

20 Filtermodule

Nachfolgend werden die insgesamt neun unterschiedlichen Filtermodule vorgestellt, die im Rahmen dieser Arbeit entwickelt wurden. Zunächst werden kurz die übergreifenden Gestaltungskriterien dargestellt (Kap. 20.1). In den folgenden Kapiteln (Kap. 20.2 bis 20.10) werden die einzelnen Filtermodule beschrieben (vgl. die Übersicht in Tab. 20.1).

Filtermodule		Anforderungen und Umsetzung siehe ...
Name	Aufgabe	
SingleAttributeFilter	Definition von Selektionskriterien auf Einzelattributen	Kap. 20.2
SpatialFilter	Intuitive Raumauswahl über den interaktiven digitalen Atlas IDA	Kap. 20.3
TimeFrameSelector	Auswahl von Zeitfenstern auf Datums- sowie numerischen Datentypen	Kap. 20.4
ValueCombinator	Definition von Kombinationen aus vordefinierten Werteausprägungen	Kap. 20.5
GlobalSearchFilter	Eingabe von Selektionskriterien über einer Anzahl zeichenkettenbasierter Attribute	Kap. 20.6
HierarchyBrowser	Navigation über und Werteselektion aus hierarchischen Thesauri	Kap. 20.7
StationClassifier	Selektion von Stationen anhand vielfältiger Klassifizierungskriterien	Kap. 20.8
StatisticFilter	Selektion von Stationen anhand statistischer Kriterien der von diesen erhobenen Zeitreihen	Kap. 20.9
AttributeSelector	Attributauswahl für die Ergebnispräsentation	Kap. 20.10

Tab. 20.1 - Übersicht über die entstandenen Filtermodule.

Abschließend (Kap. 20.11) wird kurz auf die Unterstützung des Anwenders durch Überprüfung und gegebenenfalls automatische Erweiterung der aktuellen Attributauswahl vor Beginn einer Anfrage eingegangen.

20.1 Übergreifende Gestaltungskriterien

Für jedes Filtermodul wurde eine graphische Nutzerschnittstelle entsprechend seines jeweiligen Einsatzgebietes entworfen. Im Interesse einer intuitiven Bedienbarkeit wurden allerdings übergreifende Gestaltungskriterien verwendet, die auf den im iterativen Entwicklungsprozess gewonnenen Erfahrungen basieren. Bevor auf Aufgabe und Umsetzung der einzelnen Filtermodule eingegangen wird, werden daher zunächst die gewählten Regeln vorgestellt, die für den Entwurf der Oberflächen in ihrer hier dokumentierten Ausprägung eingesetzt wurden.

20.1.1 Eigene Dialogfenster

Die graphische Oberfläche jedes Filtermoduls wird jeweils in einem eigenen Dialogfenster geöffnet. Diese Designentscheidung ergab sich aus den diversen Einschränkungen, die anhand von Prototypen sichtbar wurden, bei denen die Oberflächen der Filtermodule direkt in das Hauptfenster integriert wurden. Die Präsentation der Filtermodule in eigenen Dialogfenstern besitzt demgegenüber folgende Vorzüge:

- Verfügbarer Platz
 - ▶ Bei der Integration der graphischen Oberflächen mehrerer Filtermodule in einem Fenster ist der verfügbare Platz limitiert. Durch das Öffnen in eigenen Dialogfenstern kann jedem Filtermodul der erforderliche Platz zur Verfügung gestellt werden. Für die graphi-

schen Oberflächen einzelner Filtermodule liegt zudem jeweils unterschiedlicher Platzbedarf vor, der durch Öffnen in eigenen Fenstern individuell bereitgestellt werden kann.

- Nutzerdefinierbare Größe
 - ▶ Der Anwender kann die Größe der graphischen Oberflächen einzelner Filtermodule durch Veränderung der Fenstergröße individuell an seine Bedürfnisse anpassen.

- Fokussierung des Anwenders
 - ▶ Da der Anwender zu jedem Zeitpunkt jeweils nur mit einem Filtermodul interagiert, kann er sich auf dieses konzentrieren, ohne durch die gleichzeitige Darstellung weiterer Filtermodule abgelenkt zu werden.

Abb. 20.1 gibt einen Eindruck des bei aller individuellen Ausgestaltung einzelner Filtermodule einheitlichen Design ihrer graphischen Oberflächen. Entsprechend der oben (vgl. Kap. 19.1) ausgeführten übergreifenden Gestaltungskriterien für die Fenster der Schnittstelle verfügt jedes Filtermodul über eine Informationsleiste am unteren Fensterrand, die Rückschlüsse über den ausgewählten Datenraum gibt und zugleich zur interaktiven Ausgabe kontextabhängiger Informationen verwendet werden kann (vgl. Abb. 20.1 Mitte).

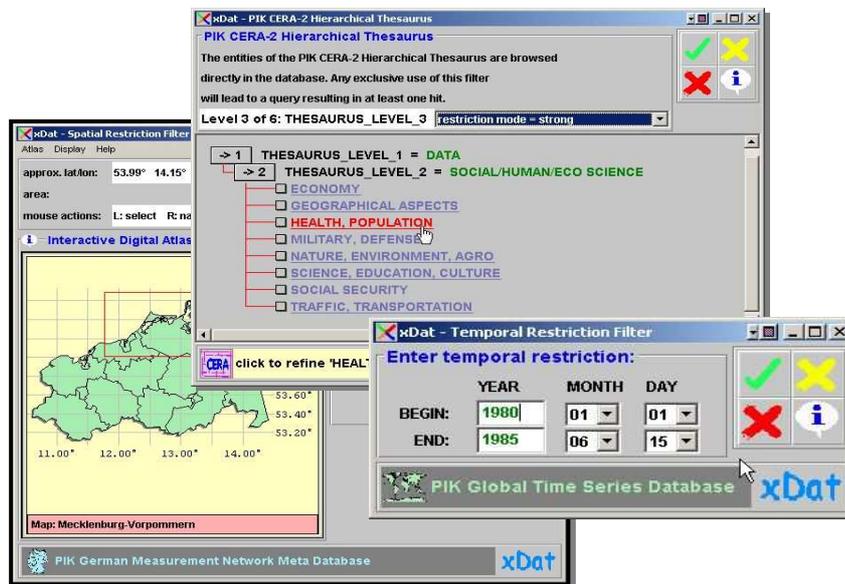


Abb. 20.1 - Die graphischen Nutzerschnittstellen einzelner Filtermodule im Einsatz: Raumselektion (hinten), Selektion von Thesaurus-Einträgen (Mitte) und Eingabe eines Zeitfensters (vorne).

Icon	Bedeutung	Aktion
	Aktivieren	Schließt das Fenster des Filtermoduls und aktiviert die entsprechende Teilbedingung für die aktuelle Anfrage
	Reset	Setzt vorliegenden Nutzereingaben zurück; das Fenster des Filtermoduls bleibt für neue Eingaben geöffnet
	Deaktivieren	Setzt vorliegende Nutzereingaben zurück und schließt das Fenster des Filtermoduls
	Kontextsensitive Hilfe	Stellt dem Anwender Informationen über die Benutzung des jeweiligen Filtermoduls zur Verfügung

Tab. 20.2 - Mögliche Anwenderinteraktionen in jedem Filtermodul.

20.1.2 Übergreifende Funktionalität

Um eine intuitive Bedienung zu unterstützen, werden dem Anwender von jedem Filtermodul unabhängig von dessen konkreter Ausgestaltung mehrere Funktionen in einheitlicher

Weise zur Verfügung gestellt (vgl. Tab. 20.2). Auf diese Weise wird sichergestellt, dass die zentralen Interaktionen - Aktivieren, Rücksetzen, Deaktivieren und Abruf der Hilfe - mit jedem Filtermodul vom Anwender jeweils in identischer Weise ausgeführt werden können.

20.1.3 Automatische Validierung beim Aktivieren

Wenn der Anwender eine Teilbedingung aktivieren möchte, führt das Filtermodul zunächst die jeweils erforderlichen syntaktischen und semantischen Überprüfungen der Nutzereingaben durch³¹³. Werden diese als fehlerhaft erkannt, verweigert das Filtermodul das Schließen seines Fensters und gibt über seine Informationsleiste entsprechende Hinweise zur Behebung des jeweiligen Fehlers. Liegen hingegen korrekte Eingaben vor, schließt sich das Fenster des Filtermoduls; die eingegebene Teilbedingung wird gemeinsam mit den bereits ausgewählten Teilbedingungen im SelectionDisplay des Hauptfensters dargestellt (vgl. Abb. 19.4b) und steht nun zur Verwendung für die aktuelle Anfrage bereit. Da die Filtermodule jeweils über Dialogfenster geöffnet werden, die das Hauptfenster solange deaktivieren, bis sie wieder geschlossen werden, kann eine Anfrage nur ausgelöst werden, wenn *alle* Filtermodule geschlossen sind. Auf diese Weise wird sichergestellt, dass sämtliche aktivierten Teilbedingungen zu jeder Zeit die Generierung einer korrekten Datenbank-anfrage erlauben.

20.2 Beliebige Einzelattribute – SingleAttributeFilter

20.2.1 Motivation und Anforderungen

Einen wichtigen und in vielen Ausprägungen auftretenden Bestandteil für die Abfrage der betrachteten Datenräume bildet die Formulierung von Teilbedingungen auf *jeweils einem* Datenbankattribut. Hierunter fallen im Kontext von Zeitreihenmetadatenbanken und Zeitreihendatenbanken bspw. Selektionen anhand von Stationsnamen oder der Höhe über dem Meeresspiegel; im Kontext von PIK CERA-2 Selektionen anhand von Titeln, Zugriffsbeschränkungen, Projekten, Datengebern u.v.m. Um ein flexibles Werkzeug für eine komfortable Anfrageformulierung basierend auf Selektionen von Wertausprägungen beliebiger Einzelattribute bereitzustellen, wurde mit dem SingleAttributeFilter ein entsprechendes Filtermodul entwickelt, dem insbesondere folgende Anforderungen zugrundegelegt wurden:

- | | |
|---------------------------------------|---|
| Attribut- und Datentyp-unabhängigkeit | ▶ Das Filtermodul soll für Anfragen gegen ein beliebiges Datenbankattribut konfiguriert werden können. Dabei sind Attribute zu unterstützen, deren Datentyp auf Zeichenketten, numerischen Werten oder Datumswerten basiert. |
| Flexible Bedingungsdefinition | ▶ Das Filtermodul soll die flexible Formulierung von Bedingungen erlauben. Dies umfasst die Unterstützung von Maskierungen (<i>Wildcards</i>) bei Anfragen gegen zeichenkettenbasierte Attribute sowie von Möglichkeiten zur nutzerdefinierten Auswahl von Wertebereichen bei Anfragen gegen Attribute mit numerischen oder Datumswerten. |
| Zugriff auf gültige Wertausprägungen | ▶ Das Filtermodul soll den Anwender durch Möglichkeiten zum interaktiven Abruf möglicher Wertausprägungen des jeweiligen Attributes aus der Datenbank (<i>Browsing</i>) und zur Selektion von Werten aus diesen unterstützen. |

Da die drei genannten Datentypen die Grundlage für die Attributdefinition in den meisten

³¹³ Die erforderliche Validierung differiert je nach Filtermodul (vgl. dazu genauer die Beschreibung der einzelnen Filtermodule in den nachfolgenden Kapiteln).

relationalen Datenbankmanagementsystemen bilden, können bereits aufbauend auf diesem Filtermodul durch entsprechende Konfiguration und Kombination graphische Abfrageoberflächen für nahezu beliebige Datenbanksichten (Views) generiert werden.

