
WWW-Anbindung von HANS-2000-Datenbanken

Anleitung

Thomas Berger <ThB@gymel.com>

21.12.2005

Inhaltsverzeichnis

1. Allgemeines und Vorbedingungen	1
2. Vorbereitung der HANS-Datenbank	2
3. Installation und Konfiguration von Avanti	2
4. Installation der CGI-Skripte und Konfiguration des Webservers	3
5. Basiskonfiguration einer eigenen Datenbank	5

1. Allgemeines und Vorbedingungen

Der WWW-OPAC (Public Online Access Catalogue) von HANS ist eine Perl/CGI-basierende Applikation, die das *Templating*-System `populo.pl` [<http://www.gymel.com/populo/>] benutzt. Der Zugriff auf die Daten erfolgt mittels der Avanti™-Datenbankservern von `allegro-C™`, die CGI-Applikation kommuniziert mit diesem über TCP/IP..

Dieser Ansatz hat zur Folge, dass insbesondere die Darstellung der Ergebnisse (Eingabemasken, Registerabschnitte, Vollanzeigen) über *Templates*, d.h. HTML-Seiten mit speziell formatierten Platzhaltern statt der konkreten Inhalte, gestaltet wird, zur Laufzeit werden dann für die Platzhalter die konkreten Werte (Benutzereingaben, Rechercheresultate) eingesetzt. Die Datenbank (bzw. genauer: Der Avanti-Server, der den Zugang zur Datenbank realisiert) kann, muss aber nicht auf dem Webserver installiert sein.

Die Systemvoraussetzungen im einzelnen:

Auf dem Webserver:

- HTTP-Server-Software mit der Möglichkeit, "eigene" CGI-Skripte in der Programmiersprache Perl ablaufen zu lassen:

<http://survey.netcraft.com/servers.html> (Marktübersicht)

- Perl™ (mindestens Version 5.3, empfohlen 5.8 oder neuer) für das Betriebssystem des Servers:

<http://www.cpan.org/ports/> (Übersicht über Distributionen)

- Das Templating-System `populo.pl` (aktuell ist Version 1.16):

<http://www.gymel.com/populo/>

Download u.a. unter:

ftp://ftp.sub.uni-hamburg.de/pub/hans/misc/hans2000/www/pop_1_16.zip
[<ftp://ftp.sub.uni-hamburg.de/pub/hans/misc/hans2000/www/>]

- Die `populo`-Templates und -Anbindung für HANS-2000-Datenbanken:

<ftp://ftp.sub.uni-hamburg.de/pub/hans/misc/hans2000/www/cgidir.zip>
[<ftp://ftp.sub.uni-hamburg.de/pub/hans/misc/hans2000/www/>]

Auf dem Datenbankserver:

- Avanti Version 2.x:

<http://www.allegro-c.de/doku/avanti/>

Clientseitige Voraussetzung ist ein (halbwegs) standardkonformer Webbrowser und eine Verbindung zum Internet. Die Ausführung von JavaScript™ sollte erlaubt werden.

2. Vorbereitung der HANS-Datenbank

Die Anbindung erwartet, dass Avanti alle benötigten Parameterdateien im Datenverzeichnis vorfinden kann. Zusätzlich benötigte Dateien sowie solche, die normalerweise im Programmverzeichnis liegen, sind im Unterverzeichnis `htm` des Programmverzeichnisses vorgehalten und müssen ins Datenverzeichnis kopiert werden.

Für die Zukunft automatisieren lässt sich dies, indem in der Konfigurationsdatei `hansconf.xml` zur entsprechenden Datenbank noch ein Konfigurationselement `WWWAdd` hinzugefügt wird:

```
<section name="databases">
...
  <vardef name="wwwdemo" value="m:\h2kdb\wwwdemo" />
...
</section>
...
<section name="m:\h2kdb\wwwdemo">
  <vardef name="DbDir" value="m:\h2kdb\wwwdemo" />
  <vardef name="DbName" value="demo" />
  <vardef name="DbTitle" value="Demodatenbank HANS mit WWW-Parametern" />
  <vardef name="Site" value="demo" />
  <vardef name="ExtraOpts" value="!Stat !debug" />
  <vardef name="WWWAdd" value="1" />
</section>
...
```

Die Section "databases" beinhaltet die Konkordanz zwischen den vom Benutzer gewählten Namen und eindeutigen, nur intern genutzten Kennungen im "name"-Attribut der späteren Einzel-Sections, die jeweils eine Datenbank beschreiben.

value ist entweder "1" oder aber eine (durch Spatien getrennte) Liste von Namen weiterer Parameterdateien, die nicht im Unterverzeichnis `htm` des Programmverzeichnisses vorliegen, jedoch aus dem Programmverzeichnis kopiert werden sollen.

