
ASIO Lyhytkurssiohjelman hakupalvelun kuvaus

Versio 1.0

Asio-Data Oy

3.1.2011

Sisällysluettelo

1. Hakupalvelun esittely	1
1.1 Yleinen toiminnallinen esittely	1
2. Hakupalvelun tekninen kuvaus	1
2.1 Tietokantakuvaus	1
2.2 Hakutoiminnon asetusten määrittäminen	2
2.3 Hakutaulun periodinen päivitys	3
2.4 Hakupalvelun ohjelmat ja kirjastot	4
3. Hakulauseiden muodostaminen	4
3.1 Haussa käytettävät parametrit	5
3.2 Hakuesimerkkejä	6
4. Hakutulosten esitysmuoto	6
4.1 XML-muotoiset hakutulokset	6
5. Muut asiat.....	7
5.1 Mahdolliset jatkokehityskohteet	7

Muutoshistoria

Nimi	Pvm	Muutoksen syy	Versio

1. Hakupalvelun esittely

1.1 Yleinen toiminnallinen esittely

ASIO Lyhytkurssiohjelman hakupalvelun tarkoituksena on mahdollistaa kurssitietoihin kohdistuvien erilaisten hakujen käyttäminen. Hakupalvelun avulla voidaan koulutustarjonnasta hakea yksittäisiä kursseja sekä erilaisten hakusanamääritysten avulla että vapaasti määriteltävien luokitteluiden avulla.

Hakupalvelu on lähtökohtaisesti tarkoitettu järjestelmien väliseen tiedonsiirtoon, jonka takia hakujen muodostaminen tuotetaan yksinkertaisilla HTTP GET-tyyppisillä pyynnöillä ja hakutulokset esitetään palvelun palauttaman XML-muotoisen dokumentin avulla. Asiakasliittymä hakujen muodostamiseen ja hakutulosten esittämiseen ei kuulu hakupalvelun ominaisuuksiin.

Hakupalvelun käyttämä kurssitiedoista luotu hakuaineisto voidaan määrittää asiakaskohtaisesti. Tämä määrittäminen sisältää tiedon siitä, mitä kurssien kuvaustekstejä otetaan mukaan sanahakuun, sekä mitä luokitteluja kursseihin liittyen voidaan haussa huomioida.

2. Hakupalvelun tekninen kuvaus

2.1 Tietokantakuvaus

Hakupalvelun teknisessä toteutuksessa hakuja varten muodostetaan hakuaineisto omaan tietokantatauluun varsinaisista kurssirekisterin tiedoista. Hakuaineistoa varten luotu tietokantataulu on nimeltään "hakuindeksi_p", jonka rakenne on seuraava:

Field	Type	Attributes	Null	Default	Extra
rowid	int(11)		No		auto_increment
tyyppi	varchar(50)		No		
avain	varchar(50)		No		
kielikoodi	varchar(25)		No		
hakuteksti	longtext		No		
p1	varchar(255)		No		
p2	varchar(255)		No		
p3	varchar(255)		No		
p4	varchar(255)		No		
p5	varchar(255)		No		
p6	varchar(255)		No		
p7	varchar(255)		No		
p8	varchar(255)		No		
p9	varchar(255)		No		

Hakutaulun sarakkeiden merkitykset ovat seuraavat:

- Tyyppi: sisältää tiedon hakuaineiston sisällöstä (esim. "course"...).

- *Avain*: yksilöivä avaintieto, joka viittaa rivillä esitetyn hakuaineiston lähteeseen (esim. *course_p*-taulusta löytyvän kurssin *rowid*-tieto).
- *Kielikoodi*: rivin hakuaineiston kielikoodi (esim. "fi", "en"...).
- *Hakuteksti*: tekstikenttä, johon on yhdistetty kaikki tekstihakuun konfiguroidut tiedot "#"-merkillä eroteltuna (esim. "kurssin sisältö # kurssin esitiedot # kurssin kuvaus").
- *P1-p9*: yleiskäyttöisiä sarakkeita konfiguroitujen luokittelutietojen tallentamiseen (esim. *p1*=kurssiryhmä, *p2*=toimipiste...).

