

Wissensbasierte digitale Unterstützung in der Pflanzenbauberatung

Anahita Nafissi¹, Fabian Weckesser², Ingmar Kessler¹, Markus Rickert¹, Matthias Pfaff¹, Sebastian Peisl² und Michael Beck²

Abstract: Obwohl die Landwirtschaft schon immer technologische Neuerungen in der Produktion einsetzt, ist gerade die landwirtschaftliche Beratung bisher noch verhältnismäßig wenig digitalisiert. Oft ist es gängige Praxis für Landwirte und Berater, gesetzliche Regelungen, Fachliteratur und Betriebsdaten in Papierform nachzuschlagen und in unstrukturierten, digitalen Dokumenten einzutragen. Zur Unterstützung der Landwirte und der Pflanzenbauberatung wird ein Entscheidungshilfesystem entwickelt, welches die Beratung in der Landwirtschaft digital unterstützen und erleichtern soll, indem es aktuelles Fach- und Expertenwissen sowie individuelle Betriebsdaten abrufen, aufbereiten und zweckgebunden auswertet. Dafür ist es notwendig, das entsprechende Fachwissen aus vielfältigen heterogenen Datenquellen in einer einheitlichen Wissensbasis verfügbar zu machen. Der hier beschriebene Ansatz verwendet Semantic-Web-Technologien wie OWL-Ontologien und SPARQL-Abfragen, um diese Daten hinsichtlich ihrer Bedeutung, d. h. semantisch, zu modellieren und abzufragen.

Keywords: Pflanzenbauberatung, Entscheidungshilfesystem, Datenintegration, RDF, OWL, SPARQL, Ontologien, Semantic Web, Linked Data

1 Einleitung und Problemstellung

Je nach Technologisierungsgrad und Persistenz der Aufzeichnung des landwirtschaftlichen Betriebes akkumulieren sich während des pflanzenbaulichen Produktionsprozesses verschiedenste Daten wie bspw. Bewirtschaftungs-, Wetter-, Boden- und Pflanzenwachstumsdaten. Diese werden automatisiert oder manuell erfasst. Dabei wird in Zukunft die zunehmende Durchsetzung der Produktion mit cyber-physischen Systemen, georeferenzierter Maßnahmenfassung und Vernetzung der Arbeitsabläufe die Qualität und Quantität der Daten noch weiter steigern. Dies gilt auch für kleinstrukturierte Regionen, wie eine Studie mit bayerischen Landwirten zeigt [RG17]. Ergänzt werden diese Informationen durch fachliches Wissen und gesetzliche Rahmenbedingungen, sowie relevante Informationen anderer Teilnehmer der Wertschöpfungskette des Ackerbaus. Dem Landwirt ist es aufgrund dezentraler Datenablage sowie inhomogener Datenstruktur und -qualität jedoch kaum möglich, das darin liegende Potenzial zur Betriebsoptimie-

¹ fortiss, Landesforschungsinstitut des Freistaats Bayern, Guerickestraße 25, 80805 München, {nafissi;ikessler;rickert;pfaff}@fortiss.org

² Hochschule Weihenstephan-Triesdorf (HSWT), Zentrum für Forschung und Wissenstransfer, Am Staudengarten 10, 85354 Freising, {fabian.weckesser;sebastian.peisl;michael.beck}@hswt.de

zung zu nutzen – weder zur ökonomischen, noch zur ökologischen oder gesellschaftlichen Zielerreichung.

Viele Betriebe in kleinstrukturierten Regionen mit heterogener Flächenausstattung, vielseitiger Produktionsausrichtung und hoher Einkommensdiversifikation profitieren von den Empfehlungen von Pflanzenbauberatern mit langjährigen Erfahrungsmustern und aktuellstem Fachwissen. Der Experte ist für den interdisziplinär arbeitenden Landwirt von hohem Mehrwert und hilft diesem, dem steigenden Effizienz- und Qualitätsdruck aller Stakeholder der pflanzlichen Produktion gerecht zu werden [PT17].

