Reference Data and APIs

R. Andrés Ferreyra, Ph.D., Ag Connections, LLC

To be presented at InfoAg 2016
in St Louis, MO, August 3, 2016

Introduction

- Precision agriculture has shown great promise in delivering the ability to convert vast quantities of data collected by farm equipment into actionable information in Farm Management Information Systems (FMIS) and other tools. However, it has not quite delivered on this promise, in part because different manufacturers' hardware and software are not interoperable.

- AgGateway's SPADE and PAIL projects have targeted this problem, modeling business processes, identifying the data needed to support them, and working with manufacturers and software companies to develop a toolkit, ADAPT, that will help convert seamlessly between various manufacturers' data formats and an industry standard common data model.

- A critical component of this system is a common way to identify products such as seed, chemicals and equipment, so when growers share data about crop protection applications (or other field operations) with partners (such as customers, bankers and custom applicators) the recipients of the data understand exactly what is meant.

Introduction, cont.

- The SPADE project developed a standardized, distributed mechanism to source this "Reference Data".

  - Standardized: We propose a single mechanism and format for industry partners to source data, and

  - Distributed, because each manufacturer can host its own reference data through a delivery tool ("API") themselves, and third parties can do so as well.

- This presentation focuses on three problems being addressed by the AgGateway Reference Data system:

  - Identification (we need common identifiers to convey common meaning across the industry),

  - Sourcing (How can we source those common identifiers?), and

  - Access (How can a farm management information system find those sources?)

Core Documents and their Relationships
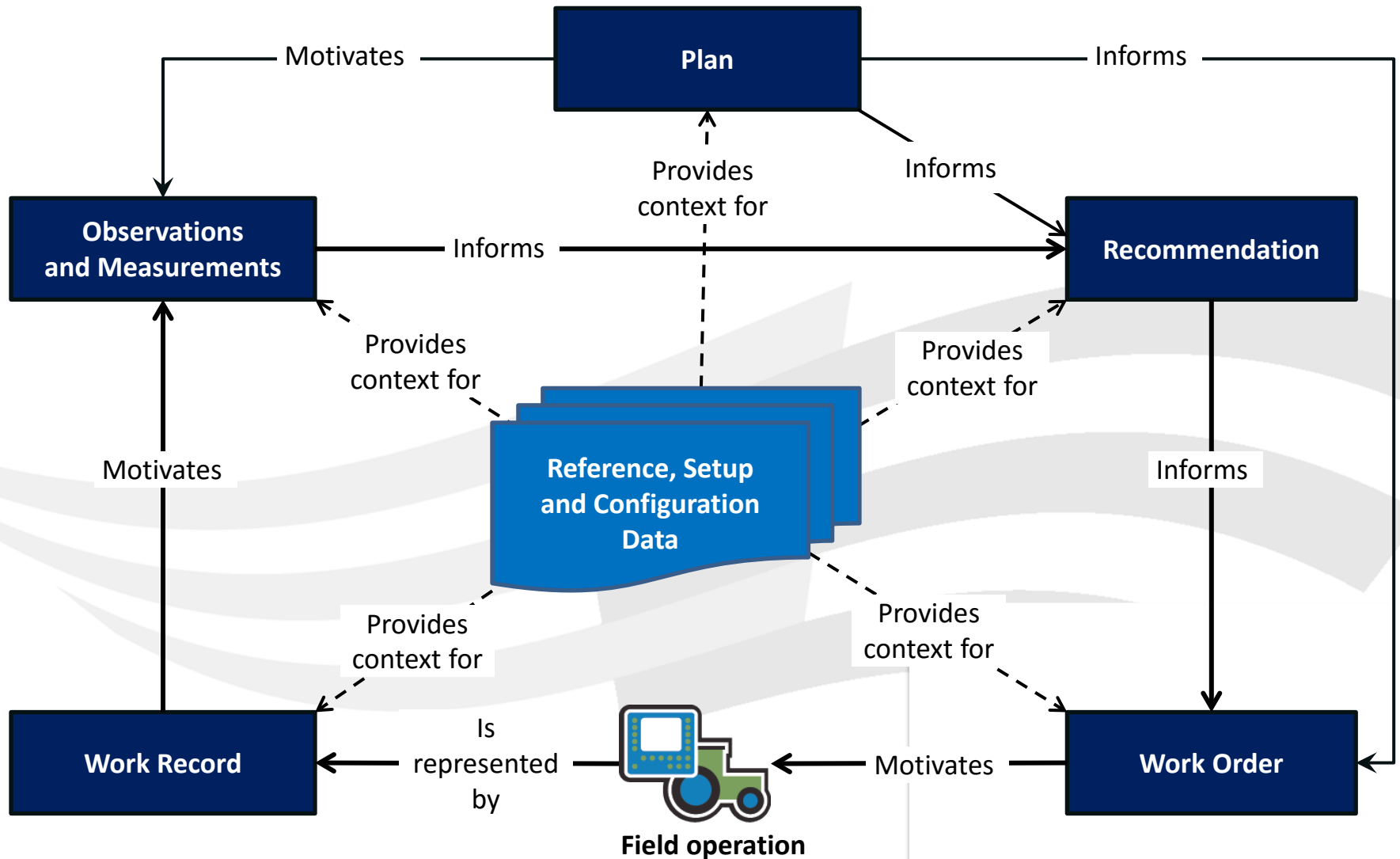
What is "Reference Data"?

