL http://ncvhs.hhs.gov/980128tr.htm.
20. Aðalsteinsson R. „... einungis eftir lögunum“. Úlfljótur 2000; 53; 569-600.

Fræðigreinar íslenskra lækna í erlendum tímaritum

Sendið heiti greinar, nöfn höfunda og birtingarstað. Miðað er við greinar sem birst hafa á yfirstandandi og síðasta ári. Til glöggvunar verður íslenskra höfunda getið með fornafni þótt svo hafi ekki verið við birtingu.

- **Ólafur Ó. Guðmundsson,** Prendergast M, Foreman D, Cowley S
  Outcome of pseudoseizures in children and adolescents: a 6-year symptom survival analysis. Dev Med Child Neurol 200; 43: 547-51.
- **Helga Ágústa Sigurjónsdóttir, Leifur Franzson,** Manhem K, **Jóhann Ragnarsson, Gunnar Sigurðsson**
  Liqurice-induced rise in blood pressure: a linear dose-response relationship. J Hum Hypertension 2001; 15: 549-52.
- Borch K, Grodzinsky E, Petersson F, Jönsson K-Å, Mårdh S, **Trausti Valdimarsson**
  *Prevalence of coeliac disease and relations to **Helicobacter pylori** infection and duodenitis in a Swedish adult population sample: a histomorphological and serological survey.* Inflammopharmacology 2001; 8: 341-50.