

EOD - Elektronické knihy na objednávku

Rostislav Krušínský

Vědecká knihovna v Olomouci
krusna@vkol.cz

Kateřina Kamrádková

Národní technická knihovna, Praha
katerina.kamradkova@techlib.cz

Věra Pospíšilíková

Moravská zemská knihovna v Brně
verkaposp@mzk.cz

**INFORUM 2010: 16. konference o profesionálních informačních zdrojích
Praha, 25. - 27. 5. 2010**

Abstrakt:

Evropské knihovny vlastní miliony knih vydaných mezi lety 1500 a 1900. Vzhledem k jejich stáří a hodnotě jsou často dostupné jen uživatelům fyzicky přítomným v dané knihovně. Se službou eBooks on Demand (EOD) budou tyto skryté poklady zpřístupněny prakticky komukoli a kdekoli. Uživatelé si knihy objednávají prostřednictvím katalogů knihoven; knihovny pak požadovanou knihu zdigitalizují, pošlou ji uživateli pomocí služby EOD. Takto zdigitalizované knihy jsou zároveň včleněny do digitálních knihoven zúčastněných institucí a tak zpřístupňovány na internetu. Služba EOD je poskytována v rámci projektu „eBooks on Demand – A European Library Network“ již 20 knihovnami v 10 evropských zemích. V České republice jsou to Vědecká knihovna v Olomouci, Moravská zemská knihovna v Brně, Národní technická knihovna v Praze a Knihovna Akademie věd ČR.

Vlastní text:

V případě zájmu o staré tisky, tisky 19. století, nebo tisky počátku 20. století jsou možnosti získání knihy dosti skrovné. Takto staré dokumenty se nedají snadno koupit, neboť nabídka antikvariátů je omezená, navíc jde často o velmi cenné a tudíž i drahé knihy. Meziknihovní výpůjční služby fungují, ale pouze v případě novějších knih. Staré tisky, ale i knihy starší sta let se obvykle knihovny zdráhají zapůjčit absenčně. Jedna varianta získání knihy je cesta do knihovny, což může být ovšem složité, pakliže je požadovaná kniha k dispozici na druhém konci republiky. Papírové kopie se ze starých knih zpravidla také nedělají, zájemce si ovšem může koupit mikrofilm – k tomu je ale zapotřebí čtečka, a tou nedisponuje každý. Další z možností je počkat na projekty masové digitalizace, u kterých ale není jisté, kdy bude žádaný tisk digitalizován. Autenticita starého tisku je samozřejmě nenahraditelná, nicméně pokud pro účely zájemce postačuje jiná forma kopie, nabízí se služba EOD.

Projekt EOD

EOD je zkratka názvu eBooks on Demand, tedy elektronické knihy na objednávku. Jedná se o službu, díky které si může zájemce objednat digitalizaci knih do formátu PDF. To je v současnosti běžně

rozšířený formát, s nímž jde pracovat prakticky na každém počítači. U mladších knih je možné zakázku zpracovat s použitím technologie Optical Character Recognition (OCR), tedy optickým rozpoznáváním znaků, umožňující fulltextové vyhledávání a kopírování zdigitalizovaného textu. Služba EOD je orientována na tzv. *volná díla*, tedy konkrétně knihy nechráněné autorským zákonem. Majetková autorská práva mají ve většině států Evropy platnost po dobu autorova života a 70 let po jeho smrti. Z toho plyne, že prostřednictvím služby EOD mohou být zpracovány veškeré staré tisky 16. - 18. století, velká část tisků 19. století, a také některá díla počátku 20. století. Zpracování v rámci EOD znamená digitalizaci kompletní knihy, včetně vazby, přídeští a také všech nepotištěných stran. Výsledný produkt tedy zachycuje veškeré podrobnosti, včetně případných exlibris, rukopisných přípisů apod.

Služba EOD vznikla jako projekt Tyrolské univerzity a zemské knihovny v Innsbrucku (UIBK), který probíhal od října 2006 do června 2008 pod názvem „Digitisation on Demand“ (Digitalizace na objednávku) a byl spolufinancován v rámci evropských programů eTEN. Projektu se účastnilo 13 knihoven z 8 evropských zemí.



