

Bude možné zvládnout proces zpracovávání vědeckých informací při jejich současném enormním nárůstu?

**Jaroslav Šilhánek a Ludmila Zetková**

Vysoká škola chemicko-technologická v Praze

[silhanek@vscht.cz](mailto:silhanek@vscht.cz)

INFORUM 2017: 23. ročník konference o profesionálních informačních zdrojích  
Praha, 30. -31. 5. 2017

### **Abstrakt**

Uvědomování si role vědy pro rozvoj lidské společnosti a její promítání jako součást státní politiky logicky vede ke stoupající produkci vědeckých informací a ke stále problematičtější otázce udržitelnosti zavedených mechanismů jejich zpracování do využitelné podoby. Instituce zabývající vytvářením referátových zdrojů jako klíčových instrumentů zveřejňování, zpřístupňování a uchovávání vědeckých informací čelí problému narůstajícího objemu primárních informací, což logicky vede k růstu nákladů na tuto činnost, zákonitě se projevující narůstajícími cenami pro konečné uživatele především v případě bibliografických bází dat. Není proto překvapující, že se setkáváme s řadou aktivit snažících se o regulaci touto či onou formou. Často se ale opomíjí skutečnost, že proces šíření, zpřístupňování a uchovávání poznatků vědy sice začíná jejich produkcí, ale cesta výsledků zpět k jejich potenciálnímu využití v dalším rozvoji vědy je vícestupňový proces, jehož jednotlivé kroky mají svůj vlastní mechanismus a problematiku. Moderní informační technologie sice naprosto zásadním způsobem zefektivnily tento proces, ale řada s ním spojených aspektů zůstává stále stejná bez ohledu na použitou technologii. Na počátku je nepochybně žádoucí jistá kontrola „kvality“, ale pokud takový krok umožní vstup informace jako trvalou součást lidského poznání, nastupuje požadavek, aby bylo možné se s takovými poznatky, seznámit se nejenom bezprostředně, ale i např. za 100 let. Referát se soustřeďuje na problematiku jednotlivých kroků procesu registrace vědeckých poznatků, a to především z hlediska jejich enormního nárůstu.

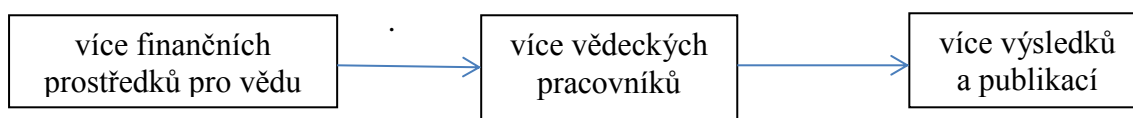
### **V čem je problém?**

Výše uvedené aspekty se v současné době stávají aktuální především vzhledem ke stále více zdůrazňované důležitosti podpory vědecké činnosti z veřejných prostředků, což sice asi nikdy a nikde nebylo zpochybňováno, ale vzhledem k celkovému vývoji vyspělé společnosti se stávají faktorem po výtce politickým. Zdůrazňování nutnosti vyšší finanční podpory vědy a výzkumu, což je dnes v programech naprosté většiny politických subjektů, je nepochybně záležitost velmi důležitá, ale má své důsledky, o kterých je logicky nutné uvažovat.

Důvodem pro položení otázky v názvu tohoto sdělení je snaha o upozornění na skutečnost, která je prakticky neustále zdůrazňována, že totiž objem vědeckých poznatků neustále narůstá, ale stala se určitým folklorem, nebo něčím, co prostě je a tak to zůstane. Problém je ovšem v tom, že jakkoliv je neustále objevování nových poznatků, souvislostí i vysvětlení

dříve nejasných faktů nepochybně vlastním smyslem vědecké a činnosti, je dále nutné tyto poznatky především zveřejňovat, což se děje, ale také zpracovávat do podoby záznamu o jejich existenci tak, aby byly nějakým způsobem vyhledatelné, a to nejenom v době jejich zveřejnění, ale za 50 a více let. A to je záležitost poněkud odlišná a představuje celou škálu konkrétních otázek a problémů, jak technických, tak logistických a hlavně nákladových, a tedy finančních. V této přednášce se pokusíme o určité zpřehlednění a především zdůraznění opomíjené skutečnosti, že cesta od vlastního nového poznatku k jeho zveřejnění a konečnému uložení do dlouhodobě a maximálně spolehlivě udržovaných archivů sestává z řady kroků, z nichž vlastní zveřejnění je jedním z nich. V dalších kapitolách se pokusíme je diskutovat sice odděleně, ale s cílem celkového výsledku.

