

4.5 Projektinformations- und Datenmanagementsystem *premis*

Ziel des Teilprojektes ist die Konzeption und die Erstellung eines Management und Informationssystem für das Verbundprojekt *pre agro*. Dieses System wird als **preMIS** bezeichnet. Als Hauptkommunikationsmittel wurde das Internet gewählt. Dazu wurde eigens eine Domain mit einer Web-Site <http://www.preagro.de> eingerichtet. Die Web-Site erscheint nach aussen mit einem externen Bereich für die Öffentlichkeit und einem durch Benutzername und Passwort geschützten internen Bereich für die Projektmitarbeiter. Das Konzept der Informationsbereitstellung beruht auf den drei Ebenen *preagro Web*, *preagro Meta* und *preagro Map*.

Die Ebene **preagro Web**, im wesentlichen durch die externen Seiten repräsentiert, dient der Darstellung des Projektes und seiner Ergebnisse nach aussen hin, sowie als Plattform für die Diskussion zum Thema precision agriculture. Des weiteren zählen dazu Hilfsmittel zur Organisation der kooperativen Zusammenarbeit. Zu diesem Zweck wurden ein BSCW-Server integriert und verschiedene eigene Skripte geschrieben.

Zum **preagro Meta** zählt das Metainformationssystem (MIS), mit dem die GIS-Daten beschrieben werden.

Die Ebene **preagro Map** umfasst die tatsächlichen GIS-Daten, die als Grundlage für die Arbeit im precision agriculture dienen sowie Hilfsmittel zur Visualisierung und zur Bearbeitung. Für die Speicherung der Daten wurde eine verbindliche Dateinamenskennung und eine Verzeichnisstruktur festgelegt. Der Zugriff auf die Daten erfolgt über die darüberliegende Ebene *preagro Meta*. Im folgenden wird besonders auf das methodische Vorgehen bei der Erstellung des Metainformationssystem eingegangen.

Im precision agriculture werden schon auf einem einzelnen Schlag viele Daten erhoben, um die Inhomogenitäten des Bodens und des Bestandes zu erfassen. Im gesamten Projekt wird eine nahezu unüberschaubare Menge an Daten erwartet. Deshalb ist es erforderlich, Metainformationen zu den einzelnen Daten zu erfassen und für eine strukturierte Recherche zur Verfügung zu stellen.

Zum besseren Verständnis sollen hier zwei Begriffe definiert werden. Als Daten werden im *pre agro Meta* die Daten verstanden, die in Dateien oder in Datenbanken ohne explizite Metadaten vorliegen. Als Datensätze werden die von den Projektpartnern erhobenen Dateien verstanden, die inhaltlich oder programmtechnisch zusammengehören. Die Regelung, welche Dateien zusammengehören, ist jeweils abhängig von der Software und dem Bearbeiter, der sie angelegt hat. Teilweise sind sie gekennzeichnet durch gleiche Dateinamen mit unterschiedlichen Extensionen im gleichen Verzeichnis stehend, andere sind nur dadurch gekennzeichnet, dass sie in einem gleichen Verzeichnis stehen. Manche zusammengehörige Dateien sind durch gleiche Teile im Namen gekennzeichnet, und bei anderen bestehen Beziehungen zu den Verzeichnisnamen. Es können aber auch beliebige Kombinationen aus diesen Zugehörigkeitsbeziehungen auftreten. Die Zugehörigkeit von Dateien über den Namen und die Verzeichnisstruktur, in der sie stehen, ist nur beim automatischen Einlesen der Metainformationen wichtig. Nach der Erfassung von Dateien sind die Zugehörigkeiten in der Metadatenbank gespeichert.