20.2.2 Umsetzung

Zur Bereitstellung einer jeweils geeigneten graphischen Oberfläche in Abhängigkeit von der Konfiguration des Filtermoduls auf unterschiedliche Datentypen wurde diese adaptiv gestaltet. Das Filtermodul passt seine graphische Bedienoberfläche entsprechend des Typs des zugeordneten Datenbankattributes automatisch an, so dass durch Konfiguration unterschiedliche Filterinstanzen generiert werden können (vgl. Abb. 20.2a und 20.2b). Unabhängig vom Datentyp des zugeordneten Attributes werden jeweils Möglichkeiten zur freien Eingabe von Bedingungen, zu ihrer nutzerdefinierten Verknüpfung anhand jeweils geeigneter logischer Operatoren (vgl. Abb. 20.2c) sowie zum Browsen gültiger Attributwerte in der Datenbank und zur Übernahme in die Anfragebedingung bereitgestellt.

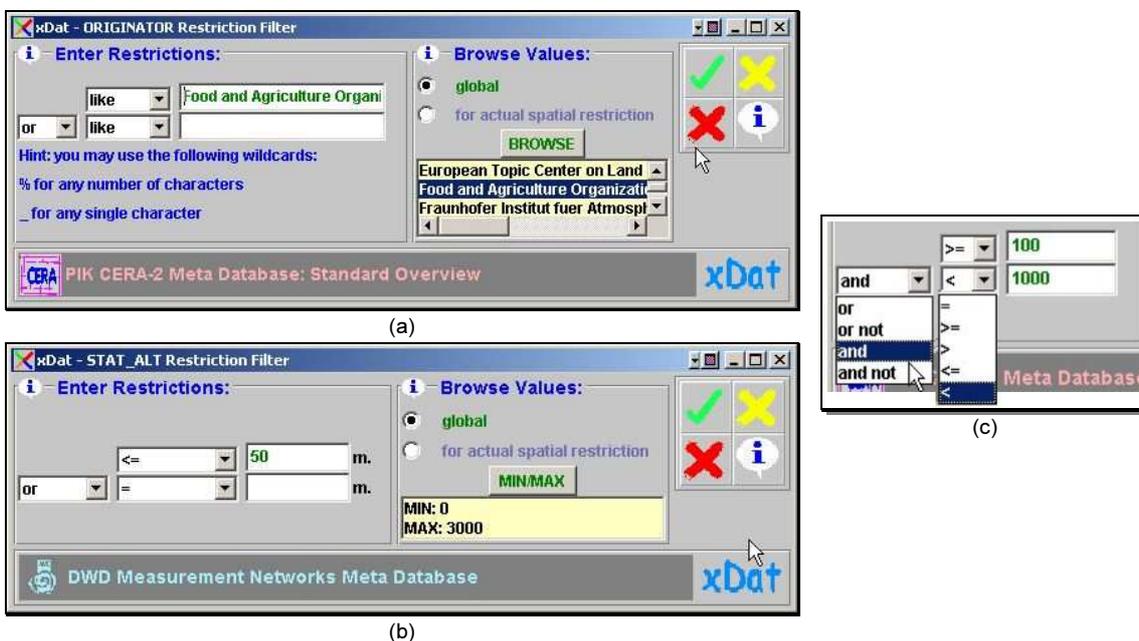


Abb. 20.2 - Der SingleAttributeFilter: (a) Konfiguration für ein zeichenkettenbasiertes Attribut: Selektion von Metadaten aus PIK CERA-2 anhand von Datengebern; (b) Konfiguration für ein numerisches Attribut: Selektion von Stationen aus einer Zeitreihenmetadatenbank anhand der Höhe über dem Meeresspiegel; (c) Detail: logische Operatoren für ein numerisches Attribut aus PIK CERA-2.

▪ Bedingungeingabe und -verknüpfung

Zur freien und flexiblen Eingabe von Bedingungen für das jeweilige Attribut stehen zwei Eingabefelder zur Verfügung, die in beliebiger Kombination verwendet werden können. Nutzerdefinierte numerische und Datumswerte werden dabei vor der Generierung einer Teilbedingung einer syntaktischen Validierung unterzogen. Zur flexiblen Verknüpfung der eingegebenen Bedingungen stellt der SingleAttributeFilter entsprechende Operatoren zur Verfügung. Die jeweils generierten Operatoren richten sich dabei nach dem Datentyp des zugeordneten Attributes und erlauben Konjunktion, Disjunktion und Negation der Einzelbedingungen sowie die Verwendung von datentypspezifischen Vergleichsoperatoren. Abb. 20.2c gibt einen Eindruck von der Bandbreite der so formulierbaren Teilbedingungen: Dargestellt sind die für ein numerisches Attribut generierten Operatoren; die aktuellen Eingaben entsprechen der Teilbedingung *„selektiere diejenigen Datensätze, deren Werteausprägung für dieses Attributes größer als oder gleich 100 und zugleich kleiner als 1000 ist“*.

▪ Mengen- und raumsensitives Browsen

Sowohl bei sporadischer Nutzung wie bei der Arbeit mit einer Vielzahl unterschiedlicher

Datenräume kann nicht davon ausgegangen werden, dass Anwender über Details der jeweiligen Inhalte - verwendete Bezeichner, Schreibweisen, Minimal- und Maximalwerte numerischer Attribute etc. - informiert sind. Um eine einfache und zielgenaue Anfrageerstellung zu erleichtern und zugleich Möglichkeiten zur schnellen Vorabinformation bereitzustellen, wurde in den SingleAttributeFilter ein interaktives Browsing integriert, das es erlaubt, auf einfache Weise jeweils gültige Werteausprägungen in der Datenbank einzusehen. Das Filtermodul löst auf Wunsch eine entsprechende Datenbankanfrage aus; die Werte werden in einer Liste zur Verfügung gestellt und können nachfolgend durch Anwahl in die Eingabefelder übertragen werden. Für Attribute, die numerische oder Datumswerte enthalten, können so jeweils die minimale und die maximale Werteausprägung eingesehen werden; für zeichenkettenbasierte Attribute ist ein Zugriff auf individuelle Werteausprägungen möglich. Da bei zeichenkettenbasierten Attributen jeweils sehr unterschiedliche Ergebnismengen auftreten können - so ist bspw. die Anzahl der in den eingebundenen Zeitreihenmetadatenbanken dokumentierten Stationsnamen jeweils um ein Vielfaches größer als etwa die Anzahl der in PIK CERA-2 dokumentierten Kontaktpersonen - kann das Browsing so konfiguriert werden, dass entweder

- ▶ sämtliche unterschiedlichen Werteausprägungen, die jeweils mit einem nutzerselektierbaren Buchstaben beginnen, oder
- ▶ sämtliche unterschiedlichen Werteausprägungen, deren Anfangsbuchstabe jeweils zu einer nutzerselektierbaren Buchstabengruppe gehört, oder
- ▶ sämtliche unterschiedlichen Werteausprägungen

aus der Datenbank ausgelesen werden (vgl. Abb. 20.3a bis 20.3c).

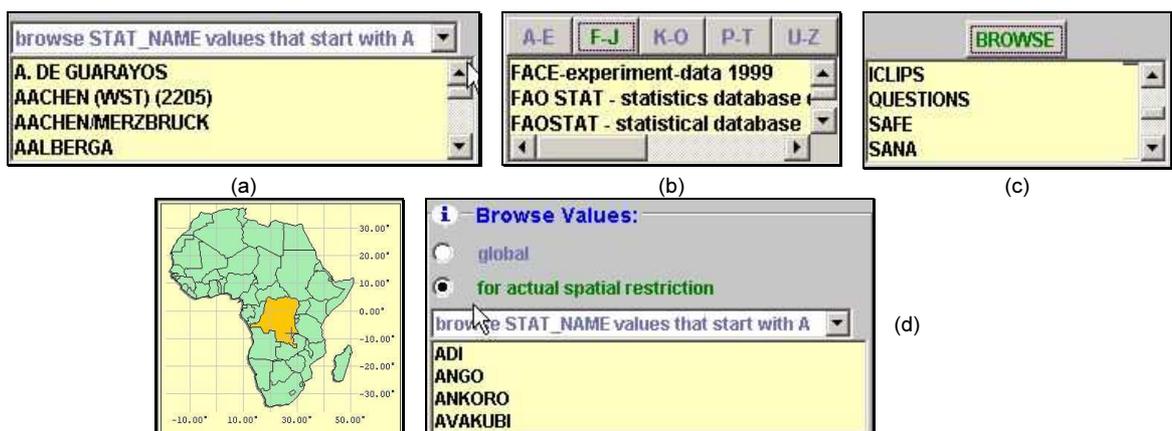


Abb. 20.3 - Beispiele für kontextabhängig konfiguriertes Browsing: (a) Stationsnamen aus der globalen Messnetz-Metadatenbank; (b) und (c) Titel von Datensätzen sowie Projektnamen aus PIK CERA-2; (d) Browsing von Stationsnamen unter Einbeziehung einer Raumauswahl.

Um die Flexibilität des Browsers noch einmal deutlich zu erhöhen, kann bei diesem zudem eine zuvor getroffene Raumauswahl (vgl. Kap. 20.3) berücksichtigt werden. Diese Funktionalität stellt der SingleAttributeFilter unabhängig vom Datentyp des adressierten Attributes zur Verfügung; Abb. 20.3d zeigt das Ergebnis eines Browsers von Stationsnamen in der globalen Messnetz-Metadatenbank des PIK, eingeschränkt auf die Raumauswahl *Zaire*.

20.3 Raumauswahl – SpatialFilter mit IDA

20.3.1 Motivation und Anforderungen

Das Filtermodul SpatialFilter wurde entwickelt, um die Raumauswahl über den eingebundenen Datenräumen in jeweils komfortabler und flexibler und zugleich möglichst konsistenter Form zu unterstützen. Da in den einzelnen adressierten Datenräumen unterschiedliche Varianten der Georeferenzierung verwendet werden, wurden hier insbesondere folgende Anforderungen zugrundegelegt:

- | | |
|---|---|
| Attribut-unabhängigkeit | <ul style="list-style-type: none"> ▶ Das Filtermodul soll für Anfragen gegen diejenigen Datenbankattribute, die im jeweiligen Datenraum zur Georeferenzierung verwendet werden, konfiguriert werden können. Dazu sind sowohl seine Adaption an dort dokumentierte Boundingboxen (vier geographische Koordinaten in vier numerischen Attributen), an Punktverortungen (zwei geographische Koordinaten in zwei numerischen Attributen) sowie an hierarchische räumliche Klassifikatoren (ein oder mehrere zeichenkettenbasierte Attribute) zu ermöglichen. |
| Graphisch-interaktive Selektion mit IDA | <ul style="list-style-type: none"> ▶ Das Filtermodul soll den interaktiven digitalen Atlas IDA (vgl. Kap. 17) bereitstellen, um über diesen die Navigation über Hierarchien graphisch-interaktiver Karten sowie die Selektion von Raumbezügen zu ermöglichen. |
| Boundingbox-referenzierte Daten | <ul style="list-style-type: none"> ▶ Es sind nutzerdefinierte Teilbedingungen zur Selektion von Daten zu unterstützen, deren Georeferenzierung jeweils über eine Boundingbox anhand von vier geographischen Koordinaten beschrieben ist. Eine entsprechende Selektion soll sowohl durch die Auswahl einer beliebigen Boundingbox wie durch die Auswahl der Boundingbox eines vordefinierten Gebietes anhand der über IDA bereitgestellten Kartenhierarchien erfolgen können. |
| Punktdaten | <ul style="list-style-type: none"> ▶ Es sind nutzerdefinierte Teilbedingungen zur Selektion von Punktdaten zu unterstützen, d.h. von Daten, die über ein Paar geographischer Koordinaten georeferenziert sind. Eine entsprechende Selektion soll ebenfalls sowohl durch die Auswahl einer beliebigen Boundingbox wie durch die Auswahl der Boundingbox eines vordefinierten Gebietes anhand der über IDA bereitgestellten Kartenhierarchien erfolgen können. |
| Räumlich klassifizierte Daten | <ul style="list-style-type: none"> ▶ Es sind nutzerdefinierte Teilbedingungen zur Selektion von Daten zu unterstützen, die über hierarchische räumliche Klassifikatoren (vgl. Kap. 18.3.1) georeferenziert sind. Eine entsprechende Selektion soll durch die Auswahl eines oder mehrerer vordefinierter Gebiete anhand der über IDA bereitgestellten Kartenhierarchien erfolgen können. |
| Räumlich klassifizierte Punktdaten | <ul style="list-style-type: none"> ▶ Es sind nutzerdefinierte Teilbedingungen zur Selektion von räumlich klassifizierten Punktdaten zu unterstützen, d.h. von Daten, die sowohl über ein Paar geographischer Koordinaten wie zusätzlich über einen oder mehrere hierarchische räumliche Klassifikatoren georeferenziert sind. Eine entsprechende Selektion soll sowohl durch die Auswahl einer beliebigen Boundingbox wie durch die Auswahl eines oder mehrerer vordefinierter Gebiete anhand der über IDA bereitgestellten Kartenhierarchien erfolgen können. |

20.3.2 Umsetzung

In die Nutzerschnittstelle des SpatialFilter (vgl. Abb. 20.4) wurde der interaktive digitale Atlas IDA integriert, so dass intuitive Navigation und Raumselektion anhand von Hierarchien maussensitiver graphischer Karten unterstützt werden.

