

Výzvy, hrozby a úzká místa informatiky.

Jaroslav Král, *MFF UK Praha*

Abstrakt

Informatika je inženýrský a vědecký obor, pro nějž je typický neobyčejně rychlý vývoj velmi často s prakticky okamžitými dopady na společenské procesy a zájmy lidí. Ti se často snaží působení IT ovlivňovat, např. tím, že prosazují pravidla práce s daty, která vedou k podstatnému omezování pozitivních efektů IT a vedou pod záminkou ochrany dat a soukromí k přímému ohrožení základních lidských práv, jako je právo na zdraví, informace či na kvalitní vzdělání.

Rychlý vývoj informatiky je hlavní příčinou publikačních zvláštností v informatice. To zpochybňuje současné zásady hodnocení inforatického výzkumu. Tempo změn a vliv informatiky na společenské procesy klade specifické nároky na vzdělávání inforaticků.

IT vede k prudkému nárůstu složitosti administrativních procesů obvykle bez významných pozitivních efektů. IT odborníci se budou proto muset více věnovat i prosazování změn v prostředí, ve kterém pracují jimi vytvořené systémy.

IT umožňuje globalizaci a tím mimo jiné i vznik daňových rájů a snazší prosazování strategie vítěz-prohrávající. To však neznamená, že je tím vyloučen výsledek „prohrávají všichni“. Pokusíme se navrhnout principy možných řešení výše uvedených výzev.

Úvod

V tomto článku se pokusíme zviditelnit některé opomíjené vlastnosti informačních technologií a opomíjené aspekty jejich využívání. Ukážeme, že mnohé existující problémy používání informačních technologií lze odvodit z jejich nezpochybnitelných vlastností – a to jak přímo, tak nepřímo přes jejich působení na společnost a společenské prostředí.

Rozhodujícím problémem IT jsou omezení na využívání a dokonce zákaz existence důležitých dat. Ukážeme kritické důsledky současné praxe v oblasti vzdělávání, ochrany lidských práv, v oblasti zdraví a možnosti pracovat a v omezování možnosti výzkumu. Prudký vývoj informatiky omezuje použitelnost obvyklých metod hodnocení vědecké práce na inforatické publikace.

Působení inforatických systémů skrývá některé významné hrozby, které v globální ekonomice ohrožují stabilitu. IT přitom z řady důvodů nemůže využívat své možnosti při predikci a prevenci nepříznivých jevů.

Specifika informatiky.

Informatika (IT) je inženýrský obor a proto má typické atributy technických oborů: vývoj a údržba artefaktů poskytujících určité funkce, potřeba určitých řemesel, norem, atd. Zvláštností je, že vytvářené artefakty jsou často nehmotné povahy. Důsledky této vlastnosti se často přeceňovaly, dnes se spíše nedoceňují.

Informatika je obor vyžadující značný rozsah vývojové práce. Je to zároveň obor, který se rozvíjí nikdy dříve neviděnou rychlostí. Výkon HW se podle Mooreho zákona zdesetinásobí během pěti let. Za zhruba stejnou dobu zastará více než polovina informatických znalostí. Rozsah informatických znalostí se navíc během této doby podstatně rozšíří. Na to by měl být brán ohled při hodnocení kvality informatického výzkumu a v přípravě a celoživotním vzdělávání IT odborníků.

Informatika zásadním způsobem přímo ovlivňuje společenské procesy a zájmy lidí. Tím se informatici stávají aktivními účastníky společenských procesů. Někdy si to ani neuvědomují, leckdy si neuvědomují důsledky. Velmi často potřebu znalostí společenských procesů a znalostí obecně humanitních podceňují, což je závažný problém při vzdělávání informatiků. Určitě ale pociťují vliv společenské situace v nejrůznějších omezeních, které zásadním způsobem omezují pozitivní efekty informatiky, aniž by významně omezovaly záporné důsledky aplikací informatiky.

Problém hodnocení kvality informatického výzkumu

Rychlost vývoje informatiky má zásadní dopady na různé indikátory kvality výzkumu informatiků, jako jsou impaktfaktory. Valná většina poznatků dosažených v IT nemůže být obvykle včas publikována v časopise a včas i citována, protože je pravidlem, že termíny zveřejnění publikací a reakcí na ně vyžadují několik let, takže při rychlosti vývoje informatiky jsou citace v časopisech často možné až v době, kdy jsou citované práce již zastaralé. Poločas rozpadu informatických znalostí je totiž pět let. Jedná se o průměrnou hodnotu. Odtud plyne, že v některých, především nových, informatických oborech je poločas rozpadu podstatně kratší než pět let.

A to nemluvíme o tom, že mnohé nově získané poznatky jsou v době objevu natolik neobvyklé, že je zprvu obtížné najít časopis, kde by mohly být včas publikovány. Pak nezbyvá, než aby byly publikovány jako příspěvek ve sborníku či dokonce jako výzkumná práce nebo kniha.

Navíc neustále dochází ke změnám v možnostech publikovat. Některé časopisy vznikají, jiné mění náplň, či zanikají. To je důležitý důvod proč se v informatice většina poznatků *musí* publikovat ve sbornících konferencí či dokonce v knihách. Mnohé poznatky vyjdou z tohoto důvodu jen ve výzkumných zprávách nebo jsou pouze součástí dokumentace softwaru.