Jak rychle vědecké poznatky narůstají, lze demonstrovat na řadě konkrétních numerických dat, zde se spokojíme s konstatováním, že celkový objem známých poznatků se v současné době přibližně zdvojnásobuje každých 10 let<sup>1</sup>! Vzhledem k výše zmíněným vyšším finančním dotacím bude evidentně platit jednoduchá relace:



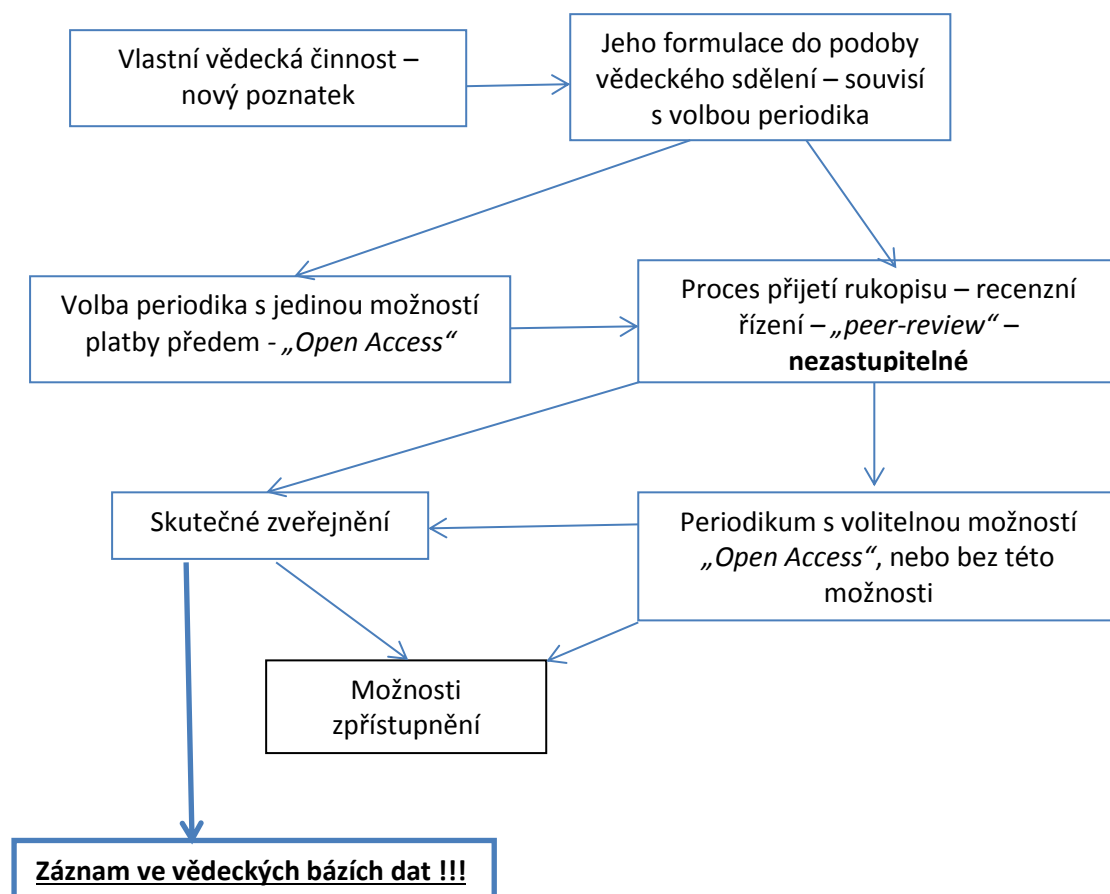
### **Cesta od nového poznatku k jeho zařazení do objemu lidského vědění.**

Na počátku musí být v pravém slova smyslu něco, co nebylo dosud známo, což v obecném slova smyslu by otevřelo obsáhlou a ne příliš smysluplnou diskusi. V každém případě je přirozenou snahou prakticky kohokoliv ve vědecké komunitě výsledky své činnosti dokumentovat jejich sepsáním a zveřejněním, ať už konkrétní důvody či motivace jsou jakékoliv. Cesta od vlastního nového poznatku začíná jeho formulací, předání do procesu jeho vydání, který nutně musí obsahovat jistý kontrolní mechanismus, typicky recenzi, přes přijetí a rozhodnutí o formě skutečného zveřejnění až po způsob, jak se k jeho úplně podobě dostane jiný zájemce. Tato cesta sice vypadá přímočaře, ale především není a ani nemůže být příliš rychlá a hlavně nemusí vždy vést k žádanému cíli. A také v sobě zahrnuje řadu vedlejších aspektů.