Obrázek č. 1: Logo programů eTen

Na pilotní projekt navazuje projekt „eBooks on Demand – A European Library Network“ (Elektronické knihy na objednávku – Evropská knihovní síť). Tento projekt byl zahájen 1. května 2009, bude probíhat až do 30. dubna 2013 a je podpořen programem Evropské komise Kultura 2007-2013. Do projektu je zapojeno 20 knihoven z 10 evropských zemí – konkrétně jde o zástupce z České republiky, Dánska, Estonska, Francie, Maďarska, Německa, Portugalska, Rakouska, Slovenska a Slovinska.

Cílem projektu je rozšíření služby EOD do většího počtu zemí a institucí v Evropě, a to nejen do knihoven, ale i muzeí a archivů – podmínkou je, aby instituce mohla nabídnout dostatek autorsky volných knih, které je možné zájemcům poskytnout v elektronické kopii. Zároveň tento projekt otevírá samotné instituce, zpřístupňuje jejich fondy, sblížuje pracovníky institucí se zákazníky - sice se vzájemně nepotkají tváří v tvář, ale přesto instituce lépe poznává potřeby zájemců a učí se je využít ve své propagaci a popularizaci. Dalším stupněm projektu je propagace historických knih v prostředí, kde se zatím neobjevovaly tak často - ve Web2.0 aplikacích: Wikipedia, Facebook, Goodreads atd. Navíc e-knihy jsou svým formátem vhodné i pro čtečky elektronických knih, takže elektronická kopie historické knihy se zde potká s e-knihami a e-dokumenty moderními. Nejnápadnější akcí z aktivit partnerů projektu jistě bude každoroční den otevřených dveří, který je věnován propagaci služby EOD pokud možno ve formách blízkých pro různé zájmové skupiny – mladí lidé, badatelé vyššího věku, organizace, jež využijí poznatky z e-knih pro svoji práci atd.

Mimo projekt se síť rozšiřuje každým dnem, jen v ČR se od doby začátku projektu připojila Knihovna Akademie věd ČR, ze zahraničních partnerů např. Švýcarská národní knihovna.



Obrázek č. 2: Logo programu Kultura

Objednávka digitalizace a její zpracování

Pokud jde o poskytování vlastní služby, každá knihovna se rozhodne, které části historického fondu nabídne v rámci služby EOD ke zpracování, a to v závislosti na místním autorském právu a interní politice.

Pro přijímání a vyřizování objednávek služby EOD byl vytvořen tzv. *Order Data Manager* (manažer zakázkových dat, zkratka ODM) centrální databáze s webovým rozhraním, jejíž stěžejní částí je *Digital Object Generator* (generátor digitálních cílových databází, zkratka DOG). Všechny knihovny zapojené v síti EOD mají ODM k dispozici a knihovnám jsou zde zobrazeny jednotlivé objednávky. Celý pracovní postup – od kontroly obdržené objednávky až po samotnou platbu – je realizován v ODM.

Software ODM je rozdělen do dvou částí, jednak zde probíhá automatický proces, který zajišťuje generování e-knihy, dále je poskytováno uživatelské rozhraní, jež umožňuje operátorovi upravovat data jednotlivých objednávek, včetně informací o zákaznících.

ODM přijímá objednávky prostřednictvím objednacích formulářů, které jsou integrovány v katalozích partnerských knihoven, přičemž podporovány jsou tyto typy katalogů: *OPAC (Z39.50)*, *naskenovaný listkový katalog*, *přizpůsobený webový formulář*. Do katalogů je integrováno tlačítko EOD „Objednat digitální kopii“, jež umožňuje zákazníkovi objednat elektronickou kopii požadované knihy. Knihovna má možnost poskytnout objednací formulář v různých jazycích. Výběr jazyka ovlivňuje kromě tohoto formuláře také jazyk tzv. *tracking page* (zákaznické stránky sledující objednávku), jazyk obálky zhotovené elektronické knihy, jazyk e-mailů, jež jsou posílány zákazníkovi a jazyk faktury.