Tento problém je znám, viz B. Meyer et al, *Research Evaluation for Computer Science*, Comm. ACM Vol. 52, No. 4., April 2009, 31-34, nebo D. Parnas, *Stop the Numbers Game*, Comm. ACM vol. 50, No. 11, Nov. 2007, 19-21, <http://tinyurl.com/2z652a>. Hlavní poznatky těchto článků jsou ve volném shrnutí následující:

1. Informatika je specifický obor, který má specifická pravidla publikování.
2. V informatice jsou velmi důležité vybrané konference a knihy.
3. Některé artefakty, jako fungující software, by se měly hodnotit stejně vysoko jako publikace.
4. Pořadí autorů uvedené v publikacích není v informatických publikacích významné.
5. Různé numerické indikátory kvality nesmí být mechanicky používány jako míry kvality publikací. V článcích jsou uvedeny křiklavé příklady, kam až to může vést.
6. Počty publikací měří spíše píli než kvalitu.
7. Každé hodnocení kvality výzkumu, zvláště hodnocení kvantitativní, musí být založeno na jasných publikovaných kriteriích.
8. Stejně numerické indikátory nesmí být bez dalšího používány v různých disciplínách.
9. Hodnotící publikace jako ISI Web of Science není adekvátní pro většinu informatických oborů.
10. Kritéria hodnocení musí být pravidelně hodnocena a revidována.

Výše uvedené všeobecně uznávané zvláštnosti informatiky, především rychlost jejího vývoje, přímo implikují body 1, 2, 4, 5 a do značné míry i body 8, 9, 10. Jsou tedy do značné míry nezpochybnitelné. Tento fakt není dostatečně zdůrazněn ani ve Myersově článku. Ve zmíněném článku se doporučují alternativy k Web of Science, např. Google Scholar. My dodáváme, že ten je plně automatický, zato ale nemá prostředky na opravu případných nedostatků.

Rychlý vývoj informatiky způsobuje, že se v informatice na rozdíl od mnoha klasických oborů nemohou vytvářet mezinárodní skupiny dlouhodobě spolupracujících vědeckých pracovníků společně pracujících na určité dlouho existující často dosti úzké oblasti společného zájmu. Členové takových skupin se přirozeně dlouhodobě vzájemně citují a mají i svoje často dosti úzce zaměřené časopisy. Vznik takových skupin v informatice, zvláště v aplikované informatice, je z výše uvedených důvodů velmi ztížen.

Všimněme si, že kritéria uplatňovaná v ČR při hodnocení informatického výzkumu, ale také při hodnocení pedagogické způsobilosti informatiků na universitách, velmi málo zohledňují výše uvedená fakta.

Informatika ve společenských procesech

Informatika zásadním způsobem ovlivňuje společenské procesy a zájmy lidí. To vyvolává protitlak projevující se v různých omezeních na využívání dat, především těch osobních. Děje se tak bez ohledu na možné ztráty, k nimž dochází tím, že se informatika nemůže využívat pro posilování zpětných vazeb ve společnosti a pro zabezpečování oprávněných zájmů občanů.

Příkladem této situace je analýza kvality vzdělávání, nebo informace o potřebě určitých profesí a jejich příjmech v závislosti na typu vzdělání a tom, kde a jak bylo získáno. Výsledky takových analýz by mohly být použity i při úpravách studijních plánů, jako je podpora studia potřebného pro technická či vědecká místa, včetně těch v inženýrských oborech. To by mohlo být podstatně účinnější, než využívání různých pseudořešení založených na jiných kritériích, než je hodnocení výstupů a výsledků škol založené na úspěšnosti jejich absolventů.

Prudký vývoj IT a oblastí jejich aplikací znamená, že je pravděpodobně nutné

- zvýšit kvalitu výuky prováděné na nižších stupních škol v oblasti STEM (science, technology, engineering, mathematics) a zároveň cvičit schopnost soustavné práce¹,
- na vysokých školách předávat znalosti nutné pro vývoj aplikací IT určených pro použití ve společenských procesech,
- zajistit kvalitní průběžné celoživotní zdokonalování znalostí IT odborníků, především však i klíčových uživatelů IT.

IT by mohly pomoci zjistit, proč vlastně není zájem o STEM studium, zda to není nízkými platy inženýrů a vědců a zda je situace opravdu vážná.

Ve všech těchto případech jsou značné mezery, které nejenom ztěžují situaci informatiky, ale vedou ke vzniku velmi silných předsudků o vzdělávání ve společnosti a dokonce k něčemu, co lze nazvat zneužívání škol k tunelování veřejných i soukromých prostředků. Z tohoto důvodu jsou nutná další opatření, aby nedošlo k opakování případu státních maturit (zamítnutí odklady).

V případě zdravotních dat se nevyužívají takové možnosti, jako jsou přímé kontroly správné medikace. To má v nejednom případě fatální následky pro pacienty. Touto cestou lze navíc ušetřit miliardy na zdravotnických výdajích a usnadnit i medicínský výzkum. Především lze takto ztížit produkci některých drog, především Pervitinu.

Klíčovým problémem jsou ve všech výše uvedených případech omezení práce s daty, které vedou k hroživým důsledkům pro kvalitu dat a k blokování efektivního využívání IT. Častou záminkou je tzv. ochrana osobních údajů.

