

# Digitálna veda – východiská, problémy a princípy

Jela STEINEROVÁ

steinerova@fphil.uniba.sk

**Súčasnú komunikačnú procesy vo vede a ich informačné zabezpečenie sa transformujú do elektronického prostredia veľkou rýchlosťou. Vznikajú nové modely a nové vízie vedeckej komunikácie nazývané rôznymi termínmi, napríklad digitálna veda, e-veda, veda 2.0. Informačná veda sleduje tieto zmeny už dlhší čas a prispieva k rozvoju modelov a praxe digitálnej vedy v oblastiach nových informačných služieb a filozofie otvoreného prístupu a otvorenej vedy (open access, open science). Dôraz sa kladie na veľké množstvo zhromaždených vedeckých empirických dát, nové metódy a postupy pri ich získavaní a spracúvaní a nové možnosti ich zverejňovania, ale aj komunikácie a kolaborácie. Spoločnou myšlienkou v rámci týchto koncepcií je zabezpečenie rozširovania vedeckých výsledkov pre spoločnosť, zabezpečovanie okamžitého prístupu k novým poznatkom, a najmä podpora inovácie a kreativity. Aktuálnymi otázkami je nielen zlepšovanie tzv. otvorenej kolaborácie a prístupu k otvoreným veľkým dátam (dátovo orientované vedy), ale aj hľadanie nových nástrojov na posudzovanie vedeckých výstupov ako vedecká metrika a systémy vedeckého renomé.**

Cieľom tohto príspevku je formulovať vybrané koncepčné východiská digitálnej vedy a možnosti jej informačného zabezpečenia na úrovni výskumu informačnej vedy. Stručne sa naznačujú terminologické východiská, zložky a modely digitálnej vedy a vybrané poznatky informačnej vedy v oblasti informačného správania vedcov, informačných interakcií a digitálnych sociálnych a humanitných vied. Uvádzame aj príklady z bibliometrických a citačných analýz. V kontexte týchto trendov interpretujeme výsledky vlastného výskumu doktorandov a v závere naznačujeme konceptuálnu schému digitálnej vedy.

## Terminologické východiská a kontexty

Digitálna veda sa vymedzuje ako premena podstaty vedeckých informačných interakcií a inovácia postupov pod vplyvom vývoja informačných a komunikačných technológií, najmä v súvislosti s otvorenosťou prístupu k dátam a publikáciám a komunikáciou a kolaboráciou v elektronickom prostredí. Niektorí autori využívajú aj termín otvorená veda, ktorý súvisí s približovaním vedeckých výstupov občanom v spoločnosti, s prepojením vedeckých inovácií a priemyslu s inováciami v prienikoch disciplín. Predstavy o nových modeloch vedeckej komunikácie majú v informačnej vede svoju históriu, napríklad termín „veľká veda“ použil už D. Price v roku 1963, keď určil zmeny vo vedeckej práci smerom k veľkým kolaboratívnym projektom v informačne intenzívnom prostredí (Brown 2010). Veda je však v zásade sociálnym prostredím s množstvom komunikácií a informačných interakcií (štúdium, diskusia, citovanie, experimentovanie, formálna a neformálna komunikácia, publikovanie ap.).

Digitálna veda obsahuje informačnú infraštruktúru, ale aj novú kultúru otvorenosti a kolaborácie s víziou nových typov vedeckého výskumu a tvorivých inovácií. Týka sa tak prírodných a technických vied, ako aj humanitných a sociálnych vied s dôrazom na metódy experimentov a simulácií v digitálnom prostredí. Vytvárajú sa aj novšie digitálne priestory, ako napr. kolaboratóriá (Lassi, Sonnenwald 2013) pre vedeckú interdisciplinárnu spoluprácu. Umožňujú ekologické využívanie raz spracovaných dát a publikácií uložených v digitálnych repozitároch.

Na koncepcnej úrovni sa digitálna veda spája s víziou informačnej vedy smerujúcej k znalostnému obratu a novým službám s pridanou hodnotou. Prejavuje sa v takých informačných interakciách, ktoré súvisia s manažmentom veľkých digitálnych dát. Príkladom sú nielen existujúce digitálne knižnice ako NSDL (Národná vedecká digitálna knižnica), Alexandria Digital Library, California Digital Library, Perseus či Europeana (Borgman 2007), ale aj procesy spracovania veľkých súborov sociálnych dát, množstvo simulácií a experimentov. Novšie modely digitálnej vedy naznačujú hľadanie nových metód – napríklad „sledovanie pohybu očí“ v kognitívnych výskumoch, vytváranie špeciálnych typov vizualizácií z množstva rôznych dát (geografické, sociálne – demografické, meteorologické, astronomické a i.).

Niektorí autori používajú aj termín veda 2.0 (Steinerová 2009, Shneiderman 2008, Burgelman et al. 2010) s cieľom naznačiť využitie nástrojov web 2.0, najmä kolaborácie a participácie, pri informačných interakciách s elektronickým informačným prostredím. Digitálna veda je však všeobecnejší a viac využívaný pojem, ktorý sa objavuje aj v niektorých víziách európskych výskumných programov (Digital Science 2013). Na európskej úrovni sa týmto problémom zaoberá systém CRIS (Current Research Information Systems) s formátom CERIF (Common European Research Information Format) na vytváranie repozitárov vo výskume.

Kontexty digitálnej vedy však nesúvisia len s vplyvom informačných a komunikačných technológií. Ukazuje sa, že veľkými výzvami sú skôr organizačné, právne, ekonomické a sociálne kontexty pri komunikácii poznatkov v jednotlivých

disciplínach. Najvýraznejšími zmenami v reťazci komunikácie sú snahy o rýchlosť publikovania v určitých disciplínach – napríklad vo fyzike sa najviac využívajú preprinty a eprinty, v chémii a molekulárnej biológii články z časopisov a v informatike dominujú konferenčné publikácie. Do hry však vstupujú aj nové zdroje a hráči – agregátori/integrátori, webové stránky vydavateľov (napr. Cambridge University Press), preprintové služby a producenti elektronických zdrojov (napr. Elsevier's Chemistry Preprint Services). Digitálny repozitár arXiv.org sa delí na časti – pre astrofyziku – astro-ph, pre matematiku či informatiku (arXiv.org 2014). Digitálnu vedu môžeme vymedziť v troch vrstvách: sociálno-komunikačnej, obsahovej a technologickej, ktoré spolu tvoria zložité socio-technologické systémy.

## Piliere a zložky digitálnej vedy v kontexte spracovania informácií

Zmeny tradičného výskumu pod vplyvom elektronického prostredia znamenajú aj zmenu tradičnej kultúry práce s informáciami. Pritom nejde len o budovanie infraštruktúry a zodpovedajúcich služieb s pridanou hodnotou vrátane inštitucionálnych digitálnych repozitárov. Dôležitým pilierom je práve manažment digitálnych dát rôznych druhov vo veľkom rozsahu a na základe toho vývoj nových modelov, metód a nástrojov na spracovanie informácií. Vo väčšine prípadov sú tieto nové metódy poznačené kolaboráciou rôznych disciplín a rôznymi interakciami medzi spoločnosťou a vedeckou komunitou. Už dávnejšie sa tiež zdôrazňuje, že noví informační profesionáli majú byť súčasťou takýchto tímov pri ich podpore vo vzájomnej komunikácii a kolaborácii. Prvým vizionárom bol práve Vannevar Bush, ktorý predstavil „Memex“ ako hľadanie nových metód komunikovania, spracovania, prepájania a rozširovania vedeckých informácií už v roku 1945.