Obrázek č. 3: Logo EOD

Většina zapojených knihoven v současnosti používá OPAC. Největší výhodou je, že zákazník data o požadované knize nevyplňuje manuálně, ale jsou do ODM z katalogu automaticky převedena pomocí protokolu Z39.50. Navíc lze s použitím skriptu nastavit, v jakém případě se tlačítko EOD „Objednat digitální kopii“ u dokumentů objeví – např. podle roku vydání zvoleného pro autorsky volná díla.

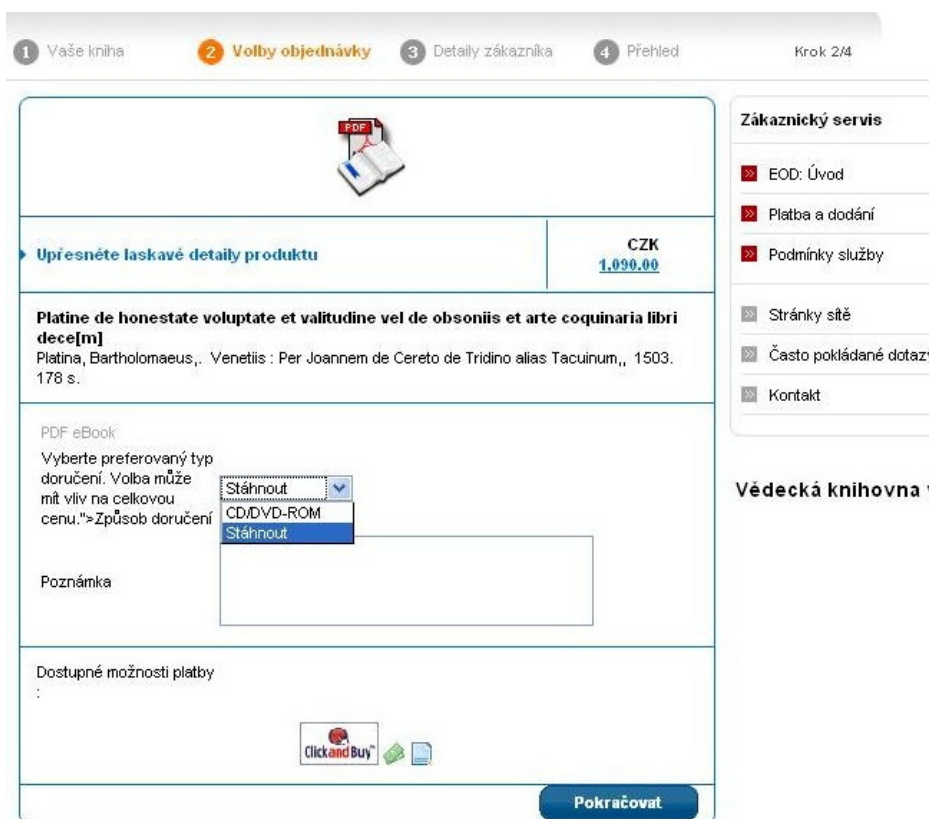
Může nastat situace, kdy zákazník pro objednání služby EOD nechce použít objednacího formuláře. V tom případě je možné učinit objednávku telefonicky nebo poslat žádost e-mailem. Operátor ODM pak musí objednávku zadat manuálně – v ODM je k tomu uzpůsoben formulář *Nová objednávka / Nový zákazník*.

Tracking page (zákaznická stránka)

V závislosti na tom, kde byla objednávka vytvořena se na zákaznické stránce a obálce e-knihy objeví logo knihovny, jež zakázku zpracovala. To zároveň slouží zákazníkovi jako záruka, s kterou knihovnou uzavřel smlouvu. Pro dokončení procesu objednávky musí zákazník akceptovat Obchodní podmínky služby EOD té které knihovny, ty mohou být dané knihovně přizpůsobeny na míru.