Úzké místo IT – procesy ochrany dat

Současná praxe ochrany osobních dat vede k závažným ztrátám. Může totiž vést k ohrožení životů a zdraví, k masivním přímým finančním ztrátám různého typu a v dlouhodobé perspektivě ke ztrátě konkurenceschopnosti států, např. postupným a není vyloučeno, že i dosti rychlým, úpadkem kvality vzdělávacích procesů a následně i kvality pracovní síly. Jde skutečné úzké místo (bottleneck ve smyslu manažerských

-
- ¹ Bohužel je to odmítáno pod záminkou, že dril na školy nepatří. Přitom se pod termínem dril míní i každý trénink schopnosti pracovat a výcvik dovedností vůbec. To se záporně projevuje nejenom ve STEM, ale také ve znalosti cizích jazyků a hlavně češtiny. Nedostatečná schopnost používat češtinu má nepříznivý dopad na mnohé inženýrské profese, především na analytiku.

věd) informatiky a nejen jí. Je paradoxní, že současná brutální praxe ochrany dat jejich likvidací či jejich zneprístupněním nezajišťuje dostatečnou ochranu dat, např. kvůli únikům dat ze sociálních sítí, e-komerce, nebo legálním využíváním různých státních registrů. Domníváme se, že lepší výsledky by umožnily promyšlenější postupy ochrany dat a nemusely by mít tolik negativních důsledků.

Kritickým problémem je skutečnost, že některé zákony a některé úřady provádějí činnosti, které ve svých důsledcích znamenají ztrátu dat nebo podstatné zhoršení jejich kvality. Příkladem jsou takové dimenze kvality dat, jako je dostupnost a použitelnost dat [WZL01,PLW02]. Vychází se přitom z nevysloveného, zato ale naprosto scestného předpokladu, že omezení přístupu k datům má očekávané efekty a přitom neimplikuje žádné podstatné záporné důsledky (např. zhoršení možností řízení státu či zhoršení ochrany zdraví).

Budeme diskutovat nejzávažnější důsledky současné praxe a pokusíme se navrhnout možná řešení. To může být založeno na existenci chráněných datových úložišť a zavedení pravidla, že každý má mít až na jasně vymezené výjimky právo na anonymní (ze zákona otevírané) informace, např. o kvalitě škol, ať jsou vypočteny z jakýchkoliv dat, které nejsou v soukromém vlastnictví.

Nesprávné či nedostatečně účinné procesy ochrany dat přispívají k tomu, že se nedaří včas analyzovat společenské procesy a, jak jsme nyní svědky, včas detekovat blížící se ekonomické otřesy a řádně analyzovat jejich průběh a předpovídat jejich budoucí vývoj

Meze a důsledky současné praxe ochrany dat

Současná praxe ochrany dat mlčky předpokládá, že z hlediska potřeb občanů, tedy z hlediska zabezpečování jejich soukromí, stačí omezit rozsah dat ve správě státu a tato data zabezpečit a omezovat shromažďování osobních dat v nestátních rukou tak, aby nemohla být zneužita. Tento předpoklad nemůže být splněn.

Data ve správě státu musí být rozsáhlá. Soukromá data jsou však legálně či pololegálně shromažďována na mnoha jiných místech: ve zdravotnictví, bankách, sociálních sítích atd. Existuje tedy mnoho kanálů úniku dat, v důsledku nedostatečné kontroly přístupu k citlivým datům. Často je ten přístup velmi snadný.

1. Mobilní telefony nejsou imunní vůči odpolechu a prozrazují i svou polohu. Obsah komunikace může být v databázích operátora po poměrně dlouhou dobu. Operátor má snad dokonce povinnost po dlouhou dobu (min. 6 měsíců) udržovat informace o tom, s kým se určitý klient bavil, s kým se pokoušel spojit a kdo se pokoušel spojit s ním. Není ani problém zaznamenávat i polohu mobilu a to i ze satelitu. Všechna tato data nejsou zcela imunní vůči zneužití. A určitě je předpokládáno, že jsou operátoři důvěryhodné instituce a data ochrání.
2. Elektronická komunikace mezi podniky a mezi občany a úřady nebývá dostatečně kódována. Například formuláře obsahující citlivé údaje jsou mnohdy posílány v nezakódované podobě (prostřednictvím http místo https).
3. Obrovské množství osobních údajů státní správy je ze zákona veřejně dostupných (katastr nemovitostí, seznamy členů různých spolků, obchodní rejstřík, atd.).

4. Mnoho služeb (hotely, banky, atd.) shromažďuje data o klientech a není zajištěno, aby se data včas likvidovala. Osobní data jsou vyžadována u mnohých blogů, serverů, databází a elektronických obchodů.
5. Mnoho citlivých dat je dostupných neopatrným využíváním sociálního softwaru (blogy, systémy jako Facebook, atd.). Filosofie mnohých takových sociálních systémů je taková, že je neobyčejně obtížné být dostatečně opatrný.
6. Citlivá osobní data jsou poskytována státům v rámci vstupních (imigračních) procedur.
7. Zabezpečení serverů a klientského software proti cíleným útokům není bez problémů. Je otázka, zda je v dostatečné míře za daných podmínek vůbec možné.

Za této situace se lze oprávněně domnívat, že nejčastější opatření ochrany dat založené na formálním znepřístupnění dat nebo jejich fyzické likvidaci není pro reálnou ochranu dat dostatečné. Zato škody tím způsobené jsou určitě více než dostatečné. Ztráty plynoucí z toho, že se důležitá data nesmí využívat ani eticky nezávadným způsobem, jsou možná dokonce mnohem větší, než jsme si vůbec ochotni připustit.

Přehlížené efekty současné praxe ochrany dat

Současná praxe ochrany dat vede ve svých důsledcích k nevyužívání dat a to i tehdy, kdy je možnost jejich zneužití velmi nepravděpodobná. To vede k virtuální skartaci obrovských objemů dat s kritickými dlouhodobými důsledky na kvalitu rozhodnutí a rozhodovacích procesů státu, podniků i občanů. A také k ohrožení zdraví a životů a rozsáhlým finančním ztrátám.