Mikäli tulevaisuudessa hakupalvelua käytetään muille tiedoille kuin kurssikuvauksiin, voidaan Tyyppi-tietoon kirjata tämä uusi tietotyyppi. Lisäksi erikielisille hauille voidaan omat hakutaulun rivit, sekä hakupyynnöt.

Koska hakuja ei tehdä suoraan kurssirekisteriin kirjattuun aineistoon, tulee hakutaulu – jonka sisältö tuodaan kurssirekisteristä – päivittää määrätyin aikaväleihin hakuaineiston pitämiseksi ajantasalla.

2.2 Hakutoiminnon asetusten määrittäminen

Hakupalvelun käyttöönoton yhteydessä tulee perustaa asiakaskohtaiset asetukset hakuaineiston muodostukselle. Nämä asetukset määrittävät sanahaussa käytetyn tekstiaineiston koostamisessa käytetyt kurssien tietoihin liittyvät kentät, sekä *p1-p9*-parametreihin kerätyt luokitustiedot. Pääosin kyse on *course_p*, sekä *course_info_p* -tauluista löytyvistä tiedoista. Lisäksi, koska hakupalvelun hakujen muodostuksessa oletetaan *p*-parametrien arvojen olevan numeerisia tunnuksia (positiivisia kokonaislukuja), voidaan asetuksissa viitata kentän tunnukselle parametri_ *p* -taulun määrittäykseen.

Hakupalvelun asetukset määritetään asiakaskohtaisia konfiguraatioita sisältävässä *server_settings.php* -tiedostossa, joka löytyy "/inc"-kansioista.

Hakupalvelun asetukset määritetään seuraavien muuttujien avulla:

Nimi	Arvon tyyppi	Selitys
<code>\$course_search_data_{kielikoodi}</code>	Array[taulu][kenttä]	Sanahaun aineiston keräyksessä käytetyt kentät taulukohtaisesti. Kielikoodin avulla voidaan muodostaa keräilyasetukset usealle hakukielelle.
<code>\$course_search_params</code>	Array[<i>p1-p9</i>][taulu, kenttä, luokka, koodi, arvo, Array[kielikoodi => tunnus], "tokenize"]	Luokitusparametrien määrittäminen. Pakollisia tietoja ovat taulu ja kenttä. Loput määrittävät yhteydet parametri_ <i>p</i> -tauluun, mikäli arvon tunnus määräytyy sen avulla. Tokenize ilmaisee, että arvoja on useita, ja ne tulee liittää yhteen seuraavalla tavalla: "arvo1 # arvo2 # arvo3..."
<code>\$course_search_search_params</code>	Array("p1","p2","p3"...)	P-luokitusten arvot voidaan liittää myös sanahaun

		tekstiaineistoon. Tässä listataan kaikki liitettävät p-luokat.
<code>\$course_search_database</code>	String	Hakupalvelun käyttämän Lyhytkurssiohjelman tietokannan nimi.

Seuraavassa on esimerkki hakupalvelun asetuksista:

```
// Kurssien hakutoimintojen konfiguraatio:
// Kurssitietojen sijainti:
// fi
$course_search_data_fi = array("course_p"=>array("course_name", "course_name_eng", "notes", "other_info", "goal", "description", "prerequisite", ""),
                              "course_info_p"=>array("kuvaus", "target_group"));
// en
$course_search_data_en = array("course_p"=>array("course_name_eng"),
                              "course_info_p"=>array("accomplishment_eng", "attachments_eng",
                                                    "bibliography_eng", "cancellation_terms_eng", "description_eng", "diploma_text_eng",
                                                    "internal_notes_eng", "kuvaus_eng", "notes_eng", "other_info_eng", "prerequisite_eng", "target_group_eng"));
// Kurssitietojen parametrit:
$course_search_params = array("p1"=>array("course_info_p", "marketing_U_A", "U", "A", "arvo", array("fi"=>"arvo", "en"=>"arvo"), "tokenize"),
                              "p2"=>array("course_info_p", "marketing_U_B", "", "", "", array()),
                              "p3"=>array("course_p", "verkko_opinto", "", "", "", array()),
                              "p4"=>array("course_p", "viikonloppuopinto", "", "", "", array()),
                              "p5"=>array("course_p", "iltaopinto", "", "", "", array()),
                              "p6"=>array("course_info_p", "2days_course", "", "", "", array()));

// Kurssitietoihin yhdistetyt parametritiedot:
$course_search_search_params = array("p2");
// Kurssien hakujen tietokanta:
$course_search_database = "asiakas12";
```