Tím nejdůležitějším je skutečnost, že v případě kladného výsledku celého publikačního procesu, tedy skutečnému zveřejnění vědeckého sdělení, musí být tato skutečnost zaregistrována jeho začleněním do vhodného sekundárního informačního zdroje, dříve referátových časopisů, dnes odpovídajících vědeckých bází dat. Pokud k tomu nedojde, je jen minimální šance, že se se zveřejněnými poznatky vůbec někdy někdo seznámí. A právě v tomto bodě je situace se současným enormním nárůstem zveřejňování nových poznatků v pravém slova smyslu kritická, protože je pracná a tudíž finančně nákladná. Je pravdou, že je možné se spokojit s automatizovaným vytvářením bibliografických záznamů do vědeckých bází dat z volně dostupných názvů dokumentů a textů abstraktů, ale zatím „ruční“, tedy vytváření bibliografických záznamů specialistou v daném oboru vede logicky k přesnějším záznamům, které se pak vrátí v možnostech relevantnějších rešerší a tedy časové úspoře bez ohledu na další aspekty. A v případě, kdy jsou přebírány z primárních dokumentů konkrétní data, ať již při vytváření specifických faktografických bází jako samostatných sekundárních zdrojích, nebo jako o součásti bází bibliografických, je role specialisty příslušného oboru

nepostradatelná. A doplňme, že se musí jednat o specialistu na vysoké úrovni s příslušným finančním oceněním.

Pokusme se schematicky načrtnout jednotlivé kroky celého procesu zveřejňování vědeckých poznatků pro následnou diskusi o jednotlivých krocích zvlášť.



## Rozhodnutí o publikování a volba periodika

Především je nutné si uvědomit, že rozhodnutí o zveřejnění svých poznatků je obecně individuální, pokud se nejedná o povinnost, která je součástí přidělení finančních prostředků. Ale v naprosté většině situací bude začínající vědec nějakým způsobem vyzván, např. svým učitelem, konsultantem, vedoucí výzkumné skupiny, katedry apod. Dále pak si bude vytvářet na základě postupně nabývaných zkušeností svůj pohled a orientaci ve světě publikačního průmyslu, přičemž důležitou roli bude hrát vlastní sebereflexe a tedy odhad, kam má smysl se pokoušet zaslat svůj rukopis. Zpravidla bude převládat snaha o publikaci v periodiku s pokud možno dobrým, až velmi dobrým renomé ve vědeckém světě, v současné situaci tedy s vyšším impakt faktorem. Realisticky uvažováno, otázka, zda zveřejněný dokument bude volně dostupný či bude jen součástí předplatného, nebude stát na prvním místě. Přesto bude dnes otázka, zda publikovat v režimu OA hrát důležitou roli proto, že už existuje relativně velký objem titulů periodik, které jinak, než v režimu OA rukopisy nepřijímají a s finálním přijetím rukopisu ke zveřejnění budou tedy spojeny finanční náklady. Do jaké míry k této

situaci přispěla relativní snadnost a i lákavost úhrady nákladů na zpracování a vydání rukopisů předem, tedy jako OA, včetně extensivní propagace, se můžeme jen dohadovat. V každém případě se tak vychází vstřícně pochopitelnému zájmu vědeckých pracovníků publikovat své výsledky.