Für die Pflanzenbauberatung ergeben sich durch integrative und datengetriebene Ansätze neue Chancen und Herausforderungen in der Beratung. Die Aufgabe, über den Einzelbetrieb hinaus bei der überbetrieblichen Analyse der zeit- und ortsgebundenen Daten betriebsindividuell zu beraten, erfordert eine gute Vernetzung. Der Pflanzenbauberater ist daher auf einen themen- und fragenspezifischen schnellen Zugriff in einer gesicherten Datenqualität angewiesen. Dies setzt Datenkonnektivität bei gesichertem Datenschutz voraus und ist damit Grundlage der Nutzung dieser verteilten Daten in Echtzeit [Wa17]. Diese Interoperabilität ist Grundlage für die Entwicklung eines Entscheidungshilfesystems, um Effizienzsteigerungspotenziale zu realisieren.

Eine aktuelle Problemstellung ist, dass Landwirte und Berater bei der Entscheidungsfindung oft noch von Hand Fachwissen in verschiedenen Datenquellen und -formaten von unterschiedlichen Herausgebern nachschlagen und miteinander kombinieren müssen. Daher bietet es sich an dieser Stelle an, die Landwirte und Berater mittels eines entsprechenden Werkzeugs zu unterstützen, welches verschiedenste Datenquellen fachlich und technisch in eine einheitliche Wissensbasis integriert, um in einer Abfrage auf sämtliches Wissen gleichzeitig zugreifen und entsprechende Zusammenhänge herstellen zu können.

Zum Aufbau eines solchen Entscheidungshilfesystems bieten sich Semantic-Web-Technologien wie Ontologien an, welche bereits in verschiedenen Kontexten in der Landwirtschaft eingesetzt werden. In [Ma15] wird die Architektur eines Linked-Open-Data-Services mit Suchserverkomponenten beschrieben. Darüber hinaus beschreibt [Ma18] eine Infrastruktur mit Semantic-Web-Technologien, um eine semantische Suche über heterogene KTBL-Planungsdaten zu ermöglichen. Im Vergleich dazu ist der Fokus dieses Ansatzes die Integration von fachlich und technisch unterschiedlichen Datenquellen verschiedenster Herausgeber, um übergreifende Fragestellungen aus diversen Teilgebieten zu beantworten. [NBF18] beschreibt einen Ansatz zur Datenintegration von mehreren pflanzenphänomischen Datenbanken, bei dem aus Termen einer relationalen Datenbank eine lokale Ontologie definiert wird und ihre Konzepte auf solche von öffentlichen Ontologien abgebildet werden. Im Vergleich dazu werden in diesem Ansatz sowohl Semantik als auch Daten integriert, um in domänenspezifischen Anwendungen landwirtschaftliche Fragen zu beantworten. Außerhalb der Landwirtschaft zeigt [Pe16] für die Domänen Montage und Holzverarbeitung, dass semantische Prozessbeschreibungen auf Basis von Ontologien über Produkt- und Produktionsdaten hinweg eine automatische Generierung von ausführbaren Roboterprogrammen ermöglichen.

2 Methodik

Eine einheitliche Wissensbasis für landwirtschaftliches Fachwissen sollte nicht nur auf Basis von (semi-)strukturierten Rohdaten aufgebaut werden, sondern auch die ihnen zugrundeliegende Bedeutung, d. h. Semantik, formal und strukturiert umfassen. Hierfür eignen sich Semantic-Web-Technologien wie OWL (Web Ontology Language), da sie ein formales, menschen- und maschinenlesbares Vokabular bieten. Dies erlaubt einerseits eindeutige Definitionen in sogenannten OWL-Ontologien und ermöglicht andererseits logische Inferenzen, um aus bestehendem (faktischem) Wissen neues (abgeleitetes) Wissen zu erschließen [Hi12]. In einem einfachen Beispiel aus der Landwirtschaft ist Winterweichweizen ein Weizen und somit, wenn entsprechend modelliert, auch automatisch, d. h. implizit, ein Getreide. An diesem Beispiel wird bereits ein Vorteil dieser Art der semantischen Wissensmodellierung deutlich: Es lassen sich Hierarchien formulieren und gewisse Regelungen für Oberklassen (hier Getreide) definieren, welche so auch automatisch für Individuen (Objekte) von Unterklassen gelten (hier Winterweichweizen).

In OWL-Ontologien werden einzelne Konzepte mittels global eindeutiger Identifikatoren bezeichnet, sogenannten IRIs (Internationalized Resource Identifier), welche für Internetadressen üblichen URLs ähneln [SR14]. Dadurch können Definitionen auch außerhalb einer Ontologie oder innerhalb anderer Ontologien von Menschen und Maschinen referenziert werden. Dies erleichtert die Vernetzung, Vereinigung und Wiederverwendung von internem und öffentlichem Wissen.