Die beschreibenden Metadaten werden als Informationen über die eigentlichen Daten verstanden. Die Trennung der Daten von den Metadaten macht die Beschreibung unabhängig vom For-

mat der Daten. Die Daten werden in einer Verzeichnisstruktur in den originalen Dateien gespeichert, so dass mit den entsprechenden Programmen auch direkt darauf zugegriffen werden kann. Die Metadaten werden in einer relationalen Datenbank, der Metadatenbank (MDB), gespeichert. Der Dateiname, der relative Pfad und die Beschreibung des Datenträgers mit Angabe des absoluten Pfades werden als beschreibende Information in die Metadatenbank übernommen und dienen der Identifizierung bzw. als Bindeglied zwischen den Daten und den dazugehörigen Metadaten. Innerhalb der MDB beziehen sich dann alle Datei-relevanten Angaben auf die Identifikatoren der Dateientabelle. Über das Metadatenmodell können also auch ausgelagerte Daten beschrieben und ein Zugriff darauf bereitgestellt werden.

Insbesondere werden Fernerkundungsdaten auf externen Datenträgern gehalten, weil die Menge dieser Daten die Kapazität des Servers schnell sprengen würde. Es wird aus diesem Bereich mit Daten von ca. 50 Gigabyte pro Jahr gerechnet.

Um eine Strukturierung der weiteren Metadaten vornehmen zu können, musste ermittelt werden, was für Daten überhaupt erfasst werden, danach mit welchen Daten sie sich beschreiben lassen und letztlich, wie sich die Metadaten für einen strukturierten Zugriff präsentieren lassen.

Die Metadaten lassen sich zunächst einteilen in administrative Daten, Daten, die Dateien beschreiben, den Raumbezug beschreibende Daten und Expertenwissen.

Zu den administrativen Angaben gehören:

Daten zu Personen, Institutionen, Projekten, Zugehörigkeiten und Zugriffsrechte

Zu den Angaben zu Dateien und Dateigruppen gehören:

Dateiname und Pfadangabe, Dateigrößen, Speicherbezeichnung, Dateityp, Dateigruppenzugehörigkeit, Zeitpunkte der Archivierung, der Erfassung und der Bearbeitung der Dateien, Angaben, welche Personen die Dateien erhoben, archiviert und bearbeitet haben, um welche Art von Dateigruppen es sich handelt, welchen Raumbezug sie haben und welche speziellen Merkmale aus der entsprechenden Fachsicht die allgemeinen Informationen ergänzen.

Die Festlegung der Attribute erfolgt auf Basis der Anforderungen, der am Projekt beteiligten Fachdisziplinen.

Zu den Angaben für den Raumbezug gehören:

Bezeichnung von Regionen, Betriebsflächen und Schlägen, Angaben über Flächen, Linien und Punkte, Angaben über Koordinatensysteme, Relationen zwischen den Objekten und Definitionen von Arten der Objekte

Zum Expertenwissen werden gezählt:

Literaturhinweise, Begriffserklärungen, Abkürzungen, Methodenbeschreibungen und die Relationen dazwischen

Zur Beschreibung der Daten wurden Merkmale gesucht, die allen gemein sind. Es wurde festgestellt, dass allen Daten ein Raumbezug, ein Sachbezug und ein Zeitbezug zugeordnet werden kann. Diese Merkmale werden als zentraler Kern des *pre agro* Meta betrachtet und sind

Pflichtangaben. Damit ist also sichergestellt, dass alle Datensätze räumlich, zeitlich und sachlich wiedergefunden und bereitgestellt werden können.

Daten, die nicht in Dateiform sondern in separaten Datenbanken abgespeichert sind, z. B. Wetterdaten der pre agro-Stationen, werden in ihrem Raum-, Zeit- und Sachbezug wie in der MDB gespeichert. Dadurch lassen sich die gleichen Suchbegriffe verwenden wie bei den GIS-Daten in Dateiform. Die Einbeziehung einer anderen Datenbank in das pre agro Meta erfolgt also dadurch, dass die Pflichtelemente der MDB in der anderen Datenbank mit abgebildet sind.

Die Merkmale sind im pre agro Meta entsprechend nachfolgender Darstellung unterteilt und nach diesem Schema erweiterbar.