Z hľadiska repozitárov sa v súčasnosti zdôrazňuje najmä otvorený prístup k publikáciám v elektronickom prostredí, ale aj modely otvoreného výskumu obsahujúce napríklad softvérové nástroje a akademické portály či kolaboratória. Otvorený prístup znamená možnosť voľného prístupu k využívaniu digitálnych zdrojov, najmä k recenzovaným článkom elektronických vedeckých časopisov pre výskum a vzdelávanie. Všeobecne ide o minimálne obmedzenia pri prístupe k dátam, zdrojom a publikáciám. Celé hnutie open access sa vynorilo v polovici 90. rokov 20. stor. a zakladá sa na Budapeštianskej iniciatíve otvoreného prístupu (BOAIS), ktorá podporuje slobodu vedeckej komunikácie a nové modely fungovania vedeckých inštitúcií a publikovania, napríklad aj v rámci programu Scholarly Publishing and Resource Coalition (SPARC) ako súčasť Asociácie amerických vedeckých knižníc.

Na úrovni sociálnych aktérov sa v rámci digitálnej vedy rozlišujú takí aktéri ako výskumníci a výskumné organizácie, ale aj vedecké spoločnosti a sponzori, riadiace orgány, grantové agentúry, občianske a vzdelávacie inštitúcie, priemysel a firmy podporujúce výskum a vedu. Významní teoretici vedy, ako je Bruno Latour alebo H. Etkowitz, zdôraznili význam pozorovania vedy v akcii a modely vzťahov v triáde – trojitej závitnici: akademická sféra, priemysel, štátna správa (Suchá 2011). Robert Merton, známy sociológ vedy, vysvetlil možnosť koexistencie konkurencie a kooperácie vo vedeckej komunikácii a je zakladateľom teórií sietí vzťahov medzi vedcami a citácií až po normy praxe a stratégie vedeckého výskumu (Borgman 2007).

Na úrovni výskumných procesov sa zdôrazňuje premena týchto výskumných a informačných procesov, prepájanie vedy s globálnymi informačnými systémami, prepájanie vedy a spoločnosti (občanov) a prepájanie vedy a inovácií v nových partnerstvách firiem a inštitúcií (Digital Science 2013).

Pri transdisciplinárnych prepojeniach vied, v rámci ktorých má svoju úlohu aj informačná veda, je však potrebné vychádzať zo špecifik a odlišností jednotlivých vied. Tieto odlišnosti sa týkajú dokumentov a informačných artefaktov, ale aj organizácie informácií a informačných postupov. Výrazné odlišnosti súvisia aj s motiváciou sociálnych, prírodných a technických vedcov pri budovaní obsahu a pri zdieľaní informácií vo virtuálnych priestoroch. Rozdiely vychádzajú práve z metodologickej odlišnosti technických a prírodných a sociálnych a humanitných vied. To sa potom prejavuje aj v budovaní samotnej infraštruktúry a v hodnotení výsledkov/výstupov týchto vied. Tieto odlišnosti mapuje aj bibliometria a webometria a dokonca ich dobre vidieť aj v používanej terminológii. Prvé úspešne fungujúce tezaury sa realizovali práve v tých disciplínach, v ktorých sa terminológia tak rýchlo nemenila (napr. tezaurus MeSH či INSPEC). Preto aj prvé inteligentné a expertné systémy modelovali poznatky z oblasti medicíny či geológie.

Informačná veda má dlhú tradíciu v organizácii poznania a delení štruktúr poznania v jednotlivých disciplínach (Fry 2013). Ako príklad môžeme spomenúť delenie disciplín podľa štruktúr poznania na čisté a aplikované a „tvrdé“ a „mäkké“, ktoré tvoria dve osi delenia vied podľa Bechera (Fry 2013). K čistým a „tvrdým“ disciplínam sa priraduje fyzika a chémia, k čistým a „mäkkým“ sociológia a ekonómia, medzi tvrdé a aplikované patria napríklad biomedicínske vedy a strojárstvo a medzi mäkké a aplikované vzdelávanie a obchod. Iný spôsob delenia disciplín podľa stupňov vzájomnej závislosti a neurčitosti problémov rozlišuje konceptuálne integrované byrokracie (napr. fyzika), rozčlenené byrokracie (napr. ekonómia) a profesionálne štandardizované disciplíny (napr. biomedicínske vedy) (Fry 2013 podľa Whitleyho).

Rôznorodosť manažmentu záznamov a dát môže spôsobovať problémy s identifikáciou dát (kedy a ako vznikajú dáta?). Inak sa meria kvalita vedy v environmentálnych vedách, inak sa merajú štatistické a demografické údaje o správaní človeka. Tu sa niekedy rozlišujú tzv. „vysoké“ disciplíny (napríklad molekulárna biológia), v ktorých sa používajú osvedčené tradičné a štandardné nástroje na spracovanie a manažment dát a „nižšie“ disciplíny ako sociológia, informatika či literárna veda, v ktorých autori spolu s výskumom vyvíjajú aj nové metódy a nové nástroje (Borgman 2007). Pritom odlišnosti v odboroch vznikajú už na začiatku výskumu pri formulovaní problémov. Napríklad problémy možno