Všem operátorům ODM v dané knihovně je zaslán upozorňující e-mail se zprávou, že byla učiněna nová objednávka. Tento e-mail obsahuje všechny důležité informace týkající se objednávky digitalizace, spolu s odkazem na konkrétní webovou aplikaci v ODM.

Zákaznická stránka zobrazuje všechny aktivní objednávky daného zákazníka. Používáním této stránky může zákazník např. změnit požadavek na odhad ceny e-knihy v přímou objednávku digitalizace, zvolit způsob platby, stáhnout finální e-knihu v závislosti na tom, jaký způsob doručení e-knihy si vybere (v PDF nebo v RTF verzi v případě, že je e-kniha vytvořena pomocí rozpoznávacího serveru ABBYY), stáhnout fakturu apod.



Obrázek č. 4: Ukázka zákaznické stránky

ODM – Manažer zakázkových dat

Vlastní aplikace ODM, přístupná přes <http://books2ebooks.eu/odm/>, je rozdělena do několika částí: ODM workflow (proces tvorby e-knihy krok po kroku), administrace uložených dat a přehled dokončených objednávek.

Aplikace nabízí řadu možností, jak práci operátorům ODM usnadnit a zpřehlednit (např. vyhledávací pole, seznam aktivních objednávek včetně typu objednávky a zobrazení dalších podrobností). Operátoři mohou rovněž upravovat detaily knihy – název, místo a rok vydání atd. v případě, že nejsou správně načteny z katalogu knihovny.

Po ověření, zda je kniha fyzicky dostupná, mohou operátoři v ODM vytisknout PDF dokument o skenovaném titulu a vložit jej do požadované knihy. Dokument, jenž obsahuje také čárový kód s unikátním číslem, slouží jako fyzický doklad objednávky. Je automaticky ukládán na FTP serveru ve složce „/misc“ konkrétní objednávky a je kdykoliv později přístupný.

Obecně existují dva způsoby, jak vygenerovat základní e-knihu. Buď ji může zhotovit knihovna lokálně, nebo použít centrálního systému. Výhodou lokálního zpracování je přímá kontrola tvorby e-knihy a jejích parametrů, minimalizace přesunu souborů mezi lokálním a centrálním uzlem, kde je e-

kniha generována a svobodná volba software pro vytvoření e-knihy. Nevýhodou je pak zvýšená míra manuální práce, chybějící standardizace pracovního procesu a úložiště ovládané pouze místně.

Bez ohledu na to, kde je e-kniha generována, na FTP serveru je prostřednictvím ODM zformováno několik složek, jež jsou umístěny na centrálním uzlu. Každá partnerská knihovna má svůj vlastní FTP účet, který je používán pro přenos dat mezi centrálním a lokálním uzlem.

V případě, že má finální verze skenů nějaké nedostatky, operátor ODM má možnost je sám opravit. Upravená verze pak musí být zkopírována do složky „*ebook_final*“ dané objednávky před tím, než je dokončen proces generování e-knihy.