Raumbezug

- *Regionen*
(Kassow, Groß-Twülpstedt, Thumby, Landshut, Raesfeld, Raguhn, Baasdorf, Zeilitzheim)
- *Betriebe*
(Ksg, Qei, Gtw, Baa, Akn, Wul, Rod, Rep, Pet, Lan, Sel, Que, Rag, Ros, Rae, Bec, Thy, Gun, Ess, Ham, Gpl, Rei, Zei)
- *Schläge* (mit/ohne precision agriculture bewirtschaftet)
ohne Aufzählung

Sachbezug

- *Kategorien*
(Ackerschlagkartei, Betriebsinformationen, Fernerkundung, geogr. Basisinformationen, Boden, Pflanze, Ökologie)
- *Datenarten*
(Luftbild, TK ohne Maßstab, Reichsbodenschätzung, Schlagumringe, TK 25, TK 200, DGK 5, DGM, Satellitenbild, Bodendaten, Bohrungen, Thermaldaten, Standortkartierung, TK 50, TK 100, ECa, Aussaatkarten, NIR Daten...)
- *Personenbezug* (Erfasser, Bearbeiter, Archivierer)

Zeitbezug

- *Erfassungsdatum*
- *Bearbeitungsdatum*
- *Archivierungsdatum*

Die Einteilung ist hierarchisch, damit eine Strukturierung vom "Grossen ins Kleine" möglich ist. Die Zuordnung der Datensätze innerhalb des Begriffsraumes für Raum-, Zeit- und Sachbezug kann in einem Koordinatensystem veranschaulicht werden. Dieser Raum wird Metaraum genannt.

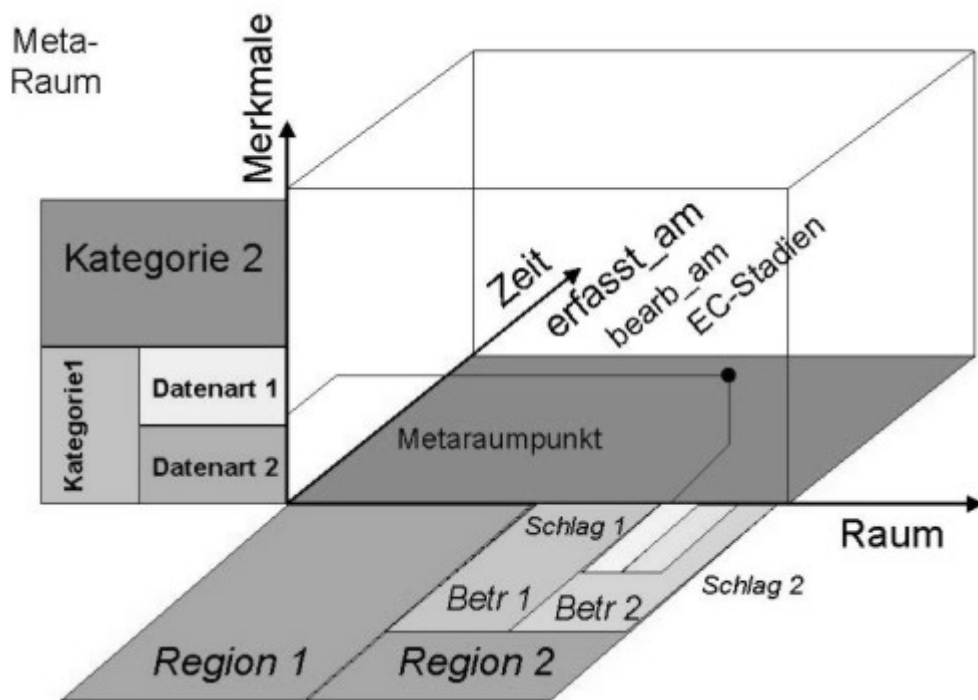


Abb. 5.-1: Die Veranschaulichung des Raum-, Zeit- und Sachbezuges durch den Metaraum

Da sich Datensätze auch auf mehrere Betriebe oder gar Regionen, z. B. Luftbilder und Satellitenbilder, beziehen lassen, können einem Datensatz auch mehrere Punkte im Metaraum zugewiesen werden. Die Mindestangaben für Datensätze im *pre agro* Meta werden Koordinaten im Metaraum genannt und ähnlich wie räumliche Koordinaten in einer Tabelle gespeichert. Die Punkte im Metaraum werden als Metaentitäten bezeichnet.