- Information needed to **unambiguously identify a product**, and to enable its use in a farm management information system (FMIS) or a field operation.

- It changes rarely, and a great part of its value lies in giving access to it to the different actors in the business process.

- Think of this as a set of controlled vocabularies (like pick list items) that can be shared across the industry.

What isn't Reference Data?

Reference Data

**All instances of a thing
("The ACME MaxSuperTron 200")**

GENERAL

Grower (Setup) Data

**One particular instance of a thing,
INDEPENDENT of its state:
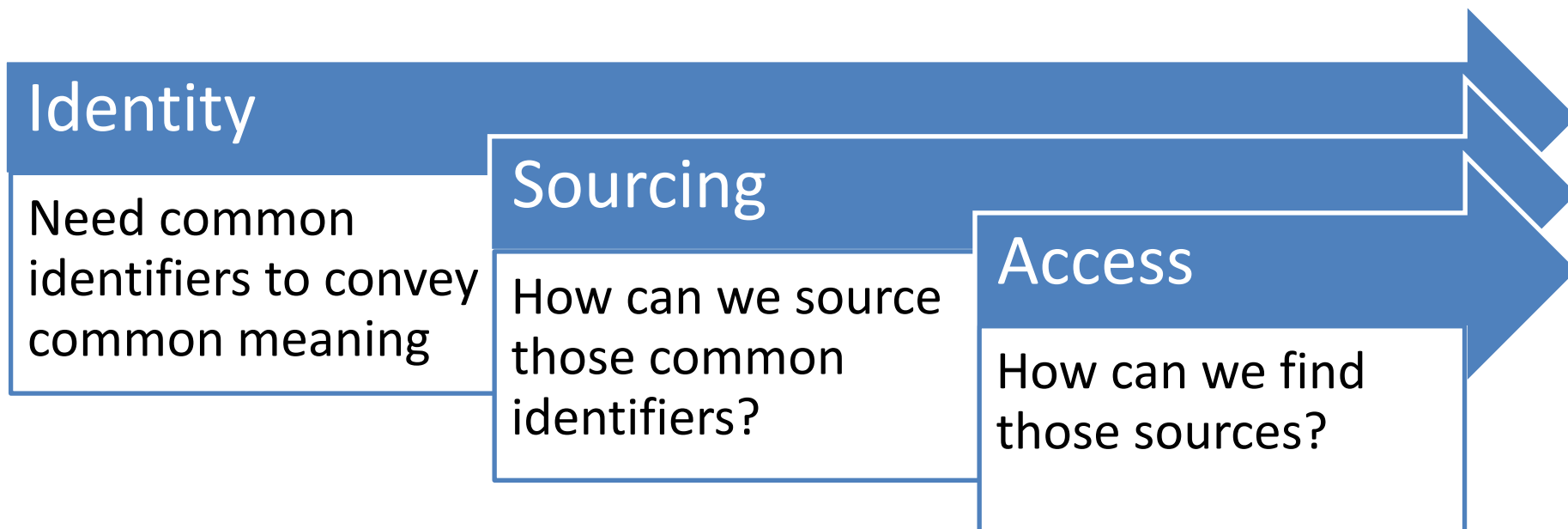("The MST200, serial #12345")**