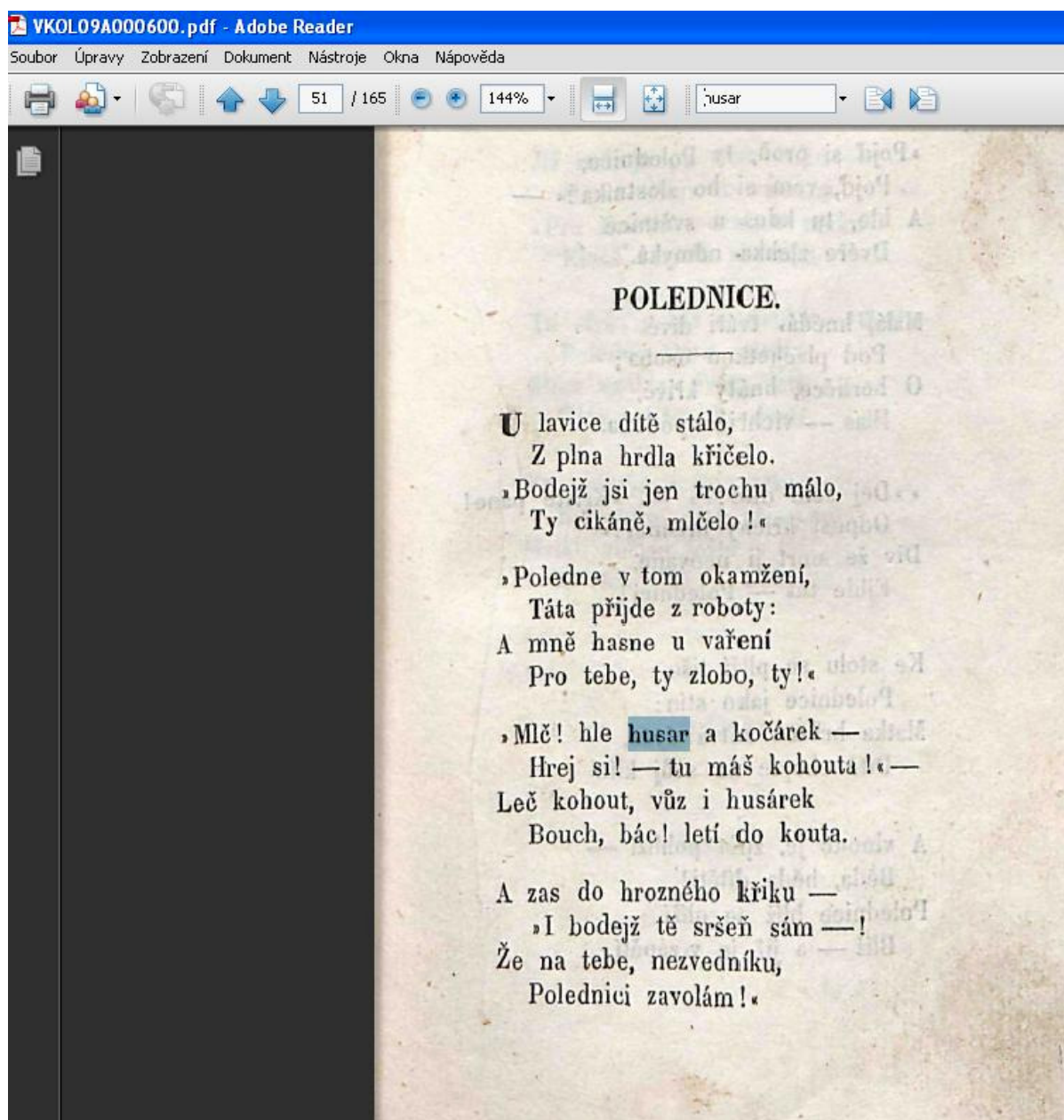
Mimo jiné existuje v ODM tlačítko „*Undo eBookGeneration*“, které může zrušit jednotlivé kroky pracovního postupu. Dají se zde nastavit speciální ceny pro určité skupiny zákazníků (např. registrovaní uživatelé knihovny, zrakově postižení) a ODM nabízí také generování přehledů a statistik týkajících se vyřizovaných objednávek (přehled o příjmech, e-knihách, počtu objednávek a zákaznících).

OCR - Automatické rozpoznání znaků

Jestliže je při výrobě e-knihy zvoleno OCR, systém potřebuje vědět, v jakém jazyce byla kniha napsána. Operátoři mají možnost vybrat jazyk knihy ze seznamu tvořícího součást DOG. Při zpracování s OCR je používán rozpoznávací server ABBYY, který podporuje určitou sadu jazyků. Ty však ne vždy souhlasí s těmi, co byly vybrány v prvním kroku workflow. Jazyk OCR pro psané písmo neexistuje. Operátor ODM buď vybere jiný jazyk (generování e-knihy jinak nemůže začít) nebo odebere možnost zpracování zakázky s OCR. Pokud je v rámci objednávky využito OCR, jsou v závěru dostupné tyto formáty e-knihy: PDF, RTF a XML.