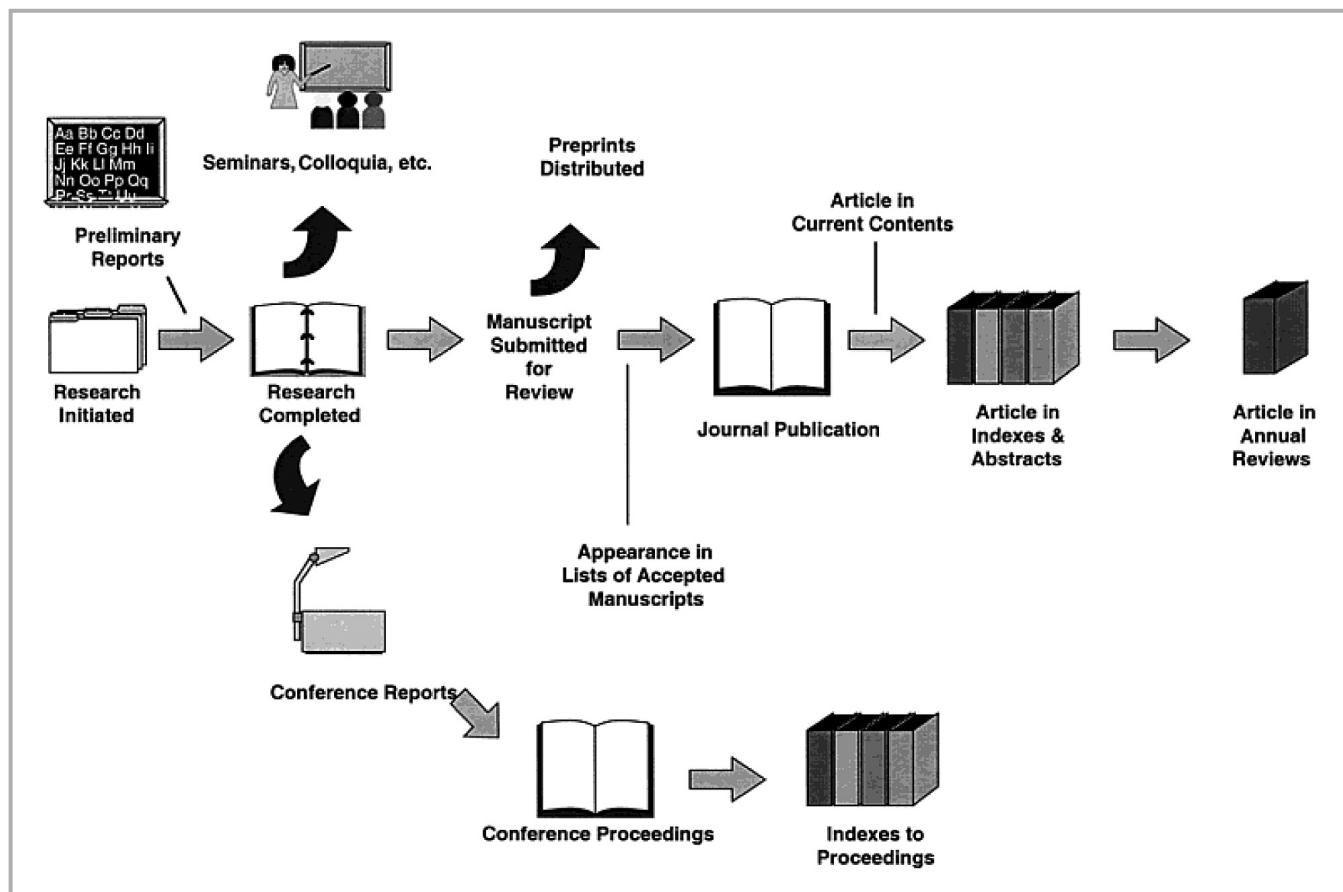
deliť na zle štruktúrované (najmä v humanitných a sociálnych vedách) a vysoko formálne štruktúrované (v prírodných a technických vedách). Vo výskumoch informačnej vedy sa však dokazuje práve integračná úloha informácií a informačných profesionálov.

Situácia sa môže skomplikovať, keď sa tie isté dáta využívajú na iné účely a pre inú komunitu či disciplínu. To je pritom jedným zo základných princípov ekologického využívania digitálnych dát. Napríklad hodnotenie kvality vody môže byť inak interpretované v odbornom environmentálnom časopise ako v interpretácii vládnych orgánov. Štatistické demografické údaje možno využiť na mnohonásobné interpretácie podľa typu otázky a uhla pohľadu. Odlišnosti v dátach sa môžu týkať aspektov ich marketingovej hodnoty alebo citlivosti (napríklad dáta vedúce k patentom alebo seizmologické dáta). V oblasti medicíny, farmácie alebo v elektronickom zdravotníctve vzniká množstvo etických dilem pri spracovaní citlivých dát o ľuďoch.

V procesoch vedeckej komunikácie tiež možno vidieť zmeny vo vzniku nových žánrov a zdrojov, najmä laboratórne webové stránky, vedecké blogy, vydateľské webové sídla, servery elektronických preprintov, diskusné fóra či open access sídla/portály alebo kolaboratória. Vo všetkých prípadoch ide o novú interaktívnu funkčnosť digitálnych knižníc ako infraštruktúry pre digitálnu vedu, tzv. interaktívne „obývané“ informačné priestory. To má veľký vplyv na zvyšovanie dostupnosti vedeckých dokumentov, ale aj na citovanie a nové indikátory naznačujúce nové možnosti bibliometrie transformovanej do webometrie (Thelwall 2009). Komunikačné princípy vedeckej komunikácie sa často opierajú o preberanie súborov (článkov) a online publikovanie článkov v repozitároch a iných digitálnych priestoroch.

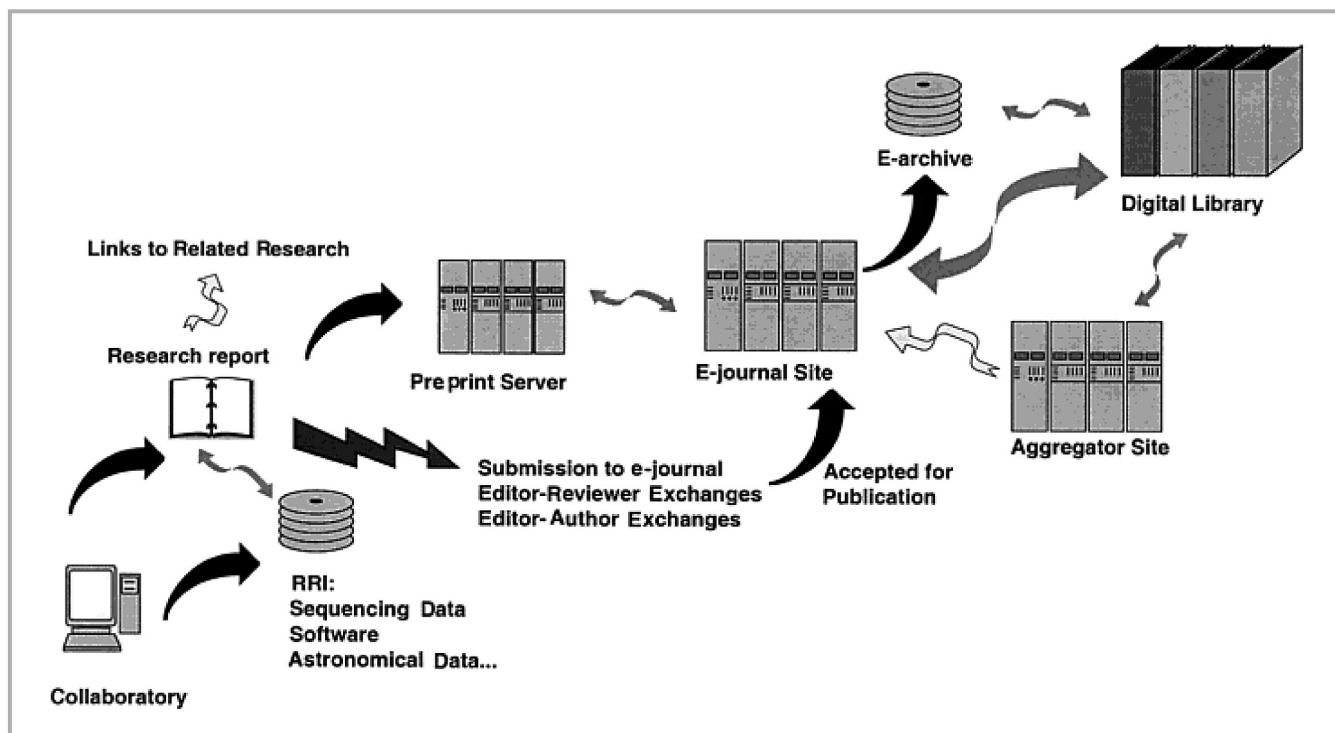
## Modelovanie vedeckej komunikácie a digitálna veda

Pôvodný model vedeckej komunikácie priniesli autori Garvey a Griffith v 70. rokoch 20. stor. na základe výskumov informačného správania a vedeckej komunikácie psychológov (obr. 1) (Hurd 2000). Piliermi modelu vedeckej komunikácie sú správy a projekty, konferencie (zborníky) a vedecké časopisy s dôrazom na procesy kolaborácie (sieťovanie, recenzovanie).