[Ab Build 24-009:] Für eine vorhandene Datenbank kann dieser Eintrag auch über einen Befehl erzeugt werden:

```
%-H2k%\config\flagdb --setvar wwwdemo WWWAdd "e-oaidc.hpr e-hnsxml.hpr"
```

3. Installation und Konfiguration von Avanti

Die Installation von Avanti erfolgt gemäß der Avanti-Dokumentation <http://www.allegro-c.de/doku/avanti/>. Unter Win32 empfiehlt es sich dringend, den Avanti-Server als Systemdienst zu aktivieren.

In der Konfigurationsdatei `avanti.conf` des Avanti-Servers (im Unterverzeichnis `etc` der Avanti-Installation) ist eine neue Section für die zur Verfügung zu stellende Datenbank einzufügen, z.B.:

```
[hansplay]
directory = d:/hans/h2kdb/wwwdemo
access = 0
konfiguration = hnono
indexparameter = demo
opac=OPAC:0
```

Dies ist der Avanti-Name der Datenbank, im Beispiel bewusst abweichend gewählt vom Namen "wwwdemo", mit dem die Datenbank in der HANS-Installation registriert ist.

Es empfiehlt sich, den Avanti-Namen so zu wählen, wie auch das CGI-Skript auf dem Webserver heißt (jedoch ohne Extension und abschließende Ziffern im Namen).

Verzeichnis, auf dem die Datenbank *auf dem Server* liegt. Dieser weicht typischerweise vom Pfad aus Sicht der Netzwerk-Clients in der Bearbeitungsumgebung ab.

Maximale Zugriffsberechtigung für alle symbolischen Benutzer der Datenbank. 0 = kein Schreibrecht

hnono.cfg wird standardmäßig vorgehalten

"Echter" Name der Datenbank, identisch mit dem Schalter -b aus der inihans.bat bzw. dem Eintrag DbName in der a99-ini-Datei bzw. hansconf.xml.

Wiederholbare Zeile der Form *symbolischer Benutzername* "=" *Avanti-Passwort* ":" *Zugriffsrecht*

Der Avanti-Dienst bzw. -Daemon ist nach dem Ändern der Konfiguration ggfls. neu zu starten.

Weitere Erläuterungen in der Avanti-Dokumentation <http://www.allegro-c.de/doku/avanti/>.

4. Installation der CGI-Skripte und Konfiguration des Webservers

Webserver ordnen der hierarchischen Gliederung von Pfaden in URLs (virtuelle Pfade) reale Verzeichnisse im Dateisystem des Servers (reale Pfade) zu und sind dabei relativ frei konfigurierbar. Allerdings unterscheiden sie normalerweise stark zwischen Verzeichnissen, aus denen Dateien geliefert werden dürfen und jenen, in denen Skripte oder Applikationen ausgeführt, jedoch nicht im Quelltext gezeigt werden dürfen (CGI-Verzeichnisse).

Die CGI-Skripte von HANS gehören in ein CGI-Verzeichnis, die CSS-Stylesheets und eventuelle Graphiken (für die Buttons) gehören in reguläre virtuelle Verzeichnisse. Da alle Dateien für den HANS-OPAC in einem einzigen ZIP-Archiv ausgeliefert werden (mit Unterordnern), muss entweder durch Konfiguration des Webservers oder durch Umkopieren einzelner Verzeichnisse die richtige Zuordnung zu "CGI" bzw. "regulär" vorgenommen werden.

Da die HANS-Anbindung unter populo.pl aus mehreren Dateien und Unterverzeichnissen besteht, empfiehlt es sich aus Gründen der Übersicht, diese in ein separates Verzeichnis zu legen und nicht mit bereits vorhandenen CGI-Applikationen zu mischen. Sofern nicht andere Überlegungen dagegen sprechen, kann aber dazu geraten werden, mehrere Datenbankverbindungen unter populo.pl durchaus in dasselbe Verzeichnis zu installieren.

Der Administrator des Webservers muss unabhängig voneinander folgende Entscheidungen treffen:

1. Soll der virtuelle Pfad der Anwendung einen bestehenden Pfad verlängern ("/cgi-bin/abteilungxy/allerhans/") oder eine neue Hierarchie eröffnen ("/cgi-hans/")?
2. Soll der reale Pfad innerhalb der bestehenden realen Verzeichnishierarchie für den Webserver existieren oder ausserhalb? Beides hat Vor- und Nachteile: Die Chance zur "Vererbung" der Konfigurationseinstellungen für virtuelle Pfade im ersten Fall, ist abzuwägen gegen eventuelle Interferenzen mit Software zur Pflege und Synchronisation der Website und auch gegen den Vorteil, alle Dateien für den OPAC an einem kontrollierten Ort zusammenhalten zu können
3. Sollen die statischen Komponenten (Stylesheets, ggfls. Graphiken und oder Introtexpte) an dem Ort verbleiben, an den sie entpackt werden (=> erfordert zusätzliches Einkonfigurieren von Pfaden) oder manuell umkopiert werden (=> vergrößert die Kohärenz von virtuellen und realen Pfaden)?