V tomto směru lze litovat, že pravděpodobně jako jistá reakce na stoupající objem rukopisů se poskytovatelé finančních dotací přiklání k hodnocení na úrovni počtu publikovaných dokumentů, místo na jejich kvalitu. Jako tzv. indikátory hodnocení se dnes vyskytuje počet publikací v časovém období, což samozřejmě musí vést k tomu, že místo soustředění poznatků v rozsáhlejší publikaci, která argumentuje více a navzájem se doplňujícími fakty, publikuje vědec raději více kratších publikací, často se do jisté míry překrývající, nicméně snadněji splňující zadaná kritéria. Tento aspekt je plně v rukách grantových i jiných agentur a nelze je svalovat na snahu o seberealizaci samotných vědců. Připomeňme ještě, že počty publikací nehrají roli jen pro získávání dalších prostředků na vědeckou činnost, ale uplatňují se v celé řadě dalších procesů, jako je kariérní postup, získání pracovních posic, principiálně nejenom vědeckých a dalších situací.

Nelze-li omezovat objem publikovaných prací administrativně, zbývá jako nejdůležitější nástroj recenzní řízení, a tím je recenzní řízení, dnes běžně známé jako „*peer-review*“. Podívejme se na něj podrobněji.

### **Recenzní řízení - „*peer-review*“**

Pro hodnověrné recenzní řízení by měly být splněny následující podmínky:

- recenzent by měl být vědec, velmi dobře obeznámený s danou problematikou, prakticky tedy v úzkém, až velmi úzkém oboru,
- v žádném případě by neměl být jakkoliv zainteresován na kladném nebo záporném hodnocení,
- měl by být naprosto nezávislý
- v zájmu objektivity i nezávislosti by se recenzent a autor neměli navzájem znát, recenze by měly být anonymní,
- v žádném případě by oba neměli pracovat na jedné instituci.

Podíváme-li se na tyto podmínky realisticky, dojdeme snadno k závěru, že se navzájem prakticky vylučují. To je ale známo od samého počátku principu hodnocení vědeckých publikací a není důvod očekávat, že tomu bude v budoucnu jinak. Jinak řečeno, jakékoliv snahy a návrhy „nového“ způsobu recenzních procesů nemohou fungovat. Jelikož obecné stížnosti na nespravedlivé recenzní řízení existují ve vědecké komunitě uplynulých cca 200 let, budou trvat i v dalších desetiletích. Připomeňme samozřejmý fakt, že pravděpodobně jinak bude probíhat recenzní řízení u časopisu, jehož přísun rukopisů jen málo převyšuje objem pravidelného naplnění cca 12 čísel titulu v běžném roce, a na druhé straně časopisu, který musí velkou část, někdy většinu, odmítat.

Tyto pesimistické závěry ale v žádném případě nemohou znamenat, že na jakékoliv hodnocení rukopisů zasílaných ke zveřejnění je nutné resignovat. Zkrátka a dobře, publikovat každý zaslaný rukopis bez jakékoliv recenze je vyloučené a vedlo by nejenom k záplavě publikovaných textů, ale hlavně k naprosté ztrátě důvěry v jakoukoliv spolehlivost publikovaných informací. Jak tedy tuto situaci řešit?

Místo návodů na řešení upozorníme na dlouhou dobu běžnou praxi, která je ale zastíněna kritizovanou a široce diskutovanou problematikou klasických představ o recenzním řízení. Prakticky každé periodikum má kromě svého vydavatele také nějakou vědeckou radu, nebo poradní radu a také (ve skutečnosti vždy) vědeckého redaktora, ať tuto práci vykonává z entusiasmů, nebo jako více či méně honorovanou činnost. Čím se jedná o časopis s vysokým renomé, a tedy vysokým IF, tím je tato funkce důležitější a zpravidla vyžaduje práci na plný úvazek. Někteří renomovaní vědci jej přijímají na určitou omezenou dobu, včetně předpokladu rovněž kvalifikovaných spolupracovníků. V tomto ohledu je velmi zajímavý článek, který nedávno uveřejnil Diederich, který působil cca 20 let jako člen redakční rady a následně 10 let jako její předseda, tedy *de facto* jako výkonný vědecký redaktor - *Scientific Editor*, vysoce respektovaného časopisu *Angewandte Chemie*, vydávané vydavatelstvím Wiley<sup>2</sup>.

Počty zaslaných rukopisů a jejich recenzní zpracování v roce 2012:

Počet zaslaných rukopisů k recenzi	Počet vyřízených recenzních posudků	Počet nerecenzovaných rukopisů
28 800	17 100	9 700

Z těchto čísel se zdá, že relativně velká část zaslaných rukopisů nebyla recenzována. To je ale jen zdánlivý závěr, ve skutečnosti se jedná o rukopisy, které byly odmítnuty na základě např. nedostatečného formálního zpracování, nízké úrovně prezentace výsledků, mimo oblast publikovanou v daném časopise, nebo z dalších důvodů. Hlavním hráčem je zde právě role vědeckého redaktora, resp. redaktorů.