Před započatím procesu OCR je možné zvolit vyrovnání obrazu a korekci linek skenů a nastavit rozlišení obrázků ve finální EOD e-knize (standardně 300 dpi). Upravit se dá rovněž kvalita obrázků, zvolená hodnota v procentech definuje velikost obrázků začleněných v PDF (standardně 50%). Operátoři mohou definovat i typ tištěného písma, jehož správná volba může ovlivnit výslednou kvalitu OCR. U každé úlohy může být označeno více typů tisku. Nejčastěji užívanými jsou: normální, psáno na psacím stroji a písmo novogotické.

OCR přináší novou kvalitu do digitalizovaného textu, umožňuje práci s textem - fulltextově vyhledávat a kopírovat. Díky OCR je také novogotické písmo převedeno do běžné latinky.



Obrázek č. 5: Ukázka elektronické knihy s použitím OCR

Normy kvality pro EOD

Při skenování v rámci služby EOD by měla každá knihovna dodržovat tzv. *normy kvality*. Účelem takového zpracování je, aby grafická a sémantická informace, které originál zahrnuje, byly obsaženy také v elektronické podobě knihy.

Knihy musí být především naskenovány kompletně. Pokud se jedná o starý nebo vzácný tisk s obálkou se speciální hodnotou, měla by být rovněž skenována, stejně tak i všechny volné strany, které za

obálkou následují (často obsahují rukopisné přípisky, exlibris apod.). V jiném případě by mělo skenování začít na první počítané stránce. Chybějící strany v knize mohou být popsány souborem metadat nebo uvnitř souboru, kdy ve standardním obraze strany bude uveden text např.: „Tato strana chybí také v originálu knihy.“.

Rozlišení skenů by mělo být minimálně 300 dpi, u velmi malých písem pak 600 dpi. Hloubka barvy je požadována 1 bitová pro strany obsahující text a půltónové ilustrace, 8 bitová šedá škála pak pro fotografie a ilustrace v šedi nebo barvě. Požadován je formát TIFF skupiny IV pro 1 bitové obrázky a TIFF RGB a/nebo LZW, JPEG, JPEG2000, PNG pro 8 bitové nebo 24 bitové skeny. Neměla by se používat ztrátová komprese a v případě JPEG by měl být použit kompresní stupeň 90%.

Při skenování je nutné rozdělit dvojstrany na jednotlivé strany s výjimkou map, článků, které přesahují na druhou stranu a rukopisů, kde jsou skenovány dvojstrany za účelem zachování celého obrazu strany. Samozřejmostí je oříznutí stránek, vyrovnání obrazu a korekce linek. Skenované soubory musí být číslovány vzestupně. Všechny knihy publikované po roce 1800 by měly být zpracovány a doručeny s OCR, které knihovna může zpracovat i lokálně.

Soubory jsou dodávány přes FTP do centrálního serveru ODM. V případě větších souborů (více než 3-4 MB u jednoho obrázku), měly by být před přenosem zmenšeny dodávající knihovnou - velké soubory TIFF LZW nebo RGB jsou nejlépe převedeny do souborů JPEG se 75 nebo 90% kompresí.

Obvykle po 7 dnech zákazník obdrží samostatnou e-knihu ve formátu PDF (v případě vícesvazkových děl dostane sérii souborů) obsahující kompletní svazek knihy. Rozlišení PDF je minimálně 150 dpi b/w, ve stupních šedé nebo v barvě, spíše však 300 dpi a s 50% kompresí (která ovlivní jen 8 a 24 bitové obrázky).

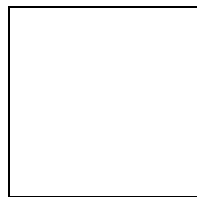
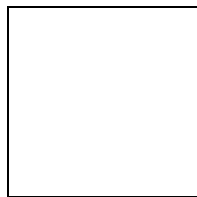
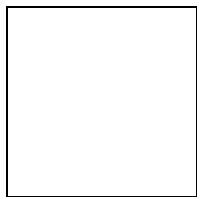
Finance

Členství v síti EOD je zpoplatněno, roční členský poplatek pro instituce je 1000 EUR, pokud instituce rozšíří své služby ještě o Print on Demand (tištěnou verzi e-knihy), zvýší se roční poplatek o 500 EUR. Existuje několik finančních modelů, které si mohou členské instituce vybrat pro platbu za zprostředkování služby EOD a Print on Demand. Patrně nejčastěji využívaný je model all inclusive – kdy z každé zakázky patří 23 % z ceny zprostředkovateli – tedy Tyrolské univerzitní a zemské knihovně v Innsbrucku, která zabezpečuje kompletaci naskenovaných stran do jednoho PDF dokumentu, OCR a přijímá platbu platební kartou od zákazníka.