Anhand dieser Entscheidungen wird der Administrator ein reales Verzeichnis für die Installation festlegen und dafür sorgen, dass Perl-Skripte in diesem Verzeichnis unter einer bestimmten URL aufrufbar sind.

Achtung

Obacht für Apache-Anwender: Skripte, die `populo.pl` nutzen, sind CGI-Applikationen, keine `mod_perl`TM-Anwendungen!

Anmerkung

Im IIS von Microsoft und anderen Webservern ist darauf zu achten, dass dieses virtuelle Verzeichnis ein "Wurzelverzeichnis für Web-Applikationen" oder "Ausgangspunkt für Anwendungen" ist. Falls sich das nicht einfach einstellen lässt, können Sie alternativ in den CGI-Skripten folgenden kryptischen Code als zweite Zeile der Datei einfügen:

```
BEGIN { ($0 =~ m:^(.*)[/\\][^/\\]+) && chdir $1 };
```

In dieses Verzeichnis muss nun `populo.pl`TM installiert werden sowie die eigentliche HANS-Anbindung:

Achtung

Unter U**X-Betriebssystemen ist mit "**unzip -a**" zu entpacken, damit die Zeilenumbrüche in den Skripten auf die richtige Konvention eingestellt werden.

1. Installieren Sie `populo.pl` entsprechend der Anleitung in der Readme-Datei (<http://www.gymel.com/populo/readme.html> ; Kurzfassung: Das Archiv `pop_1_16.zip` in ein *separates* Verzeichnis ausserhalb der Kontrolle des Webserverns entpacken, den Inhalt des Unterverzeichnisses `base` (zumindest `populo.pl` und `popdebug.pm`) in das gewählte CGI-Verzeichnis kopieren.

Alternativ: auf einer Win32-Maschine entpacken und `populo.pl` und `popdebug.pm` im Textmodus per FTP auf die Zielmaschine transportieren.

2. Extrahieren Sie das Archiv `cgidir.zip` mit den Dateien der HANS-2000-Anbindung in das gewählte CGI-Verzeichnis und passen die ersten Zeile von `hansdemo.pl` und `hansdemo2.pl` ggfs. auf den Pfad des Perl-Interpreters auf Ihrer Maschine an (Readme-Datei beachten!). Unter U**X-Betriebssystemen bitte die Execute-Berechtigungen von `hansdemo.pl` und `hansdemo2.pl` überprüfen und ggfls. setzen.

Alternativ: auf einer Win32-Maschine entpacken und inklusive aller Unterverzeichnisse per FTP auf die Zielmaschine übertragen, dabei muss dann "text" als Übertragungsmodus eingestellt sein, nur für das Unterverzeichnis `hansdemo_cf/images` muss es "binary" sein (Viele ftp-Client-Programme unterstützen eine Einstellung "auto", die die unterschiedlichen Übertragungsmodi automatisch berücksichtigt).

3. Testen Sie den Aufruf von der Kommandozeile:

Eingabe von

```
perl -c popolo.pl <Enter>
perl -c hansdemo.pl <Enter>
```

im CGI-Verzeichnis sollte jeweils "Syntax OK" liefern. Dann Eingabe von:

Unter U**X:

```
./hansdemo.pl <Enter>
Strg+D
```

Unter Win32:

```
perl hansdemo.pl <Enter>
Strg+Z
```

Es sollte dann vielzeiliger HTML-Code über den Schirm rauschen.