Současná legislativa a praxe ochrany dat nemůže podstatně snížit škody, ke kterým dochází v důsledku kompromitace osobních dat, protože nemůže – díky existenci mnoha cest kompromitace – rozsah kompromitace dat a tím vyvolaných škod podstatně omezit.

Možnost zneužití dat v informačních systémech se dá snížit různými technickými opatřeními. To umožňuje omezit překážky ve využívání dat. To by mohlo přinést významné společenské efekty. Situaci si osvětlíme na problému hodnocení vzdělání a na problému ochrany medicinských dat.

Jakým požadavkům musí vyhovovat systém pro evaluaci škol

Evaluace škol a vzdělání je proces vyhodnocování kvality určitých objektů či artefaktů. Podle ISO 8402 je kvalita multidimenzionální koncept vycházející z potřeb uživatele resp. hodnotitele. Podle ISO 9000:2005-Plain English Dictionary (<http://www.praxiom.com/iso-definition.htm>) je kvalita pospána takto:

"The *quality* of something can be determined by comparing a set of inherent characteristics with a set of requirements. If those inherent characteristics meet all requirements, high or excellent quality is achieved. If those characteristics do not meet all requirements, a low or poor level of quality is achieved.

Quality is, therefore, a question of degree. As a result, the central quality question is: How well does this set of inherent characteristics comply with this set of requirements?

In short, the *quality* of something depends on a set of inherent characteristics and a set of requirements and how well the former complies with the latter.

According to this definition, *quality* is a relative concept. By linking quality to requirements, ISO 9000 argues that the *quality* of something cannot be established in a vacuum. *Quality* is always relative to a set of requirements."

Kvalita určité školy, respektive určitého vzdělání, může skutečně být pro různé hodnotící subjekty různá, poněvadž záleží na jejich situaci a jejich požadavcích.

Použitelnost žebříčků vypracovaných různými státními orgány, novinami či institucemi nemusí odpovídat potřebám jednotlivých občanů. To opomíjíme fakt, že tyto žebříčky mohou být úmyslně zkresleny. Zlovyk hodnocení bez jasného vymezení kritérií hodnocení (zvolených atributů kvality) je velmi rozšířen. To významně snižuje účinnost informačních technologií.

Vzdělanost a talenty

Moderní společnost nemůže prosperovat bez talentovaných a kvalitně vzdělaných lidí a dnes také bez správného využití prostředků informačních technologií. Oba tyto aspekty vzájemně souvisí. Nedostatečné využívání informačních technologií (IT) v e-governmentu i jinde je do značné míry způsobováno nedostatkem některých profesí – především těch, kde jsou potřeba talent a hluboké a správně zaměřené vzdělání. Týká se to i schopnosti soustředěně pracovat, což je dovednost, která by se také měla cvičit ve škole a která je velmi třeba i v informatice.

Nedostatek kvalitně vzdělaných odborníků omezuje vývoj nových prostředků informačních technologií a možnosti efektivního využívání nástrojů využívajících IT. Je třeba umožnit, aby si jednotlivci či organizace vypracovali zjednodušeně řečeno vlastní virtuální "žebříčky" kvality oborů či škol podle dynamicky a individuálně volitelných kritérií.

To lze zajistit například vhodnými dotazy nad databází umožňující hodnotit úspěšnost absolventů podle různých kritérií – pokud možno těch, které většina subjektů považuje za relevantní. Může to být průměrný plat pracovníků určité profese v závislosti na věku či nějaké kritérium společenské prestiže. Nejběžnější dotazy by mohlo umožňovat jádro systému, výpočty specifických informací požadovaných některými uživateli by mohly být za vhodných bezpečnostních opatření prováděny prověřenými aplikacemi poskytnutými (vyvinutými) samotnými uživateli.

Moderní IT takové řešení v principu umožňuje. Jedno z možných řešení je popsáno v [KŽ06]. Háček je v tom, že je nutné dlouhodobě udržovat informace o absolventech jednotlivých škol či oborů včetně jejich příjmů a jiných osobních dat.

Všimněme si, že informace, které jsou relevantní jako kritérium kvality, jsou zpravidla anonymní, neidentifikuje konkrétní osoby, ke kterým se daná data vztahují. Občané by k nim měli mít ze zákona přístup. Přesto to naráží na legislativní bariéry a na nich založená nedomyšlená až primitivní pravidla ochrany dat.

Na druhé straně je možné specifikovat obecná kritéria (požadavky na kvalitu) hodnocení škol a vzdělání relevantní pro širokou veřejnost. Zřejmě by to mělo být jasné stanovení primárních úkolů vzdělanosti. Samotné jasné formulování takového požadavku nejen umožní návrh vhodného softwarového systému, ale je důležité pro formulaci obecných společenských postojů ke vzdělávání a vybudování žádoucích kontrolních mechanismů. Je ale důležité i pro efektivní využívání informatiky.

Informatika zde může velmi pomoci, ale jen tehdy, podaří-li se překonat předsudky o ochraně osobních dat a zlomit odpor mocných lobby a přesvědčí veřejnost o potřebě kontroly a evaluace vzdělávacích systémů. To je významná bariéra vazužívání IT vůbec. Příklad státních maturit je dostatečně výmluvný.

Cíle vzdělávání a informatika

Domníváme se, že prvotním cílem studia je vytvoření předpokladů k tomu, aby se byl absolvent schopen uživit. Pokud se neuživí, nemá cenu mluvit o tom, jak se během studia podařilo rozvinout jeho osobnost. V tomto směru převládají předsudky a nepodložené dohady.