Das Modell des Metaraumes lässt sich auch für die Darstellung der Relationen zwischen den Datensätzen, den Benutzern des *pre agro* Meta und den Aktionen, z. B. lesen, schreiben, löschen, anwenden. Hier steht die Frage: „wer darf womit was tun?“, wobei mit wer immer eine Gruppe gemeint ist und eine Gruppe aus mindestens einem Benutzer besteht. Diese Metaräume sind in Tabellen einer relationalen Datenbank abgebildet.

Zu jeder Art von Datensatz werden noch zusätzliche fachbezogene Metadaten abgespeichert. Dazu ist jedem Datensatz eine Tabelle zugeordnet.

Bei der Definition der zusätzlichen Attribute wurde darauf geachtet, dass keine direkten Redundanzen auftreten. Indirekt können sie jedoch vorkommen, z. B. bei der Überschneidung von Koordinaten im übergeordneten Koordinatensystem mit der Angabe einer Region. Die Verknüpfung zwischen den Datensätzen, dem Metaraum und den Tabellen mit den speziellen Metadaten erfolgt über die Datensatz_ID (siehe Abb. 5.-2).

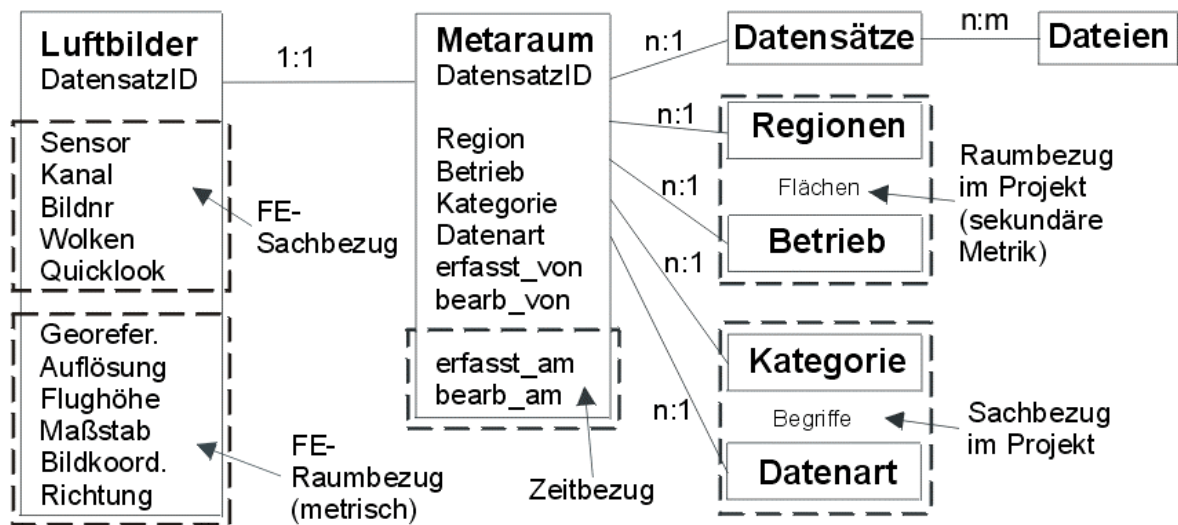
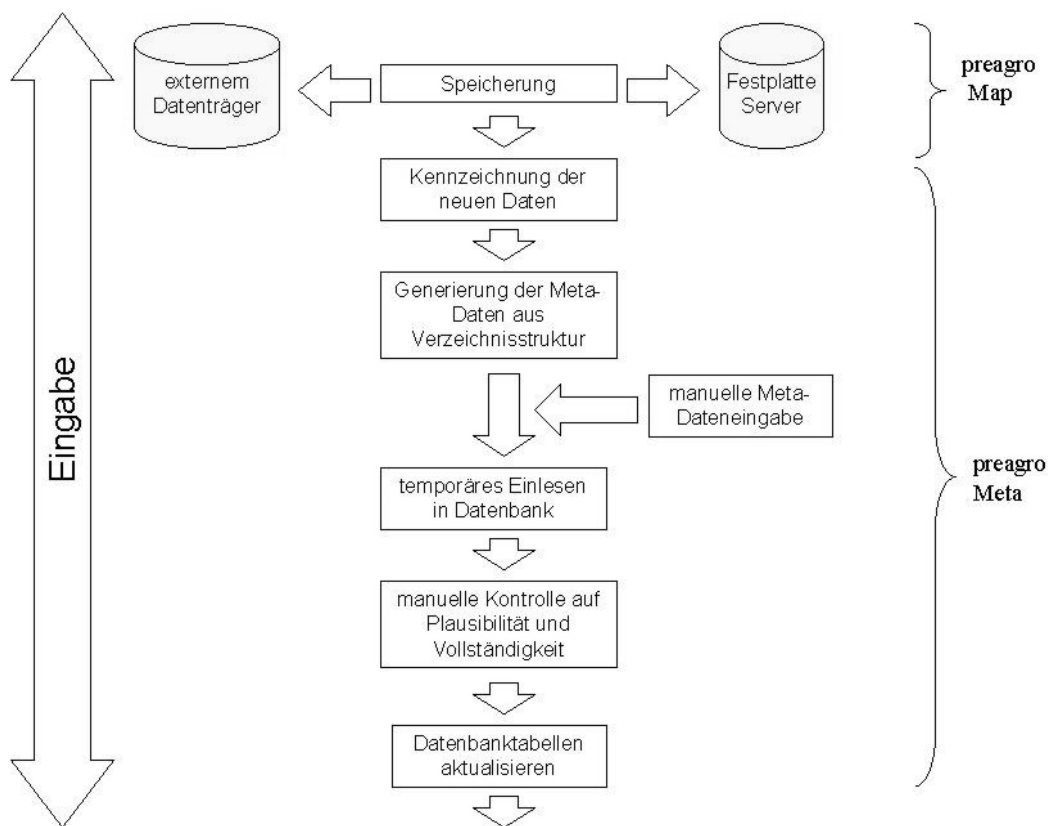
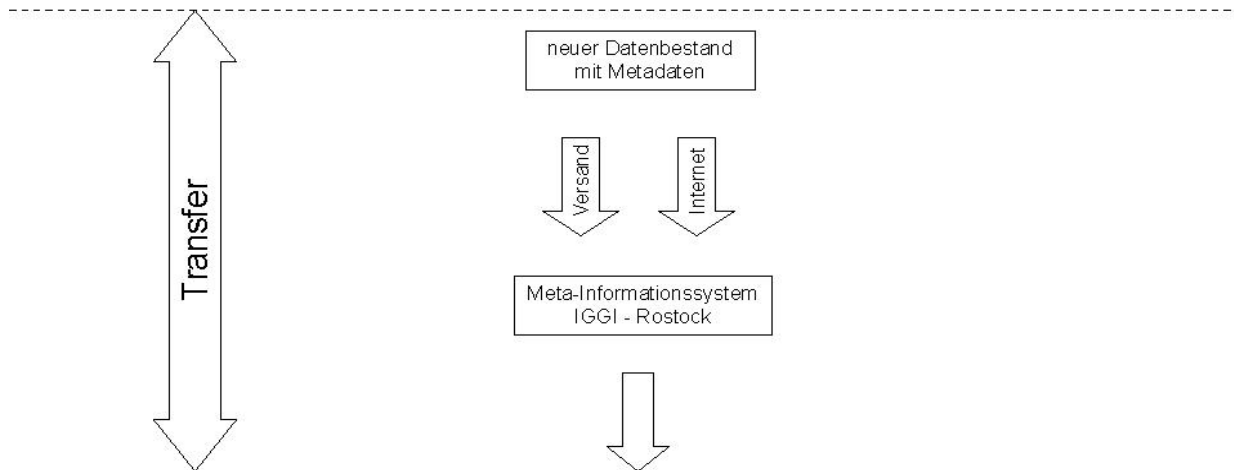