Configuration (Setup) Data

**One particular instance of a thing,
in the context of its current state:
("MST200, #12345, now installed
at Lat,Lon X,Y, using Widget Z")**

SPECIFIC

Problems to solve

Identity

Need common identifiers to convey common meaning

Sourcing

How can we source those common identifiers?

Access

How can we find those sources?

AgGateway

Problem 1: Identity "What product was used?")

- How do systems communicate what product should be used, and was actually used, in a given operation?

Sends ⭐

Has ⭐

⭐, 🔺, 🔷, represent the same product!

Already has 🔺
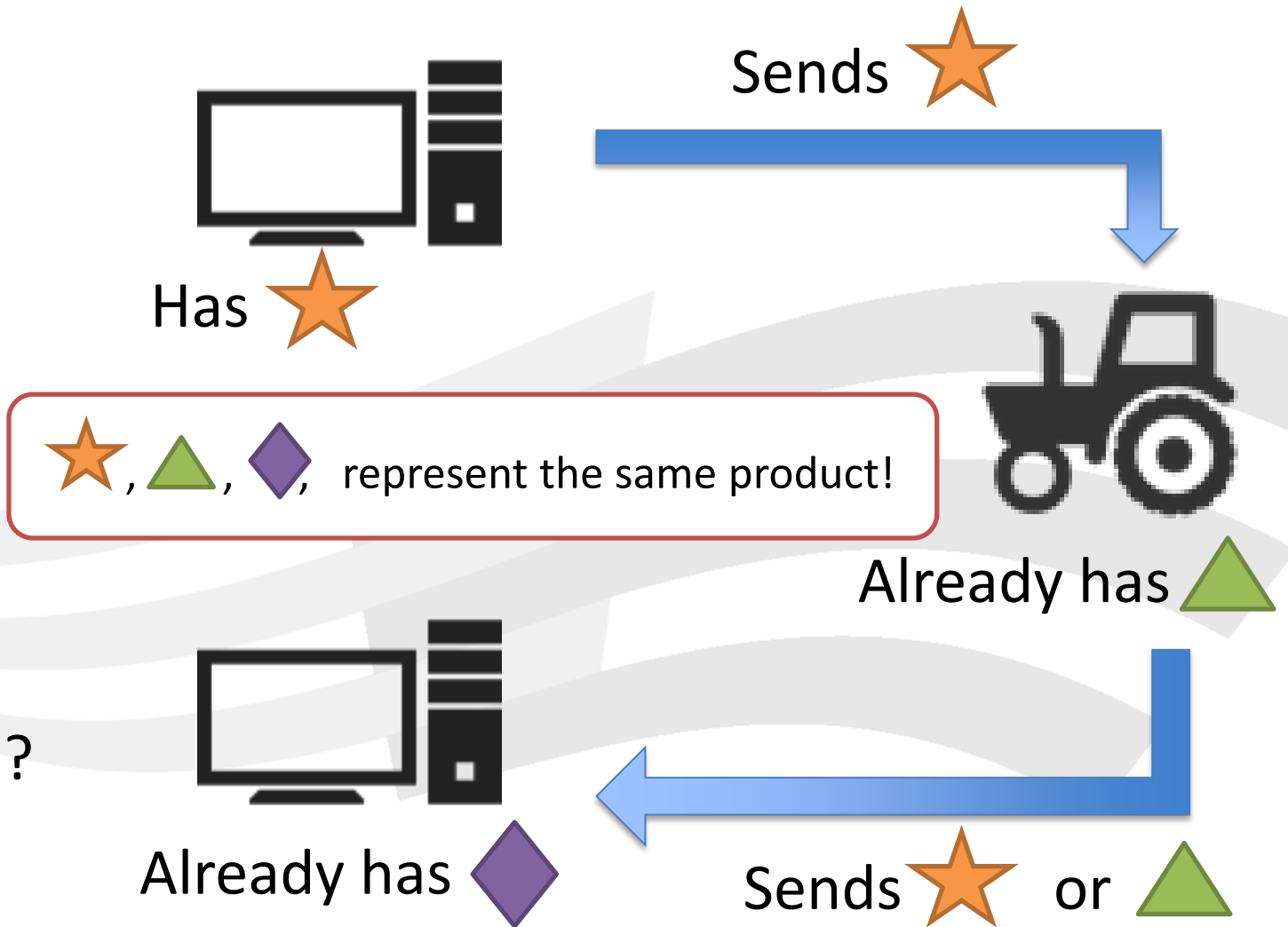
- What gets recorded in the example?