V průzkumech personálních agentur z posledních let, např. agentury Manpower, bije do očí nedostatek technických a technologických profesí a kvalifikovaných dělnických profesí. U nás to asi souvisí i s tím, že se bezhlavě omezovaly nebo rušily některé studijní obory a učební obory.

Nedostatek dělnických a technických profesí a také vývojářů a vědců se považuje za problém až hrozbu i pro USA [Jam07,Spe06]. Je s podivem, že tento fakt nijak neznepokojuje naši veřejnost ani naše ministerstvo školství. To dokonce v podstatě podporuje omezování výuky přírodních věd, technických oborů a matematiky.

V USA se nedostatečná podpora výuky těchto oborů označovaných zkratkou STEM (science, technology, engineering, mathematics) považuje za závažný problém (viz www.stemeducation.com a [Jam07]). I tam není náprava snadná, možná že i z podobných příčin jako u nás. Nedostatky výuky STEM ohrožuje perspektivy všech technických oborů, včetně informatiky.

Naši studenti se v mezinárodních srovnáních STEM znalostí a také informatiky propadli až někam mezi země, kde je problém zajistit povinnou školní docházku. Přitom stále trvají snahy dále redukovat výuku STEM znalostí a dovedností na středních i základních školách. Bohužel jsou i příznaky toho, že ani v ostatních předmětech, například jazycích, se dosahované výsledky nezlepšují, i když právě zlepšení v těchto směrech má být údajnou motivací pro redukcii STEM předmětů. Zvládání jazyků je totiž také dřina a dril.

To vše silně ohrožuje budoucnost informatiky – a nejen jí – u nás. IT je totiž inženýrský obor, který je navíc blízký matematice. Hrozí, že nebude dostatek odborníků, kteří by byli schopni nejen vyvíjet IT systémy, ale ani je instalovat a udržovat.

Na straně uživatelů bude pak problém konceptuálně složité IT systémy vůbec používat. Naši absolventi budou nakonec schopni provádět práce na úrovni utahování šroubů u montážního pásu. Možná ani to ne – nenaučí-li se soustavné práci, bude jim i tato primitivní práce činit nemalé potíže. Výcvik píle a soustavná práce, včetně např. zvládání matematiky a vlastně i jazyků, je něco, co se pod záminkou vyloučení drilu z výuky odsuzuje. Informatika by mohla poskytovat vhodnou zpětnou vazbu. To ale naráží na nedostatečný zájem veřejnosti i státních orgánů.

Ochrana zdravotních dat jako zdravotnická hrozba.

Pro ochranu zdravotních dat existují velmi přísné předpisy. Tak přísné, že se často preferuje fyzická likvidace dat. Přesně to se stalo při pokusu vytvořit systém S, který měl ztížit možnosti produkce drogy pervitinu. Ve výrobě pervitinu jsme velmoc a lze jen obtížně odhadovat, kolik převážně mladých lidí pošle pervitin na onen svět a kolik kriminálních aktivit se na toto odvětví nabaluje.

Princip řešení byl v principu jednoduchý. Léky mohly být vydávány jen proti předložení občanského průkazu a bylo on-line kontrolováno, zda na daný průkaz nebylo vydáno v poslední době příliš mnoho léků použitelných k výrobě pervitinu. Systém S se osvědčil. Výroba pervitinu významně poklesla.

Zároveň se objevily další možné významné aplikace. Systém umožňoval on-line kontrolu a blokování evidentně chybných medikací, které mohou mít a nezdá se májí fatální následky. Znamenalo by to záchranu nemalého počtu životů. K takovým případům dochází především v důsledku nedostatku informací o právě probíhajících medikacích (podávaných lécích) nebo o všech onemocněních daného pacienta.

Existovaly i možnosti podstatného snížení léčebných nákladů a využití systému v epidemiologické operativě a zdravotnickém výzkumu.

Výsledek je znám. Úřad pro ochranu osobních údajů činnost systému zakázal a sebraná data musela být zničena. Výroba pervitinu se zase naplno rozeběhla. Pokud bude systém vůbec provozován, budou jeho možnosti podstatně omezeny a to přesto, že pravděpodobnost vyžrazení dat byla velmi malá a také k ní nedošlo.

Je zřejmé, že akcí Úřadu na ochranu osobních údajů byly přímo ohroženy životní zájmy pacientů i zdravých občanů. Lze mít za to, že možné poškození zájmů občanů systém S prakticky neumožňoval. Skutečné zájmy občanů však byly faktickým zákazem používání systému významně ohroženy.

Ochrana dat a copyright.

Informační technologie mají pozitivní, ale leckdy naopak velmi nežádoucí efekty při ochraně autorských práv. IT umožňují efektivní kontrolu plagiátorství. Především avšak usnadňují vyhledávání informací, často jsou i nutnou podmínkou pro efektivní vyhledávání informací po celém světě. IT nabízí technické prostředky usnadňující v dříve nepředstavitelném rozsahu editaci a tisk a také kopírování dokumentů. To vše jsou věci, které lze jen přivítat. Bohužel nejsou bez stínů.

Autorské právo omezuje přístup k informacím a má proto při necitlivém uplatňování řadu nepříjemných důsledků, které zesiluje používání IT. Velcí vydavatelé si často vymínají převod autorských práv na sebe, aby zajistili dostatečný prodej svých produktů. To silně omezuje prostor pro budování centrálních úložišť článků a knih známých jako digitální knihovny. Právo kopírovat dokumenty je mnoha případech striktně omezováno, ale bez toho nelze moderní IT rozumně používat. Mimo jiné proto, že kromě bariér daných legislativou jsou při vyhledávání informací i problémy související s tím, že je na webu stále větší množství balastu.

