

## Pregled in analiza izbranih podatkovnih zbirk ter predmetnih oznak za področje lesarskih znanosti in tehnologije

Tomaž BARTOL<sup>1</sup>

### Izvelek

Uporabniki morajo znati identificirati najrelevantnejše vire kakovostnih znanstvenih in tehničnih informacij. Razumeti morajo strukturo zbirk, iskalne tehnike, kontroliran besednjak, terminologijo in klasifikacijske sheme. V članku podajamo glavne značilnosti bibliografskih podatkovnih zbirk (Agricola, Agris, Cab Abstracts (CABI), Compendex, Inspec, Iconda), statističnih zbirk (Eurostat, Faostat) in zbirk tehničnih dokumentov - standardov in patentov (ISO, WIPO) glede na značilnosti lesarstva oz. predelave gozdnih proizvodov. Opisane so razlike v posameznih iskalnih poljih in iskalna načela (platforme, Boolovi logični operatorji, predpone, pripone, znaki za krajsanje ali maskiranje). Posebno pozornost namenjamo predmetnim oznakam: ožje ključne besede (*Descriptors, Heading Words, Controlled Terms, Keywords*) in širše predmetne kategorije (*Subject Categories, Concept Codes, CABI Codes, Classification Codes*). Podani so seštevki besed oz. indeksnih gesel ali pojmov z najvišjo pojavnostjo. S tem pregledom želimo spodbuditi rabo dodatnih podatkovnih zbirk ter rabo strukturiranih besednjakov, kar naj se izrazi v boljših iskalnih rezultatih.

**Ključne besede:** lesarstvo, predelava gozdnih proizvodov, podatkovne zbirke, kontrolirani besednjaki, predmetne oznake, indeksiranje, standardi, patenti, statistika, značilnosti iskanja

## *Review and analysis of selected databases and subject headings with regard to wood sciences and technology*

### Abstract

*End-users should be able to identify the most relevant scientific and technical information resources. They must understand database structure, query techniques, controlled dictionaries, terminology and indexing schemes. In the present article, we review principal characteristics of the following information systems: bibliographic databases (Agricola, Agris, Cab Abstracts (CABI), Compendex, Inspec, Iconda, statistical collections Eurostat and Faostat, and technical standards and patents ISO and WIPO with regard to the field of wood sciences or processing of forest products. Differences in database fields and search characteristics (search platforms, Boolean logic, prefixes, suffixes, wildcards) are described. Special attention is placed on subject headings in each respective database: narrow keywords (Descriptors, Heading Words, Controlled Terms, Keywords) and broader categories (Subject Categories, Concept Codes, CABI Codes, Classification Codes). Word counts are shown for the terms with the highest frequency of occurrences. This review promotes the use of additional databases, encourages employment of structured vocabularies, which should result in improved search results.*

**Keywords:** *wood science, processing of forest products, databases, controlled dictionaries, thesauri, subject headings, indexing, standards, patents, statistics, search characteristics*

## 1 Uvod

### 1 Introduction

V svetu obstajajo različne zbirke za posamezne znanstvene discipline. Številni primeri iz scientometrične literature pričajo o tem, da se znanstveniki kar pre pogosto omejijo na eno samo osrednjo zbirko za svoje področje in pri tem zanemarijo vedno večjo razkropitev informacij. Pri tem lahko spregledajo in izgubijo marsikateri relevantni vir. Predvidevamo, da takšno informacijsko vedenje verjetno velja tudi za naše domače strokovnjake s področja lesarstva, ki sicer s pridom uporabljajo zbirke, kakršni sta

*CAB Abstracts in Web of Science*, pri tem pa morda pozabljajo na druge zbirke, ki so jim prav tako brezplačno dostopne z lokacij naših raziskovalnih in izobraževalnih ustanov. Pogosto se izkaže, da strokovnjaki marsikateroga vira niti ne poznajo, ali pa informacijskega sistema ne znajo ustrezno uporabljati in optimizirati iskalnih zakonitosti. Razkropitev informacij je na področju lesarstva še večja kot pri drugih biotehniških disciplinah, saj gre za dokaj interdisciplinarno disciplino, ki sega od bioloških osnov drevesa kot rastline pa vse do predelave, analize materiala in do končnega oblikovanja izdelka, zato je tu še posebej pomembno, da se strokovnjaki usposobijo za

<sup>1</sup> prof. dr. T. B., Biotehniška fakulteta, Oddelek za agronomijo, Katedra za informatiko, Jamnikarjeva 101, SI-1000 Ljubljana, tomaz.bartol@bf.uni-lj.si

pridobivanje kakovostnih informacij. Za to pa je potrebno določeno informacijsko znanje.

Uporabniki se včasih tudi premalo zavedajo, da so lahko strokovni kontrolirani besednjaki standardiziranih podatkovnih zbirk precej raznoliki, prav tako kot so lahko raznoliki besednjaki pri naravnih jezikih. Uspešna raba nekega izraza v posamezni zbirki pa še ne pomeni, da se bo isti izraz podobno vedel tudi v neki drugi zbirki. Tako lahko kak uporabnik zaradi majhnega števila zadetkov razočaran opusti iskanje, čeravno bi moral malo pobrskati tudi po alternativnih pojmi, ki bi verjetno priklicali precej več zadetkov.

Namen tega prispevka je večplasten. Podati želimo pregled izbranih zbirk, ki zajemajo podatke za področje lesarstva. Pregledati želimo nekatere značilnosti in posebnosti teh zbirk. Še posebej se bomo osredotočili na pregled sistemov kontroliranih besednjakov oz. predmetnih oznak (*deskriptorjev, vsebinskih predmetnih kategorij*). Predvsem pa želimo pri naših uporabnikih s področja lesarstva spodbuditi rabo še kakšne druge zbirke, ki tudi zajema visoko relevantne dokumente, ki jih pri že poznanih zbirkah niso našli. Hkrati želimo pri tem spodbuditi rabo bolj kompleksne iskalne strategije. Strokovnjaki bodo tako pridobili bolj osredotočene in bolj kakovostne informacije. Prispevek naj rabi tudi kot pregled nekaterih relevantnejših splošnih angleških izrazov s področja lesarstva.

Tudi v Sloveniji se nekateri avtorji lotevajo problematike lesarske terminologije. Obstaja nekaj terminoloških geslovnikov (TORELLI 1979, EUMABOIS 1980) ali člankov, ki obravnavajo specializirano tematiko, npr. posamezne drevesne vrste (TORELLI 2000) ali postopke (ŠEGA 2005). V splošnem pa skoraj ne obstajajo tuji članki, ki bi se usmerjeno lotevali problematike lesarske scientometrije. Pri tem še omenimo, da sistematične scientometrične oz. bibliometrične raziskave segajo v sredino 20. stol. Scientometrija (ali asociativno bibliometrija), ki analizira značilnosti znanosti oz. znanstvenih objav, se v svojih začetkih povezuje z delom znanstvenikov, kot so Garfield, ki je utemeljil Science Citation Index (GARFIELD 1964) ali de Solla Price (1963). Vsebine lesarske informatike se morda skrivajo v drugih širših scientometričnih temah, ki pa jih v članku ne obravnavamo. Bolj kot pregled in primerjavo obstoječe teoretične literature bomo zatoj podali primerjavo obstoječih informacijskih sistemov in njihove iskalne dokumentacije oz. navodil, ki so na voljo prek svetovnega spleta.

Prispevek združuje pregled splošnih značilnosti zbirk in informacijskih sistemov ter vsebinsko in frekvenčno analizo predmetnih oznak posameznih zbirk glede relevantnosti pojmov za področje lesarstva oz. predelave lesa. Članek torej temelji tako na oceni iskalnih algoritmov pri posameznih zbirkah, ki so včasih lahko pomanjkljivi, kot na lastnem eksperimentu, ki naj omogoči primerjavo podatkovnih zbirk in oceno njihove primernosti in specifik za področje lesarstva.