Obr. 1 Zjednodušený model vedeckej komunikácie podľa Garvey – Griffitha (Hurd 2000)

Všeobecný model zmien vo vedeckej komunikácii neskôr sformulovala napr. Hurdová (Hurd 2000). Tento model sa často cituje ako základ pre uvažovanie o digitálnej vede a je naznačený na obr. 2.



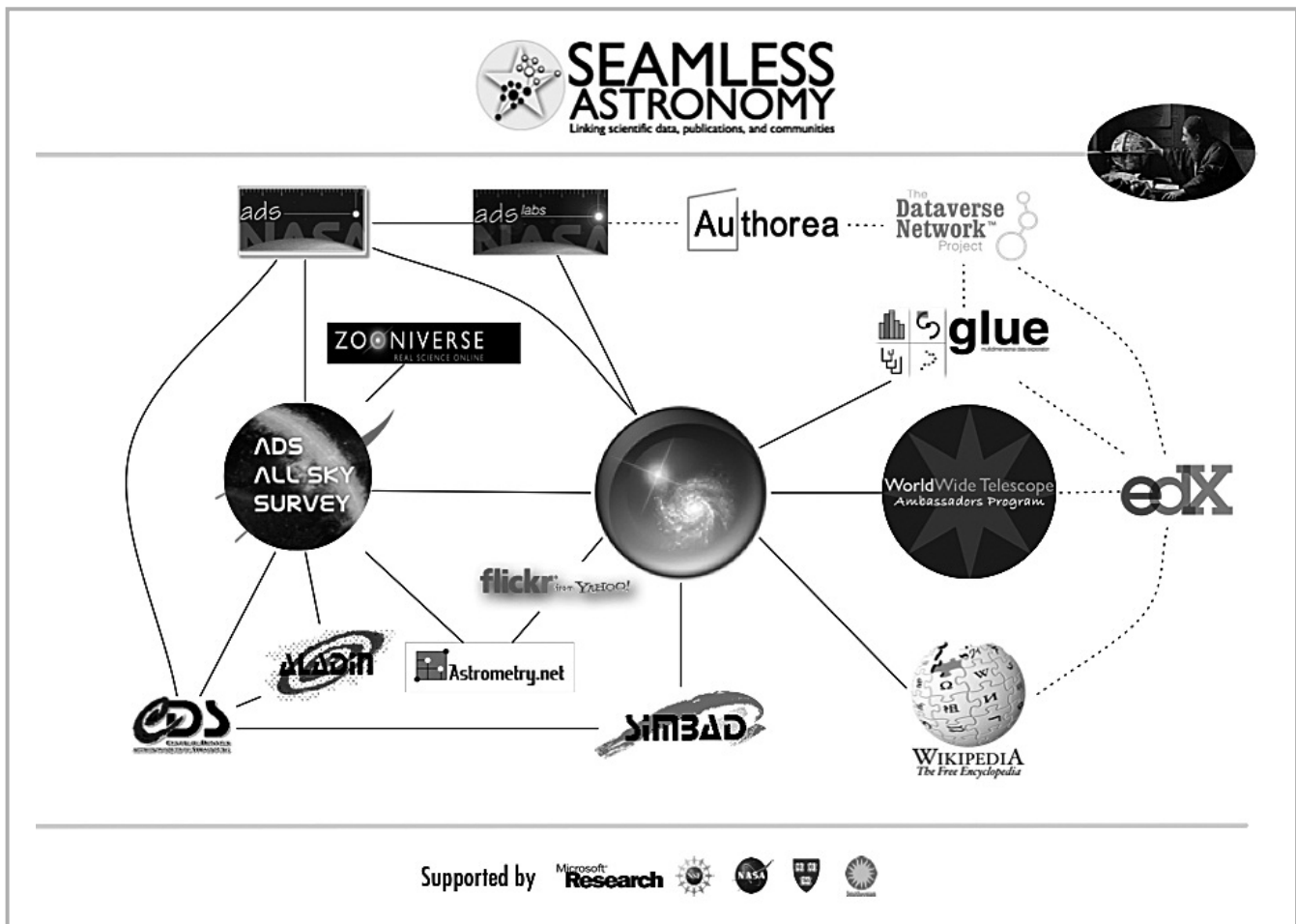
Obr. 2 Model vedeckej komunikácie (Hurd 2000)

V novších modeloch, ako napr. v modeli Hurdovej, vidieť aj úlohu digitálnych knižníc a oddelenie množstva dát od ich využitia v produkcii vedeckých poznatkov. Zložitost týchto modelov je spôsobená prepájaním poznatkov o informačnom správaní vedcov a vedeckej komunikácii s informačnými systémami. Tak vznikajú zložité sociotechnické systémy ako súčasť digitálnej vedy a súvisia aj so sociálnou informatikou zameranou na vplyv technológií na sociálne komunikačné procesy. Prepojené dáta, tzv. LinkedData, znamenajú ďalší rozvoj webu smerom k prepojeniam a kontextu. Pôvodná koncepcia vznikla vo vedeckej komunite (Hall 2014). Linked Open Data je nová paradigma podporujúca publikovanie, zdieľanie dát a znalostí vo webe, pričom sa vytvára štruktúra dát z prepojení zdrojov a tým sa zabezpečuje aj integrácia dát pre znalostné systémy.

Mnohí autori z oblasti informačnej vedy priniesli koncepcie budovania novej informačnej infraštruktúry a kolaboratórií pre vedeckú komunikáciu (napr. Atkins et al. 2003, Sonnenwald 2013, Brown 2010, Lynch 2008, Lassi and Sonnenwald 2012). V zložitom systéme sa prepájajú aspekty merania a hodnotenia výsledkov (bibliometria, citačné indexy), komunikácie v disciplínach, informačného správania vedcov, využívania elektronických časopisov, online publikovania aj modelovania a navrhovania digitálnych knižníc.