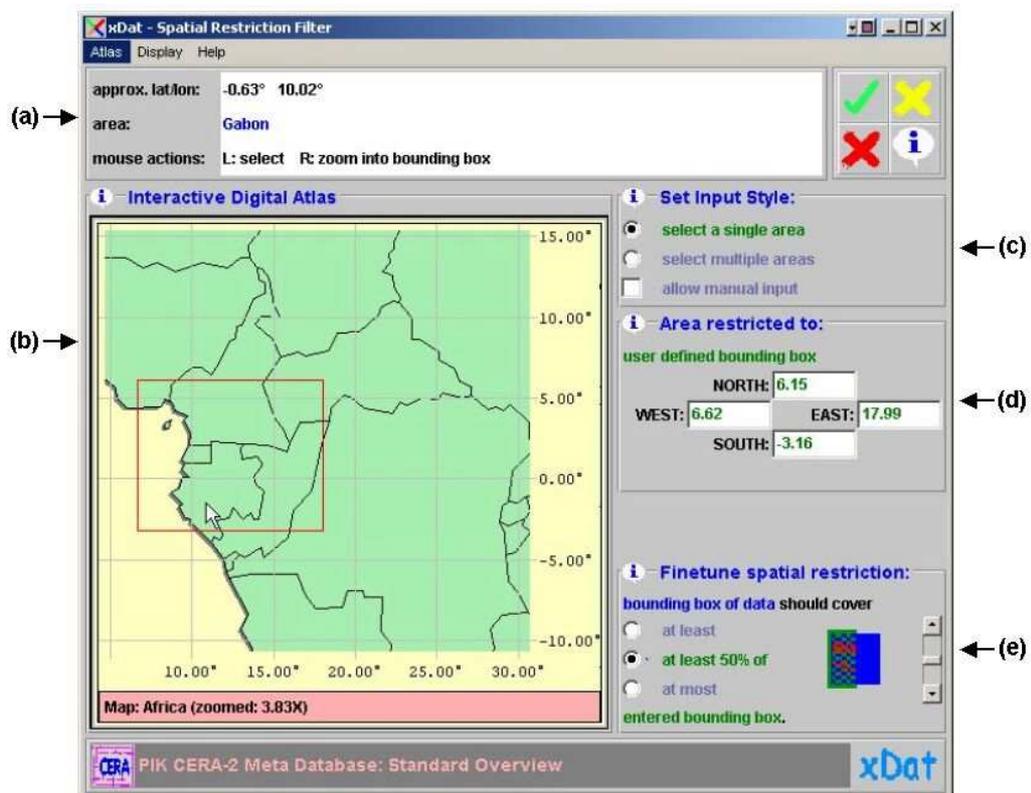


Abb. 20.4 - Graphische Oberfläche des SpatialFilter: (a) interaktives Feedback von IDA; (b) IDA mit nutzerdefiniertem Gummiband; (c) Auswahl des Selektionsmodus; (d) Darstellung der von IDA übertragenen Selektionsparameter; (e) Feinabstimmung der Teilbedingung.

Für eine komfortable Raumauswahl stehen dem Anwender damit die in Kap. 17.5.4 beschriebenen Selektionsformen zur Verfügung, auf die IDA über den SpatialFilter jeweils entsprechend dessen Konfiguration für den zu adressierenden Datenraum adaptiert wird. Eine nutzerdefinierte Auswahl wird anhand von Selektionsparametern von IDA an den SpatialFilter übertragen und von diesem in eine entsprechende Teilbedingung umgesetzt. Nachfolgend werden die über den SpatialFilter möglichen Formen der Raumselektion überblicksartig dargestellt.

▪ Freie Definition einer Boundingbox

Stellt der adressierte Datenraum entweder Daten mit Boundingbox- oder mit Koordinatenpaar-Referenzierung bereit, kann der Anwender einen beliebigen Raumbezug jeweils durch Aufziehen eines Gummibandes in IDA definieren (vgl. Abb. 20.4b). Die vom SpatialFilter generierte Teilbedingung basiert in diesen Fällen auf einem Abgleich der nutzerdefinierten Boundingbox mit den Boundingbox- bzw. Punktkoordinaten des Datenraumes. Bei Punktdaten, die zusätzlich über einen entsprechenden hierarchischen räumlichen Klassifikator georeferenziert sind, wird in die Teilbedingung zusätzlich der Klassifikator der aktuellen Karte einbezogen, um nur solche Daten zu identifizieren, deren Georeferenzierung sie zugleich innerhalb der nutzerdefinierten Boundingbox und innerhalb der auf der Karte dargestellten Gebiete einordnet.

▪ Selektion eines vordefinierten Gebietes

Auch die Selektion eines vordefinierten Gebietes kann für verschiedene Datenräume jeweils einheitlich durch Anwahl seiner entsprechenden graphischen Darstellung in IDA durchgeführt werden. Stellt der adressierte Datenraum Daten mit Boundingbox- oder mit ausschließlicher Koordinatenpaar-Referenzierung bereit, legt der SpatialFilter der Teilbedingung die vier geographischen Koordinaten der das Gebiet umschließenden Boundingbox zugrunde (vgl. Abb. 20.5a); Daten, die über einen entsprechenden hierarchischen

räumlichen Klassifikator referenziert sind, werden hingegen durch Abgleich der Klassifikatoren identifiziert (vgl. Abb. 20.5b).

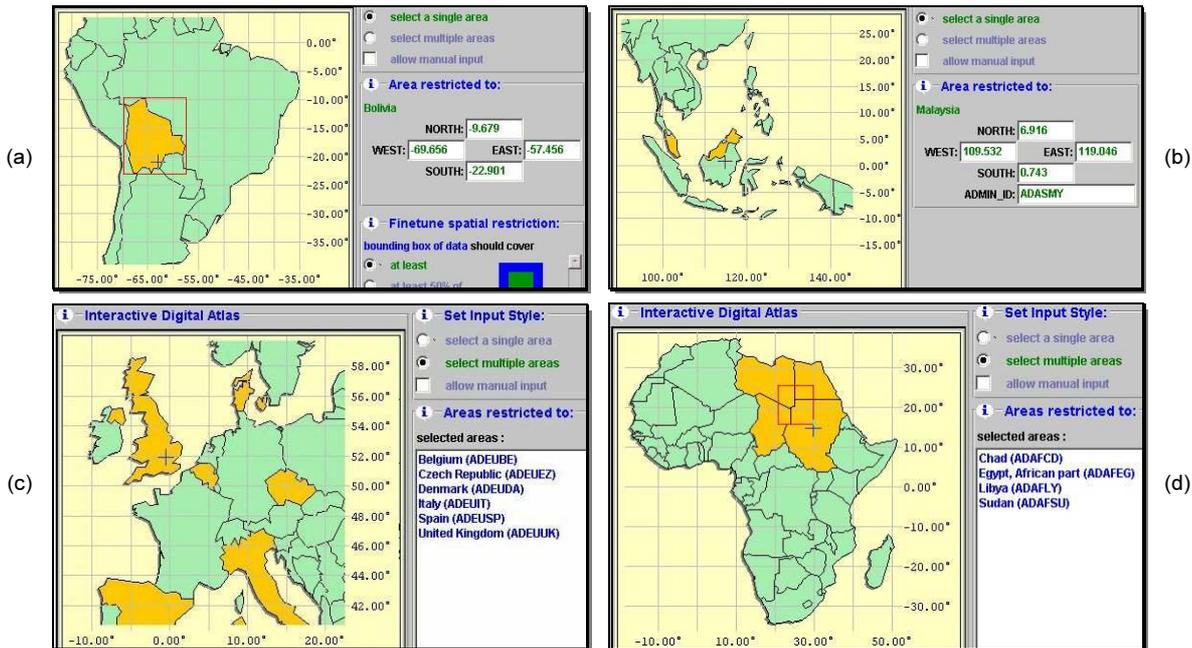


Abb. 20.5 - Formen der Raumselektion: (a) Selektion boundingbox-referenzierter Daten; (b) Selektion von Daten anhand hierarchischer räumlicher Klassifikatoren; Selektion mehrerer vordefinierter Gebiete durch sukzessive Anwahl (c) sowie durch Gummiband (d).

▪ Selektion mehrerer vordefinierter Gebiete

Die Selektion von Daten, die über einen entsprechenden hierarchischen räumlichen Klassifikator georeferenziert sind, kann zusätzlich durch die nutzerdefinierte Auswahl mehrerer vordefinierter Gebiete in IDA durch sukzessive Anwahl (vgl. Abb. 20.5c) oder Aufziehen des Gummibandes (vgl. Abb. 20.5d) erfolgen. In beiden Fällen listet der SpatialFilter zur besseren Orientierung die Namen der jeweils ausgewählten Gebiete in alphabetischer Reihenfolge auf; die Teilbedingung wird jeweils basierend auf den entsprechenden Klassifikatoren generiert.

▪ Manuelle Abänderung des Raumbezuges

Um eine textuelle Eingabe des Raumbezuges oder eine nutzerdefinierte Abänderung der von IDA übertragenen Werte zu ermöglichen, kann der SpatialFilter auf manuelle Eingabe umgeschaltet werden. In diesem Falle können die einzelnen Eingabefelder, in denen die von IDA übermittelten Koordinaten dargestellt werden, vom Anwender editiert werden; damit besteht beispielsweise die Möglichkeit, eine Boundingbox mit offenen Grenzen zu definieren, etwa um solche Stationen zu selektieren, die südlich und westlich nutzerdefinierter Koordinaten liegen. Manuell geänderte Koordinatenwerte werden vor der Generierung einer Teilbedingung automatisch auf syntaktische und semantische Korrektheit überprüft³¹⁴.

▪ Nutzerdefinierter Abgleich von Boundingboxen

Ist der SpatialFilter für Selektionen aus einem Datenraum konfiguriert, dessen Daten anhand von Boundingboxen georeferenziert sind, wird die flexible Integration einer nutzerdefinierter Boundingbox in die zu generierende Teilbedingung unterstützt. Auf diese Weise können zum einen Daten selektiert werden, deren Boundingbox die nutzerdefinierte Boundingbox jeweils vollständig abdeckt; diese Bedingung kann ferner durch Auswahl eines Prozentsatzes, zu dem die Abdeckung erfolgen soll, justiert werden. Alternativ hierzu können

³¹⁴ Die semantische Validierung dient zur Verhinderung „unmöglicher“ Koordinatenwerte und Wertekombinationen.

solche Daten selektiert werden, deren Boundingbox ihrerseits von der nutzerdefinierten Boundingbox vollständig umschlossen wird (vgl. Abb. 20.4e).