Koncentrace a fúze editorských firem vyvolává nebezpečí, že některé firmy, které požadují převod autorských práv na sebe, získají časem monopol na určité typy znalostí.

Tato situace je velmi nepříznivá a má tendenci se dále zhoršovat. Možná řešení jsou založena částečném oslabení síly autorského práva, protože hrozí, že země které se současné praxe ochrany dat drží, budou vůči ostatním zemím postupně v nevýhodě. Problém je i v oblasti lidských práv, neboť to ohrožuje veškerou tvůrčí činnost, i tu ve vědě. Nakonec už není naprosto vyloučeno, že nebudeme moci volně psát formule popisující přírodní zákony.

Opětně jsme v situaci, kdy lze dosáhnout velké přínosy z používání IT (např. digitální knihovny), není to ale možné, pokud se nezmění legislativa. Navíc i zde vlastně dochází k omezování některých práv, které považujeme za samozřejmé. Příkladem jsou nadměrná omezení pro přístup k informacím.

Částečným řešením může být použití obdoby otevřených licencí známých z oblasti softwaru. Pokusem o řešení v oblasti kultury je licence Creative Commons. Tento problém a řada dalších problémů spjatých s IT, především s Internetem je studována v knize Lawrence Lessig: *Free Culture, How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, český překlad lze nalézt na <http://wiki.root.cz/Main/FreeCulture>. V této knize jsou uvedeny mnohé další hrozby způsobované tím, že se na webu používá či spíše zneužívá copyright vhodný spíše pro 19. století. Samotná kniha je volně dostupná na Internetu a lze zakoupit i její tištěnou verzi, viz [ISBN 0-14-303465-0](http://www.isbn.org/0-14-303465-0).

I v tomto případě dochází k nevyváženému přístupu. Současná praxe ochrany autorských práv skrytě i zjevně preferuje řešení, která hrozí, že bude v oblasti vědeckého výzkumu výrazně omezována možnost výzkumu mimo obří korporace.

Podpora zpětných vazeb ve společnosti

Proč k tomu dochází? Můžeme za to i my – odborná veřejnost. Prakticky jsme se nepokusili zabránit tomu, aby legislativa omezovala používání dat natolik, že i seznam nejvíce žádaných profesí jsme pro účely tohoto článku museli zjišťovat v materiálech firmy Manpower, protože jsme nemohli snadno využít informace od nějakého českého zdroje, například od Ministerstva školství nebo Ministerstva práce a sociálních věcí.

Nemůžeme ani dostatečně přesvědčivě vyloučit, zda omezování vědecké, technické, technologické a matematické (STEM) výuky na školách není nakonec správné. Nevíme dostatečně přesně, kde jsou kteří odborníci potřeba a jaký průměrný plat tam mají a jaké jsou tam profesionální vyhlídky. Moderní IT umožňuje řešení, které zjišťování takových informací technicky umožňuje. O legislativě platí spíše opačné tvrzení.

Jedno z možných řešení je popsáno v [KŽ06] a je ve zkrácené podobě popsáno níže. Háček je v tom, že je nutné dlouhodobě udržovat informace o absolventech jednotlivých škol či oborů včetně jejich příjmů a jiných osobních dat. To naráží na legislativní bariéry. K prozrazení citlivých dat potřebných k výpočtu zveřejnitelných informací by při vhodném technickém řešení docházet nemuselo.

Výchova informatických odborníků

Výchova informatiků je ohrožena klesající kvalitou STEM vzdělávání na nižších stupních škol. Doby, kdy se ve světových soutěžích programátorů umísťovali naši

studenti na předních místech², jsou dávno pryč. Možné cesty ke zlepšení využívající možnosti současné informatiky jsou uvedeny výše.

Na vysokých školách je problém optimální kombinace úzce technických („programátorských“) znalostí a dovedností se znalostmi „humanitního“ typu, jako je schopnost práce v týmu, dovednost vyjednávat, schopnost chápat a dovést i měnit sociální procesy a základy managementu potřebné pro specifikaci požadavků na softwarové artefakty a pro řízení vývojových prací.

Hlavní výzvou pro universitní studium je potřeba kombinace vysoce technických (programátorských) znalostí a dovedností se sociálními znalostmi a dovednostmi. Podobné požadavky existují i pro jiné technické obory. V informatice je vazba mezi technickými a netechnickými znalostmi užší. Mladí informatici hledí na humanitní znalosti s velkým despektem. To je zvláště silné u lidí, kteří mají tendenci se uplatnit jako hekři.

Rychlý vývoj informatiky vyvolává potřeby neustálého celoživotního studia informatiků. Na tom by měly spolupracovat všechny instituce, které v tomto směru působí: vysoké školy, firmy i konsultační a školící firmy. Jednotlivé instituce spolu až dosud málo kooperují.

IT může významně zlepšit podmínky a výsledky celoživotního vzdělávání. Celoživotní vzdělávání a trénink je nutnou podmínkou pro vykonávání stále širší palety zaměstnání, IT obory jsou jen jedním příkladem z mnoha, viz Karolina Bartsch, the employment projections 2008-2018, US Bureau of Labor Statistics, November 2009, stránka 10; viz www.bls.gov/opub/mlr/2009/Art1full.pdf.