## 2 Nekateri značilnosti iskanja po podatkovnih zbirkah

### 2 Some database search characteristics

S pomočjo vsebinskih pojmov lahko pri standardnih podatkovnih zbirkah iščemo po poljih *naslov, izvleček* in *predmetne oznake*. Za predmetne oznake obstajajo različna imena, same oznake pa so lahko hierarhično in vsebinsko različno strukturirane. Tehnika iskanja oz. iskalna sintaksa se lahko oblikuje na podlagi menijev, kjer uporabnik izbira in kombinira zahteve s pomočjo že prej oblikovanih polj, ali pa uporabnik zahtevo oblikuje sam s pomočjo *iskalnih (Boolovih) operatorjev* in ustreznih *predpon* ali *pripon*, če želi iskanje omejiti na specifična polja. Slednja tehnika zahteva višjo raven uporabniške pismenosti, ponuja pa več možnosti glede oblikovanja natančnejše iskalne zahteve.

Iskalna orodja žal niso popolnoma standardizirana, zato morajo uporabniki za uspešno rabo zbirk natančneje poznati strukturo in iskalne zakonitosti pri vsaki zbirki posebej. K temu sodijo predvsem poimenovanje *polj, pripone/predpone*, znaki za *krajšanje* oz. *maskiranje* in raba ločil ali posebnih znakov. Poleg različnih zbirk obstajajo tudi različni *namenski iskalni programi*, imenovani tudi *portali* ali *platforme*, ki lahko gostijo več zbirk, dostopnih po podobnem načelu.

Pri vsebinski hierarhični globini oz. podrobnosti iskanja obstajajo med zbirkami in portali razlike. Pri nekaterih sistemih obstaja možnost popolnoma neomejenega (prostega) iskanja po vseh poljih hkrati. S tem prikličemo vse možne pojavnosti nekega izraza. Seveda lahko v takem primeru dobimo veliko šuma. Če prosto iščemo z izrazom *wood*, bomo tako priklicali tudi vse dokumente, kjer se *Wood* pojavlja kot avtor. Da bi se tak šum oz. balast sistematično izločil, nekatere zbirke prosto iskanje pogosto omejijo na izbrana polja. Pri zbirki *CAB Abstract*, dostopni prek portala *Ovid*, se prosto iskanje hkrati opravi le po poljih *naslov, ključne besede* in *izvleček*. Če želimo pojavnost izraza *Wood* iskati npr. po avtorskem polju, moramo uporabiti ustrezno pripono (*au*) za polje (*wood.au*). Seveda lahko vsebinsko iskanje omejimo tudi zgolj na posamezna polja, kot so npr. *naslov, izvleček* in *ključne besede*.

#### 2.1 Iskalna sintaksa in iskanje po specifičnih poljih

##### 2.1 Search query and field specific searching

Izraz iskalna sintaksa se v informatiki oz. pri iskanju informacij nanaša predvsem na medsebojno razvrščanje iskalnih kriterijev ali gesel, ki jih s pomočjo (Boolovih) iskalnih operatorjev ter *predpon* oz. *pripon* medsebojno povežemo tako, da poiščemo oz. prikličemo za nas najbolj optimalne dokumente, ki jih vsebuje neka podatkovna zbirka.

Pri iskanju s pomočjo Boolovih operatorjev ne smemo pozabiti, da ima operator **AND** (pa tudi **NOT**) vedno prednost pred operatorjem **OR**. Iskalna gesla po potrebi

povežemo z oklepaji.

Iskanje s sintakso (**A or B**) and **C** priključuje kombinacije **AC**, **BC**, **ABC**; medtem ko iskanje s sintakso **A or B and C** (brez oklepaja) priključuje **A**, **BC**, **ABC**.

Pri tem primeru vidimo, da izpustitev oklepajev pripelje do drugačnega iskalnega rezultata.

V preglednici 1 navajamo primere za iskalno sintakso po izbranih podatkovnih zbirkah oz. iskalnih programih, in sicer s pomočjo pojmov v naslovu dokumenta, omejenih z letom publikacije. Načela povezave iskalnih izrazov in Boolovih operatorjev ostajajo enaka, nekoliko pa se razlikujejo poimenovanja oznak za polja, ki se lahko pojavljajo kot predpone ali pripone. Ta načela niso standardizirana, kar morajo uporabniki upoštevati.

**Preglednica 1:** Iskalna sintaksa, predpone/pripone in oznake za polja pri izbranih zbirkah

**Table 1:** Query, prefixes/suffixes, field names in selected databases

Zbirka Database	Iskalna sintaksa Query
Agricola/NAL	<b>tkey</b> (product? and (wood or timber?)) <sup>a</sup>
Agris/FAO	<b>title:</b> (product* <b>AND</b> <sup>b</sup> (wood <b>OR</b> timber)) and <b>date:</b> 2005
CAB Abstracts/Ovid	(product\$ and (wood or timber).ti and 2005.yr
Compendex & Inspec	(product* and (wood or timber)) <b>wn ti and 2005wn yr</b>
Iconda	(product* and (wood or timber)) <in> <b>title and year=2005</b>
Web of Science	<b>ti=(product* and (wood or timber)) and py=2005</b>

<sup>a</sup> leta pri zbirki *Agricola* ni moč omejiti s pripono

<sup>b</sup> *Agris* zahteva pisanje operatorjev z velikimi črkami

Izpostavimo lahko obveznost rabe velikih črk pri zbirki *Agris*. Pri primeru zbirke *Agricola* vidimo, da leta ni moč omejiti s pripono. Leto je možno izbrati v posebnem polju, kar lahko naredimo tudi pri drugih zbirkah, ki pa vendarle omogočajo iskanje s predpono/pripono. Ta možnost olajša iskanje ekspertnim iskalcem, ki lahko enako sintakso uporabljajo v več zbirkah hkrati, če seveda prej sistematično uredijo obliko predpon oz. pripone.

## 2.2 Predmetne oznake (deskriptorji, predmetne kategorije)

### 2.2 Subject headings (descriptors and subject categories)

Za indeksiranje dokumentov, t.j. opisovanje vsebine s predmetnimi oznakami, večina zbirk predvideva sistem predmetnih kategorij in deskriptorjev. **Predmetne kategorije** se uporabljajo za indeksiranje in posledično iskanje dokumentov na širši ravni **makro-kategorizacije** (širši vsebinski koncepti), **deskriptorji** (oz. standardizirane ključne besede) pa so namenjeni indeksiranju in iskanju dokumentov na ravni **mikro-kategorizacije** (ožji vsebinski koncepti). Sezname standardiziranih deskriptorjev se

imenujejo **tezavri**. Obravnavane zbirke, z izjemo zbirke Iconda, uporabljajo tako tezavre (deskriptorje) kot predmetne kategorije. Poimenovanje teh predmetnih oznak ni popolnoma standardizirano.

V preglednici 2 predstavljamo nekaj možnih načinov iskanja z deskriptorjem **wood** po različnih zbirkah/portalih za različne tipe predmetnih oznak. V zadnjem stolpcu lahko vidimo povezavo tako deskriptorjev (wood) kot širših predmetnih kategorij z ustrežno predpono/pripono. Izpisana imena šifriranih kategorij so prikazana v opombi pod črto. Vidimo lahko, da so poimenovanja precej različna. O tem bo več govora v nadaljnjih poglavjih. V tretjem stolpcu so predstavljena izvorna poimenovanja za deskriptorje in za predmetne kategorije, in sicer v povezavi z ustrežno iskalno sintakso za polja.

Sezname predmetnih kategorij za vse obravnavane zbirke so na voljo brezplačno prek spleta. Pri tezavrih je več omejitev, saj so tezavri pogosto samostojne in celostne podatkovne zbirke. Dostop do **tezavra CAB** je na voljo le posredno prek plačila za licenčni dostop do same zbirke. Zbirka Iconda nima ne kategorij ne deskriptorjev, pač pa predvideva le možnost rabe prostih nestandardiziranih ključnih besed. V preglednici 2 (v tretjem stolpcu) podajamo tudi informacijo (referenco) o spletnih virih oz. dostopu do predmetnih oznak in tezavrov, če so ti dostopni javno in brezplačno.