In diesem Ansatz wird heterogenes, landwirtschaftliches Fachwissen aus verschiedenen Datenquellen wie bspw. dem Leitfaden der Düngung für Acker- und Grünland [We18] semantisch in einer einheitlichen Wissensbasis modelliert, um die Zusammenhänge in der vereinten Wissensmenge für intelligente Abfragen aufzuzeigen. Bei den Datenquellen handelt es sich in der Regel um primär für menschliche Leser gedachte Formate, wie PDF-Dokumente oder HTML-Webseiten, und um (semi-)strukturierte Daten, wie CSV-Tabellen, relationale Datenbanken (SQL, Microsoft Access) oder RESTful APIs. Aus diesen Datenquellen wird das Fachwissen, d. h. Semantik und Daten, modelliert bzw. importiert und soweit zweckmäßig als OWL-Ontologien gespeichert, um in einer Abfrage mit nur einer Abfragesprache direkt auf vielfältiges Fachwissen zugreifen zu können.

Die gemeinsame Verwendung von verschiedenen und meistens nicht als OWL-Ontologien verfügbaren Datenquellen wird oft schon durch leichte Unterschiede beim Gebrauch von fachlichen Begriffen, Definitionen und Einheiten sowie technischen Identifikatoren und Schlüsseln erschwert. Daher werden die impliziten Strukturen und Bedeutungen aus diesen Datenquellen in gemeinsamen OWL-Ontologien explizit definiert. Bei äquivalenter Semantik können einheitliche Klassen und Eigenschaften (wieder)verwendet werden und bei abweichenden Bedeutungen können diese fallweise umgewandelt oder ausdrücklich als eigenständig bzw. verwandt modelliert werden. Dies bildet einerseits die Verbindung zwischen Wissen aus verschiedenen Datenquellen ab

und macht andererseits das Fachwissen leichter für Menschen und Maschinen zugänglich.

Das Schema des Entscheidungshilfesystems wird in Abb. 1 dargestellt. Die Datenquellen bzw. Quelldokumente wie bspw. das Gelbe Heft der LfL fließen mit ein bei der Modellierung der Strukturen und der Semantik in den Ontologien für landwirtschaftliches Allgemeinwissen (Upper Ontology farmexpert) und in den Ontologien für die einzelnen Teilgebiete/Domänen (z. B. fieldrecord, fruit, weather). Letztere enthalten Definitionen von Klassen, Arten von Eigenschaften und häufige Begriffe. Zusätzlich werden die Quellen in maschinenverständlichen Datenontologien (z. B. bsv-seed-2018, bvl-pesticide-2018) abgebildet, welche die Rohdatenwerte enthalten und durch die Domänenontologien interpretiert werden können. In den unteren Anwendungsontologien (Betriebs-Ontologie mueller-gbr) werden bspw. Betriebsdaten und Feldmaßnahmen semantisch gespeichert und stehen so gemeinsam mit dem Fachwissen den Abfragen von Anwendungen wie etwa einer benutzerfreundlichen Web-GUI zur Verfügung.

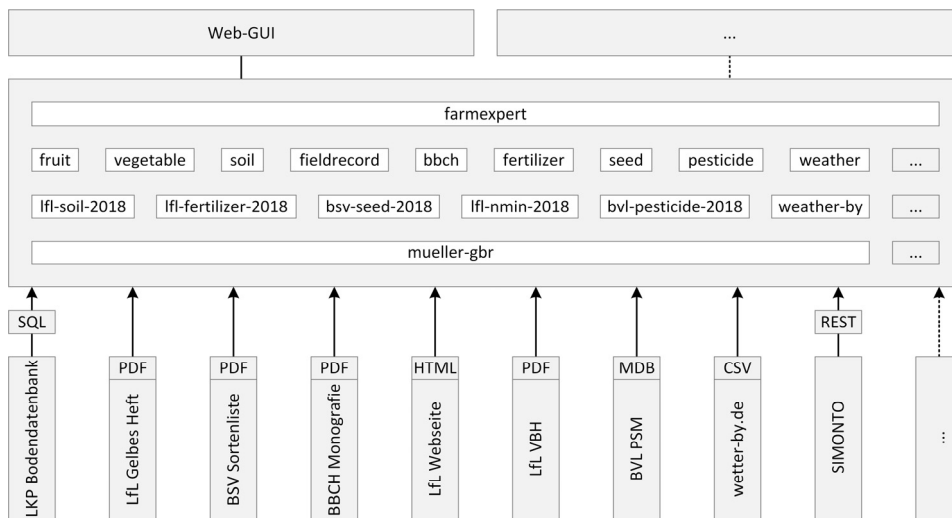


Abb. 1: Schematischer Aufbau des Entscheidungshilfesystems; von unten nach oben: Datenquellen (mit Formaten bzw. Schnittstellen), Wissensbasis (bestehend aus Ontologien zu Anwendungen, Daten, Domänen und Allgemeinwissen) und Anwendungen