20.4 Zeitfenster – TimeFrameSelector

20.4.1 Motivation und Anforderungen

Analog zur Raumauswahl bildet die Definition von Zeitfenstern im gegebenen Kontext ein zentrales Selektionskriterium; so ist sowohl für PIK CERA-2 wie für die einzelnen Zeitreihenmetadatenbanken und Zeitreihendatenbanken jeweils die Auswahl von Metadaten anhand einer nutzerdefinierten zeitlichen Abdeckung zu unterstützen. Um die komfortable Formulierung entsprechender Teilbedingungen zu ermöglichen, wurde das Filtermodul TimeFrameSelector entwickelt, dem insbesondere folgende Anforderungen zugrundegelegt wurden:

- | | |
|---------------------------------------|---|
| Attribut- und Datentyp-unabhängigkeit | ▶ Das Filtermodul soll für Anfragen gegen Datenbankattribute, die im jeweiligen Datenraum zur Dokumentation von zeitlichen Abdeckungen verwendet werden, konfiguriert werden können. Dabei sind Attribute zu unterstützen, deren Datentyp auf numerischen ³¹⁵ oder Datumswerten basiert. |
| Skalierbare Genauigkeit | ▶ Das Filtermodul soll auf die Definition <i>jahresgenauer</i> (Selektion von Jahren), <i>monatsgenauer</i> (Selektion von Jahren und Monaten) sowie <i>tagesgenauer</i> (Selektion von Jahren, Monaten und Tagen) Zeitfenster konfigurierbar sein. |
| Flexible Bedingungen | ▶ Das Filtermodul soll die flexible Formulierung von Bedingungen durch Möglichkeiten zur Definition offener wie geschlossener Zeitfenster unterstützen. |

20.4.2 Umsetzung

Abb. 20.6 zeigt die graphische Nutzerschnittstelle des TimeFrameSelector. Für die Auswahl von Tagen und Monaten stehen vordefinierte Werte zur Verfügung, die eine schnelle Selektion durch Anwahl erlauben; die Definition von Jahreszahlen erfolgt durch freie Eingabe.



Abb. 20.6 - Die graphische Oberfläche des TimeFrameSelector, konfiguriert auf die Definition tagesgenauer Zeitfenster.

Der TimeFrameSelector unterstützt die Auswahl teilloffener oder geschlossener Zeitfenster, so dass die Auswahl von Datensätzen mit Zeitbezug flexibel an die Anforderungen des jeweiligen Anwenders angepasst werden kann. Zu unterscheiden sind:

- **Zeitfenster mit unterer Grenze.**

Zeitfenster mit unterer Grenze werden durch Eingabe eines Anfangsdatums definiert. Auf

³¹⁵ In CERA-2 werden numerische Attribute für die Dokumentation von Zeitbezügen verwendet, um auch paläontologische Zeiträume beschreiben zu können (vgl. Kap. 8.1).

diese Weise werden diejenigen Datensätze ausgewählt, deren zeitliche Abdeckung anhand eines Anfangsdatums dokumentiert ist, das vor dem eingegebenen Datum liegt oder gleich diesem ist.

- **Zeitfenster mit oberer Grenze.**

Zeitfenster mit oberer Grenze werden durch Eingabe eines Enddatums definiert. Auf diese Weise werden diejenigen Datensätze ausgewählt, deren zeitliche Abdeckung anhand eines Enddatum dokumentiert ist, das nach dem eingegebenen Datum liegt oder gleich diesem ist.

- **Zeitfenster mit unterer und oberer Grenze.**

Zeitfenster mit unterer und oberer Grenze werden durch Eingabe eines Anfangs- sowie eines Enddatums definiert. Auf diese Weise werden diejenigen Datensätze ausgewählt, deren zeitliche Abdeckung sowohl anhand eines Anfangsdatums dokumentiert ist, das vor dem eingegebenen Anfangsdatum liegt oder gleich diesem ist, sowie anhand eines Enddatums, das nach dem eingegebenen Enddatum liegt oder gleich diesem ist.

- **Validierung der Nutzereingaben**

Die jeweiligen Nutzereingaben werden vor der Generierung einer Teilbedingung validiert. Eine syntaktische Überprüfung ist dabei nur für die frei editierbaren Jahreszahlen erforderlich. Tages- bzw. monatsgenaue Zeitfenster werden auf Vollständigkeit überprüft; dabei wird bspw. sichergestellt, dass bei Auswahl eines Tages auch Monat und Jahr ausgewählt werden. Wurde hingegen nur ein Jahr eingegeben, werden tages- bzw. monatsgenaue Zeitfenster automatisch ergänzt, wobei bspw. bei tagesgenauen Zeitfenstern das Anfangsdatum auf den 1. Januar und das Enddatum auf den 31. Dezember gesetzt wird. Tag-Monat-Kombinationen werden ferner auf semantische Korrektheit überprüft, um „unmögliche“ Datumsangaben wie einen 30. Februar oder einen 31. November auszuschließen; ebenfalls wird sichergestellt, dass ein 29. Februar nur für ein Schaltjahr ausgewählt werden darf. Werden sowohl Anfangs- und Enddatum eingegeben, wird zudem sichergestellt, dass das Anfangsdatum vor dem Enddatum liegt oder gleich diesem ist, um die unabsichtliche Definition solcher Zeitfenster auszuschließen, die in keinem Fall zu einer Selektion von Datensätzen führen.

20.5 Vordefinierte Werte – ValueCombinator

20.5.1 Motivation und Anforderungen

Für bestimmte Entitäten wie Variablen, zeitliche Auflösungen oder Formen der Datenverfügbarkeit ist eine freie Eingabe und Verknüpfung von Bedingungen, wie sie der SingleAttributeFilter (vgl. Kap. 20.2) erlaubt, wenig attraktiv. Geeigneter für eine komfortable Definition von Teilbedingungen ist hier vielmehr eine direkte Präsentation aller gültigen Wertausprägungen in Verbindung mit der Möglichkeit, aus diesen nutzerdefinierte Kombinationen erstellen zu können. Für solche Aufgaben wurde das Filtermodul ValueCombinator entwickelt, dem folgende Anforderungen zugrundegelegt wurden:

- | | |
|---------------------------------------|---|
| Attribut- und Datentyp-unabhängigkeit | <ul style="list-style-type: none"> ▸ Das Filtermodul soll für Anfragen gegen ein beliebiges Datenbankattribut konfiguriert werden können; dabei sind solche Attribute zu unterstützen, deren Datentyp auf numerischen Werten³¹⁶ oder auf Zeichenketten basiert. |
|---------------------------------------|---|

³¹⁶ Sowohl in den Zeitreihenmetadatenbanken wie den Zeitreihendatenbanken finden numerische Attribute bspw. für Variablen oder zeitliche Auflösungen Verwendung. In diesem Fall finden konfigurierbare Transformationen zwischen diesen Werten und ihrer textuellen Repräsentation für den Anwender durch den DynamicValueMapping (DVM)-Mechanismus der Schnittstelle (vgl. Kap. 16.2.2)

- Übersichtliche und flexible Selektion
- Das Filtermodul soll die jeweils selektierbaren Werteausprägungen darstellen und nutzerdefinierbare Kombinationen aus diesen durch direkte Anwahl erlauben.

Eines der vorgesehenen Einsatzgebiete des Filtermoduls ist die Unterstützung des Anwenders bei der Selektion von Variablen aus Zeitreihendatenbanken. Um eine komfortable Auswahl von Zeitreihenmetadaten durch die Selektion von Variablen zu ermöglichen, soll das Filtermodul zusätzlich auf folgende Eigenschaften konfigurierbar sein:

- Auswahlmodus
- Der Anwender soll sowohl Teilbedingungen formulieren können, die zur Selektion derjenigen Stationen führen, an denen *mindestens eine* der selektierten Variablen erhoben wird, wie Teilbedingungen, die zur Selektion derjenigen Stationen führen, an denen jeweils *sämtliche* selektierten Variablen erhoben werden.
- Anzeige von Maßeinheiten
- Zur Unterstützung des Anwenders soll das Filtermodul auf die zusätzliche Darstellung der zugehörigen - bspw. physikalischen, chemischen oder phänologischen - Maßeinheit jeder selektierbaren Variable konfiguriert werden können.

20.5.2 Umsetzung

Der ValueCombinator wurde so entworfen, dass er zur Laufzeit an die dem Anwender jeweils zu offerierenden Werteausprägungen adaptiert werden kann. Die Werte werden zur Laufzeit eingelesen und entsprechend zur Verfügung gestellt; eine zusätzliche Hilfefunktion eröffnet dabei jeweils Zugriff auf konfigurierbare Informationen über die Semantik der selektierbaren Werteausprägungen.

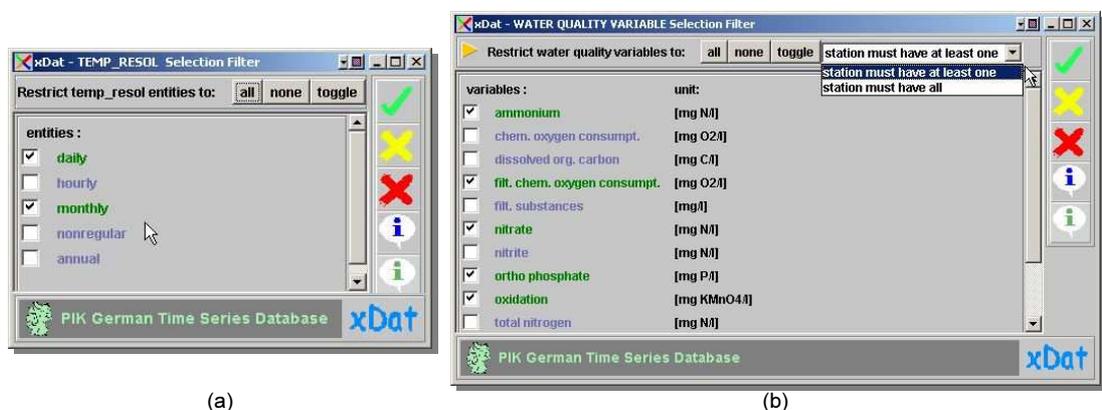


Abb. 20.7 - Beispiele für den Einsatz des ValueCombinator³¹⁷: Präsentation und Auswahl von zeitlichen Auflösungen (a) und Variablen (b) einer Zeitreihendatenbank.

Abb. 20.7 zeigt zwei durch unterschiedliche Konfigurationen generierte Varianten des ValueCombinator. Um eine komfortable Bedienung zu ermöglichen, werden neben einer direkten Anwahl einzelner Werteausprägungen über Schaltflächen zusätzlich Möglichkeiten zur automatischen (De)Selektion aller angebotenen Werteausprägungen sowie zur Invertierung der aktuellen Selektion angeboten.

statt.

³¹⁷ Die in Abb. 20.7 gezeigten Screenshots geben zugleich einen Eindruck von der Arbeitsweise des DVM-Mechanismus (vgl. Kap. 16.2.2): Sowohl die zur Auswahl angebotenen Messfrequenzen (vgl. Abb. 20.7a) wie die Variablen (vgl. Abb. 20.7b) sind in der unterliegenden Datenbank in Form *numerischer* Identifikatoren abgelegt; die Umsetzung in eine menschenlesbare textuelle Darstellung erfolgt transparent für den Nutzer durch den konfigurierbaren Abbildungsmechanismus des Client.

Der ValueCombinator wird in der gegenwärtigen Konfiguration der Schnittstelle nicht als eigenständiges Filtermodul, sondern in diversen Konfigurationen als Bestandteil eines speziellen Filtermoduls zur Stationsklassifikation (StationClassifier, vgl. Kap. 20.8) eingesetzt, das seinerseits für jede der eingebundenen Zeitreihenmetadatenbanken und Zeitreihendatenbanken verwendet wird.