**Predmetne kategorije** navadno predstavljajo sezname, sestavljeni od nekaj deset do nekaj sto pojmov, zato so ti sezname razmeroma kratki. **Tezavri deskriptorjev** so precej bolj obsežni, zajemajo pa lahko več deset tisoč pojmov. Tezavri lahko rabijo kot precej prikladno terminološko orodje za identifikacijo in, posledično, rabo specializiranih pojmov tudi za druge namene, npr. pisanje člankov. Kot verjetno najkvalitetnejši mednarodni tezaver s področja biotehnike lahko omenimo tezaver **Agrovoc** zbirke *Agris* oz. informacijskega sistema *Waicent*, čeprav sicer vsebuje manj deskriptorjev kot **tezaver CAB**. V nasprotju s Cabovim tezavrom omogoča tezaver **Agrovoc** precej natančno identifikacijo hierarhičnih razmerij med deskriptorji na podlagi natančno strukturiranih drevesastih struktur. Poleg tega je Agrovoc dostopen javno, trenutno pa zajema izraze v že kar deset jezikih (angleščina, arabščina, češčina, francoščina, japonščina, portugalščina, slovaščina, španščina, tajščina), zato se lahko uporablja tudi kot precej natančno strokovno prevajalsko orodje med temi jeziki.

Kot brezplačno dostopni večjezični strokovni tezaver lahko omenimo tudi tezaver zbirke **Pascal** znotraj širšega informacijskega portala **INIST-CNRS / ConnectSciences** (*Institut National d'Information Scientifique et Technique - Centre National de la Recherche Scientifique*), ki poleg naravoslovja in tehnike zajema tudi druge vede, npr. družboslovje in humanistiko. Ta portal prek iste spletne strani omogoča dostop tako do tezavra oz. deskriptorjev (**INIST Vocabulary**) kot do predmetnih kategorij (*Classification of Standards*). Delovanje tega mednarodnega portala se

**Preglednica 2:** Poimenovanje posameznih predmetnih oznak pri različnih bibliografskih zbirkah in predpone / pripone za predmetne oznake

**Table 2:** Terminology subject heading in respective databases and subject heading prefixes/suffixes

Zbirka Database	Pred. oznaka Subject head.	Orig. ime oznake Orig. name of the head.	Pripona Suffix	Sintaksa Query
AGRICOLA/NAL	deskriptorji kategorije	Subjects * Subject Categories **	skey 072a	<b>skey</b> wood <b>072a</b> K510 <sup>a</sup>
AGRIS/Waicent	deskriptorji kategorije	Descriptors *** Subject Categories ****	subject: subject:	wood K50 <sup>b</sup>
Biosis/Ovid	deskriptorji kategorije	Descriptors Concept Codes *****	hw cc	wood.hw 53500.cc <sup>c</sup>
CABA/Ovid	deskriptorji kategorije	Heading Words CabiCodes *****	hw cc	wood.hw KK520.cc <sup>d</sup>
Compendex & Inspec	deskriptorji kategorije	Ei Controlled Terms Ei Classification Codes *****	cv cl	wood <b>wn cv</b> 811.2 <b>wn cl</b> <sup>e</sup>
Iconda	deskriptorji kategorije	Keywords kategorij NI	KEYWORD -	(wood) <in> <b>KEYWORD</b> <sup>f</sup> -
Web of Science	prosto	topic	ts	<b>ts=wood</b>

- \* National Agricultural Library Services: National Agricultural Library Agricultural Thesaurus. <http://agclass.nal.usda.gov/agt/agt.shtml> (15. maj 2007)
- \*\* AGRICOLA Subject Category Code Scope Notes. Agriculture Network Information Center. (2000). <http://permanent.access.gpo.gov/Cheddar%20Cheese%20Prices/cheddar/www.agnic.org/cc/index.html> (15. maj 2007)
- \*\*\* Agricultural Information Management Standards Web site (Agrovoc). [http://www.fao.org/aims/ag\\_intro.htm](http://www.fao.org/aims/ag_intro.htm) (15. maj 2007)
- \*\*\*\* Agris Subject Categories. <http://www.fao.org/scripts/agris/c-categ.htm> (15. maj 2007)
- \*\*\*\*\* Biological Abstracts Concept Codes. <http://etextb.ohiolink.edu/help/bioa-concept.html#535> (15. maj 2007)
- \*\*\*\*\* CABICODE Alphabetical List. [http://support.dialog.com/searchaids/dialog/f50\\_cabicodes\\_list.shtml](http://support.dialog.com/searchaids/dialog/f50_cabicodes_list.shtml) (15. maj 2007)
- \*\*\*\*\* EELS. Engineering E-Library, Sweden (Browse engineering resources classified using EI classification). <http://eels.lub.lu.se/> (15. maj 2007)

Razlaga nekaterih okrajšav za pripone/predpone:

<b>s key</b>	subject key
<b>072a</b>	programska šifra za kategorije pri zbirki Agricola
<b>hw</b>	heading word
<b>cc</b>	category code
<b>wn cv</b>	within controlled vocabulary
<b>wn cl</b>	within classification

<sup>a</sup> K510=Forest Products (Wood)

<sup>b</sup> K50=Processing of forest products

<sup>c</sup> 53500=Forestry and Forest Products

<sup>d</sup> KK520 = Wood Utilization and Engineered Wood Products

<sup>e</sup> 811.2 = Wood Products

<sup>f</sup> Iconda zahteva velike črke pri rabi pripone **KEYWORD**, ne pe tudi pripone **title**

koordinira v Franciji, deskriptorji pa so na voljo v angleščini, francoščini in španščini.

### 3 Izbrane podatkovne zbirke in informacijski sistemi

#### 3 Selected databases and information systems

##### 3.1 Bibliografske informacije

###### 3.1 Bibliographic information

Med številnimi obstoječimi zbirkami predstavljamo nekaj pomembnejših zbirk, ki vsebujejo večje število dokumentov, relevantnih za področje lesarstva in sorodnih vsebin. Zbirke *Agricola*, *Agris* in *CAB Abstracts* so velike splošne biotehniške zbirke in zajemajo vse biotehniške vidike, torej od lesarstva v tehnološkem do gozdarstva v biološkem smislu. Zbirka *Biosis* poudarja biološke vsebine. Zbirki *Compendex* in *Inspec* indeksirata predvsem tehniške oz. tehnološke vidike, zbirka *Iconda* pa gradbeništvo in arhitekturo. Zbirke *Web of Science*, ki smo jo za primerjavo uvrstili v uvod, nismo posebej analizirali, saj je splošno poznana kot najbolj avtoritativna zbirka za vsa področja znanosti, vendar pa pri tej zbirki ni možno iskanje zgolj s kontroliranim besednjakom, zato taka pri-

merjava z drugimi zbirkami ni možna. Več podatkov o obravnavanih pa tudi številnih drugih zbirkah je na voljo prek spletnih strani velikih mednarodnih informacijskih posrednikov, kot so npr. *OVID* (2005), *STN* (2005) ali *Thomson Dialog* (DIALOG 2005). Podatki o kumulativnem številu zapisov in letnem prirastu se med posredniki nekoliko razlikujejo.

##### Izbrane biotehniške, biološke in tehniške bibliografske zbirke

**AGRICOLA.** Zbirka zajema vsa podpodročja biotehnike oz. kmetijstva v širšem smislu, vključujoč gozdarstvo in lesarstvo. Vsebuje okrog 3.800.000 zapisov od leta 1970 in letno prirašča za okrog 110.000 novih enot. Gradi jo ameriška nacionalna knjižnica NAL (National Agricultural Library). Zadnja leta se osredotoča predvsem na zbiranje ameriških dokumentov. Dostopna je prek komercialnih namenskih iskalnih programov pa tudi brezplačno prek spletnih strani knjižnice NAL (2007).