Technická řešení pro zabezpečení a využití dat

Technické řešení výše uvedených problémů může být založeno na následujících zásadách:

1. Data potřebná pro výpočet otevřených informací se anonymizují, tj. data umožňující jednoduchou identifikaci osob, např. osobní čísla, se zakódují tak, aby byly zachovány vazby mezi údaji pro danou osobu pocházející z různých časových okamžiků a různých zdrojů.
2. Pro zvýšení bezpečnosti se data uloží výhradně do datového úložiště spravovaného důvěryhodnou institucí *I*.
3. Veškeré výpočty informací provádí *I*. Je ale možné a z praktických důvodů žádoucí, aby některé potřebné aplikace vypracovávali uživatelé. Aplikace ale musí být institucí *I* prověřovány a používány až po tom, co se prověří, že s velkou pravděpodobností nenarušují soukromí.
4. Všechny informace vypočtené v *I* jsou na výstupu prověřovány, zda nekompromitují osobní data. Toto opatření je žádoucí jako další bariéra proti úniku osobních dat. Opatření z předchozích bodů mohou být velmi účinná, nemohou ale být absolutně účinná.

² Platí to pro STEM soutěže a zdá se, že se to nevyhýbá ani kolektivním sportům.

5. Každý uživatel se zavazuje nezneužít případné selhání ochranných mechanismů. Toto opatření je pro úplnost, protože pravděpodobnost, že všechna ostatní opatření selžou, je prakticky nulová.

Anonymizace dat vznikajících v dlouhém časovém intervalu

Hlavním problémem informačních systémů pracujících s osobními daty, které vznikaly v průběhu mnoha let, je fakt, že vztahy mezi daty musí být založeny na jednoznačné identifikaci subjektů (např. osob) a ta nemůže být zcela nezávislá na identifikátorech, kterými se prokazuje totožnost daných subjektů v reálném světě. Tyto identifikátory nazveme otevřené. V oblasti e-governmentu je takových systémů mnoho. Systém hodnocení škol je jen jedním z nich.

Částečným řešením tohoto problému je požadavek, aby byla taková data používána jen na akreditovaných serverech [KŽ06] a jejichž výstupem jsou anonymizované informace vyžadované uživateli.

Zabezpečení dat může být dále zlepšeno tímto obratem: Každému subjektu se při vstupu dat, které se ho týkají, do systému se přiřadí nový identifikátor (nazvěme jej uzavřený) a tento identifikátor zůstává pevný po celou dobu, kdy systém pracuje s daty vztahujícími se k danému subjektu.

Technicky se uzavřený identifikátor přiřadí náhodně při prvním výskytu otevřeného identifikátoru ve vstupních datech a přiřazení otevřený identifikátor – uzavřený identifikátor se udržuje pomocí dobře zabezpečené tabulky. V samotné databázi se otevřený identifikátor vůbec nevyskytuje. Toto řešení je možné modifikovat tak, aby se uzavřený identifikátor v zájmu bezpečnosti pravidelně měnil. Tuto proceduru nazveme anonymizací dat.

Anonymizace dat by teoreticky mohla umožnit, aby s daty mohl pracovat kdokoliv, nejen pouze pověřená instituce. Problém je s bodem 4. Striktně vzato se může *I* bez anonymizace obejít. Důvodem anonymizace je to, že se tímto způsobem zvyšuje odolnost systému vůči útokům hackerů i zaměstnanců instituce *I* (inside attacker).

Vybudování dostatečného aparátu zpřístupňujícího data za dostatečné úrovně bezpečnosti není tedy technicky neřešitelný problém. Problém je zlomení nedůvěry občanů a zlomení odporu lobby těch, kterým současný stav vyhovuje.

Informatické hrozby

S informatikou jsou spojeny významné hrozby. Nejdůležitější není široce medializovaná a velmi rozšířená obava z Velkého Bratra (zneužívání osobních dat politiky a státem). Tato obava vedla k výše popsáním nevhodným procesům ochrany dat. Nejdůležitější je pravděpodobně podpora raketového růstu složitosti administrativních procesů, která zvyšuje administrativní náročnost téměř všeho. Tento fenomén znevýhodňuje běžné občany, které si nemohou často dovolit pomoc poradců.

Nepřímým důkazem jsou hejna výškových administrativních budov v Praze i okolí a stále složitější zákony a také přetrvávající pomalost a jiné nedostatky administrativy.

Tyto hrozby mohou být fatální, srv. knihu Tainter, A, Kolapsy složitých společností, Dokořán, Praha, 1998. V této knize se mj. uvádí, že ve složitých civilizacích je obvyklé, že se společenská struktura a administrativní procesy stávají stále složitější, ale investice do nich přinášejí postupně stále nižší efekt. To pak dosti často vyústí do kolapsu takových civilizací.

Dokonce se tvrdí, že úpadek hospodářství v Turecké říši byl kdysi urychlen tím, že se Turci od čínských válečných zajatců naučili vyrábět papír a záhy byly jejich správní orgány zavaleny horami papíru. Ostatně i z vlastní zkušenosti víme, že se naše státní správa i přes masivní investice do e-governmentu nijak nezjednodušuje.

Klesající výnosy investic do informatiky jsou skutečností i když z výše uvedeného vyplývá, že to není nutností. Tedy pokud se nedomníváme, že moc partikulárních lobby, kterým vyhovuje současný stav, je absolutní. Změny vyžadují společné úsilí IT odborníků a uživatelů IT, tedy spolupráci se širokou veřejností, která je hlavním uživatelem veřejných informačních systémů, zvláště v e-governmentu.