Abb. 5.-2: Raum-, Zeit- und Sachbezug für die Fernerkundung (FE) im Metadatenmodell

Datenfluss

In Abbildung 5.3 ist der Datenfluss schematisch dargestellt. Nachdem die datenerfassenden Stellen ihre Daten an das *pre agro* MIS geschickt haben, werden sie dort zusammen mit ihren Metadaten erfasst und gespeichert. Die Erfassung der Metadaten erfolgt teilweise automatisch, soweit sie sich aus den Daten ermitteln lassen, und teilweise durch manuelle Eingabe. Nach der Recherche, die durch die Metadatenabfrage möglich wird, kann der Datennutzer über eine Bestellung und den Transfer auf die Daten zugreifen.

Datenfluß-Diagramm im preagro Management- und Informationssystem



Fortsetzung

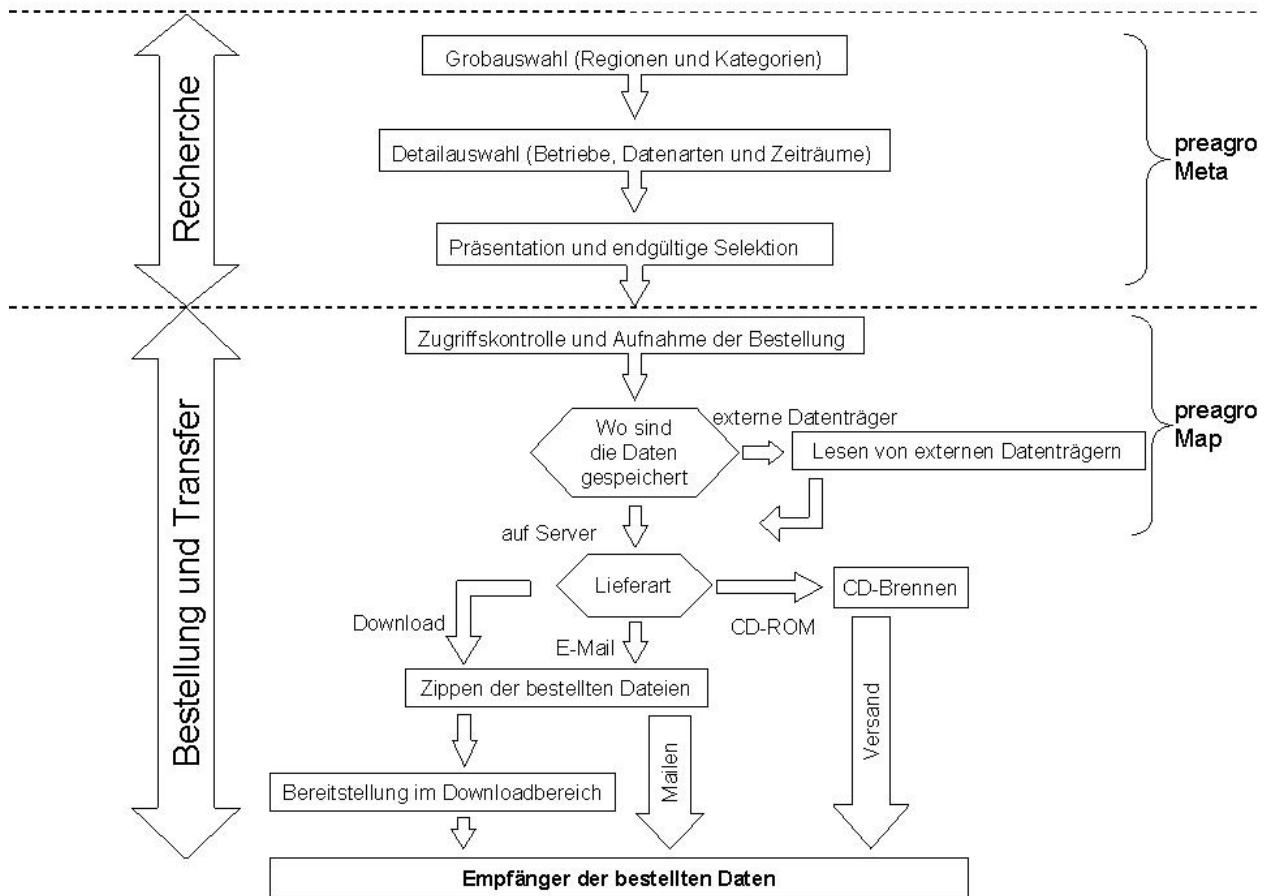


Abb. 5.-3: Datenablaufplan im pre agro-Management- und Informationssystem (pre agro MIS)

Zur Recherche über die Daten wurde eine graphische Oberfläche erarbeitet, in der über vorgegebene Begriffe aus den Raum-, Zeit- und Sachbezügen ausgewählt werden kann. Am Anfang wählt der Benutzer in einer Grobauswahl die Region und die Kategorie aus (siehe Abb. 5.-4). Danach erscheint eine Detailauswahl über alle Betriebe und Datenarten, zu denen Daten vorliegen. Darin erfolgt eine Auswahl der gewünschten Betriebe, der Datenart und der Zeitangabe (siehe Abb. 5.-5). Bei Region, Kategorie, Betriebe und Datenart können immer mehrere Auswahlkriterien gleichzeitig gewählt werden. Bei der Zeit kann zwischen den Datumsangaben für Archivierung, Bearbeitung und Erhebung unterschieden werden, wobei auch hier eine Mehrfachauswahl möglich ist. Des Weiteren können einzelne Intervalle