4. Der Datei hansdemo.README entnehmen Sie bitte die Erläuterung zur realen und virtuellen Verzeichnisstruktur für die Datenbank "hansdemo". Unbedingt vorzunehmen ist folgendes:
 - Für die CSS-Stylesheets im realen Unterverzeichnis hansdemo_cf/styles : Eingestellt ist, dass auf diese über den virtuellen Pfad /hansstyle zugegriffen wird. D.h. Sie müssen den Webserver so konfigurieren, dass diese Zuordnung erfüllt ist, oder den Inhalt von hansdemo_cf/styles umkopieren und in der externen Konfigurationsdatei hansdemo.conf den von Ihnen gewünschten virtuellen Pfad einstellen (nach "CSSURLS" suchen).
 - Für die Imagedateien der graphischen Buttons im realen Unterverzeichnis hansdemo_cf/images : Eingestellt ist, dass auf diese über den virtuellen Pfad /hansimag zugegriffen wird. D.h. Sie müssen den Webserver so konfigurieren, dass diese Zuordnung erfüllt ist, oder den Inhalt von hansdemo_cf/images umkopieren und in der externen Konfigurationsdatei hansdemo.conf den von Ihnen gewünschten virtuellen Pfad einstellen (nach "ImagePath" suchen) oder die Nutzung der Graphischen Buttons unterbinden (nach "FancyButtons" und "GraphicButtons" in hansdemo.conf suchen)
5. Unter der URL `http://<IhrServer>/<Ihr/Virtueller/Pfad>/hansdemo.pl` sollte die Anbindung der Demodatenbank nun aufstartbar sein. Die Recherche funktioniert jedoch nur, wenn Sie Avanti für eine Datenbank "hansdemo" aufgesetzt haben. Um mit einer bereits existierenden Demo-Datenbank zu testen: Ergänzen Sie testweise am Ende von hansdemo.conf die folgende Zeile:

```
$AvantiHost = "avanti.hans-support.de";
```

5. Basiskonfiguration einer eigenen Datenbank

Es ist gängige Praxis, parallel zu einem Populo-Interface `xy.pl` noch ein "diagnostisches" Interface `xy2.pl` aufzusetzen. Die beiden Skripte sollten identisch sein, bis auf die Setzung von `$Debug` und das Aktivieren der Belegung von `@WantDebug` (vgl. die mit "\$\$" gekennzeichneten Stellen in `hansdemo2.pl` . Gerade beim Aufsetzen einer neuen Datenbank erleichtern die diagnostischen Funktionen die Fehlersuche ungemein. Die Empfehlung ist nun, ausgehend vom diagnostischen Interface `xy2.pl` alle Einstellungen und Anpassungen zu testen und erst abschliessend `xy2.pl` auf `xy.pl` zu kopieren und dort dann erst **`$Debug=0`** (und **`$Border=0`**) einzustellen.

Wenn Sie eine Kopie des Skripts `hansdemo.pl` unter dem Namen `xy.pl` (oder – je nach Betriebssystem und Fähigkeiten des Webservers – `xy.cgi` oder `xy` ohne Extension) erstellen, so wird es bei Ausführung versuchen, die Avanti-Datenbank `xy` unter dem Avanti-Server auf seinem Standardport 4949 auf dem lokalen Rechner anzusprechen, dabei wird der Standard-Avanti-Benutzer `opac` mit Password `opac` benutzt. Unter günstigen Bedingungen ist Ihre Datenbank dann also sofort ansprechbar.

Falls Konfigurationseinstellungen notwendig sind und (dies ist die Regel) weitere Anpassungen anfallen, gehen Sie bitte wie folgt vor: Kopieren Sie die "externe" Konfigurationsdatei `hansgenerisch.conf` nach `xy.conf` und nehmen dort Ihre Anpassungen vor. Es wird empfohlen, nach Abschluss der Anpassungsarbeiten aus der Datei `xy.conf` alle nicht genutzten (also Kommentar gebliebenen) Einstellungen zu löschen, damit die tatsächlich vorgenommenen Abänderungen zum Standard besser sichtbar sind.

Nehmen Sie möglichst keine Änderungen am Stylesheet `hans-opac.css` vor, sondern kreieren Sie eine eigene `css`-Datei, die als erste Anweisung(en) ggfls. Ihr globales Stylesheet und anschliessend `hans-opac.css` importiert (vgl. `hansdemo.css` für die Demo-Datenbank), Die URL dieses Stylesheets muss dann in `xy.conf` angegeben werden (vgl. die Anbindung von `hansdemo.css` in `hansdemo.conf`). Alternativ können Sie in `xy.conf` auch das Einbinden mehrerer Stylesheets konfigurieren.

Der Einführungstext in `hansdemo_cf/text/intro.htm` muß gewiss angepasst werden. Alternativ können Sie einen völlig freien, eigenen Text entwerfen, dessen URL Sie dann in `xy.conf` in den Variablen `$ForceStaticHSIntro` und `$ForceStaticHSUsage` vermerken. Da dieser Text dann nicht mehr unter Kontrolle von `populo` präsentiert wird, müssen Sie in eventuellen Verlinkungen auf den Katalog dessen URL (im Beispiel `<Ihr/Virtueller/Pfad>/xy.pl`) explizit eingetragen werden.