⭐? 🔺? 🔷?

Already has 🔷

Sends ⭐ or 🔺

Common Identifiers

- Manufacturers (or third-party data providers) define a single identifier for a given product or equipment and make it available along with the product's other information.

- Communication using this identifier will be consistent and error-free.

- Any system wanting more information about a given product can use this identifier to retrieve it from the appropriate reference data API implementation

- Mappings between the unique identifier and other, proprietary identifiers that have the same meaning can be accommodated.

Problem 2: Sourcing

- Users usually have to manually enter or look up label information critical to the proper and safe usage of a product.

- This can result in:
  - Frustration when trying to find information, and
  - Errors in the data entry

- The user needs **easy access** to this information, and farm management software providers want that access in a **consistent** manner to reduce complexity and redundancy.

Solution: Reference Data APIs (RDAPIs)

- A standard defining:
  - A system for finding products and
  - The information about each product that can be communicated
- A producer or a third-party host implements the API and through it provides accurate, up-to-date information about the products to the consumer

What's an API?

"Application Programming Interface," a software mechanism to deliver data from a web server to users. To date, the SPADE projects have defined APIs to use in delivering seed, crop protection, and equipment

AgGateway

RDAPIs make it easier for everyone

- Growers planning to use the products and agronomists developing recommendations for them can do so with the latest and most accurate information available, all within their existing farm management software

- Management software providers are able to use a common specification to obtain accurate, machine-readable label information for products

Product API

- An API was designed for crop inputs, such as crop varieties and crop protection products
- Allows searching for products by various attributes
  - Brand
  - Crop
  - Manufacturer
  - Crop protection type (e.g. Insecticide, Herbicide)
  - Trait
- Has a sophisticated data model, compatible with ADAPT, that enables representing complex products that include multiple, already-labeled products, that combine seeds and crop protection and/or crop nutrition ingredients, etc.

Equipment API

- An API was designed for equipment products, such as tractors, combines and planters

- Allows searching for equipment by various attributes
  - Brand
  - Type (e.g. tractor, combine)
  - Manufacturer
  - Model
  - Series