über mehrere Jahre oder der Gesamtzeitraum (von - bis) ausgewählt werden. Bei den Kategorien Ackerschlagkartei, Fernerkundung, Pflanze und Ökologie können zusätzlich das Pflanzenentwicklungs-Stadium (von - bis) und die Fruchtarten ausgewählt werden. Anschliessend werden die gesuchten Datensätze mit ihren Metadaten angezeigt, und der Benutzer kann die Auswahl ggf. weiter einschränken.

Grobauswahl von GIS-Daten

Übersichtskarte

Zur Zeit noch keine graphische Auswahl möglich!

Regionen:

- Groß-Twülpstedt
- Kassow
- Landshut
- Seda
- TAS-Raguhn
- Thumbby
- Wimex
- Zeilitzheim

mehrere möglich

Kategorien:

- (Ackerschlagkarte)
- Betriebsinformation
- Fernerkundung
- geor. Basisinformation
- Boden
- Pflanze
- (Ökologie)

mehrere möglich

Zu den Kategorien, die in Klammern gesetzt sind, liegen noch keine Daten vor.

[weiter >>](#)

Abb. 5.-4: Grobauswahl für die GIS-Datenrecherche

Für die Bestellung der Daten wird der Nutzer aufgefordert einen Grund der Bestellung (Stichwort) und einen Verwendungszweck (ausführlich) einzugeben. Diese Angaben dienen dem Projekt zur späteren Auswertung. Es werden dadurch Erkenntnisse über die Datenverwendung und die Abläufe beim Management von precision agriculture erwartet. Nach der Angabe der Bestellart (per E-Mail oder per CD-ROM) wird im *pre agro* Meta automatisch abgefragt, für welche Datensätze der Besteller die Zugriffsrechte hat. Nachdem der Besteller diese Liste der ausgewählten Daten bestätigt, werden die zugehörigen Dateien je nach Bestellart weiterverarbeitet und entweder zum Download bereitgestellt oder als CD-ROM per Post verschickt.