**AGRIS.** Zbirka zajema podobna podpodročja biotehnike oz. kmetijstva kot Agricola. Vsebuje okrog 3.200.000 zapisov od leta 1975 dalje in letno prirašča za okrog 130.000 enot. Gradi se kooperativno prek nacionalnih centrov Agris, ki ga imamo tudi v Sloveniji, podatki pa

se v enovito bazo združujejo pri FAO v Rimu v okviru informacijskega sistema Waicent (World Agricultural Information Centre). Dostopna je prek komercialnih namenskih iskalnih programov pa tudi brezplačno prek spletnih strani FAO/Waicent (AGRIS/CARIS 2007).

**BIOSIS Previews** (tudi kot **Biological Abstracts**). Zajema vsa podpodročja bioloških znanosti. Gradi jo organizacija Thomson Scientific. Vsebuje okrog 13.000.000 zapisov od leta 1969 dalje in letno prirašča za okrog 560.000 enot. Zajema tudi gozdarstvo in lesarstvo s poudarkom na bioloških vidikih (BIOSIS 2007).

**CAB Abstracts (CABA)**. Tudi ta zbirka zajema vsa področja biotehnike. Velja za najkvalitetnejšo biotehniško podatkovno zbirko. Vsebuje okrog 4.000.000 zapisov od leta 1973 dalje in letno prirašča za okrog 180.000 enot. Gradi jo organizacija CAB International s sedežem v Veliki Britaniji (CAB International 2005). Zbirka je dostopna le prek komercialnih namenskih iskalnih programov. Iz podatkov te zbirke se napajajo tudi številne manjše specializirane zbirke, ki se trži-jo kot samostojni izdelek. Taka zbirka je npr. Forest Science Database (FSD 2007a, b), bivši TREECD, ki s 650.000 zapisi hrani tudi arhivske podatke od 1939 dalje. Skupna zbirka CABA v novjšem segmentu večinoma zajema tudi zapise iz FSD.

**Compendex**. Zajema vsa področja tehnike. Vsebuje okrog 8.000.000 zapisov od leta 1969 dalje, z letnim prirastom okrog 500.000 enot. Gradi se v okviru informacijskega sistema Ei-Engineering Information, ki je v lasti založbe Elsevier. Iskalni sistem omogoča hkratno iskanje po zbirkah Compendex ter Inspec z avtomatskim izločanjem dvojnikov (EI 2007).

**Inspec**. Zajema tehniko s poudarkom na elektrotehniki. Vsebuje okrog 8.000.000 milijonov dokumentov od leta 1969 dalje, z letnim prirastom okrog 400.000 zapisov. Gradi se v okviru informacijskega sistema Ei-Engineering Information, tako kot zbirka Compendex.

**Iconda**. The International CONstruction DATabase. Zajema vsa področja gradbeništva, vključno z lesenimi konstrukcijami in stavbnim pohištvo. Vsebuje okrog 600.000 zapisov od leta 1976 dalje. Gradi se v okviru centra *Fraunhofer Information Centre for Regional Planning and Building Construction (IRB)* (Iconda 2007).

### 3.1.1 Biotehniške informacije

#### 3.1.1.1 Agricultural / biotechnical information

##### CAB Abstracts

Kot terminološko osnovo za vse nadaljnje zbirke smo uporabili zbirko **CAB Abstracts (CABA)**, ki sicer velja za najkvalitetnejšo splošno biotehniško zbirko. S pomočjo drevesastih struktur asociativnih (*RT - Related Term*) in hierarhičnih relacij (*NT - Narrower Term*, *BT - Broader Term*) iz tezavra CAB smo identificirali pomembnejše predmetne oznake (*predmetne kategorije in deskriptorje*),

ki se v tej zbirki uporabljajo za indeksiranje in iskanje dokumentov z lesarskimi vsebinami.

##### Predmetne kategorije (**CabiCodes**)

Vseh predmetnih kategorij, ki se pri zbirki CABA imenujejo **CABI Codes**, je okrog 250. Sezname so prosto dostopni na spletu. Skupina oz. kategorija *KK (Forestry, Forest Products and Agroforestry)* zajema vse širše vidike gozdarstva in lesarstva. Vseh dokumentov, ki so bili v tem obdobju indeksirani z lesarskimi podkategorijami, je bilo 17.000. To število na zajema vsote posameznih pojavnosti iz preglednice 3, pač pa iskalno unijo vseh dejansko različnih dokumentov, saj se posameznemu dokumentu lahko priredi več kategorij iz iste skupine pa tudi iz drugih skupin.

##### Deskriptorji (**CAB Descriptors**)

Deskriptorjev, ki se uporabljajo za indeksiranje na ravni mikro-kategorizacije, je precej več. CAB Thesaurus vsebuje kar 48.000 deskriptorjev, poleg tega pa še 10.500 nedeskriptorjev, ki kot premetne kategorije uporabnika usmerjajo k rabi najustrežnejšega (kontroliranega) izraza za neki pojem. Tezaver je prek gesla dostopen tistim uporabnikom, ki imajo do zbirke licenčni dostop, npr. prek namenskega programa *Ovid*.

Za osnovo smo vzeli terminološke korene: **board-, furniture, timber-, veneer-, -wood-** in zbrali deskriptorje, ki te izraze vsebujejo v smislu lesarskega konteksta, npr. *particleboards, plywood, veneers, structural timbers* ipd. Pri tem smo pazili na možnost sestavljenega pisanja ter na možnost tako leve kot desne pojavnosti, npr. *plywood* ali *woodworking*. Druge deskriptorje smo pridobili s pomočjo analize asociativnih in hierarhičnih odnosov med deskriptorji. Zajeli smo tudi nekaj deskriptorjev s področja gozdarstva, in sicer predvsem za vsebine spravila in obdelave lesa (prim. *fellings, logging, saws ...*), ki se v zbirki CABA asociativno vežejo tudi na lesarstvo. Pri tem poudarjamo, da je potencialno zanimivih deskriptorjev še več, zato so podatki v preglednici 3 namenjeni le informaciji o izbranih deskriptorjih. Koreni so poudarjeni s krepkim tiskom; pojavnost je razvrščena padajoče.

Tu je treba omeniti nekatere značilnosti procesa indeksiranja oz. določanja deskriptorjev posameznih dokumentov. Iskanje s pomočjo širših deskriptorjev ni zanesljivo, saj v zbirki CABA ne zajamejo ozkih in natančnih pojmov. V preglednici 3 lahko vidimo, da se samostojen deskriptor *wood* tako pojavi le pri 1470 bibliografskih zapisih, deskriptor *wood properties* pa kar pri 2114 zapisih. Te značilnosti indeksiranja narekujejo pazljivost pri rabi posameznih deskriptorjev. Uporabnik lahko namreč pri neprecizni rabi zgreši množico relevantnih dokumentov.

**Pomembno:** če želimo priklicati vse dokumente, kjer se deskriptor *wood* pojavlja sam, pa tudi v vseh možnih kombinacijah, moramo iskalno sintakso zasnovati kot ***wood \$.de***, kjer pa je treba paziti, da bo med izrazom *wood* in znakom za dolar (ki ponazarja desno krajšanje) dejansko *presledek*. S sintakso ***wood\$.de*** (brez presledka pred \$) prikličemo tudi izraze, kot so *woodlands, woody plants* ipd.