Ceny knih pro zákazníky se liší podle zvyklostí jednotlivých knihoven, zpravidla kopírují cenu podobných služeb knihovny. Jednotlivé instituce si mohou určovat cenovou politiku samy. Veškeré elektronické knihy vzešlé z této služby jsou zpravidla po 2 měsících zpřístupněny na Internetu a všem dalším zájemcům jsou k dispozici zdarma.

Rozšíření a budoucnost

Jak již bylo řečeno, uživatelské prostředí je stejné u všech poskytovatelů. Zájemci o službu si tedy mohou, po seznámení se s objednávacím systémem v jedné instituci, obdobně jednoduše objednat knihu ve všech institucích účastnících se projektu. V České republice je to kromě Vědecké knihovny v Olomouci ještě Moravská zemská knihovna, Národní technická knihovna a Knihovna Akademie věd ČR. Jen katalogy těchto čtyř knihoven obsahují statisíce knih, které je možno zpracovat pomocí služby EOD.



Obrázek č. 6: Loga knihoven ČR zapojených do projektu EOD

Síť EOD se bude rozšiřovat až do roku 2013. Zájemci o poskytování služby EOD by měli splňovat několik podmínek. Předně je to připojení k internetu. Dalším neméně důležitým faktorem je vlastnictví nebo správa velkého množství autorsky volných děl. Samozřejmostí je také knižní skener schopný dosáhnout kvality nejméně 300 dpi. Dále je zapotřebí elektronický katalog a personál k jeho nastavení. Stejně tak je nutné mít personál pro skenování a komunikaci se zákazníky.

V případě zájmu stačí, aby instituce kontaktovala koordinátora projektu (UIBK). Poté získá balíček s potřebnými dokumenty a přístup k testovacímu rozhraní, kde je možné simulovat celý proces objednávky a výroby elektronické knihy. Následným krokem je podepsání smlouvy a připojení se k síti EOD. Nakonec stačí jen upravit vlastní on-line katalog, nastavit ceník a čekat na objednávky.

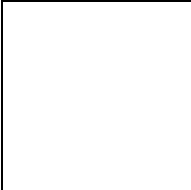
Instituce, které se zapojí do sítě EOD rozšíří své portfolio o novou službu, která je žádaná a perspektivní. Digitalizované knihy patří jim, a mohou je přidat do své digitální knihovny. Zároveň získají přístup k aplikaci pro správu požadavků a komunikaci s uživatelem, vytváření a doručování elektronických knih a evidenci příjmů. V rámci sítě EOD získá zapojená instituce také propagační materiály, které může použít pro své potřeby. Síť EOD slouží také jako poradní skupina a poskytuje návody, standardy a doporučení pro pracovní postupy.

V současné době bylo digitalizováno již přes 3200 knih, což představuje 840 000 stran. Služby EOD využilo přes 2000 zákazníků z celého světa a oblíbenost a rozšíření služby pomalu narůstá. Do budoucna se počítá s rozšířením služby o nadstavbu Print on Demand, tedy tisk na objednávku, který je již některými knihovnami poskytován. Dále se např. plánuje vytvoření jednotného rozhraní umožňujícího vyhledávat všechny knihy zpracované v rámci služby EOD. Vzhledem k akceleraci rozvoje internetu v posledních letech se dá očekávat také rozšíření možností pro využití této služby.