20.6 Transparente Suche – GlobalSearchFilter

20.6.1 Motivation und Anforderungen

Typischerweise werden Teilbedingungen einer Anfrage eindeutig konkreten Attributen zugeordnet. Dieses Konzept ist immer dann nicht optimal, wenn dem Anwender nicht bekannt ist, welchen Attributen eine Bedingung geeignet zuzuordnen ist oder wenn diese Zuordnung für eine Anfrage nur eine untergeordnete Rolle spielt. Bei Metadatenbanken wie PIK CERA-2 mit ihrer Vielzahl von Attributen können solche Situationen leicht auftreten; hier kann es hilfreich sein, zunächst *alle* Datensätze zu selektieren, in denen ein nutzerdefinierbarer Begriff in der Wertausprägung eines beliebigen Attributes enthalten ist, um diese Datensätze nachfolgend weiter auswerten zu können. Das Filtermodul GlobalSearchFilter wurde entwickelt, um in solchen Fällen eine komfortable Bedingungsformulierung zu erlauben. Dazu soll die Eingabe eines Suchbegriffs zu einer Teilbedingung führen, die eine Vielzahl von Attributen einbezieht und diejenigen Datensätze selektiert, für die der Suchbegriff in der Wertausprägung mindestens eines dieser Attribute enthalten ist. Insbesondere soll auf diese Weise auch neuen oder sporadischen Anwendern, die mit den semantischen Konzepten oder Inhalten eines Datenraumes nicht oder nur unzureichend vertraut sind, ein schneller und unkomplizierter Zugriff ermöglicht werden. Dem Filtermodul wurde insbesondere folgende Anforderung zugrundegelegt:

- Attribut-
unabhängigkeit
 - Das Filtermodul soll für Anfragen gegen eine beliebige Menge von Attributen mit jeweils zeichenkettenbasiertem Datentyp konfiguriert werden können.

20.6.2 Umsetzung

Die graphische Oberfläche des GlobalSearchFilter wurde bewusst einfach gehalten (vgl. Abb. 20.8). In der Standardeinstellung werden Groß- und Kleinschreibung des Suchbegriffes ignoriert; möchte der Anwender diese berücksichtigt wissen, kann er dies durch Auswahl des Modus *case sensitive* erreichen.



Abb. 20.8 - GlobalSearchFilter, konfiguriert für die PIK CERA-2-Metadatenbank, mit nutzerdefiniertem Suchbegriff „elbe“.

Der GlobalSearchFilter wird gegenwärtig für die Anfrageerstellung an PIK CERA-2 eingesetzt; dabei erlaubt seine aktuelle Konfiguration auch die Einbeziehung derjenigen Attribute, die dem Anwender nicht für die Eingabe von Teilbedingungen zugänglich sind, wie bspw. der ausführlicheren Beschreibung der dokumentierten Datensätze (vgl. Tab. 18.2).

20.7 Hierarchische Thesauri – HierarchyBrowser

20.7.1 Motivation und Anforderungen

Um eine thematische Zuordnung der zu dokumentierenden heterogenen Datensätze zu erlauben, beinhaltet die allgemeine Institutsmetadatenbank PIK CERA-2 einen hierarchischen Thesaurus, der über sechs einzelne Datenbankattribute beschrieben wird. Damit ist es möglich, individuelle Datensätze durch entsprechende Einträge in diesen Attributen beispielsweise als

- ▶ *Data* ⇨ *Social/Human/Eco Science* ⇨ *Social Security*
- ▶ *Data* ⇨ *Social/Human/Eco Science* ⇨ *Economy* ⇨ *Biological Field (Human Food)*
- ▶ *Data* ⇨ *Earth Science* ⇨ *Biosphere*
- ▶ *Data* ⇨ *Earth Science* ⇨ *Atmosphere* ⇨ *Precipitation* ⇨ *Acid Rain*

zu verschlagworten. Eine intuitive Navigation über und Auswahl von Einträgen aus dieser Begriffsstruktur stellt entsprechend ein wichtiges Element zur Erschließung dieser Metadatenbank dar. Um dies in komfortabler Form zu ermöglichen, wurde das Filtermodul HierarchyBrowser entwickelt, dem insbesondere die folgenden Anforderungen zugrundegelegt wurden:

- | | |
|--------------------------|---|
| Attribut-unabhängigkeit | ▶ Das Filtermodul soll für Anfragen gegen eine beliebige Menge von Attributen mit jeweils zeichenkettenbasiertem Datentyp konfiguriert werden können. Dabei sind die Attribute in eine eindeutige Rangfolge zu bringen, um so eine Hierarchie abzubilden. |
| Navigation und Selektion | ▶ Das Filtermodul soll es dem Anwender erlauben, die offerierten Attributhierarchien graphisch-interaktiv zu durchwandern und Werte aus diesen zu selektieren. |

20.7.2 Umsetzung

Abb. 20.9 und Abb. 20.10 zeigen die graphische Oberfläche des HierarchyBrowser. Auf jeder Hierarchieebene kann der Anwender zwischen Selektion und Navigation wählen. So können auf der jeweils aktuellen Hierarchieebene beliebige Kombinationen der dort dargestellten Werteausprägungen ausgewählt werden, um auf diese Weise eine Teilbedingung für die Gesamtanfrage zu generieren, die entsprechend verschlagwortete Datensätze selektiert. Alternativ kann einer der Einträge der aktuellen Hierarchieebene zur Navigation zur nächsttieferen Hierarchieebene selektiert werden, für die dann die dem bisherigen Navigationspfad des Anwenders entsprechenden Einträge verfügbar gemacht werden.

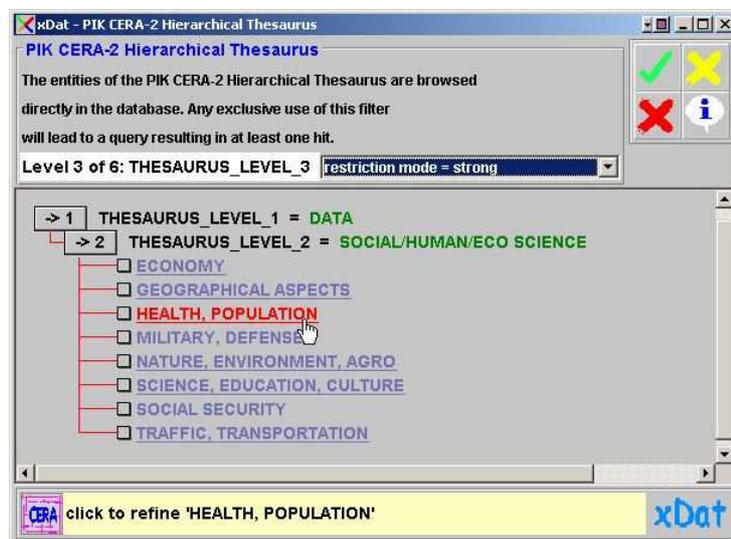


Abb. 20.9 - Die graphische Oberfläche des HierarchyBrowser.

Die Werte für jede Hierarchieebene werden dabei dynamisch (on the fly) aus der jeweiligen Datenbank ausgelesen. Bei jedem Zugriff auf eine tiefere Ebene wird in die Anfrage automatisch der durch die Navigation des Anwenders von der Wurzel bis hin zur aktuellen Ebene bestimmte Pfad integriert. Auf diese Weise wird sichergestellt, dass jeweils nur diejenigen Werte einer Hierarchieebene ausgelesen werden, die der Nutzerauswahl auf den höheren Ebenen entsprechen und dass zugleich jeweils nur solche Werte dargestellt werden, denen mindestens ein Datensatz des adressierten Datenraums zugeordnet ist. Aufgrund dieser Funktionsweise führt eine nachfolgende Selektion aus dem Datenraum, die sich ausschließlich auf eine über den HierarchyBrowser formulierte Teilbedingung stützt, stets zu Anfrageergebnissen, die aus mindestens einem Datensatz bestehen.

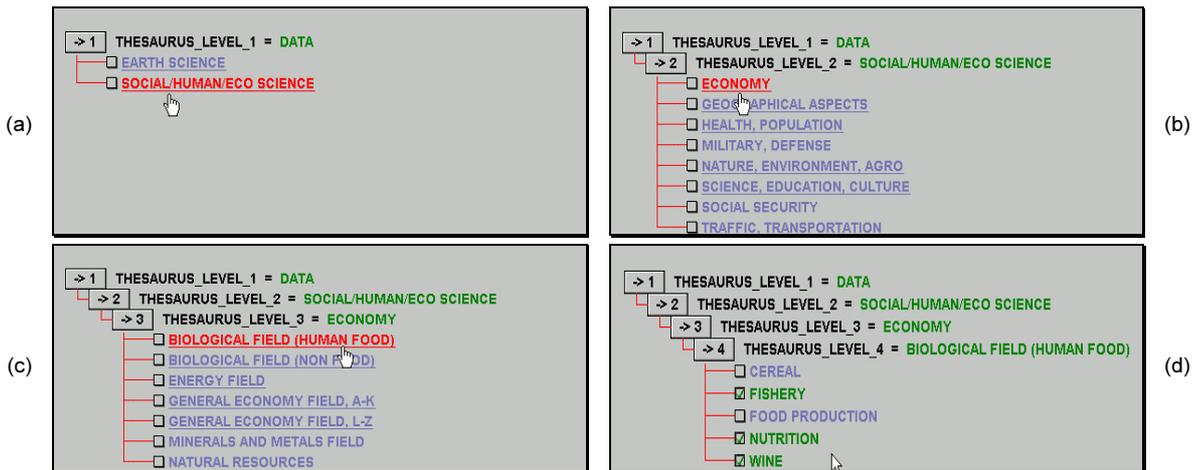


Abb. 20.10 - Interaktive Navigation durch die Hierarchie: (a) bis (c) Sukzessive Verfeinerung von Einträgen; (d) Auswahl von Werten für die Bedingungsformulierung.

Einträge der aktuellen Ebene, für die in der Datenbank Werte auf der nächsttieferen Ebene vorhanden sind, werden in Analogie zu Hyperlinks im World Wide Web unterstrichen dargestellt; die Anwahl eines so gekennzeichneten Eintrags führt zur Navigation zur nächsttieferen Ebene und zur Anzeige der entsprechenden Einträge. Abb. 20.10 gibt einen Eindruck von der Bedienung des Filtermoduls.

Zur Generierung der Teilbedingung werden zwei unterschiedliche Modi zur Verfügung gestellt. Im strengen Modus (restriction mode = *strong*) werden nur diejenigen Datensätze selektiert, deren Wertausprägungen in sämtlichen relevanten Attributen den durch den HierarchyBrowser ausgewählten Werten entsprechen, deren Verschlagwortung also ebenfalls aus dem so definierten Pfad durch die Hierarchie besteht oder mit diesem beginnt. Im schwächeren Modus (restriction mode = *weak*) werden zusätzlich auch jene Datensätze ausgewählt, deren Verschlagwortung aus einem kürzeren Pfad besteht, der einem von der Wurzel beginnenden Teilpfad des vom Anwender selektierten Pfades entspricht.

Pfad:		Data	Earth Science	Atmosphere		
Selektion im strengen (▲) und im schwächeren (▼) Auswahlmodus:						
▲	▼	Data	Earth Science	Atmosphere		
▲	▼	Data	Earth Science	Atmosphere	Air Quality	
▲	▼	Data	Earth Science	Atmosphere	Precipitation	Acid rain
	▼	Data	Earth Science			
	▼	Data				
		Data	Earth Science	Biosphere		
		Data	Social/Human/Eco Science			

Tab. 20.3 - Auswirkung von strengem (▲) und schwächerem (▼) Auswahlmodus auf die Selektion von Datensätzen.

Tab. 20.3 verdeutlicht den Effekt der beiden Auswahlmodi beispielhaft anhand des Pfades *Data* ⇒ *Earth Science* ⇒ *Atmosphere*. Während im strengen Auswahlmodus nur Datensätze ausgewählt werden, deren Verschlagwortungspfad ebenfalls mit dem vollständigen vorgegebenen Pfad beginnt oder gleich diesem ist, kommen im schwächeren Auswahlmodus zusätzlich solche Datensätze hinzu, die entweder nur mit dem Pfad *Data* ⇒ *Earth Science* oder nur mit *Data* verschlagwortet sind.