Der Administrator von *premis* bekommt bei einer Bestellung von Daten, die auf externen Datenträgern vorliegen, automatisch eine Aufforderung zur Bereitstellung der Daten. Nach dem Bereitstellen dieser Daten wird der Besteller informiert und hat 7 Tage Zeit, sich das bestellte Paket auf seinen Rechner zu laden bzw. bekommt eine CD-Rom per Post zugeschickt. Die Bestellinformationen können von der Projektleitung eingesehen und ausgewertet werden.

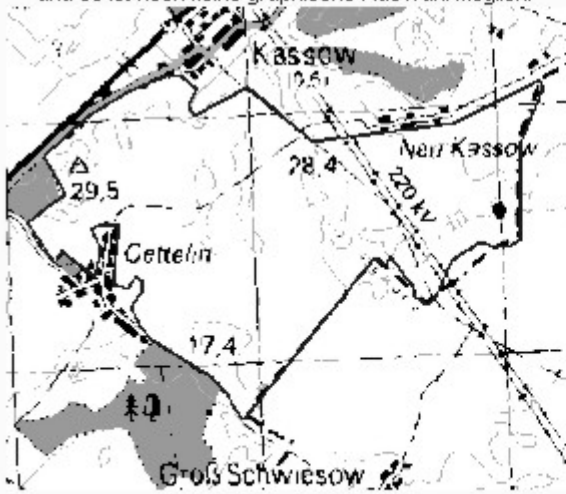
GIS-Daten-Detailauswahl

Ausgewählte Regionen: Kassow, Landshut
Ausgewählte Kategorien: geogr. Basisinformationen, Boden

Betrieb:

Schlagauswahl:

Zur Zeit wird nur die erste gewählte Region angezeigt
und es ist noch keine graphische Auswahl möglich!



Datenart:
 TK ohne Maßstab
 Reichsbodenschätzung
 TK 25
 DGM
 Bodendaten

Zeitangabe:
 Archivierung
 Bearbeitung
 Datenerhebung (Mehrfachauswahl möglich)

Suchzeitraum:
 von: 01 Januar 1999
 bis: 16 August 2000

Zeitintervalle
 Daten aus allen Zeiträumen

Kleine Vorschaubilder anzeigen
 Parameter anzeigen

Abb. 5.-5: Detailauswahl für die GIS-Datenrecherche

Für die Recherche von Fernerkundungsdaten wurde ein Java basiertes Tool entwickelt, um die textuelle Metadatenabfrage mit einer graphisch interaktiven Auswahl zu ergänzen.

Zur Unterstützung der wissenschaftlichen Arbeiten im Verbundprojekt ist somit eine erste Basis geschaffen. Das *pre agro* MIS (*premis*) wurde dabei so konzipiert, dass es für die ständig wachsenden Anforderungen des Projektes kontinuierlich weiterentwickelt wird.

Die Metadatenbank wurde mit MS-ACCESS unter NT auf einem Arbeitsplatzrechner entwickelt und in eine MySQL Datenbank über die ODBC-Schnittstelle exportiert. Access mit ODBC-Zugriff dient neben phpMyAdmin mit PHP3-Internetzugriff als Datenbankclient zur Weiterentwicklung und Pflege der Datenbank. Die dynamischen Seiten der *preagro* Web-site werden mit PHP3 realisiert.

Zur Zeit werden Skripte geschrieben, die es den Projektteilnehmern ermöglichen ihre Daten selbständig über das Internet auf den Server zu laden und die Metadaten dazu zu erzeugen und gegebenenfalls zu ändern. Darüber hinaus werden Hilfsmittel entwickelt, die die Administration der Datenbestände und der Inhalte der Web-site über das Internet ermöglichen.