2.3 Hakutaulun periodinen päivitys

Kaikki hakupalveluun tehdyt haut tehdään hakupalvelun omaa hakuindeksi_p-taulua vasten, siksi on tärkeää, että hakutaulu päivitetään säännöllisin väliajoin vastaamaan viimeisintä kurssirekisterin tilaa. Hakutaulun päivittäminen ei ole useissa tapauksissa erityisen raskas operaatio, joten se voidaan yleensä tehdä useita kertoja päivässä.

Hakutaulukon automaattinen periodinen päivitys voidaan ajastaa Unix-pohjaisissa ympäristöissä Cron-ajastuspalvelulla, jolloin konfiguraatiossa käsketään käyttöjärjestelmää tekemään http-kutsu wget-komennon avulla seuraavaan palveluun osoitteeseen: "{palvelimen ja asennuskansion osoite}/lyhytkurssi/haku_paivita.php". "haku_paivita.php"-ohjelman ajo suorittaa hakuindeksi_p -taulun sisällön päivittämisen vastaamaan viimeisintä kurssirekisterin tilaa.

Esimerkki Cronin konfiguraatiosta (päivitys kerran tunnissa):

```
*/60 * * * * wget -q -O - 'http://asp.asio.fi/lyhytkurssi/haku_paivita.php' >/dev/null
```

Hakutaulukko voidaan päivittää tarvittaessa myös manuaalisesti Lyhytkurssiohjelman ylläpitovalikosta löytyvässä "Hakutoimintojen ylläpito"-näkyvässä. Tähän näkymään voidaan tarvittaessa tulevaisuudessa siirtää hakujen konfiguraatiota, mikäli se muokataan käyttöliittymän kautta tehtäväksi.

2.4 Hakupalvelun ohjelmat ja kirjastot

Hakupalvelun ohjelmallinen toteutus koostuu seuraavista tiedostoista:

Sijainti ja tiedosto	Toiminnot
/kalenterit2/kalenteri_hakutoiminnot.php	Hakutoimintojen ylläpito näkymä. Vaatii autentikoinnin.
/lyhytkurssi/haku.php	Hakujen vastaanotto, käsittely ja hakutulosten palautus. Ei vaadi autentikointia.
/lyhytkurssi/haku_paivita.php	Hakutaulukon päivittämisen käynnistys. Ei vaadi autentikointia.
/inc/hakuindeksi.php	Hakutaulun sisällön muodostamiseen käytetyt funktiot.

3. Hakulauseiden muodostaminen

Kun hakupalvelun asetukset on määritelty ja hakutaulu on päivitetty asetusten mukaisesti, voidaan hakupalvelun avulla tehdä erilaisia hakuja. Kaikki haut tehdään HTTP GET -tyyppisellä kyselyllä /lyhytkurssi/haku.php -ohjelmaan, siten että hakutekijät (hakusanat ja p-luokitukset) annetaan ns. "query string"-parametreilla.

Hakupalvelun hakutekijät jakaantuvat kahteen luokkaan: sanahaut ja luokitushaut. Molempia voidaan käyttää erikseen siten, että toista ei tarvitse käyttää.

Seuraavassa esitellään käytössä olevat parametrit ja niiden käyttö.

3.1 Haussa käytettävät parametrit

Seuraava taulukko listaa kaikki hakulauseiden muodostuksessa käytettävissä olevat parametrit, jotka annetaan haun HTTP GET -kutsussa seuraavassa muodossa: `"/haku.php?[parametri1]=[arvo1]&[parametri2]=[arvo2]&..."`.