20.8 Stationsklassifizierung – StationClassifier

20.8.1 Motivation und Anforderungen

Für die Unterstützung des Anwenders bei der Selektion von Stationen aus Zeitreihenmetadatenbanken oder Zeitreihendatenbanken anhand der jeweils verfügbaren Klassifikationskriterien (vgl. Kap. 18.3.2) wurde das Filtermodul StationClassifier entworfen. Da diese Kriterien in einzelnen Metadatenbanken in unterschiedlichen Ausprägungen vorliegen und entsprechend der jeweiligen Retrievalinteressen verwendbar sein sollen, bestehen hier besondere Anforderungen sowohl bezüglich der *Orientierung* des Anwenders über die jeweils verfügbaren Klassifizierungskriterien wie bezüglich der *Auswahl* von Stationen anhand von nutzerdefinierbaren Kombinationen dieser Kriterien. Dem Filtermodul wurden daher die nachfolgend beschriebenen Anforderungen zugrundegelegt.

▪ Orientierung über die Klassifizierungskriterien

Die effiziente Nutzung der jeweils aktuellen Ausprägung der Klassifizierungskriterien des ausgewählten Datenraumes durch den Anwender setzt zunächst voraus, dass eine schnelle Orientierung über diese unterstützt wird. Darunter fallen:

- | | |
|-----------------------|---|
| Stationstypen | ▶ Das Filtermodul soll den Anwender in geeigneter Weise darüber informieren, welche Stationstypen im ausgewählten Datenraum zur Selektion zur Verfügung stehen. |
| Subtypen | ▶ Das Filtermodul soll den Anwender in geeigneter Weise darüber informieren, ob die im adressierten Datenraum dokumentierten Stationen zusätzlich anhand von Subtypen klassifiziert sind, und, falls dies der Fall ist, anhand welcher Subtypen jeder Stationstyp untergliedert wird. |
| Variablen | ▶ Das Filtermodul soll den Anwender in geeigneter Weise darüber informieren, ob im adressierten Datenraum eine variabelgenaue Selektion möglich ist, und, falls dies der Fall ist, welche Variablen jedem Stationstyp zugeordnet sind. |
| Zeitliche Auflösungen | ▶ Das Filtermodul soll den Anwender in geeigneter Weise darüber informieren, ob im ausgewählten Datenraum eine Selektion anhand von zeitlichen Auflösungen möglich ist und, falls dies der Fall ist, welche unterschiedlichen Ausprägungen verfügbar sind. |
| Datenverfügbarkeit | ▶ Das Filtermodul soll den Anwender in geeigneter Weise darüber informieren, anhand welcher Formen der Datenverfügbarkeit die dokumentierten Stationen klassifizierbar sind. |

▪ Selektion von Klassifizierungskriterien

Die intuitive Orientierung über die jeweiligen Klassifizierungskriterien bildet die Grundlage für eine auf diesen basierende nutzerdefinierte Stationsauswahl. Hierfür sind flexible Kombinationsmöglichkeiten auf funktionale Weise bereitzustellen:

- | | |
|---------------|---|
| Stationstypen | ▶ Das Filtermodul soll die Selektion von im ausgewählten Daten- |
|---------------|---|

- raum dokumentierten Stationen anhand nutzerdefinierbarer Kombinationen von Stationstypen unterstützen.
- Subtypen ▶ Liegt im ausgewählten Datenraum eine Klassifizierung von Stationen anhand von Stationstypen sowie von Subtypen vor, so soll das Filtermodul die Selektion von Stationen anhand nutzerdefinierbarer Kombinationen aus Stationstypen und Subtypen unterstützen.
- Variablen ▶ Liegt im ausgewählten Datenraum eine variabelgenaue Beschreibung vor, so soll das Filtermodul die Selektion von Stationen anhand nutzerdefinierbarer Kombinationen aus Stationstypen und Variablen unterstützen.
- Zeitliche Auflösungen ▶ Liegt im ausgewählten Datenraum eine Beschreibung der zeitliche Auflösungen vor, so soll das Filtermodul die Selektion von Stationen anhand nutzerdefinierbarer Kombinationen aus diesen unterstützen.
- Datenverfügbarkeit ▶ Dokumentiert der ausgewählte Datenraum Stationen mit unterschiedlichem Status der Datenverfügbarkeit, so soll das Filtermodul die Selektion von Stationen anhand nutzerdefinierbarer Kombinationen aus diesen unterstützen.

Um jedem Anwender einen möglichst hohen Freiheitsgrad bei der Umsetzung seiner Anforderungen zu erlauben, soll er frei entscheiden können, welche der im ausgewählten Datenraum verfügbaren Klassifizierungskriterien bei einer Selektion jeweils Verwendung finden sollen. Ermöglicht werden sollen daher Teilbedingungen, die sich aus der *beliebigen Kombination* der oben aufgeführten Auswahlkriterien ergeben. Um dies zu illustrieren, werden an dieser Stelle stellvertretend zwei Beispiele anhand konkreter Anforderungen aufgeführt. Im ersten Fall werden die Klassifizierungskriterien Stationstyp, Subtyp und Datenverfügbarkeit verwendet; die zweite Teilbedingung basiert auf einer Kombination der Klassifizierungskriterien Stationstyp, Variable und zeitliche Auflösung:

- Beispiel 1 ▶ Selektiere aus dem ausgewählten Datenraum Meteorologie-Stationen sowie Phänologie-Stationen aus den Bereichen Agrarwirtschaft und Forstwirtschaft; beschränke die Auswahl dabei jeweils auf solche Stationen, deren Daten im Institut entweder über eine Datenbank oder in einer anderen digitalen Form verfügbar sind.
- Beispiel 2 ▶ Selektiere aus dem ausgewählten Datenraum Meteorologie-Stationen, die mindestens eine der drei Variablen Niederschlag, Luftdruck oder Bewölkung messen, sowie beliebige Wasserqualitäts-Stationen sowie solche Hydrologie-Stationen, an denen sowohl Wasserstände wie Durchflussmengen erhoben werden; beschränke die Auswahl dabei jeweils auf solche Stationen, die ihre Messungen täglich durchführen.

20.8.2 Umsetzung

Um einen möglichst anschaulichen und einheitlichen Zugang zu den unterschiedlichen Zusammensetzungen und Ausprägungen der Klassifizierungskriterien in den einzelnen Zeitreihenmetadatenbanken und Zeitreihendatenbanken zu unterstützen, wurde die graphische Oberfläche des StationClassifier in drei logische Ebenen aufgeteilt, die jeweils

getrennt konfigurierbar sind. Jede Ebene ist spezifischen Klassifikationskriterien vorbehalten:

- Ebene 1: Stationstyp
 - ▶ Entsprechend seiner Bedeutung als übergreifend verwendetes Klassifikationskriterium bilden Möglichkeiten zur Auswahl aus den jeweils vorliegenden Werteausprägungen des Stationstyps die zentrale Ebene dieses Filtermoduls.

- Ebene 2: Typabhängige Kriterien
 - ▶ Die zweite Ebene bilden Klassifikationskriterien, deren Werteausprägungen sich in Abhängigkeit von jeweils einem konkreten Stationstyp ergeben. Hierzu zählen Variablen oder Subtypen; der Zugriff auf solche typabhängige Kriterien erfolgt jeweils getrennt für jeden angebotenen Stationstyp.

- Ebene 3: Typunabhängige Kriterien
 - ▶ Die dritte Ebene bilden schließlich allgemeine Klassifikationskriterien, deren Einschränkung für sämtliche Stationen unabhängig von deren Typ gelten sollen; hierunter fallen bspw. zeitliche Auflösungen oder der Status der Datenverfügbarkeit.

Die Konfigurierbarkeit des StationClassifier erlaubt ohne Reprogrammierung die Erstellung unterschiedlicher Klassifikatoren, so dass in Zusammenarbeit mit Datenadministratoren und Anwendern schnell verschiedene Kombinationen bereitgestellt und auf ihre Eignung hin ausgewertet werden können. Abb. 20.11 zeigt unterschiedliche Stationsklassifikatoren, die auf diese Weise erzeugt wurden. Der StationClassifier stellt die jeweils auswählbaren Stationstypen zur Verfügung und grenzt ihre Darstellung zusätzlich durch die hierfür vorgesehenen farbigen Symbole voneinander ab (vgl. Kap. 19.5).

(a)

- ▶ Auswahl von Hydrologie-, Meteorologie-, Phänologie- und Wasserqualitätsstationen, zugehörigen Variablen sowie zeitlichen Auflösungen.

(b)

- ▶ Auswahl von Meteorologie- und Phänologiestationen, zugehörigen Subtypen sowie Verfügbarkeitskriterien.

(c)

- ▶ Auswahl von Meteorologie- und Hydrologiestationen sowie Verfügbarkeitskriterien.

Abb. 20.11 - Verschiedene Konfigurationen des StationClassifier: Klassifikationskriterien für die deutsche Zeitreihendatenbank (a), die DWD-Messnetz-Metadatenbank (b) sowie die globale Messnetz-Metadatenbank (c).

Sind zusätzlich typabhängige und / oder allgemeine Klassifikationskriterien verfügbar, werden entsprechende Schaltflächen generiert, die den Zugriff auf diese ermöglichen. Hierfür erzeugt der StationClassifier entsprechend konfigurierte Instanzen des ValueCombinator (vgl. Kap. 20.5), die interaktiv bereitgestellt werden. Abb. 20.12 zeigt die Verwendung des StationClassifier zur Einschränkung auf Wasserqualitäts-Stationen (vgl. Abb. 20.12a), tägliche zeitliche Auflösung (vgl. Abb. 20.12b) sowie die drei Variablen *nitrate*, *nitrite* und *total nitrogen* (vgl. Abb. 20.12c).

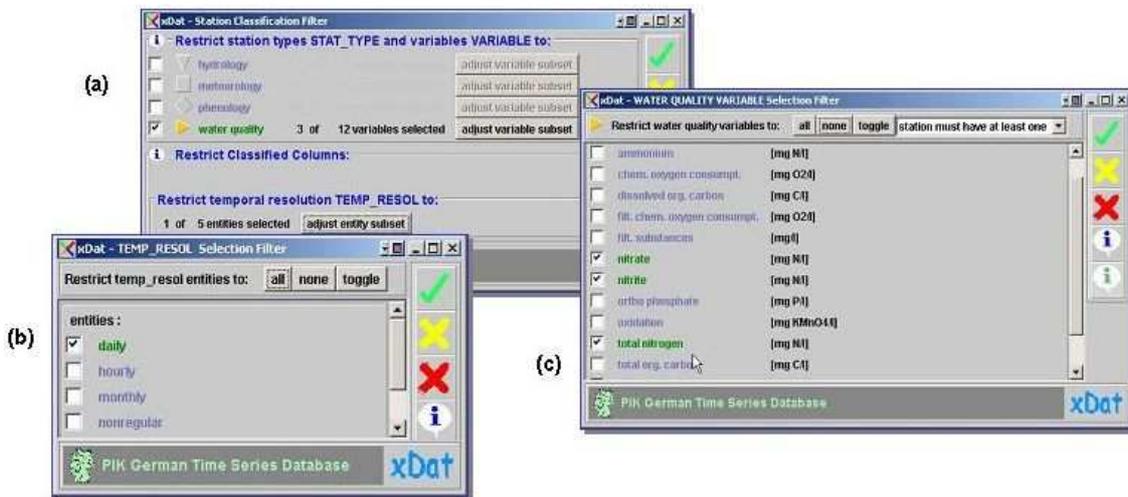


Abb. 20.12 - Orientierung über und Auswahl aus den Klassifizierungskriterien einer Zeitreihendatenbank: (a) Stationstypen; (b) zeitliche Auflösungen; (c) Variablen.