- Is currently being extended to include sensors and irrigation systems.
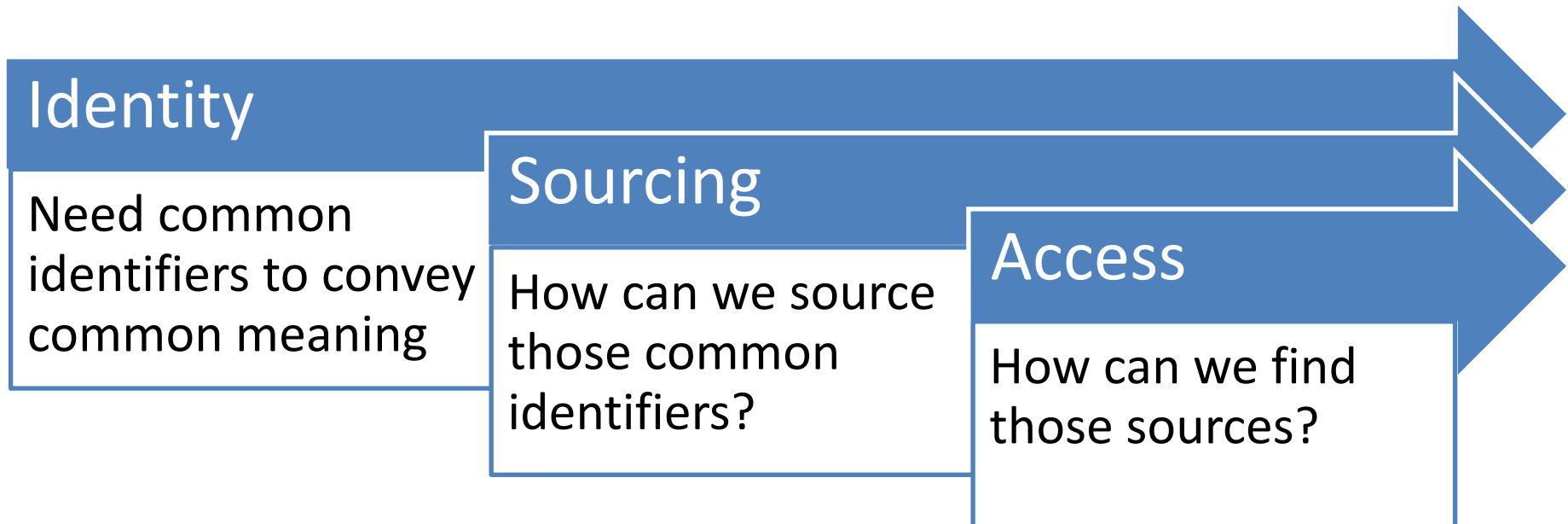
AgGateway

Problem 3: Access

- How does a system wanting information about a product find the appropriate API for that product?

- Since the system is distributed, how do we avoid every management software provider having to maintain a list of API hosts, and being forced to release updates to include new hosts?

Directory service

- A central, searchable location containing resource locators for all the API hosts and a description of their content

- Initial entry of a host must be done manually, but discovery of content can be automated by scraping the various APIs and storing the results.

- Allows searching for APIs by various attributes
  - For crop products: seed variety vs crop protection, trait, moa, chemical class, brand
  - For equipment: manufacturer, brand, series, type

AgGateway

Problems, revisited

Identity

Need common identifiers to convey common meaning

Sourcing

How can we source those common identifiers?

Access

How can we find those sources?

Provide feedback!

- The current API and model designs are not set in stone and could use feedback from experts.
  - Are the discovery points presented sufficient for users to find your products you source or use?
  - Do the properties presently in the models properly reflect the product information you need or source?
- Your input will help us ensure the solution meets the needs of all industry, and facilitate adoption.

AgGateway

SPADE Reference Data Team

- Arren Mund – amund@agleader.com

- Andres Ferreyra - andres.ferreyra@agconnections.com

- Stuart Rhea - stuart.rhea@agconnections.com

- Ben Craker - Ben.Craker@agcocorp.com

- Jim Wilson - jim.wilson@aggateway.org

- Sharon Lewis - sharon.lewis@syngenta.com

Questions?

- Resources
  - https://www.youtube.com/watch?v=u_sJx75VDkk&feature=youtu.be

  - https://aggateway.atlassian.net/wiki/download/attachments/66486443/RDAPI_overview.pptx?version=1&modificationDate=1458763395850&api=v2