Digitálne repozitáre a registre dát ako sociotechnické systémy na univerzitách a v disciplínach sú čoraz aktuálnejšie. Ešte v roku 1991 založil Paul Ginsparg online repozitár preprintov pre fyzikov, arXiv.org. Tradíciu majú aj CogPrints (psychológia a kognitívne vedy), RePec (ekonomické vedy) či PubMedCentral (v medicíne). S produkciou množstva dát v rôznych disciplínach sa začali vytvárať nielen repozitáre prác, ale aj repozitáre zhromaždených dát, ktoré sa vo výskume analyzujú. Tradíciu má chemická informatika (repozitár Chemical Abstracts Service, CAS). Tieto digitálne repozitáre sústreďujú veľké množstvá empirických dát z výskumov. Niektorí autori hovoria o novej revolúcii/paradigme vo vede (dátovo intenzívna 4. paradigma vedy) práve pri navrhovaní takýchto repozitárov a nástrojov a metód na prácu s nimi. V týchto činnostiach zohrávajú dôležitú úlohu noví informační profesionáli a nové metódy analýzy textu (text miningu). Príkladom môžu byť dáta a dátové banky z oblasti genetiky, pričom ešte pred ukončením projektov a publikovaním si výskumníci riešia otázky intelektuálneho vlastníctva analyzovaných dát. Príkladmi inštitucionálnych repozitárov sú napr. e-Scholar (Purdue University, escholar.lib.purdue.edu), Digital Resources Commons (OhioLINK, drc.ohiolink.edu) či MIT OpenCourseWare (web.mit.edu/ocw). Z európskych projektov sa ako príklad uvádza HumanBrainProject (<https://www.humanbrainproject.eu/>) alebo projekt Graphene (<http://graphene-flagship.eu/>). Príklady prepojených dát a digitálnych portálov sa objavujú aj v astronómii, napríklad systém Seamless Astronomy na Harvarde (obr. 3).

Príklady nájdeme aj v geológii (<http://geoknow.eu/Welcome.html>), ale aj v oblasti vzdelávania sa systémy e-learningu integrujú do webu Linked Open Educational Data (Hall 2014). Pre univerzity sa odporúča budovanie stratégií manažmentu veľkých dát z výskumov v repozitároch aj v súvislosti s princípmi informačnej ekológie a ekosystémov (Erway 2013, Lynch 2008, Research Data Management 2013, Steinerová et al. 2012).



Obr. 3 Príklad digitálneho vedeckého systému v astronómii (Seamless astronomy)  
<https://www.cfa.harvard.edu/~agoodman/seamless/>

## Informačné správanie vedcov a vedecká komunikácia

Zásadnejšie sa mení aj informačné správanie vedcov pri využívaní informácií (online čítanie článkov) a pri online publikovaní. Tu sa tiež formulujú nové modely akademického publikovania (Whitworth, Friedman 2009). Mnoho výskumov informačného správania sa začalo práve v oblasti prírodných a technických vied, najmä v súvislosti s explóziou vedeckej literatúry po 2. svetovej vojne. Neskôr, od 80. rokov 20. stor., sa objavili aj výskumy informačného správania vedcov v sociálnych a humanitných vedách. Palmerová skúmala poľnohospodárskych vedcov a identifikovala niekoľko štýlov informačného správania na základe psychometrických výskumov a rozhovorov (Case 2002). V iných výskumoch sa ukázalo, že geológovia a astronómovia najčastejšie využívajú profesionálne kontakty ako zdroje informácií (Case 2002). Rozdiel medzi prírodnými vedami a technickými vedami dokázali ďalšie výskumy. Napríklad prírodní vedci viac píšú (články) a technickí vedci viac čítajú, ale menej publikujú. Základom aj v digitálnom prostredí ostáva recenzovaný vedecký časopis. Rozsiahlejšie prieskumy Tenopirovej (Brown 2010, Fry 2013) napríklad dokázali, že v prírodných vedách prevažujú preferencie vo využívaní elektronických časopisov, ale astronómovia vo vyššej miere vyhľadávajú v elektronických databázach.

Zaujímavé poznatky z výskumov informačného správania vedcov sú obsiahnuté v mnohých modeloch informačného správania. Napríklad model Davida Ellisa vychádza z výskumov vedcov z prírodných aj sociálnych vied. V mnohých ďalších výskumoch informačného správania vedcov sa zdôrazňujú profesionálne kontakty a vedecký časopis ako základ vo využívaní informačných zdrojov (Case 2002). Sociometrické výskumy naznačili význam neformálnej komunikácie v pracovnom prostredí vedcov a úlohu vedúcich výskumníkov ako selektorov zdrojov, smerov a stratégií výskumu. Napríklad Erdelezová a Means (2005) skúmali kolaboráciu medzi prírodnými vedcami na University of Missouri-Columbia a potvrdili význam personálnej komunikácie, neformálnych interakcií a sieťovania vo výskume. Intelektuálne a sociálne väzby v informačnom správaní pretrvávajú aj v digitálnej vede.

Osobitné vzorce sa objavili v oblasti sociálnych a humanitných vied s mnohonásobnými analýzami a interpretáciami dát získaných napríklad z rozhovorov, s neustálymi prechodmi medzi centrálnymi a okrajovými tematickými časťami. Na-

príklad Foster (2005) na základe výskumov vedcov z humanitných vied vypracoval nelineárny model informačného správanie s dôrazom na dynamiku štádií (otvorenie, orientácia, konsolidácia) a holistický princíp práce s informáciami. Niektorí autori tiež dokazujú, že v humanitných vedách, napríklad v histórii, dominujú najmä kognitívne procesy vyvodzovania záverov a viacnásobnej interpretácie, v interdisciplinárnych oblastiach sa často prejavuje „preklad“ pojmov medzi disciplínami (Brown 2010).

Informačné správanie vedcov v elektronickom prostredí najviac ovplyvňujú možnosti využívať otvorený prístup k elektronickým časopisom a článkom. K novším procesom sa priraduje online publikovanie a zdieľanie dát a spolupráca v kolaboratóriách. Výskumy tu tiež dokazujú zaujímavé vzorce, súvislosti aj prepojenia a odlišnosti medzi odborními. Zatiaľ čo vo fyzike sa filozofia eprintov (preprintov) výrazne osvedčila a dominuje využívanie kvalitných článkov sprístupnených ešte pred recenzovaním a publikovaním, v chémii fungovala služba preprintov len obmedzene (Elsevier's Chemistry Preprint Server) (Brown 2010). Dlhá tradícia medicíny podporujúca online sprístupňovanie publikácií prostredníctvom open access zabezpečuje najvyššie počty citácií v experimentálnej medicíne. Tiež je zaujímavé, že online konferenčné príspevky z informatiky získavajú až o približne 300 percent vyššiu citovanosť ako iné publikácie (citačná analýza Lawrenca) (Brown 2010).

Využívanie novších analytických nástrojov na odkrývanie vzorcov správania vedcov, prepojení, citácií, ale aj počtov prevzatých článkov a „čítaní“ vedie k novým výzvam pre digitálnu vedu pri hľadaní nových modelov vedeckej komunikácie aj hodnotenia jej produkcie. Komunikácia cez prepojenia webových dát vytvára nové možnosti aplikácií webometrie. Trendom sú analýzy citácií v Google Scholar a Scopus (Elsevier), ale aj špecializované citačné databázy ako CiteSeer pre informatiku a vizualizácia domén poznania z týchto analýz (Thelwall 2009). Objavuje sa napríklad kolaboratívne tagovanie či zdieľanie rôznych druhov dát a tzv. „reťazové interakcie“: čítanie, citovanie, komentovanie, hodnotenie, prieskum, extrakcia, verifikácia ap. Zaujímavým príkladom je napr. systém kolaboratívneho tagovania Connotea.