Přehlížené hrozby globalizace

Globalizace ekonomických procesů, především finančních toků, umožněná IT je podmínkou existence daňových rájů a ztráty pocitu odpovědnosti mnohých občanů vůči společnosti, ve které žijí. Začíná se z jejich strany stále častěji vůči většině společnosti uplatňovat taktika vítěz-ztráta. Preferují snižování vlastního příspěvku na fungování společnosti a státu bez ohledu na ohrožení jeho stability. Tvrdí, zčásti oprávněně, že se státními penězi plýtvá. To je skutečnost způsobená i nekvalitním návrhem informatických systémů ve veřejných informačních systémech. Je to ale ve skutečnosti výmluva. Problém je, že se nejen neplatí daně potřebné nejen na důchody, ale také na silnice a obecně na fungování státu.

Chování těchto skupin je někdy přirovnáváno k chování kočovníků, kteří napadali a ničili staré civilizace, aby po dílu zkázy zmizeli kdesi ve své původní domovině. Dnes kamsi do neznáma mizí finance a rostou nejrůznější finanční bubliny. I v tomto případě by se našly informatické prostředky, jak škody omezit. Není k tomu ale opětně vůle.

Je ale velmi pravděpodobné, že i dnes stále platí poznatek, že situace „vítěz-prohrávající“ nakonec vyústí do situace „prohrávají všichni“, např. díky globální nestabilitě. Informační systémy by mohly přispět ke snížení této hrozby. Je to ale velmi obtížný úkol směřující proti krátkodobým zájmům velmi vlivné vrstvy občanů a institucí.

IT otevírá velký prostor pro nejrůznější spekulativní aktivity probíhající ve milisekundách. Takové činnosti jsou výhodně k vytváření nejrůznějších spekulativních bublin a pro zvyšování nestability globální ekonomiky.

Závěr. Změna úkolů informatiky

Domníváme se, že informatici se musí více starat o podmínky, ve kterých pracují jejich technologie a zahrnovat je do specifikace požadavků na informační systémy a také se pokoušet je nepřijímat pasivně. Je nejvýše žádoucí mít základní znalosti o společenských procesech ovlivňovaných informatickými systémy.

Současná nevyhovující praxe ochrany citlivých dat vede k velmi kritickým důsledkům. Jen s trochou nadsázky ji lze přirovnat k rozhodnutí zahazovat to, co je ve vztahu mezi státy obdobou "competitive intelligence" nebo "business intelligence" ve vztahu mezi podniky.

Prvým krokem k nápravě by mělo být prosazení změny v přístupu k datům. Nemělo by se vycházet z úvah typu: Data nemohou být použita, protože mohou být případně (s jakkoli malou pravděpodobností) zneužita.

Při rozhodnutích implikujících překážky v použití dat by se mělo ze zákona stanovit, jak to ovlivní možnost, že data budou zneužita (lépe řečeno: jak se změní pravděpodobnost jejich zneužití s uvážením všech existujících bezpečnostních hrozeb), a vyhodnotit, zda neexistuje cesta, jak data upravit a bez překážek využívat. Mělo by se také vyhodnocovat, jaké ztráty jsou spojeny s faktickým zákazem využití dat. To vše by mělo být zahrnuto do procesů e-governmentu. Možná není přehnané tvrzení, že je to nutný předpoklad toho, abychom mohli pomýšlet na dlouhodobé členství v klubu rozvinutých ekonomik.

Problematická praxe ochrany dat je částí širšího problému. Je to stav společenského prostředí, které je málo vstřícné a často plné předsudků vůči používání informatiky tam, kde by mohla přinést největší efekty. To se týká mikrosociálních i makrosociálních procesů.

Problémy k řešení v budoucnu:

1. Studium důsledků stávající praxe práce s daty, jako je pozastavení kontroly výdeje léčiv v ČR.
2. Možné změny pravidel práce s daty, např. pomocí akreditovaných institucí a pomocí změn legislativy. Řešení souvisejících technických problémů jako je účinná filtrace vypočítaných informací.
3. Návrh a pokud možno i prosazení pravidel hodnocení škol podle dlouhodobé úspěšnosti jejich absolventů. Řešení souvisejících technických problémů.
4. Modernizace pravidel ochrany autorských práv.
5. Studium možností jak upravit metody hodnocení výzkumu v informatice i v jiných oblastech.
6. Optimalizace a modernizace studia IT odborníků, především jeho celoživotních forem.
7. Cesty prosazování potřebných legislativních změn.
8. Cesty využívání informatiky pro zefektivnění procesů analýzy společenských procesů.

Reference

[Com07] Committee on Engineering Education, National Academy of Engineering: *Education of Engineer of 2020*. The National Academies Press, Washington, DC, USA, 2007.

[Jam07] Jamieson, L.: **Engineering education in a changing world** *The Bridge*, Spring 2007, 2007.

[KŽ06] Král, J., Žemlička, M.: **Kvalita dat a informací - základní omezení IT ve veřejné správě** V: Pour, J., Voříšek, J (Eds.) *Systems Integration 2006*. VŠE Praha, Praha, 2006. ss. 215-222.

[PLW02] Pipino, L., Lee, Y.W., and Wang, R.Y.: Data quality assessment. *Comm. of ACM*, 45(4):211-218, 2002.

[Spe06] Spellings, M.: Answering the challenge of a changing world. 2006.

[WZL01] Wang, R.Y., Ziad, M., and Lee, Y.W.: *Data Quality*, vol. 23 of *The Kluwer International Series on Advances in Database Systems*. Springer, 2001.