**Preglednica 3:** Izbrani deskriptorji za področje lesarstva iz zbirke CAB Abstracts in število pripadajočih dokumentov v obdobju 2000-2005

**Table 3:** Selected wood science related descriptors in CAB Abstracts and occurrence of respective documents between 2000-2005

Deskriptor Descriptor	n.	Deskriptor Descriptor	n.	Deskriptor Descriptor	n.
wood properties	2114	dendroclimatology	272	timber connectors	34
wood	1470	wood panels	271	skidding equipment	28
forest products industries	1360	laminated wood	253	wood liquefaction	28
wood anatomy	1060	skidding	239	timberyards	27
wood destroying fungi	1010	pulpwood	237	wood flour	27
clear felling	1006	waste wood	233	budwood	25
wood density	955	earlywood	220	liberation felling	25
fuelwood	928	sawmills	211	wood tar	25
wood strength	783	woodland grasslands	211	sanitation fellings	24
sawdust	775	old and fossil wood	201	shortwood logging	22
timbers	774	wood ash	199	stored wood	22
wood preservatives	749	reaction wood	172	branchwood	19
sawnwood	728	fibreboards	157	clear strip felling	19
wood preservation	723	wood extracts	150	log size	19
wood products	717	joints (timber)	147	blockboard	16
dendrochronology	711	wood technology	144	sleepers	16
dead wood	707	structural timbers	142	veneering	16
forest products	678	juvenile wood	113	wood carving	14
heartwood	619	stumpage value	113	improvement fellings	12
timber trade	604	beams	110	wood chemical industry	12
wood chips	593	pulpwood production	107	wood wool	12
wood moisture	579	wood hydrolysis	107	bentwood	10
sapwood	578	chainsaws	99	frame saws	10
logging effects	572	wood borers	99	wood conservation	8
wood fibres	565	wood dust	96	wood gas	8
logs	560	skidders	95	waferboards	6
wood chemistry	554	wood cement	94	paintings on wood	4
particleboards	509	wood shavings	91	log cabins	3
wood defects	459	log grade	89	sour felling	3
sawmilling	443	salvage felling and logging	79	wood handling	3
wood residues	423	poles	78	wood odour	3
felling	415	smallwood	74	carved wood	2
hardwoods	401	illicit felling	68	clearing saws	1
logging machines	384	wood plastic composites	65	frozen wood	1
roundwood	375	composite wood assemblies	58	hand saws	0
plywood	371	saws	57	late wood	0
wood pulp	363	circular saws	54	log splitters	0
selective felling	348	whole tree logging	48	logging shears	0
wood utilization	337	flakeboards	46	logs with inclusions	0
timber supply	329	group fellings	46	refactory wood	0
softwoods	328	log breakdown methods	45	timbers, tropical	0
veneers	326	wood smoke	42	wood gum	0
furniture	308	tree length logging	37	wood molasses	0
composite boards	307	wood workers	37	wood sugar	0
woodworking	284	recycled wood	36	wood sulphite liquor	0
decayed wood	276	bandsaws	35	woodlot forestry	0

**Preglednica 4:** Izbrane predmetne kategorije za področje lesarstva v naštetih zbirkah ter število pripadajočih dokumentov v obdobju 2000-2005

**Table 4:** Selected wood science related subject categories in relevant databases and occurrence of respective documents between 2000-2005

Zbirka Database	Koda Code	Ime kategorije Name of the category	n.
CABI Codes	KK500	Forest Products and Industries (General)	2194
	KK510	Wood Properties, Damage and Preservation	6846
	KK515	Logging and Wood Processing	4509
	KK520	Wood Utilization and Engineered Wood Products	3632
	KK530	Chemical and Biological Processing of Wood	2289
Agris Subject Categories	K50	Processing of forest products	1537
AGRICOLA - Subject Category Codes	K510	Forest products (wood)	1841
	K520	Forest prod. (composite and reconstituted wood	498
	K530	Forest products (pulp and paper)	1386
	K540	Forest products (chemicals)	91
	K590	Forest products (miscellaneous)	215
Ei Classif. Codes (Compendex/Inspec)	415.3	Wood Structural Materials	1883
	811.1.1.5	Paper and Board	9183
	811.1.1.6	Paper and Board Specialties	6527
	811.1.1.7	Paper and Board Manufacture	3785
	811.1.1.8	Paper and Board Treatment	1097
	811.1.2.3	Paper and Board Machines	2402
	811.2	Wood & Wood Products	8165
	811.2.1	Wood and Wood Fibers	4968
	811.2.2	Pulpwood and other Fibrous Raw Materials	503
	811.2.3	Wood Waste, Bark and Agricultural Residues	841
	811.2.4	Wood Extractives and Silvichemicals	266
	811.2.5	Particleboards and Related Structural Products	334

### 3.1.2 Tehniške informacije

#### 3.1.2 Technical information

##### Compendex + Inspec

Po obeh komplementarnih tehniških zbirkah *Compendex* in *Inspec* je možno iskati skupaj prek informacijskega sistema *Engineering Village 2 (Ei)*. Tudi ta sistem uporablja kontrolirane predmetne oznake, in sicer tako deskriptorje na ravni mikro-kategorizacije (**Ei controlled terms**) kot predmetne kategorije na ravni makrokategorizacije (**Ei classification codes**).

Geslovník *Ei Thesaurus* zajema približno 16.000 deskriptorjev, predmetne kategorije *Ei* pa so precej bolj podrobne kot pri zbirki CABA, saj je kategorij *Ei* kar okrog 800. Temeljijo na štirih hierarhičnih ravneh, v nasprotju s CABOM, kjer sta le dve ravni.

Kot pri številnih drugih zbirkah tudi tukaj pri iskanju ne gre brez programskih pomanjkljivosti oz. posebnosti. Pri sestavljenih deskriptorjih je pri tej treba nujno

iskati na način fraznega iskanja, t.j. z rabo narekovajev, npr. "wood products". Iskanje z izrazom *wood products* brez oklepajev priključuje tudi kombinacije *wood chemicals*, *drug products* ipd.

**Pomembno:** Ta zbirka ima specifiko, da je mogoče s pomočjo narekovajev kot samostojno "frazo" zajeti tudi posamezno besedo. Če besedo *wood* postavimo med narekovaje kot "*wood*" *wn cv*, priključimo tiste dokumente, kjer se *wood* pojavlja natanko kot samostojni deskriptor, in ne kot del fraze (6.120-krat). Če pa iščemo kot *wood wn cv* (brez narekovajev pri *wood*), dobimo 9.132 pojavnosti, torej tako vse samostojne pojavnosti kot pojavnosti v frazah *wood laminates*, *wood preservation* ipd.

Pri tovrstnem navajanju frekvenc oz. pojavnosti je treba upoštevati, da posamezni deskriptorji lahko predstavljajo tudi precej širši kontekst in ne le tistega v povezavi z lesarstvom oz. lesom (npr. *laminated composites*, *sawing of metal...*)

V preglednici 5 predstavljamo deskriptorje in njihovo pojavnost v teh dveh zbirkah, v preglednici 4 pa pojavnost zapisov, indeksiranih s predmetnimi kategorijami. Kot pri prejšnji zbirki tudi tu predstavljamo le najpogostejše predmetne oznake. Osnovne deskriptorje smo najprej poiskali s pomočjo *tezavra CABA*, potem pa smo jih v zbirkah Compendex/Inspec nadalje poiskali na asociativni in hierarhični način. Pri predmetnih kategorijah smo uporabili hierarhični klasifikacijski sistem Ei, in sicer vrhno kategorijo 400 ter predvsem kategorijo 800, ki zajema večino pojmov s področja lesarstva. Zanimivo je, da področje lesarstva zajema širša kategorija *Chemical Engineering*, znotraj katere se nato razvrščajo ustrezne lesne podkategorije.