20.9 Statistische Anfragen – StatisticFilter

20.9.1 Motivation und Anforderungen

Die nutzerdefinierte Selektion von Stationen aus Zeitreihendatenbanken kann verfeinert werden, wenn in diese statistische Eigenschaften der an den Stationen erhobenen Zeitreihen einbezogen werden können. Dies soll zunächst an zwei Beispielen verdeutlicht werden.

▪ Beispiel 1 – Vollständigkeit der Messwerte

In der Regel bedingt die Weiterverwendung von Zeitreihen Mindestanforderungen an die Vollständigkeit der innerhalb eines bestimmten Zeitraumes erhobenen Werte. In der Praxis kann jedoch nicht davon ausgegangen werden, dass alle Zeitreihen jeweils komplett vorliegen; so kann beispielsweise der zeitweilige Ausfall von Messgeräten dazu führen, dass einzelne Reihen sporadische Lücken aufweisen. In der Folge sind nicht alle Stationen, die gegebene Variablen in einem bestimmten Zeitraum erhoben haben, gleichermaßen für eine Selektion von Zeitreihen geeignet; entsprechend ist die Möglichkeit, die Vollständigkeit von Zeitreihen für einen nutzerdefinierten Zeitraum als Kriterium für die Selektion von Stationen einzubeziehen, von hoher Relevanz.

Niederschlag täglich China	Mindestvollständigkeit für jedes Jahr zwischen 1970 und 1980, an dem Daten erhoben wurden										
	90%	91%	92%	93%	94%	95%	96%	97%	98%	99%	100%
Anzahl Stationen	212	209	198	180	152	130	116	108	107	107	107

Tab. 20.4 - Anzahl der über die globale PIK-Zeitreihendatenbank verfügbaren Stationen mit täglichen Niederschlagsmessungen in China und spezifischen Anforderungen an die Vollständigkeit für jedes Jahr zwischen 1970 und 1980, an dem Daten erhoben wurden (Abfrage im Mai 2004).

Tab. 20.4 zeigt am Beispiel von Niederschlagsmessungen aus China für die Jahre von 1970 bis 1980 in der globalen Zeitreihendatenbank des PIK, wie sich das Vollständigkeits-

kriterium auf die Anzahl der jeweils geeigneten Stationen auswirkt – durch ein Anheben der geforderten Mindestvollständigkeit von 90% auf 100% wird die Anzahl der selektierten Stationen von 212 auf 107 um nahezu die Hälfte reduziert.

▪ Beispiel 2 – Messwerteigenschaften

Hilfreich für eine zielgenaue Selektion ist ebenfalls die Identifikation von Stationen, an denen die Messwerte selber bestimmte Eigenschaften aufweisen, beispielsweise um Stationen auszuwählen, deren Zeitreihen typisch für aride, humide oder besonders warme bzw. kalte Regionen sind. Fragestellungen, die in diese Kategorie fallen, sind beispielsweise

- die Identifikation derjenigen Stationen, die in Brandenburg liegen, von 1980 bis 1995 tägliche Niederschlagsmessungen durchgeführt haben *und* an denen die jährliche Niederschlagssumme jeweils zwischen 500 und 550 mm liegt
- die Identifikation derjenigen Stationen weltweit, an denen die erhobenen durchschnittlichen Tagestemperaturen im Jahresmittel unter -10°C liegen.

Durch Einbeziehung solcher Kriterien kann die Zahl der selektierten Stationen oft drastisch reduziert werden. Tab. 20.5 verdeutlicht dies am Beispiel der Selektion von Stationen anhand der Jahresmittelwerte von Tagesdurchschnittstemperaturen.

Jahresmittel der Tagesdurchschnittstemperatur	1970-1990								
	maximal					minimal			
	-10°C	-5°C	0°C	5°C	10°C	10°C	15°C	20°C	25°C
Anzahl Stationen	9	31	85	208	666	379	169	34	11

Tab. 20.5 - Anzahl der über die globale PIK-Zeitreibendatenbank verfügbaren Stationen mit spezifischen Anforderungen an die jährlichen Mittelwerte von Tagesdurchschnittstemperaturen. Einbezogen wurden Stationen unabhängig von ihrem Standort, deren Zeitreihen für jedes Jahr von 1970 bis 1990 eine durchschnittliche Vollständigkeit von mindestens 90% aufweisen (Abfrage im Mai 2004).

Um eine entsprechende Unterstützung bei der Selektion von Stationen aus Zeitreibendatenbanken zu ermöglichen, wurde das Filtermodul `StatisticFilter` entworfen, das die nutzerdefinierte Formulierung von Anforderungen an statistische Eigenschaften der an den Stationen erhobenen Zeitreihen erlaubt.

20.9.2 Umsetzung

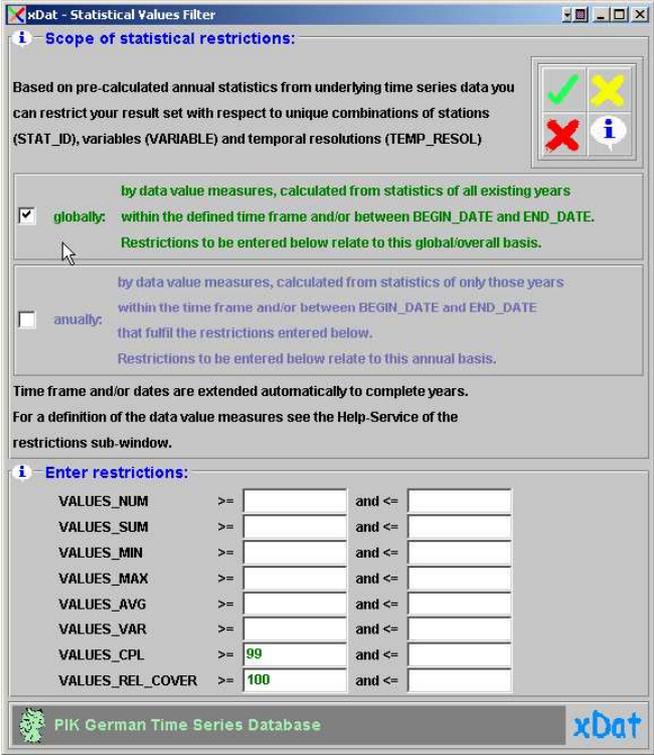
Abb. 20.13 zeigt die graphische Oberfläche des `StatisticFilter`. Anders als die anderen Filtermodule wurde der `StatisticFilter` für spezifische Voraussetzungen der abzufragenden Datenräume konzipiert. Er generiert seine Teilbedingungen für Datenbanktabellen, die vorberechnete Jahrestatistiken in einer bestimmten Form bereitstellen.

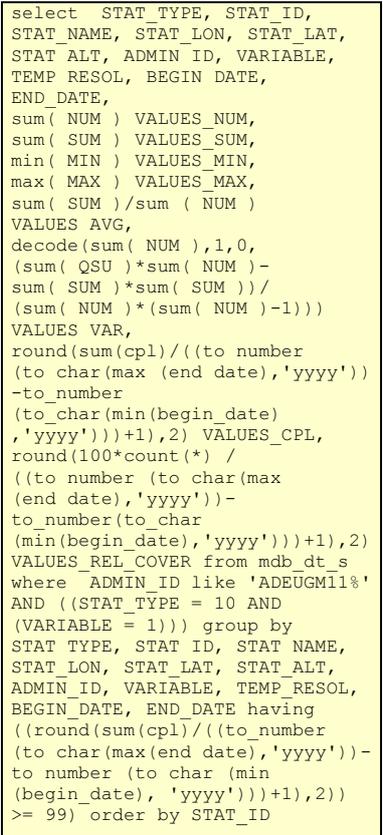
Diese Jahresstatistiken werden von jeder der eingebundenen Zeitreibendatenbanken über eine zusätzliche Datenbanktabelle bereitgestellt, die Kennwerte (Durchschnitte, Summen, Minima und Maxima, Varianzen, Vollständigkeit) für jede auftretende Kombination aus Station, Variable, zeitlicher Auflösung und Jahr enthält. Die Abbildung einer Anfrage auf die jeweils geeignete Tabelle erfolgt dabei transparent durch serverseitige Analyse der in die Anfrage involvierten Attribute, so dass Anfragen, die eine vom `StatisticFilter` generierte Teilbedingung enthalten, automatisch auf die entsprechende Jahresstatistik abgebildet werden (vgl. Kap. 16.2.1, Tabellenkaskaden).

Der `StatisticFilter` erlaubt nutzerdefinierte Einschränkungen bezüglich der statistischen Kennwerte der Jahresstatistiken (vgl. Tab. 18.11); die so definierten Bedingungen werden mit den über den `StationClassifier` ausgewählten Variablen und einem über den `TimeFrameSelector` definierten Zeitfenster in Bezug gesetzt. Verzichtet der Anwender auf die Angabe einer unteren und / oder oberen Grenze für das Zeitfenster, werden statt dessen untere und / oder obere Grenze der jeweiligen zeitlichen Abdeckung der Zeitreihen zugrun-

degelegt. Zusätzlich kann festgelegt werden, welche Jahre des ausgewählten Zeitraumes in eine Berechnung einzubeziehen sind. Gegenwärtig stehen hierfür zwei Modi zur Auswahl:

- Modus *globally*
 - ▶ berücksichtigt sämtliche Jahre innerhalb des vorgegebenen Zeitfensters, für die Daten vorliegen (Vollständigkeit für dieses Jahr größer 0 Prozent). Aus den entsprechenden Werten wird zunächst ein Gesamtwert berechnet, auf den anschließend die gesetzte Bedingung angewendet wird.
- Modus *annually*
 - ▶ berücksichtigt nur diejenigen Jahre innerhalb des vorgegebenen Zeitfensters, für die Daten vorliegen und die vorgegebene Bedingung erfüllt ist.

(a) 

(b) 

```

select STAT_TYPE, STAT_ID,
STAT_NAME, STAT_LON, STAT_LAT,
STAT_ALT, ADMIN_ID, VARIABLE,
TEMP_RESOL, BEGIN_DATE,
END_DATE,
sum( NUM ) VALUES_NUM,
sum( SUM ) VALUES_SUM,
min( MIN ) VALUES_MIN,
max( MAX ) VALUES_MAX,
sum( SUM )/sum( NUM )
VALUES_AVG,
decode(sum( NUM ),1,0,
(sum( QSU )*sum( NUM ) -
sum( SUM )*sum( SUM ))/
(sum( NUM )*(sum( NUM )-1)))
VALUES_VAR,
round(sum(cpl)/((to number
(to char(max(end date),'yyyy'))-
to number
(to_char(min(begin_date)
,'yyyy')))+1),2) VALUES_CPL,
round(100*count(*) /
((to number (to char(max
(end date),'yyyy'))-
to number(to_char
(min(begin_date),'yyyy')))+1),2)
VALUES_REL_COVER from mdb_dt_s
where ADMIN_ID like 'ADEUGM1%'
AND ((STAT_TYPE = 10 AND
(VARIABLE = 1))) group by
STAT_TYPE, STAT_ID, STAT_NAME,
STAT_LON, STAT_LAT, STAT_ALT,
ADMIN_ID, VARIABLE, TEMP_RESOL,
BEGIN_DATE, END_DATE having
((round(sum(cpl)/((to number
(to char(max(end date),'yyyy'))-
to number (to char (min
(begin_date), 'yyyy')))+1),2))
>= 99) order by STAT_ID

```

Abb. 20.13 - Der StatisticFilter: (a) Graphische Oberfläche; selektiert ist ein Vollständigkeitswert von 99%, der für 100% der Jahre des aktuell ausgewählten Verfügbarkeitszeitraumes erfüllt sein soll. (b) Eine basierend auf diesen Kriterien von der Schnittstelle generierte SQL-Anweisung.

Ergänzend kann der Anwender Bedingungen bezüglich der relativen Abdeckung definieren, in dem er prozentuale Vorgaben für das Verhältnis zwischen der Anzahl der einbezogenen Jahre und der durch das Zeitfenster festgelegten Anzahl von Jahren festlegt. Die generierten Teilbedingungen führen jeweils zu einer dynamischen Berechnung von Kenngrößen aus den im adressierten Datenraum vorgehaltenen Jahresstatistiken in Abhängigkeit des vom Anwender ausgewählten Zeitfensters; die so erzeugten Werte werden mit den nutzerdefinierten Vorgaben verglichen, um die entsprechenden Stationen zu identifizieren.