Z podrobnejších bibliometrických analýz vyplynuli aj obmedzenia dvoch najčastejšie používaných faktorov hodnotenia významu vedeckých časopisov – IF – impakt faktora, faktora vplyvu časopisu (alebo JIS – Journal Impact Factor) a II (immediacy index, indexu rýchlosti citácií). Tieto indexy vznikli v Inštitúte vedeckých informácií (the Institute of Scientific Information), založenom E. Garfieldom v 60. rokoch 20. stor., v súčasnosti Thomson Reuters. IF identifikuje frekvenciu citovania priemerného článku z určitého časopisu v danom roku, II meria, ako často je priemerný článok citovaný v tom roku, keď bol publikovaný. Obmedzenia týchto faktorov sa týkajú práve odlišných citačných postupov a práce s informáciami v odlišných odboroch (Brown 2010). Faktory vplyvu a iné indikátory (impakt faktory, h-index a iné) preto treba považovať len za jednu zo súčastí komplexného posudzovania kvality vedeckej činnosti v jednotlivých odboroch (Thelwall 2009).

Pri informačnom správaní vedcov je tiež zaujímavé zisťovanie vplyvu obsahu článku na citovanie. Tu sa zistilo, že najviac citovanými článkami boli práve metodologické články v odboroch ako biológia, chémia či astronómia. V niektorých odboroch však väčšiu úlohu ako prístup k vedeckej literatúre zohrávajú obmedzenia v prístrojovom či materiálom vybavení. Aj keď niektoré výskumy dokazujú význam online publikovania pre zvyšovanie viditeľnosti výskumu a počtu citácií, mnohí zdôrazňujú dôležitosť kontroly kvality časopisov prostredníctvom recenzovania v elektronickom prostredí.

Špeciálnym procesom digitálnej vedy je práve zdieľanie informácií, dát a výsledkov výskumov, pričom sa prelína neformálna aj formálna komunikácia, napríklad výmena nástrojov, diskusie nápadov, laboratórne záznamy, názory kolegov a expertov. Tieto komunikácie a interakcie sa vyznačujú voľnejšou štruktúrou v otvorenom sieťovom priestore, v ktorom môžu vznikať nové projekty, kontakty, neformálne posudzovanie výsledkov. V otvorenosti a kolaborácii vznikajú aj nové metódy analýz, dolovania poznatkov a vizualizácie, zdieľania nástrojov a zdrojov, publikovania a vedeckej metricky. Otvorenosť vedie aj k lepším možnostiam riešiť zložité globálne transdisciplinárne problémy ako klimatické zmeny, finančné krízy, energetická bezpečnosť a iné.

## Informačné interakcie a digitálne sociálne a humanitné vedy

Informačné interakcie v informačnej vede znamenajú informačné aktivity pri spracovaní a využívaní informačných zdrojov a produkcii nových poznatkov. Informačná veda hľadá nové modely informačných interakcií na podporu digitálnej vedy. Významné centrá digitálnej vedy sa historicky sformovali v USA a UK. V oblasti digitálnych sociálnych vied je to najmä centrum v Oxforde – Oxford e-Social Science Project v rámci Oxford e-Research Centre (2014).

V informačnej vede a praxi sa čoraz viac pozornosti venuje práve modelovaniu informačných interakcií v oblasti digitálnych sociálnych a humanitných vied a kultúry. Interpretačný a individualistický prístup v práci s dátami naznačuje preferencie prezerania, analýz, reťazenia a autorského filtrovania. Často sú digitálne objekty priamo prepojené s analýzami a môžu viesť k novým príbehom, vzorcom alebo modelom (napríklad digitálna knižnica Perseus z oblasti archeológie). Sociálne vedy sa zameriavajú na skúmanie správania človeka, dáta sa získavajú kvantitatívnymi aj kvalitatívnymi metódami (experimenty, rozhovory, prieskumy, pozorovania). Analýzy sú poznačené zložitou kontextov a snahou premeniť implicitné informácie na explicitné poznatky. Preto sú potrebné špeciálne nástroje na organizáciu informácií, opis informačných objektov, zdieľanie informácií. V digitálnych sociálnych a humanitných vedách možno dáta rozdeliť na dva druhy – (1) vlastné dáta získané originálnymi nástrojmi z prieskumov, experimentov, pozorovaní a (2) dáta dostupné z vládnych a iných inštitúcií (demografické, ekonomické, sčítania obyvateľov – bývanie, vzdelanie a i.). Tieto dáta súvisia aj so špeciálnymi etickými opatreniami pri ich analýzach. Okrem toho sa pri ich spracovaní musia využívať procesy čistenia, analýz, verifikácie a ochrany súkromia a manažment autorského práva.

Mnoho kontextuálnych informácií sa skrýva aj v spracovaní archívnych zdrojov či špeciálnych digitalizovaných artefaktov. Nejasnosti pri zdieľaní a agregácii takýchto digitálnych objektov sa týkajú nielen autorských práv, ale aj tradície publikovať v sociálnych a humanitných vedách skôr interpretácie ako dáta či digitálne objekty. Rozdiely sú aj medzi repozitármi v sociálnych vedách. Pri repozitári CogPrints v psychológii sa využíva samoarchivácia textov príspevkov, pri repozitári RePec v ekonómii sa postupuje inštitucionálne a obsahuje len prepojenia na plné texty. Pritom v humanitných vedách vznikajú rôzne simulácie historických objektov, miest či udalostí (napr. repozitár obrazov v jaskyniach v Dunhuangu v Číne, the Mellon International Dunhuang Archive ako súčasť Artstor Digital Library) (Borgman 2007). Pre modelovanie objektov v čase a priestore sa tu tiež využívajú geografické informačné systémy.

Informačné interakcie v digitálnych sociálnych a humanitných vedách vedú k používateľským inováciám a podpore kreatívneho procesu v priebehu využívania dát a digitálnych objektov. Aj keď existujú zaujímavé slovníky v disciplínach ako psychológia, sociológia a pedagogika, terminologická práca je v týchto vedách zložitejšia ako v prírodných vedách či medicíne kvôli interdisciplinárnymi presahom. V analýzach sociálnych dát možno využívať metódy text miningu a digitálne technológie na predpovedanie politických trendov či tvorbu sociálnych a kultúrnych modelov.

## Vybrané výsledky výskumu doktorandov

Výsledky výskumu informačného správania doktorandov v rámci širšieho projektu modelujúceho vybrané kognitívne aspekty webu potvrdzujú rozdiely v metódach aj postupoch jednotlivých disciplín. Podrobnejšie informácie o tomto kvalitatívnom výskume boli vysvetlené v iných prácach (Steinerová 2013). Na identifikáciu informačného správania vybraných doktorandov sme využili pološtruktúrované rozhovory s 19 doktorandmi z UK a STU. Obsahovali 28 otázok v niekoľkých aspektoch zameraných na výskumné správanie, informačné stratégie, typy zdrojov, triedenie zdrojov, využívanie sociálnych médií a publikovanie. Výsledky analýz naznačujú, že odlišnosti v informačných potrebách vyplývajú z odlišností v metódach a type prieskumu (experiment, teória, simulácia). Spoločnou motiváciou je prirodzená zvedavosť, overovanie informácií, riešenie problémov a argumentácia. Sociálne médiá sa využívajú najmä pasívne, ale aj na zdieľanie výsledkov či distribúciu dotazníkov, pritom sa potvrdilo prepájanie formálnej a neformálnej komunikácie. Často vznikajú aj nové typy informačných produktov pri opätovnom využívaní dát, nástrojov a stratégií. Potvrdila sa aj potreba podpory mladých vedcov pri budovaní terminológie a metodologických znalostí a tiež pri sieťovaní s expertmi. V informačnom správaní v elektronickom prostredí v sociálnych sieťach sa zaujímajú o nové uhly pohľadu na problémy (inšpirácie) a o overené najlepšie postupy (praktické znalosti a skúsenosti kolegov).