**Preglednica 5:** Izbrani deskriptorji za področje lesarstva iz sistema Ei (Compendex+Inspec) in število pripadajočih dokumentov v obdobju 2000-2005

**Table 5:** Selected wood science related descriptors in CAB Abstracts and occurrence of respective documents between 2000-2005

Deskriptor Descriptor	n.
wood	9132
"wood"	6120
laminated composites	3543
timber	2670
softwoods	2397
hardwoods	2382
wood products	1999
lumber	1141
logging (forestry)	1119
sawmills	903
charcoal	803
plywood	486
sawing	472
furniture	449
wooden construction	369
sawdust	364
furniture manufacture	313
wood preservation	307
wood wastes	299
particle board	285
wood fuels	274
woodworking	209
veneers	209
saws	207
wood chemicals	140
wood laminates	101
wooden floors	86
wooden buildings	77
skidding	77
wooden beams and girders	73
wooden bridges	53
woodworking machinery	49
woodworking tools	31
wooden containers	15
woodrooms	7
wooden fences	2

## Iconda

Tudi zbirka Iconda daje možnost iskanja s ključnimi besedami, vendar te niso standardizirane oziroma razvrščene v tezaver, pač pa se oblikujejo prosto. Število potencialnih ključnih besed v povezavi s pojmom *wood* je torej zelo visoko in ga ni moč zajeti sistematično. Kot zanimivost pa lahko dodamo, da je Iconda ena redkih zbirk, ki omogočajo iskanje z levim krajšanjem oz. maskiranjem. Izraz *\*wood* tako npr. priključuje izraze *plywood*, *birchwood*, *oakwood* ipd., lahko pa tudi kakšna krajevna imena, npr. *Goodwood*. Pri takem načinu je torej vsekakor treba računati na nekaj balasta.

Pri tej zbirki tudi ni moč šteti pojavnosti posameznih osnovnih besed, v nasprotju z zbirkama CABA in Compendex, saj je polje za ključne besede indeksirano oz. invertirano le po načelu posameznih besed (*word-indexed*), kjer sistem poišče vse deskriptorje, kjer se taka beseda pojavi.

**Pomembno:** Iskanje z izrazom *wood* pri zbirki Iconda priključuje ne le samostojnih deskriptorjev *wood*, pač pa tudi vse pojavnosti kombiniranih deskriptorjev, npr. *solid wood*, *wood screws*, *wood wool*. Iskanja torej ni moč omejiti na dokumente, kjer se pojavlja zgolj izraz **wood** kot tak.

## 3.2 Statistične informacije

### 3.2 Statistical information

#### EUROSTAT

Evropski statistični urad *Eurostat* zbira in harmonizira podatke, pridobljene s strani nacionalnih statističnih uradov držav članic Evropske unije (EUROPA – Eurostat 2007). Posamezne kategorije zajemajo tudi podatke držav, partneric EU. Eurostat izdaja stotine statističnih publikacij, med njimi tudi evropski statistični letopis (**Eurostat Yearbook**), ki pa vsebuje le najbolj osnovne podatke držav članic EU.

Za iskanje bolj natančnih podatkov o posameznih parametrih ali dejavnostih je možno uporabljati enostavni splošni iskalnik (Quick search). Iskanje z izrazom *wood* na splošno priključuje 322 dokumentov oz. datotek znotraj sistema *Eurostat*, izraz *wood products* pa 139 dokumentov. Vendar izraz *wood products* v sistemu Eurostat sploh ne obstaja kot sistemski deskriptor. Za natančnejše iskanje po dokumentih Eurostat je možno uporabljati tudi deskriptorje oz. kategorije iz seznama **CODED (Eurostat Concepts and Definitions Database)** (CODED 2006). Seznam predmetnih oznak CODED zajema okrog 4400 ožjih kategorij oz. deskriptorjev, ki so razvrščeni v devet glavnih kategorij. Kategorija *Agriculture and Fisheries* zajema 425 deskriptorjev. V to kategorijo se razvršča večina deskriptorjev, relevantnih za področje lesarstva. Na seznamu predstavljamo le tiste deskriptorje iz sistema CODED, ki vsebujejo izraz *wood*. Za iskanje priporočamo tudi splošno oblikovana gesla.



## Seznam deskriptorjev CODED

### List of CODED Descriptors

Chemical **wood pulp**  
 Further processed sawn**wood**  
 Industrial round**wood** (w. in the rough)  
 Mechanical **wood pulp**  
 Other industrial round**wood**  
**Plywood**  
 Pulp**wood**, round and split  
 Round**wood** production  
 Sawn**wood**  
 Semi-chemical **wood pulp**  
**Wood charcoal**  
**Wood fuel**  
**Wood pulp**  
**Wood residues**  
**Wood-based panels**  
**Wooden furniture**  
**Wooden wrapping and packaging material**  
**Woody biomass**

## FAOSTAT

**Faostat** je kooperativno grajena statistična zbirka **Združenih narodov FAO** (FAOSTAT 2007). Podatke o agregatih svoje kmetijske proizvodnje prispevajo države članice Združenih narodov ter nekatere mednarodne organizacije. Ker so možnosti centralnega preverjanja in nadzora podatkov s strani osrednjega koordinatorja omejene, je pri rabi podatkov potrebna pazljivost.

Ta zbirka zajema podatke v pet glavnih kategorijah: *Agriculture, Nutrition, Fisheries, Forestry, Food Quality Control*.

Kategorija Forestry je razdeljena na kategoriji **Forestry Data** ter **Forestry Trade Flow**. Prva kategorija se deli še na: **Roundwood, Sawnwood, Wood-Based Panels ter Pulp, Paper&Paperboard**.

Na seznamu predstavljamo glavne skupine produktov, ki pa se v sami zbirki delijo bolj podrobno.

## Seznam kategorij Faostat

### List of Faostat Categories

Chips and Particles	Roundwood
Fibreboard	Sawlogs and Veneer Logs
Forest Products	Sawnwood
Hardboard	Veneer Sheets
Insulating Board	Wood Based Panels
Joinery and Carpen of Wood	Wood Charcoal
Mechanical Wood Pulp	Wood Fuel
Particle Board	Wood Pulp
Plywood	Wood Residues
Pulpwood and Particles	Wooden Furniture

Za omenjene produkte je možno podatke iskati predvsem v naslednjih statističnih kategorijah: *Production, Imports, Exports (Quantity, Value)*.

## 3.3 Tehnične informacije

### 3.3 Technical information

#### Standardi ISO

Za standarde obstajajo različne zbirke, zato predstavljamo le osrednjo mednarodno organizacijo za standarde **ISO (International Standards Organization)** in njen iskalni sistem **Catalogue Online**. Za klasifikacijo in iskanje standardov v sistemu ISO se uporablja klasifikacija **ICS (International Classification of Standards)**. Glavnih klasifikacijskih področij ICS je 40, vendar vse številke niso v rabi, tako da je zadnje področje numerirano s številko 97. Največ pojmov v povezavi z lesom in sorodnimi dejavnostmi zajema področje *Wood Technology* (79), ki sestoji še iz pet ožjih podpodročij, ki se delijo na nadaljnja podpodročja. V preglednici 6 predstavljamo le izbor.

Poleg klasifikacije ICS je možno tudi sistematično iskanje s pomočjo oznake **tehničnih odborov (Technical Committees)**, ki sprejemajo standarde. Teh odborov je okrog 200. V preglednici 6 predstavljamo tudi nekaj odborov, ki sprejemajo standarde v povezavi z lesarstvom.

Standarde je v katalogu ISO možno iskati tudi s prosto oblikovanimi ključnimi besedami, in sicer po pojavnosti iskalnega izraza v naslovu, kratkem izvlečku in polnem besedilu standardov, čeprav se brezplačno izpiše ta le naslov in izvleček (bibliografska oblika standarda), prikaz polnega besedila pa je treba plačati. Izvlečki so precej kratki in dajejo le najosnovnejšo informacijo o vsebini standarda.

Obstaja tudi slovenska izdaja ICS (2007) v prevodu **Slovenskega inštituta za standardizacijo (SIST 2007)**, vendar pa v slovenskem katalogu **SIST (Iskalo SIST)** standardov ni mogoče iskati s šiframi ICS, pač pa le s prosto oblikovanimi angleškimi in slovenskimi ključnimi besedami. Iskati je možno le standarde, ki so bili sprejeti za rabo v Sloveniji.