Mikäli haun yhteydessä ei ole määritetty kyseistä parametria, saa se arvokseen oletusarvon.

Parametrin nimi	Oletusarvo	Selite / esimerkki
Kieli	"fi"	Määrittää haun kielivalinnan. Jokainen kielivalinta tulee ottaa erikseen huomioon hakupalvelun asetuksissa (kts. Kohta 2.2). Esim: <code>"...&kieli=fi"</code> .
Hakusanat	Tyhjä	Lista sanahaussa käytettävistä hakusanoista välimerkillä eroteltuna. Mikäli kyse on lauseesta (hakutuloksissa tulee olla annetut sanat peräkkäin), ilmaistaan se paramerilla

		<p>“vertaus”.</p> <p>Esim: “...&hakusanat=kurssi opetus kevät”.</p>
Vertaus	“osa”	<p>Sanahaun tyyppi. Vaihtoehdot ovat “osa”, “lause” ja “kaikki”. Osa tarkoittaa, että ainakin yhden hakusanoista on esiinnyttävä sanahakuaineistossa. Kaikki tarkoittaa, että kaikkien hakusanojen tulee löytyä hakuaineistosta. Lause tarkoittaa, että hakuaineistosta tulee löytä annetus hakusanat samassa peräkkäisessä järjestyksessä.</p> <p>Esim: “...&vertaus=kaikki”.</p>
p1-p9	Tyhjä	<p>P-luokitusten (1-9) arvot. Luokitukselle voidaan asettaa kyllä/ei arvoja pilkulla eroteltuna. Huutomerkki (“!”) arvon edessä tarkoittaa negaatiota.</p> <p>Esim: “...&p3=4,!11,2”.</p>
p	“1a2a3a4a5a6a7a8a9”	<p>P-luokitusten lausemuoto, sekä vaihtoehtoisesti myös luokitusten arvot. Lausemuoto tarkoittaa loogisen lausekalkyylin esitystä, joka tukee seuraavia operaatioita: and, or, xor, not, sekä sulkujen (“(”, “)”) avulla esitettyjä järjestyssuhteita. Lisäksi p-luokitusten arvot voidaan antaa luokituksen yhteydessä []-sulkujen sisällä samaan tapaan, kuin p1-p9 -parametreissa.</p> <p>Operaatiot ja luokitukset on p-parametrin arvossa lyhennetty seuraavasti: a = and o = or x = xor 1-9 = p1-p9</p> <p>Mikäli p-parametria ei ole annettu ja käytetään oletusta, korvautuvat sellaiset parametrit joita ei ole annettu tautologialla “and 1=1”, jotta ne eivät riko kyselyä.</p> <p>Esim: “...&p=((2[1,!2,4]a1[!1])o(5))”, joka tarkoittaa: (p2=1 and not-2 and 4 and p1=not-1) or p5, jossa p5 on parametrilla p5 annettu lista arvoista.</p>

3.2 Hakuesimerkkejä

Seuraavat ovat oikein muodostettuja hakuja:

Hakulause	Selitys
../haku.php?hakusanat=matematiikka kevät&vertaus=kaikki&kieli=fi	Hakee kaikki kurssit, joiden hakupalveluun konfiguroiduissa tekstikuvauksissa esiintyy molemmat sanat "matematiikka" ja "kevät". Hakukieli on "fi".
../haku.php?p=1[1]a2[3]	Hakee kaikki kurssit, joiden p1-luokitus on 1 ja p2-luokitus on 3.
../haku.php?hakusanat=testi lause&vertaus=lause&p1=1,2,!3	Hakee kaikki kurssit, joiden tekstikuvauksista löytyy "testi lause", sekä joilla p1-luokitus on 1 ja 2 ja ei 3.
../haku.php?hakusanat=english spring&vertaus=osa&kieli=en&p=(3[2]a2[1])o (3[1]a2[4])&p1=1,2&p2=3,4	Hakee kaikki kurssit, joiden tekstikuvauksista löytyy ainakin toinen sanoista "english" ja "spring". Hakukieli on "en" ja luokitukset ovat (p3=2 and p2=1) or (p3=1 and p2=4). Parametreja p1 ja p2 ei tässä tapauksessa käytetä haussa mihinkään.