Einfache, häufige Fragen von Landwirten können mit der SQL-ähnlichen Abfragesprache SPARQL (SPARQL Protocol and RDF Query Language) formuliert und so weitgehend automatisch und dynamisch von der Wissensbasis beantwortet werden. Dabei wird auch die Intelligenz des Systems nicht in Programmcode festgelegt, sondern soweit möglich als Fachwissen und Logikregeln in den Ontologien oder andernfalls innerhalb der SPARQL-Abfragen modelliert.

Das SPARQL-Beispiel in Abb. 2 zeigt eine Abfrage zum Fachwissen einer Weizensorte und der zu erwartenden Qualitätsgruppe, die in verschiedenen Ontologien modelliert sind. Hierbei wird nicht nur eine statische Tabelle mit den ertragsbedingten Zu- und Abschlagskoeffizienten angezeigt. Stattdessen wird bereits der individuell zu erwartende Ertrag mit dem Referenzertrag verrechnet, um den Stickstoffbedarfswert, i. d. F. für die gesetzliche Düngebedarfsermittlung, anzuzeigen.

```
SELECT ?seedVariety ?cropType ?yieldInDtPerHa ?nitrogenRequirementInKgPerHa
WHERE {
  VALUES (?seedVariety ?yieldInDtPerHa) { (seed:JBAsanoWinterSoftWheat 75.0) }
  ?seedVariety fex:growsTo ?cropType .
  ?cropType fex:hasData [
    fex:hasRequiredNitrogenInKgPerHa ?rNInKgPerHa ;
    fex:hasYieldInDtPerHa ?yieldLevelInDtPerHa ;
    fex:hasYieldDifferenceInDtPerHa ?yDInDtPerHa ;
    fex:hasPerYieldDifferenceAdditionOfNitrogenInKgPerHa ?perYDAdditionInKgPerHa ;
    fex:hasPerYieldDifferenceReductionOfNitrogenInKgPerHa ?perYDReductionInKgPerHa ] .
  BIND (?yieldInDtPerHa - ?yieldLevelInDtPerHa AS ?yieldDeltaInDtPerHa)
  BIND (IF (0.0 < ?yieldDeltaInDtPerHa, ?perYDAdditionInKgPerHa, ?perYDReductionInKgPerHa) AS ?perYieldFactorInKgPerHa)
  BIND (?rNInKgPerHa + ?yieldDeltaInDtPerHa / ?yDInDtPerHa * ?perYieldFactorInKgPerHa AS ?nitrogenRequirementInKgPerHa)
}
```

	seedVariety ↕	cropType ↕	yieldInDtPerHa ↕	nitrogenRequirementInKgPerHa ↕
1	seed:JBAsanoWinterSoftWheat	fruit:WinterSoftWheatVarietyA	"75.0"^^xsd:decimal	"222.50"^^xsd:decimal

Abb. 2: Die OWL-Ontologien wurden in einer RDF-Datenbank-Software gespeichert und mit einer SPARQL-Abfrage (oben) gleichzeitig ausgelesen und ausgewertet für das Ergebnis (unten)

3 Fazit und Ausblick

Landwirtschaftliches Wissen aus technisch und fachlich heterogenen Datenquellen wurde in einer semantischen Wissensbasis vereint, um es in einer übergreifenden Abfrage abrufen und auswerten zu können. Dies bietet neue Möglichkeiten, um Landwirte und Berater beim Nachschlagen von Fachwissen und Betriebsdaten sowie bei der Entscheidungsfindung dynamisch zu unterstützen.

Es ist geplant, bei der Integration von bisherigen und zukünftigen Datenquellen verstärkt maschinelle Schnittstellen und formale Abbildungen einzusetzen, um größere Datenmengen zu importieren. Dadurch könnte der Abdeckungsgrad an Hintergrundwissen für häufige landwirtschaftliche Fragen wesentlich erhöht und flexibler auf (jährliche) Änderungen bei fachlichen und gesetzlichen Gegebenheiten reagiert werden.