Zajímavé jsou pak počty skutečně vydaných vědeckých článků v roce 2012 a dvou dalších letech:

2012	2013	2014
4 944	5 160	5 194

Je zřejmé, že počet vydaných článků je jen mírně vzrůstající, průměr je cca 5 000 článků za rok/svazek, z čehož vyplývá, že podíl rukopisů přijatých ke zveřejnění je cca

17,36 %

Tento podíl je značně typický pro hodnocení úrovně daného časopisu a není pak překvapující, že hodnota IF se pohybuje kolem 13!!! Stojí za upozornění, že činnost vědeckého redaktora/editora je v případě takového titulu zcela zásadní a jedná se o činnost na plný úvazek. Vědecký editor je pak zaměstnancem vydavatele a je zodpovědný za jeho úroveň.

Zbývá otázka, co se stane s rukopisy, které byly odmítnuty? Ve většině případů dochází k jisté formě recyklace, někteří autoři své rukopisy doplní, věnují více péče formálním otázkám, pokračují ve výzkumu a nejpravděpodobněji je pošlou do jiného časopisu. V současné době řada hlavně velkých vydavatelů automaticky nabízí odmítnutým autorům jejich vydání v jiném titulu vydávaném daným vydavatelem.

Na závěr připomeňme dosti diskutovaný případ „vytvoření“ počítačovým programem série zcela fiktivní vědecké publikace, navíc opakovaně přeložené počítačovými překladači do

francouzštiny a zpět do angličtiny a zpět. Z 304 takových článků bylo 157 přijato pro otištění v časopisech OA<sup>3,4</sup>.

### **Rukopis prošel recenzním řízením a je zveřejněn - co dál?**

V případě volby časopisu vydávaného jen v elektronické podobě jako OA není co řešit, pokud autor tuto cestu zvolil, pak zpravidla až v tomto stádiu dochází k finanční úhradě, aby byla eliminována principiální možnost zajistit si zaplacením nákladů vlastní zveřejnění. Nicméně, vydavatel této formy ví, že přijetí bude znamenat finanční příjem, odmítnutí nikoliv. Realizace zveřejnění, např. umístění na server vydavatele, je současně vyřešena i následná otázka jeho dostupnosti.

V případě „odloženého“ rozhodování o způsobu úhrady nákladů na vydání, musí autor právě zde rozhodnout, zda zvolí režim OA a uhradí příslušné poplatky (tyto poplatky už dostaly oficiální název a příslušnou zkratku v podobě označení *Author Processing Charge*, neboli *APC*). Ať tak, nebo tak, publikace je zveřejněna a přístupná zájemcům. Ale ještě musí být uskutečněn absolutně nejdůležitější krok, a to je vytvoření záznamu o existenci dané publikace pro nějakou vědeckou bázi dat, v každém případě obsahující její název, autory a přesně identifikovaný primární zdroj, případně předmětová hesla, nebo jiný typ věcného popisu. Rozšířené záznamy pak mohou obsahovat plné texty abstraktů, dnes přebírané takové, které vytvořil sám autor. Žádná současná vědecká báze dat nevytváří texty abstraktů nezávisle na abstraktech autora (za zmínku stojí, že tuto činnost je stále vykonávána pro profesionální patentovou bázi dat DERWENT, dnes součástí holdingu Thomson Ltd.).

Ale naprosto nejdůležitějším krokem je zpracování vydaného článků do záznamu ve vědecké bázi dat. Tento mimořádně důležitý krok byl již podrobně zmíněn v úvodní kapitole, včetně zdůraznění hlavního problému v současné době, tedy enormnímu nárůstu vydávaných vědeckých publikací. Zde narážíme na bariéry jak lidských zdrojů, tak logicky následně finančních nákladů. Přitom v tomto ohledu platí zcela obecná pravda:

**Pokud pro vydanou vědeckou publikaci neexistuje bibliografický záznam v nějaké vědecké bázi dat, tato práce s největší pravděpodobností nikdy nebude přečtena a využita jako zdroj nových vědeckých poznatků!!!**