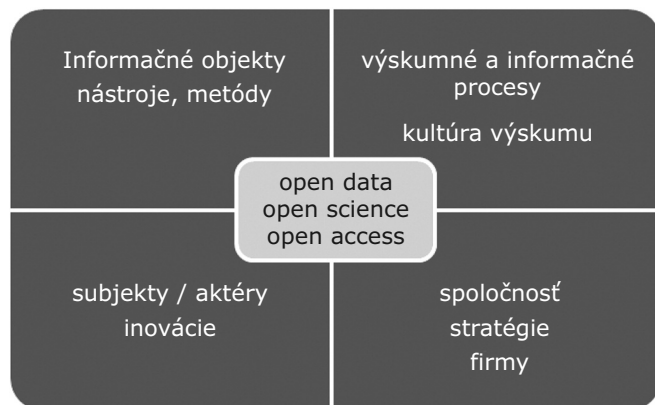
Model informačnej podpory informačných interakcií doktorandov (Steinerová, Hrkčková 2014) kladie dôraz na objavovanie znalostí pri práci s informačnými zdrojmi a objektmi ako súčasťou vedeckej práce doktorandov. Obsahuje procesy sociálneho sieťovania a zdieľania informácií, projektového a informačného manažmentu, reprezentácie poznania a vizualizácie, analýzy a syntézy, kritického posudzovania a interpretácií a prezentácií. Na základe toho sme sformulovali odporúčania na vytváranie špeciálnych informačných služieb s pridanou hodnotou pre doktorandov na princípoch dostupnosti, viditeľnosti a vhodnosti. Tieto služby môžu obsahovať metodologické nástroje na manažment informácií a projektov, terminologický rozvoj, sociálne sieťovanie, analýzy a syntézy. Na základe toho možno určiť metodologickú gramotnosť doktorandov (Steinerová 2013). Odporúčame tiež vytváranie novších produktov s pridanou hodnotou s interaktívnymi prvkami, či už tradičnejších (blogy, wiki, komentáre, recenzie), alebo nové reprezentácie dát v simuláciách, znalostných mapách alebo personalizovaných agregovaných produktoch pre jednotlivých vedcov podľa témy záujmu.

## Záver

Digitálna veda je dôležitou súčasťou výskumu v rôznych projektoch a iniciatívach aj v rámci rozvoja teórie a praxe informačnej vedy. Okrem technologickej infraštruktúry sú pre informačnú vedu zaujímavé najmä modely, metódy a nástroje práce s informáciami. Prínosom sú aj bibliometrické a webometrické analýzy. Informačná veda môže zohrať úlohu aj pri zlepšovaní manažmentu digitálnych dát, budovaní digitálnych repozitárov a podpore open access prístupu k zdrojom. Množstvo dát v rôznych formách (obrazy, senzorické dáta, simulácie, monitorovanie online správania, meteorologické dáta a i.) si vyžaduje špeciálne modely ich neskoršieho spracovania a integrácie v sociálnych kontextoch a v nových nástrojoch (virtuálne experimenty, robotika). To môže viesť k novým výskumným otázkam, ako napr. biodiverzita či genetický výskum. Z výskumov informačného správania vedcov sa môžu vynoriť novšie modely formovania vedeckej gramotnosti a interakcií vedy a spoločnosti.

Konceptuálna schéma digitálnej vedy je naznačená na obr. 4. Zobrazuje zložitú interakciu medzi informačnými objektmi, nástrojmi a metódami, ktoré využívajú rôzni aktéri (vedci, tímy, inštitúcie, knižnice, vydavatelja). Títo aktéri sú v neustálej interakcii so spoločnosťou aj firemnou sférou pri realizácii výskumných a informačných procesov v akademickom informačnom prostredí. Napriek množstvu novších neformálnych interakcií v sociálnych médiách základom vedy v digitálnom prostredí ostávajú publikácie a dáta (obsah), preto ako princípy digitálnej vedy naznačujeme v strede open access k publikáciám a dátam a princíp open science – otvorenej vedy v spoločenskom kontexte.

Digitálna veda znamená nielen návrh zložitých socioekonomických systémov a nástrojov vedeckej komunikácie, ale aj premenu vedy od hierarchickej organizácie poznania smerom k horizontálnym interdisciplinárnym prepojeniam. V nich sa vynárajú nové disciplíny a metódy prostredníctvom otvorenosti zdrojov a nástrojov. Objavujú sa aj nové formy kolaborácie a vedeckej metriky s novými publikačnými kanálmi pri verifikovaní a viacnásobnom využívaní dát a výsled-



Obr. 4 Digitálna veda – zložky a princípy

kov. Pritom sa predpokladá otvorenie vedy smerom k participácii budúcich výskumníkov vo vedeckých komunitách. Rozdiely v disciplínach vidieť už teraz, pritom za najprogressívnejšiu oblasť v rozvoji digitálnej vedy sa považuje astronómia, sociálne vedy sú zas na druhej strane spektra rozvoja. Všeobecne sa zdôrazňuje potreba zmien v kultúre správania výskumníkov, vedeckých organizácií aj súkromných firiem s podporou nových metód, participácie spoločnosti a kolaborácie.

Zložitosť digitálnej vedy ako sociotechnického systému je výzvou pre informačnú vedu. Problémom nie je ani tak technologická infraštruktúra ako množstvo rôznych záujmov a informačných interakcií medzi rozličnými aktérmi, ktorí súperia o zdroje (vedci, univerzity, vydavatelia, producenti zdrojov, grantové agentúry, knižnice). Hľadanie rovnováhy medzi týmito aktérmi je skôr sociálnou a strategickou otázkou riadenia dôvery, identity, lokácie či intelektuálneho vlastníctva. Obsah je tvorený práve dátami

a publikáciami a tu ostáva významná úloha pre nových informačných profesionálov, ktorí môžu využiť svoju dlhoročnú expertízu v organizácii zdrojov, metrikách publikácií a citácií a poznaní informačného správania používateľov.