4. Hakutulosten esitysmuoto

4.1 XML-muotoiset hakutulokset

Hakupalvelu palauttaa hakutuloksena XML-muotoisen dokumentin, joka sisältää kaikkien hakutuloksiin sopivien kurssien tiedot. Seuraavassa on esimerkki hakupalvelun vastauksesta (1 hakutekijöitä vastaava kurssi):

```
<?xml version="1.0" encoding="iso-8859-1"?>
<kurssit>
  <kurssi>
    <nimi>Digikuvauksen perusteet, 5 op</nimi>
    <nimi_eng></nimi_eng>
    <kurssitunnus>HTTS5060.1K013</kurssitunnus>
    <viikkonumero>1K013</viikkonumero>
    <alkupvm>110214</alkupvm>
    <loppupvm>110422</loppupvm>
    <viim_ilm_pvm>110124</viim_ilm_pvm>
    <lisatietoja>Opettaja Mika Karhulahti
    etunimi.sukunimi@jamk.fi</lisatietoja>
    <muu_info>Opiskelijalla on oltava digitaalinen kamera, jossa on mahdollisuus
    manuaalisäätöihin.</muu_info>
    <tavoite>Opiskelija ymmärtää kuinka digikamera toimii. Opiskelija osaa ottaa teknisesti ja
    esteettisesti laadukkaita digitaalisia valokuvia. Opiskelija ymmärtää digitaalisen
```

kuvanmuodostuksen peruseriaatteen. Opiskelija ymmärtää digitaalisen työnkulun perusteet.

Opiskelija voi parantaa digitaalisia kuvia eri ohjelmistojen avulla. </tavoite>

<sisalto>Digitaaliset kamerat sekä kuvateknologiat

Hyvän valokuvauksen peruseriaatteen

Digitaalisen kuvankäsittelyn ohjelmistot: digitaalisten valokuvien parantaminen kuvankäsittelyn avulla

Digitaalisen valokuvauksen työnkulku

Tulostustoimintojen perusteet </sisalto>

<kuvaus></kuvaus>

<kohderyhma></kohderyhma>

<esitiedot>Tietotekniikan perusteet, WWW-multimedia tai vastaavat taidot </esitiedot>

<minulle_sopii>1 </minulle_sopii>

<haluan></haluan>

<tarvitsen></tarvitsen>

<hakusanat></hakusanat>

<linkki>[https://amp.jamk.fi/asio_v930/kurssiilmo/jamk/index.php?](https://amp.jamk.fi/asio_v930/kurssiilmo/jamk/index.php?asio=Zmt1cnNzaTtjNDY0NmM=)

[asio=Zmt1cnNzaTtjNDY0NmM=</linkki>](https://amp.jamk.fi/asio_v930/kurssiilmo/jamk/index.php?asio=Zmt1cnNzaTtjNDY0NmM=)

</kurssi>

</kurssit>

Hakutuloksia tulee mahdollisesti muokata asiakaskohtaisesti, jotta parametrien tiedot saadaan lisättyä hakutuloksiin tarvittaessa. Tämä voidaan toteuttaa kirjoittamalla oma asiakaskohtainen funktio hakutuloksien muodostukselle tiedostoon haku.php.

5. Muut asiat

5.1 Mahdolliset jatkokehityskohteet

- Hakupalvelun asetukset voitaisiin määrittää käyttöliittymän kautta Lyhytkurssiohjelman ylläpidossa. Tällöin hakupalvelun asetukset voitaisiin tallentaa tietokantaan, sekä myös muodostaa useita hakupalvelukonfiguraatioita asiakaskohtaisesti.
- Ohjelma voisi näyttää listauksen eniten käytetyistä hakusanoista, jolloin hakujen optimointi helpottuisi (hakusanojen lisäys kurssikuvauksiin).
- Hakutulosten esittäminen voitaisiin konfiguroida käyttöliittymän kautta: palautettavan XML-dokumentin kenttien valinta.