Ein weiterer, geplanter Fokus ist der maschinelle Live-Zugriff auf bestehende Datenbanken und Webservices. So könnten aktuelle Daten wie bspw. Wettervorhersagen bei Entscheidungsfindungen miteinfließen und die Wissensbasis leichter parallel und in Zusammenarbeit mit bestehenden landwirtschaftlichen Software-Systemen genutzt werden. Ein weiteres Ziel dabei ist, Landwirte langfristig zu entlasten, indem Betriebsdaten und Feldmaßnahmen nur einmalig direkt bei der Ausführung eingegeben und vom gleichen oder von anderen Systemen, bei Einwilligung des Landwirts zur Weitergabe dieser sen-

siblen Daten, mehrfach automatisch wiederverwendet werden. Beispielsweise könnte zum Jahresende eine Schlagdokumentation oder zum Jahresbeginn eine Düngebedarfs-ermittlung automatisch ausgefüllt werden, sodass der Landwirt die vorausgefüllten Eingabefelder nur noch überprüfen muss.

Für Endanwender, d. h. Landwirte und Berater, ist die Anbindung des Entscheidungshilfesystems an bestehende oder neuartige, intuitive Benutzeroberflächen entscheidend. Nur so kann eine hohe Akzeptanz bei großen Benutzergruppen erreicht werden.

Ein Zugriff auf weitere, noch verschlossene Datenquellen würde eine größere Reichweite und Tiefe des modellierten landwirtschaftlichen Fachwissens ermöglichen. Das Entscheidungshilfesystem könnte dadurch bei komplexeren Fragestellungen weitere Einflussfaktoren berücksichtigen und eine bessere Unterstützung liefern.

Dieser Beitrag entstand im Kontext eines durch das Bayerische Staatsministerium für Ernährung, Landwirtschaft und Forsten geförderten Projekts.

Literaturverzeichnis

- [Hi12] Hitzler, P. et.al.: OWL 2 Web Ontology Language Primer (Second Edition). W3C, 2012, <https://www.w3.org/TR/owl2-primer>, Stand: 08.11.2018.
- [Ma15] Martini, D. et.al.: KTBL-Planungsdaten auf dem Weg in die Zukunft - Bereitstellung über Linked Open Data. In (Ruckelshausen, A. et.al. Hrsg.): Informatik in der Land-, Forst- und Ernährungswirtschaft, Geisenheim, S. 105-108, 2015.
- [Ma18] Martini, D.: Webservices auf heterogenen Datenbeständen – Methoden der Umsetzung am Beispiel der KTBL-Planungsdaten. In (Ruckelshausen, A. et.al. Hrsg.): Informatik in der Land-, Forst- und Ernährungswirtschaft, Kiel, S. 155-158, 2018.
- [NBF18] Nafissi, A.; Bruns, B.; Fiorani, F.: Ontologies for resolving semantic heterogeneity in information integration among plant phenomics databases. In (Ruckelshausen, A. et.al. Hrsg.): Informatik in der Land-, Forst- und Ernährungswirts., Kiel, S. 167-170, 2018.
- [Pe16] Perzylo, A. et.al.: Intuitive Instruction of Industrial Robots: Semantic Process Descriptions for Small Lot Production. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Deajeon, S. 2293-2300, 2016.
- [PT17] Paustian, M.; Theuvsen, L: Adoption of precision agriculture technologies by German crop farmers. Precision Agriculture 18, S. 701-716, 2017.
- [RG17] Roosen, J.; Groß, S.: Digitalisierung in Land- und Ernährungswirtschaft. Vereinigung der Bayerischen Wirtschaft e.V., München, S. 27-30, 2017.
- [SR14] Schreiber, G.; Raimond, Y.: RDF 1.1 Primer. W3C, 2014, <https://www.w3.org/TR/rdf11-primer>, Stand: 08.11.2018.
- [Wa17] Wahmhoff, W.: Einführung Prozessmanagement pflanzlicher Produkte, S. 137, 2017, <https://www.dbu.de/media/061217051349d3ga.pdf>, Stand: 25.10.2018.
- [We18] Wendland, M. et.al.: Leitfaden für die Düngung von Acker- und Grünland. Bayerische Landesanstalt für Landwirtschaft (LfL), Freising, 2018.