## Literatúra

- arXiv.org (2014). Cornell University Library. [online]. [cit. 2014-02-23]. Dostupné na: <http://arxiv.org/>
- ATKINS, D.E., DROEGEMEIER, K.K., FELDMAN, S.I., GARCIA-MOLINA, H., KLEIN, M.L. a P. MESSINA, P. (2003). *Revolutionizing Science and Engineering through Cyberinfrastructure*: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Washington, DC: NSF. [online]. [cit. 2014-02-23]. Dostupné na: <http://www.nsf.gov/cise/sci/reports/atkins.pdf>
- BORGMAN, Ch. (2007). *Scholarship in the Digital Age. Information, Infrastructure and the Internet*. Cambridge: MIT Press 2007. 336 s.
- BROWN, C. (2010). Communication in the Sciences. Chapter 7. In *Annual Review of Information Science and Technology*. ARIST. Vol. 44. Ed. B. Cronin. Medford, N.J.: Information Today 2010, 287-316.
- BURGELMAN, J.-C., OSIMO, D. a M. BOGDANOWICZ. (2010). Science 2.0 (change will happen...). *First Monday* [online]. 15 (7). [cit. 2014-02-23]. Dostupné na: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2961/2573>.
- CASE, D. (2002). *Looking for information: a survey of research on information seeking, needs and behaviour*, New York: Academic Press.
- DIGITAL Science in Horizon 2020. (2013). [online]. 14 s., 2 Annexes. [cit. 2014-02-23]. Dostupné na: <http://ec.europa.eu/programmes/horizon2020/>
- ERDELEZ, S. a T. MEANS. (2005). Measuring changes in information sharing among life science researchers. In *Knowledge Management: Nurturing Culture, Innovation and Technology*. Proc. of the 2005 ICKM. Ed. S. Hawamdeh. New Jersey, World Scientific. Publ. 2005, 29-40.
- ERWAY, R. (2013). *Starting the Conversation: University-wide Research Data Management Policy*. [online]. Dublin, Ohio: OCLC Research. [cit. 2014-02-23]. Dostupné na: <http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf>.
- FIDEL, R. (2012). *Human Information Interaction: An Ecological Approach to Information Behavior*. Cambridge, MIT Press.
- FOSTER, J. (2005). Nonlinear Information Seeking. In *Theories of information behavior*. Ed. K. Fisher, S. erdeleu, I. McKechnie. Medford: information Today 2005, 254-258.
- HALL, W. (2014). Linked Data: The Quiet Revolution. In *ERCIM News*. Special Theme: Linked Open Data. [online]. No 96, Jan. 2014, s.4. [cit. 2014-02-23]. Dostupné na: <http://ercim-news.ercim.eu/>
- HURD, Julie (2000). The Transformation of Scientific Communication: A Model for 2020. In *JASIST*, 51 (14), 2000, 1279-1283.
- FRY, J. (2013). Considerations in adopting a disciplinary analysis of scholarly communication and information behaviors. In *ASIST European Workshop 2013*. Proc. of the 2nd ASIST European Workshop. June 5-6, Turku/Abo 2013. Ed. I. Huvila. Abo, Abo Academi University 2013, 61-78.
- KROLL, S. a R. FORSMAN. (2010). *A Slice for Research Life: information Support for Research in the United States*. [online]. Dublin, Ohio: OCLC. [cit. 2014-02-23]. Dostupné na: <http://www.oclc.org/research/publications/library/2010/2010-15.pdf>



- LASSI, M. a D.H. SONNENWALD. (2013). The socio-technical design of a library and information science collaboratory. *Information Research*, [online]. **18**(2) paper 576. [cit. 2014-02-23].  
[Dostupné na: <http://InformationR.net/ir/18-2/paper576.html>]
- LYNCH, C. (2008). The Institutional Challenges of Cyberinfrastructure and e-research. In *EDUCAUSE Review* [online]. 2009, 43 (6). [cit. 2014-02-23]. Dostupné na:  
<http://connec.educause.edu/Library/EDUCAUSE+review/TheInstitutionalChallenge/47446>
- SCIENCE 2020. [online ]. Microsoft Research. Stephen Emmott et al. Cambridge, UK: Microsoft Corporation, 2006. . [cit. 2014-02-23]. Dostupné na: <http://www.microsoft.com>
- SHNEIDERMAN, B. (2008). *Science 2.0*. [online ]. In: Science, Vol. 319, March 2008, 1349-1350. [cit. 2014-02-23]. Dostupné na: <http://www.sciencemag.org>.
- SONNENWALD, D. (2013). Visioning the Future of Information and Library Science: Challenges and Opportunities. In Hobohm, H.C. *Informationswissenschaft zwischen virtuellen Infrastruktur und materiellen Lebenswelten*. (ISI 2013), 22-34.
- STEINEROVÁ, J. (2013). Information interactions as part of digital scholarship. In *ASIST European Workshop 2013*. Proc. of the 2nd ASIST European Workshop. June 5-6, Turku/Abo 2013. Ed. I. Huvila. Abo, Abo Academi University 2013, 33-49.
- OXFORD e-Research Centre. (2014). [online]. [cit. 2014-02-23]. Dostupné na: <http://www.oerc.ox.ac.uk/>
- STEINEROVÁ, Jela et al. (2012). *Informačná ekológia akademického informačného prostredia*. Záverečná správa z výskumu VEGA 1/0429/10. Editor Jela Steinerová. Bratislava: Vydavateľstvo UK, 2012. 96 s.
- STEINEROVÁ, J. (2012). Information ecology – emerging framework for digital scholarship. *LIDA 2012*. . [online ]. Vol 12. [cit. 2014-02-23]. Dostupné na: <http://ozk.unizd.hr/proceedings/index.php/lida2012/article/view/66/37>
- STEINEROVÁ, J. (2013). Methodological Literacy of Doctoral Students – an Emerging Model. In: *Worldwide Commonalities and Challenges in Information Literacy Research and Practice. ECIL 2013*. Conference Proceedings. Ed. S. Kurbanoglu et al. Cham: Springer International Publishing 2013, 148-154. CCIS 397.
- STEINEROVÁ, J. a A. Hrčková. (2014). Information support of research information interactions of PhD. students in Slovakia. In *Fifteenth International Conference on Grey Literature: The Grey Audit*, 2-3 December 2013. Ed. D. Farace, J. Frantzen. Amsterdam: TextRelease, February 2014. (GL conference series, ISSN 1386 2316, No.15). (v tlači).
- SUCHÁ, L. (2011). *Nové modely vedeckého publikovania*. Rigorózní práce. Brno: MU 2011. 122s.
- RESEARCH Data Management: Principles, Practices and Prospects. (2013). [online ]. Washington, CLIR 2013. 99 s. [cit. 2014-02-23]. Dostupné na: <http://www.clir.org>
- THELWALL, M. (2009). Bibliometrics to webometrics. In *Information science in transition*. Ed. A. Gilchrist. London: Facet 2009, 347-376.
- WHITWORTH, B. & R. FRIEDMAN. (2009). Reinventing academic publishing online Part II: A Sociotechnical vision, *First Monday*. [online ]. Volume 14, Number 9, September 2009. [cit. 2014-02-23]. Dostupné na:  
<http://firstmonday.org/ojs/index.php/fm/article/view/2642>

*Príspevok bol spracovaný v rámci projektu APVV-0208-10 TraDiCe – Kognitívne cestovanie po digitálnom svete webu a knižníc s podporou personalizovaných služieb a sociálnych sietí a projektu „Univerzitného vedeckého parku“ Univerzity Komenského v Bratislave (ITMS 26240220086).*

prof. PhDr. Jela Steinerová, PhD.  
Katedra knižničnej a informačnej vedy  
Filozofická fakulta  
Univerzita Komenského

**RECENZENTI:**

prof. Ing. Josef Basl, CSc.  
prof. RNDr. Jiří Ivánek, CSc.