#### Intelektualna lastnina - patenti (WIPO)

Osrednja mednarodna organizacija za varstvo intelektualne lastnine, kamor se uvrščajo patenti, modeli in znamke, je **WIPO (World Intellectual Property Organization)** (WIPO 2007). Za vsebinsko razvrščanje dokumentov WIPO uporablja mednarodno patentno klasifikacijo **IPC (International Patent Classification)** (IPC 2005). Zadnja izdaja IPC vsebuje okrog 70.000 različnih klasifikacijskih enot, ki se razvrščajo v osem vrhnjih kategorij oz. sekcij. Leseno pohištvo se uvršča v sekcijo **A (Human necessities)**, lesarstvo in drugi leseni izdelki pa predvsem v sekciji **B (Performing operations; Transporting)** in **E (fixed constructions)**.

Javno dostopne mednarodne informacije o dokumentih, povezanih z intelektualno lastnino, je možno iskati prek elektronske knjižnice **WIPO (Intellectual Property Digital Library)** (IPDL 2007), in sicer tako prek klasifikacije IPC kot s pomočjo prosto oblikovanih ključnih

**Preglednica 6:** Izbrane kode za področje lesarstva v sistemih ISO in WIPO**Table 6:** Selected wood science related codes in ISO and WIPO systems

Zbirka Database	Koda Code	Ime kategorije Name of the category
ICS ( <i>International Classification of Standards</i> )	71.100.50	Wood-protecting chemicals
	79	Wood technology
	79.020	Wood technology processes (Including logging and wood treatment)
	79.040	Wood, sawlogs and sawn timber
	79.060	Wood-based panels
	79.080	Semi-manufactures of timber (Incl. parquet, wood paving, handles, etc.)
	79.100	Cork and cork products
	79.120	Woodworking equipment
	91.080.20	Timber structures
ISO TC ( <i>Technical Committees</i> )	89	Wood-based panels
	136	Furniture
	165	Timber structures
	218	Timber
IPC ( <i>International Patent Classification</i> )	A47	Furniture
	B27	Working or preserving wood or similar material; nailing or stapling machines in general
	E04 C 2/10	Building elements of relatively thin form... of wood, fibres, chips, vegetable stems
	E04 C 3/12	Structural elongated elements designed for load-supporting ... of wood, e.g. with reinforcements...
	E04D 1/20	Roof covering... of plastics, fibrous materials, or wood
	E04F 11/108	Stairways, ramps, or like structures... of wood
	E04G 9/04	Forming or shuttering elements for general use... the form surface being of wood
	E04H 7/32	Construction or assembling of bulk storage containers... mainly of wood
	E06B 3/10	Window sashes, door leaves... wing frames... of wood

besed. Prav tako je mogoče iskati tudi naše domače bibliografske informacije o intelektualni lastnini, in sicer prek iskalnika Podatkovnih baz **Urada Republike Slovenije za intelektualno lastnino** (URS 2007). Gre predvsem za dokumente, kot so npr. podeljeni patenti, objavljene prijave in registrirani modeli, prijave in registrirane znamke.

## 4 Razprava in sklepi

### 4 Discussion and conclusions

V prispevku smo se osredotočili na nekatere zakonitosti znanstvenega in tehničnega informiranja na vsebinskem področju lesarstva, in sicer s pregledom nekaterih izbranih znanstvenih, tehničnih in statističnih informacijskih sistemov in podatkovnih zbirk. Pregledali smo nekaj osnovnih zakonitosti pri tvorjenju iskalne sintakse in primerjali in ocenili glavne razlike in podobnosti med različnimi zbirkami in iskalnimi navodili iz teh zbirk. Pregled sistemov kaže na nekatere pomembne razlike, ki jih mora uporabnik upoštevati, če želi s pridom izkoristiti sistemsko pravila. Izpostavili smo tudi nekaj pomanjkljivosti in specifik, ki v zbirkah in iskalnih navodilih navadno niso zadostno poudarjene, nepoznavanje pa lahko privede do pomanjkljivih ali celo napačnih iskalnih rezultatov. V

nadaljevanju smo posebno pozornost namenili predmetnim oznakam (deskriptorjem, predmetnim kategorijam), katerih poznavanje lahko pomembno vpliva na kakovost informacijskega odziva. Zbirke se pri vsebinskem in terminološkem pristopu kar precej razlikujejo, zato bi moral uporabnik dobro poznati ta vsebinska načela in razlike med zbirkami. Žal tudi tu marsikatera zbirka ne poudarja te potrebe; kreatorji zbirk imajo torej nekako nerealna pričakovanja glede ravni informacijske usposobljenosti posameznih uporabnikov.

V prispevku zbrani sezname predmetnih oznak in nadaljnje informacije, dostopne prek navedenih spletnih virov, naj torej uporabnikom olajšajo iskanje najustrežnejših iskalnih izrazov. Frekvenca pojavnosti nekaterih pojmov nakazuje aktualnost problematike. Nekaj osnovnih primerov iskalne sintakse naj pripomore k bolj preciznemu tvorjenju iskalne sintakse.

S prispevkom smo želeli tudi spodbuditi naše strokovnjake, da sežejo še po drugih podatkovnih zbirkah in informacijskih sistemih, ki so na voljo prek univerzitetne računalniške mreže. Nenazadnje so naročnine na nekatere podatkovne zbirke precej visoke, konzorcijski način nabave informacijskih virov pa omogoča več ustanovam

racionalnejši dostop do več različnih virov. Hkrati je treba vedeti, da se raba informacijskih virov statistično računalniško zabeleži, uporabniška raba pa dokazuje upravičenost konzorcijske nabave; zbirke namreč niso poceni in jih je treba tudi uporabljati, drugače nabava ni upravičena. Lesarski strokovnjaki z Biotehniške fakultete lahko torej še bolj zavzeto uporabljajo različne informacijske vire, saj bodo s tem potrjevali enakovredno sodelovanje naše ustanove v procesih domačega in mednarodnega kroženja znanstvenih, strokovnih in tehničnih informacij.

## 5 Summary

In order to retrieve high quality scientific and technical information, it is necessary to get acquainted with authoritative global information systems. Users should be able to identify the most relevant resources. They should then be capable of using each respective resource, which is not always evident owing to some particularities of database architecture, possibly resulting in inferior search results. There exist differences in query techniques, and more importantly, users should be more aware of differences and complexities of controlled dictionaries (thesauri) and indexing schemes. The article reviews principal characteristics of each information system in question and offers some advice as to the optimization of retrieval techniques.

The standard database content can be searched through *title*, *abstracts* and *subject heading*. Searches can be conducted with the help of predefined menus or more complex Boolean search queries composed of *terms*, *operators*, *prefixes* or *suffixes*. Search tools are not standardized, so end-users should be aware of characteristics of each respective database. Also, same databases can be accessed through different search portals or platforms. Among databases there exists significant difference with regard to the depth of *subject indexing*. Topical searches can be conducted in specific fields or in the full-record of a database. Some databases do not offer full-record search utility, but will restrict retrieval to the aforementioned content fields (*title*, *abstracts* and *subject headings*). Boolean query principles are standardized to a higher degree than other elements such as *prefixes*, *suffixes*, *punctuation marks* and *wildcards (truncation marks)*. We present the following different forms of prefixes/suffixes representing the **document title field**: tkey, title:, .ti, wn ti, <in> title, ti= **publication year field**: date: yr, wn yr, year=, and py=, and **descriptor field**: skey, subject:, .hw, wn cv, <in> KEYWORD, ts. Special attention needs to be given to punctuation marks.