20.10 Attribute für die Ergebnispräsentation – AttributeSelector

20.10.1 Motivation und Anforderungen

Die nutzerdefinierte Selektion von Datenbankattributen für die Präsentation von Anfrageergebnissen nimmt eine Sonderstellung unter den Teilbedingungen ein. Während alle bis-

her beschriebenen Filtermodule dazu dienen, jeweils Teilbedingungen zu definieren, die die Auswahl der zu extrahierenden *Datensätze* beeinflussen, dient eine nutzerdefinierte Attributselektion dazu, das Ergebnis auf diejenigen *Attribute* einzuschränken, die jeweils von Interesse sind. Um es dem Anwender zu ermöglichen, auf diese Weise die Ergebnispräsentation seinen Anforderungen gemäss vorzukonfigurieren, wurde das Filtermodul AttributeSelector entwickelt, an das insbesondere die folgenden Anforderungen gestellt wurden:

- | | |
|---------------------------------------|---|
| Attribut- und Datentyp-unabhängigkeit | ▶ Das Filtermodul soll auf beliebige Datenbankattribute konfiguriert werden können; dabei sind Attribute zu unterstützen, deren Datentyp auf numerischen Werten, Datumswerten oder Zeichenketten basiert. |
| Konfigurierbare Vorbelegung | ▶ Das Filtermodul soll einen konfigurierbaren Satz der jeweils eingebundenen Attribute selektieren können, um für jeden Datenraum eine jeweils geeignete Attributauswahl vorab bereitstellen zu können. |
| Vermeidung sinnloser Kombinationen | ▶ Das Filtermodul soll darauf konfiguriert werden können, eine unabsichtliche Abwahl von Attributen, deren Werteausprägungen für die sinnvolle Interpretation der Werteausprägung anderer, vom Anwender selektierter Attribute unabdingbar sind, zu verhindern. |

20.10.2 Umsetzung

Während pro Datenraum mehrere Filtermodule zum Einsatz kommen können, ist jeweils *genau ein* AttributeSelector bereitzustellen. Der AttributeSelector kann dabei auf die Bereitstellung einer beliebigen Auswahl der Attribute des jeweils zu adressierenden Datenraumes konfiguriert werden³¹⁸. Um eine höhere Unabhängigkeit von den in der jeweiligen Datenbank verwendeten Attributnamen zu erreichen, wird ihre konfigurierbare Maskierung unterstützt. Jedem Attribut kann dazu ein Bezeichner zugeordnet werden, der dann in der graphischen Oberfläche anstelle des Originalnamens verwendet wird. Zur Anfragegenerierung wird jeder Bezeichner dann transparent für den Nutzer auf den entsprechenden Attributnamen umgesetzt³¹⁹.



Abb. 20.14 - Die graphische Oberfläche des AttributeSelector, konfiguriert für eine Zeitreihendatenbank.

Die graphische Oberfläche des AttributeSelector stellt die auswählbaren Attribute in Form direkt anwählbarer Werte zur Verfügung (vgl. Abb. 20.14); dabei wird für jeden Datenraum

³¹⁸ Somit ist es im Bedarfsfall auch möglich, einzelne Attribute als *nicht* für die Ergebnispräsentation zugänglich zu deklarieren

³¹⁹ Dies kann bspw. für eine anwenderfreundlichere Präsentation kryptisch bezeichneter Datenbankattribute oder eine Darstellung von Attributnamen in unterschiedlichen Sprachen eingesetzt werden.

entsprechend der jeweiligen Konfiguration ein Satz von Attributen vorselektiert. Zusätzlich werden folgende Interaktionen³²⁰ unterstützt:

- Default ▶ Es wird nur die jeweils vorkonfigurierte Attributauswahl selektiert.
- Select all ▶ Sämtliche Attribute werden automatisch selektiert.
- Unselect all ▶ Sämtliche abwählbaren Attribute werden automatisch deselektiert.

Die Orientierung über die Bedeutung der jeweils angebotenen Attribute wird dabei durch die Möglichkeit erleichtert, interaktiv Informationen über diese abzurufen. Für jeden Datenraum kann dem AttributeSelector eine Hilfedatei zugeordnet werden, die entsprechende Beschreibungen enthält und die der Anwender auf Wunsch öffnen kann.

Die Semantik einzelner Datenbanken kann es erfordern, die Auswahl sinnvoller Attributkombinationen sicherzustellen. Abb. 20.15 zeigt drei Ergebnistabellen einer Zeitreihendatenbank, die aus jeweils unterschiedlichen Kombinationen der Attribute Stat_Name (Stationsname), Variable (erhobene Variable) und Temp_Res (zeitliche Auflösung) bestehen.

Variable	Temp_Res	Stat_Name	Variable	Temp_Res	Stat_Name	Temp_Res
precipitation	daily	Station A	precipitation	daily	Station A	daily
temperature	monthly	Station B	temperature	monthly	Station B	monthly
air pressure	daily	Station C	air pressure	daily	Station C	daily
cloudiness	daily	Station D	cloudiness	daily	Station D	daily

(a)

(b)

(c)

Abb. 20.15 - Auswirkungen von Attributkombinationen auf den Informationsgehalt von Ergebnistabellen: Tabelle (b) ist im gegebenen Kontext interpretierbar; die Tabellen (a) und (c) hingegen nicht.

Dabei wird deutlich, dass nicht jede so entstehende Ergebnistabelle interpretierbar ist. Während die Attributkombination der Tabelle in Abb. 20.15b das Ablesen sinnvoller Informationen erlaubt, ist dies für die beiden anderen Tabellen nicht der Fall. Durch Abwahl des Attributes Stat_Name entsteht die Tabelle in Abb. 20.15a, die keine Rückschlüsse darauf zulässt, an welcher Station die beschriebenen Parameter gemessen werden. Damit ist sie ebenso wenig aussagekräftig wie die Tabelle in Abb. 20.15c, in der auf das Attribut Variable verzichtet wurde – hier kann zwar abgelesen werden, wie oft an den einzelnen Stationen Werte erhoben werden, aber nicht, um welche Parameter es sich dabei jeweils handelt.

Es ist offensichtlich, dass in solchen Fällen die Selektion einiger Attribute für die Darstellung eines Anfrageergebnisses nur dann sinnvoll ist, wenn zugleich bestimmte weitere Attribute ausgewählt werden. Um Anwender davon zu entbinden, unabsichtlich zu nicht interpretierbaren Anfrageergebnissen zu gelangen, kann der AttributeSelector entsprechend konfiguriert werden. Dabei können einzelne Gruppen von Attributen definiert und für diese jeweils zentrale Attribute ausgezeichnet werden. In jeder Gruppe führt dann die Selektion eines Attributes zur automatischen Selektion der zentralen Attribute; die Deselektion zentraler Attribute einer Gruppe führt hingegen zur Deselektion aller dieser Gruppe zugeordneten Attribute. Zusätzlich können die Gruppen in eine Rangfolge gebracht werden, so dass die Selektion der zentralen Attribute höher eingestufte Gruppen jeweils sichergestellt wird. Ferner können einzelne Attribute als *nicht deselektierbar* gekennzeichnet werden. Dies wird primär dazu eingesetzt, um sicher zu stellen, dass Attribute, die für eine programmseitige Ergebnisverarbeitung³²¹ notwendig sind, im Ergebnis enthalten sind.

³²⁰ Auf Möglichkeiten, an dieser Stelle eine Vorsortierung anhand von Werteausprägungen einzelner Attribute auszuwählen, wurde verzichtet, da entsprechende Funktionalitäten von den Auswertungsmodulen (vgl. Kap. 21) interaktiv bereitgestellt werden. Über das Auswertungsmodul TableView (vgl. Kap. 21.2.3) kann ferner die Anordnung der Ergebnisattribute verändert werden.

³²¹ So erfordert beispielsweise die Auswertung von Ergebnisdaten, die sich jeweils auf eine Station beziehen, eine eindeutige Zuordnung jedes Datensatzes über ein entsprechendes Identifikatorenattribut.

Der `AttributeSelector` generiert aus den vom Anwender ausgewählten Attributen eine Teilbedingung, die zu einer entsprechenden Attributselektion für das Anfrageergebnis führt. Das Filtermodul kann dabei programmseitig aufgefordert werden, zuvor zu überprüfen, ob bestimmte Sätze von Attributen selektiert wurden und gegebenenfalls weitere Attribute zu selektieren. Diese Funktionalität wird dazu eingesetzt, um vor der Generierung einer Gesamtanfrage festzustellen, ob der Anwender Attribute abgewählt hat, die für die Bereitstellung bestimmter Formen der Ergebnisauswertung erforderlich sind, und gegebenenfalls eine automatische Nachselektion anzubieten (vgl. Kap. 20.11).

Der `AttributeSelector` wird bei sämtlichen eingebundenen Datenräumen eingesetzt. Bei den einzelnen Zeitreihenmetadatenbanken und Zeitreihendatenbanken dient er jeweils zur nutzerdefinierten Auswahl aus den jeweils verfügbaren Attributen. Eine andere Verwendung wurde für die allgemeine Institutsmetadatenbank PIK CERA-2 gewählt. Da sich für diesen Datenraum eine Präsentation der Anfrageergebnisse in Form einer vordefinierten Formularmaske als geeignet erwiesen hat (vgl. Kap. 21.1), ist hier die Auswahl von Attributen vor Beginn einer Anfrage nicht erforderlich. Entsprechend ist der `AttributeSelector` für PIK CERA-2 ausschließlich auf eine interaktive Bereitstellung von Informationen über die Bedeutung der verfügbaren Attribute konfiguriert.

20.11 Generieren und Ausführen einer Anfrage

Die internen Abläufe, die zur Generierung einer Anfrage aus den über die jeweils bereitgestellten Filtermodule nutzerdefinierten Teilbedingungen führen, wurden in Kap. 15.1 beschrieben. Hier ist noch nachzutragen, dass die Schnittstelle dem Anwender eine gewisse Unterstützung bei der Auswahl geeigneter Attribute für die Auswertung der Anfrageergebnisse anbietet. Die Konfiguration legt neben den jeweils zu verwendenden Auswertungsmodulen auch die Attribute fest, auf denen diese aufsetzen sollen (vgl. Kap. 14.3.2). So erfordert bspw. die interaktive raumbezogene Visualisierung von Zeitreihenmetadaten die Einbeziehung der Attribute, die die geographische Breite und Länge der Stationen dokumentieren, während für einen nachfolgenden Zugriff auf Zeitreihen die Werteausprägungen der Attribute zur Beschreibung von Stationsidentifikatoren, Variablen und zeitlichen Auflösungen erforderlich sind (vgl. Kap. 18.4.1). Ist die Schnittstelle für den ausgewählten Datenraum auf eine oder mehrere dieser Funktionalitäten konfiguriert, konsultiert der Filter-Manager der Kernkomponente *Selektion* (vgl. Kap. 15.1.2) bei der Anfragegenerierung den `AttributeSelector`, um festzustellen, ob der Anwender eines oder mehrere der hierfür jeweils relevanten Attribute deselektiert hat. Ist dies der Fall, wird der Anwender vor Ausführung der Anfrage darauf hingewiesen; er kann dann auswählen, ob die gewählte Attributkombination beibehalten oder ob automatisch um die fehlenden Attribute erweitert werden soll.

Die generierte Anfragebeschreibung wird an den Server der Schnittstelle weitergeleitet, der eine entsprechende Datenextraktion durchführt und die Ergebnisdaten zurück zum Rechner des Anwenders transferiert. Diese werden intern bereitgestellt und die gemäß der aktuellen Systemkonfiguration zu verwendenden Auswertungsmodule aktiviert.