### **Zpřístupnění publikované vědecké práce – koncepce *Open Access***

Je jistým paradoxem, že tento konečný aspekt budí pozornost a občas i vášně, přesto, že všichni aktéři podílející se na publikačním procesu mají přirozený zájem, aby publikace byla dostupná všem, kteří o ní mají zájem. Současně je to ale poslední místo, kdy lze kompenzovat náklady na celý publikační proces, pokud nejsou tyto náklady uhrazeny předem v podobě APC (viz výše). V klasické době tištěných zdrojů byly náklady přirozeně hrazeny nákupem do knihoven, typicky formou předplatného, a pak logicky následně byly dostupné komukoliv, kdo do knihovny přišel, a to v naprosté většině bezplatně. Jistou komplikaci přinesla nejdříve technika snadného kopírování, ale zásadní problém nastal až s prakticky úplným přechodem na elektronickou podobu s prakticky neomezenou přenositelností a téměř žádnou možností

skutečné kontroly. Přičemž naprostá většina autorů vědeckých publikací bez námitek akceptuje nutnost náklady na proces vydávání kompenzovat.

Za této situace se před více než patnácti lety (2002) objevila jako *deus ex machina* koncepce, založená především na elektronické podobě vědeckých článků a následně na myšlence úhrady nákladů na zpracování vědeckých rukopisů do podoby vědeckého článku předem, místo až *ex post* formou předplatného. Tato koncepce, která si získala širokou popularitu pod označením Open Access, byla od prvních memorand mnohokrát prezentována a popularizována, v neposlední řadě na těchto konferencích, např. už v roce 2005<sup>5</sup>. Dominantní předností je skutečnost, že úhrada nákladů předem okamžitě otevírá přístup k plnému textu vědeckého článku. Samozřejmě to znamená naprosto zásadní obrat ve financování šíření vědeckých informací, sice velice lákavý a nepochybně užitečný, ale otázkou, která je od prvních představ a koncepcí řešena je, jak tohoto obratu dosáhnout, resp. jak přesměrovat tok peněz od předplatného na platbu předem.

Jestliže ale tato nepochybně lákavá a na první pohled logická koncepce je na stole už prakticky 15 let, pak současný objem publikací existující v tomto režimu, který podle různých zdrojů se pohybuje mezi 17% až méně než 25%, nelze označit jako jednoznačný úspěch. Je proto na místě se ptát, proč dosud nedošlo k masovému nebo alespoň velmi výraznému posunu k této koncepci.

Zdá se, že problémem je výše nákladů na vydávání formou OA ze strany autorů. Již první koncepce OA, ať již ze strany zavedených vydavatelů, nebo vydavatelů nových, pracovaly s úhradou za zpracování rukopisu do vědecké publikace ve výši až 3000 USD, a tyto částky se v zásadě nezměnily, jediné se rozšířily oběma směry, tak dnes jsou požadovány částky za APC od cca 1000\$ do cca 4 500\$. Od samého počátku byla proklamovaná výše 3000\$ považována za přemrštěnou, ale, jak je vidět, původní odhad byl správný. Jestliže nyní uvážíme, že do příjmy publikace je vloženo velké množství práce nehledě už na náklady vlastního výzkumu, pak požadovat ještě dalších 3000\$ jen obtížně přijatelné pro kohokoliv, a to i v případě, že takovou částku bude hradit někdo jiný, grantová agentura, vědecká instituce autora či sponzor. Autora či autory publikace budou takové částky vždy iritovat, jedná se např. o častou výši konferenčních poplatků i velmi renomovaných konferencí. Zdá se tedy, že bude nutné celou koncepci OA prodiskutovat a řešit hlavně v kontextu celého procesu zveřejňování vědeckých poznatků, a také vzhledem k hlavnímu tématu této přednášky, mimořádně rychle narůstajícímu objemu vědeckých poznatků.

#### Literatura:

1. Bornman Lutz; Mutz, Ruediger: Journal of the Association for Information Science and Technology, 66, 2215, 2015.
2. Diederich, F.: Angewandte Chemie Int. Ed., 52, 2-4. 2013. 3.
3. Bohannon, J.: Science, 342, 60, 2013.
4. Slavíček, P.: Vesmír, 93, 63, 2014.
5. Šilhánek, J.: Sborník konference Inforum 2005.