We included the following bibliographic databases, which comprise topics related to wood sciences or processing of forest products and related subject matter: **Agricola**, **Agris**, **Cab Abstracts (CABA)**, **Compendex**, **Inspec** and **Iconda**. We selected subject headings in each respective database: narrow keywords (*Descriptors*, *Heading Words*, *Controlled Terms*, *Keywords*) and broader

categories (*Subject Categories*, *Concept Codes*, *CABI Codes*, *Classification Codes*). Special attention was given to Cab Abstracts, considered the principal general database for all fields of agricultural and related sciences. We identified the most important wood-science related descriptors, based on word-stems such as **board-**, **furniture**, **timber-**, **veneer-**, **-wood-**. We singled out some 150 primary descriptors. Specific descriptors, such as **wood properties**, **forest products industries**, **wood anatomy**, **wood destroying fungi**, **wood density**, **fuelwood**, accounted each for between 2000 and 1000 database records in the 2000-2005 period. The searches in technical databases Compendex+Inspec retrieved the following high-frequency descriptors, yielding between 3,500 and 2,000 records: **laminated composites**, **timber**, **soft woods**, **hard woods**, **wood products**. In the next step, we identified all relevant broader topical categories in the above databases. The topical category codes occur under different names, such as *Subject Categories*, *Concept Codes*, *Ei Classification Codes*, etc. Databases Compendex+Inspec provide the highest number of different codes, followed by CABA and Agricola.

We also examined statistical databases **Eurostat** and **Faostat**. Both the Eurostat's *CODED Descriptors* and Faostat's *Categories* employ some 20 different wood-related descriptors (e.g. **Mechanical wood pulp**, **Plywood**, **Sawnwood**). Finally, we investigated technical databases **ISO** (International Standards Organization), and **WIPO** (World Intellectual Property Organization). The ISO's *ICS (International Classification of Standards)* employs 10 principal wood-related categories (e.g. **Wood-protecting chemicals**, **Wood technology processes**, **Wood-based panels**). There are four ISO *Technical Committees* that deal mainly with wood or timber. The WIPO's *IPC (International Patent Classification)* also employs 10 principal wood-related categories (e.g. **Furniture**, **Working or preserving wood or similar material**, **Structural elongated elements... of wood**).

The review of the above information systems wishes to promote employment of additional databases, given that there may exist more pertinent databases than users are aware of. It encourages the employment of structured vocabularies and thus promotes better terminological choice, and offers some principal examples to this end. This should result in improved and more comprehensive utilization of information resources that are available as a consortium access to users at the Slovenian institutions of higher education.

## 6 Viri

### 6 References

AGRICOLA Subject Category Code Scope Notes, 2000. Agriculture Network Information Center, <http://permanent.access.gpo.gov/Cheddar%20Cheese%20Prices/cheddar/www.agnic.org/cc/index.html> (15. maj 2007).

- Agricultural Information Management Standards Web site (Agrovoc), [http://www.fao.org/aims/ag\\_intro.htm](http://www.fao.org/aims/ag_intro.htm) (15. maj 2007).
- AGRIS/CARIS Information Centre Homepage, <http://www.fao.org/AGRIS/> (15. maj 2007).
- Agris Subject Categories, <http://www.fao.org/scripts/agris/c-categ.htm> (15. maj 2007).
- Biological Abstracts Concept Codes, <http://etextb.ohiolink.edu/help/bioa-concept.html#535> (15. maj 2007).
- BIOSIS - Thomson Scientific, <http://www.biosis.org/> (15. maj 2007).
- CAB International, 2005. CAB International, <http://www.cabi.org/> (15. maj 2007).
- CABICODE Alphabetical List, [http://support.dialog.com/searchaids/dialog/f50\\_cabicodes\\_list.shtml](http://support.dialog.com/searchaids/dialog/f50_cabicodes_list.shtml) (15. maj 2007).
- CODED, Eurostat Concepts and Definitions Database. <http://forum.europa.eu.int/irc/dsis/coded/info/data/coded/en.htm> (2. jan. 2006).
- DE SOLLA PRICE, D., 1963. Little Science, Big Science. New York, Columbia University Press, 119 s.
- DIALOG Bluesheets, 2005. (Databases by File Number) Thomson Dialog, <http://library.dialog.com/bluesheets/html/bl0118.html> (15. maj 2007).
- EELS, Engineering E-Library, Sweden (Browse engineering resources classified using EI classification), <http://eels.lub.lu.se/> (15. maj 2007).
- EI - Welcome to Engineering Information, 2005. Elsevier, <http://www.ei.org/> (15. maj 2007).
- EUMABOIS, 1980. Večjezični terminološki priročnik s slovenskimi izrazi za stroje in opremo za obdelavo lesa, sestavljen po tehniški klasifikaciji Komiteja evropskih konstruktorjev za obdelavo lesa "EUMABOIS" 1980, 1984, Ljubljana, Zveza inženirjev in tehnikov gozdarstva in lesarstva Slovenije, 39 s.
- EUROPA - Eurostat, [http://epp.eurostat.cec.eu.int/portal/page?\\_pageid=1090,1&\\_dad=portal&\\_schema=PORTAL](http://epp.eurostat.cec.eu.int/portal/page?_pageid=1090,1&_dad=portal&_schema=PORTAL) (15. maj 2007).
- FAOSTAT, <http://faostat.fao.org/> (15. maj 2007).
- FSD, 2007a. Forest Science Database, <http://www.forestscience.info/> (15. maj 2007).
- FSD, 2007b. Forest Science Database, <http://www.ovid.com/site/catalog/DataBase/158.jsp?top=2&mid=3&bottom=7&subsection=10>
- GARFIELD, E., 1964. Science Citation Index - A New Dimension in Indexing.- Science 144, 3619: 649-654.
- Iconda, The International Construction Database, 2005. Fraunhofer Informationszentrum Raum und Bau. <http://www.irbdirekt.de/iconda/> (15. maj 2007).
- ICS, Mednarodna klasifikacija za standarde, Prva izdaja, 2003. Ljubljana, Slovenski inštitut za standardizacijo: 66 p. (prevod International Classification For Standards. ISO. Fifth edition, 2001), <http://www.sist.si/slo/z1/mednarodna-%20klasifikacija%20za%20standarde.pdf> (15. maj 2007).
- IPC, International Patent Classification (IPC), English version, Sections, 2005. WIPO. [http://www.wipo.int/classifications/fulltext/new\\_ipc/index.htm](http://www.wipo.int/classifications/fulltext/new_ipc/index.htm) (2. jan. 2006).
- IPDL, Intellectual Property Digital Library, <http://www.wipo.int/ipdl/en/> (15. maj 2007).
- NAL Catalog (Agricola), <http://agricola.nal.usda.gov/> (15. maj 2007).
- National Agricultural Library Services: National Agricultural Library Agricultural Thesaurus, <http://agclass.nal.usda.gov/agt/agt.shtml> (15. maj 2007).
- Ovid Alphabetical Listing (Databases), 2005. 2005 Ovid Technologies, [http://www.ovid.com/site/catalog/Catalog\\_DataBase.jsp](http://www.ovid.com/site/catalog/Catalog_DataBase.jsp) (15. maj 2007).
- SIST iskalo, [http://iskalo.sist.si/slo\\_strani/FMPro?-db=Orders.FP5&-format=SearchPage\\_sl1.html&-new](http://iskalo.sist.si/slo_strani/FMPro?-db=Orders.FP5&-format=SearchPage_sl1.html&-new) (15. maj 2007).
- STN Database Summary Sheets, 2005. American Chemical Society. <http://info.cas.org/ONLINE/DBSS/dbsslist.html> (15. maj 2007).
- ŠEGA, B., 2005. Preskušanje in razvrščanje lepil za les.- Les 57, 4: 101-105.
- TORELLI, N., 1979. Večjezikovni slovar izrazov, ki se uporabljajo v lesni anatomiji IAWA: slovenska verzija. Ljubljana, Biotehniška fakulteta, 33 s.
- TORELLI, N., 2000. Oljka (*Olea europaea* L.), oljkov les.- Les 52, 9: 286-289.
- URS, Podatkovne baze Urada Republike Slovenije za intelektualno lastnino, <http://www2.uil-sipo.si/> (15. maj 2007).
- WIPO - World Intellectual Property Organization, <http://www.wipo.int/portal/index.html.en> (15. maj 2007).