Evropské knihovny, archivy a muzea vlastní množství dokumentů, reprezentujících kulturní a historickou rozmanitost Evropy. Digitalizováním dokumentů se nejen zajišťuje jejich trvalé uchování, ale i snadnější přístupnost. Také v rámci iniciativy i2010 byly pro zajištění snadného přístupu k informacím stanoveny tři směry: dostupnost online, digitalizace analogových sbírek, uchování a ochrana. Projekt eBooks on Demand – A European Library Network přispívá do všech třech oblastí: zdigitalizované knihy jsou zároveň ukládány do digitálních knihoven institucí, většinou volně přístupných na internetu. Přitom je třeba obsah kvalitně popsat, aby mohl sloužit jako odkaz budoucím generacím, především však, aby byl v záplavě nestrukturovaných informací naležitelný. V projektu EOD existuje snaha o zdokonalování uchování a následného vyhledávání historického obsahu. Sjednocení postupů a spolupráce v Evropské knihovní síti pak šetří čas i peníze: digitalizace se stává cenově dostupnější i pro knihovny, které by si ji mnohdy nemohly dovolit.

Digitalizace, jak poukazuje A. Manguel, tento zdánlivě nový systém napodobuje jiný systém, který fungoval před staletími na středověkých univerzitách, kde mohli text doporučený učitelem opisovat písaři, kteří si zřídili obchod mimo prostory univerzity a nabízeli své služby studentům. Středověké rukopisy přečkaly celá staletí. Pevně doufejme, že dnešní moderní varianta bude fungovat stejně efektně a dlouho.

Zdroje:

- *DoD DokuWiki : main:eod:documents:quality_standard* [online]. 2008, 2008/03/13 [cit. 2010-04-30]. Dostupné z WWW: <[http://al-neu.uibk.ac.at/dokuwiki/doku.php?id=main:eod:documents:quality_standard&s\[\]=quality&s\[\]=standards](http://al-neu.uibk.ac.at/dokuwiki/doku.php?id=main:eod:documents:quality_standard&s[]=quality&s[]=standards)>. Přístupné pouze po přihlášení.
 - Innsbruck University Library. *ODM Administration Guide : Technical documentation* [online]. v1.10. Innsbruck : Department for Digitisation and Digital Preservation, June 2008 [cit. 2010-04-30]. Dostupné z WWW: <<http://www.books2ebooks.eu/odm/login.do>>. Přístupné pouze po přihlášení.
 - GSTREIN, S.; MÜHLBERGER, G. User-driven content selection for digitisation – the eBooks on Demand Network. In *Cultural Heritage on line. Empowering users : an active role for user communities* [online]. Fondazione Rinascimento Digitale, c2009 [cit. 2010-02-20], s. 1-6. Dostupný z WWW: <<http://rinascimentodigitale.org/eventi/conference2009/proceedings-2009/grstein.pdf>>.
 - MÜHLBERGER, G.; GSTREIN, S. eBooks on Demand (EOD) : a European digitization service. *IFLA Journal* [online]. 2009, vol. 35, no. 1 [cit. 2010-02-20], s. 35-43. Dostupný z WWW: <<http://archive.ifla.org/V/IFLAJ/IFLA-Journal-1-2009.pdf>>. ISSN 0340-0352. DOI: 10.1177/0340035208102034.
 - MANGUEL, A., *Knihovna v noci*. Brno: Host, 2009. 302, lii s. ISBN: 978-80-7294-336-4.
- 
- i2010: Digitální knihovny. *Knihovna plus* [online]. 2005, č. 2 [cit. 2010-04-19]. Dostupný z WWW: <http://knihovna.nkp.cz/knihovnaplus51/digitalni_knihovnyEU.htm>. ISSN 1801-5948.
 - *EOD : eBooks on Demand* [online]. 2008 [cit. 2010-04-30]. Dostupné z WWW: <<http://books2ebooks.eu/>>.

Poděkování:

Projekt eBooks on Demand je realizován za finanční podpory Evropské unie. Za obsah sdělení odpovídají výlučně autor a spoluautoři. Sdělení nereprezentují názory Evropské komise a Evropská komise neodpovídá za použití informací, jež jsou jejich obsahem.