JOURNAL OF THE NACAA

ISSN 2158-9429

VOLUME 17, ISSUE 1 – JUNE, 2024

Editor: Linda Chalker-Scott

Witkowski, N.[1], Buckmaster, D.[2], Basir, M.[3], Sakata, H.[4]

[1]Extension Educator, ANR, Purdue Extension-Porter County, Valparaiso, Indiana, 46383
[2]Professor of Ag and Biological Engineering, Dean's Fellow for Digital Agriculture, Purdue University, West Lafayette, Indiana, 47907
[3]Graduate Research Assistant, Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana, 47907
[4]Associate Director, Purdue University Online, West Lafayette, Indiana, 47907

Practical Database Solutions for Digital Agriculture: Simple and Private

Abstract

Quality farm data is crucial for decision support but farmers often record data incompletely and inconsistently. Records needed as contextual metadata to improve enterprise analysis can also facilitate logistics and tactical decisions and feed more advanced simulation or digital twin models. We offer several well-structured open-source templates as an on-ramp to private databases for agriculturalists with modest spreadsheet skills. These Airtable databases use simple data-validated forms and have the look and feel of customized apps to yield operational data that is tidy, machine- and human-readable, editable, and exportable for analysis in other software. This manuscript includes resources explaining, in detail, how to build databases for activity records on farms, in businesses, and in other settings. These resources should facilitate the infusion of digital agriculture principles through Extension and structured educational programming.

Abbreviations: AI – Artificial Intelligence; FAIR – findable, accessible, interoperable, and reusable; FMIS – farm management information systems; OEM – original equipment manufacturer; UAV – unmanned aerial vehicle

Keywords: activity, context, data, database, digital agriculture, metadata, operations, records, spreadsheets.

Introduction

Farmers vary tremendously in their commitment and ability to keep good records. The idea of records and notes, often done in notebooks, dates to at least the mid-1800s (Joly, 2011). Unfortunately, these notes are often cryptic, misplaced, damaged and remain unused. If these records are in a physical form, such as in notepads, sticky notes, receipts, etc., they are often in the wrong place or tossed into a file, making them practically useless. And yet, even if the records are digital (like text messages or photos), the lack of structure can render them less useful simply from a search-and-find perspective. As Pagano et al. (2013) illustrated, data is collected for a plethora of purposes via many different approaches and this makes it difficult to interpret in context. Whether big or small, data integration and exchange is an ongoing challenge. Progress will require that records be kept digitally. Additionally, storing records in the cloud logically and tidily immediately solves certain access and consistency issues and ensures that the data is FAIR (findable, accessible, interoperable, reusable). More importantly, keeping records in interoperable digital formats can also enable mining for insights and analysis in simplistic (sorting or filtering) or complex (multifactor analysis on yields, harvest date projections, etc.) manners because the data exchange works.

Farmers and agricultural advisers frequently use spreadsheets they designed or acquired from public and land grant institutions or consultants. These spreadsheets facilitate decision-making by reducing the tedium of what-if analyses and serving as a simple means of aggregating smaller entities within the farm (fields and pens) into a representation of the whole. A common weak link in using these decision aids, however,

is knowing the values for inputs to these calculations. Farm management information systems (FMIS), original equipment manufacturer (OEM) platforms, and other cloud services are making strides toward interoperability (AgGateway, 2021); yet gaps remain as detailed information that only a human knows is often overlooked, such as the exact content of the insecticide bin during planting, to give one example. To fill this void, we can leverage even limited experience with spreadsheets to introduce databases. Records populated during an operation could produce high-quality data in a structured and tidy format for use in those decision aids. Increasingly, we also use agricultural system models to improve management by estimating yield and sustainability metrics at various scales (Antle et al., 2017). Those models have data needs that can only be met with forward-looking contextual metadata collection. Antle et al. (2017) specifically call out the need for private site and farm-specific input use and outcome data for farm-level management decisions.

Although spreadsheets can be set up as databases, database software has capabilities that dramatically increase the capacity to collect and synthesize data. Google Forms, for example, can be used for data entry into a Google Sheet (Hickson, 2020). This is a workable solution in some instances, but it requires pseudo-coding and planning to ensure the data is ultimately usable and complete. Airtable (Airtable 2022; Meijer, 2022) offers similar capacities with improved data validation for form completion and the look and feel of customized apps for mobile devices. With the rapid advancement in artificial intelligence, we are also likely to see more adaptive feedback applied to data entry (Chen et al., 2010). This approach will further improve accuracy, but it requires task and context specific adaptations and sophisticated app development. A solid starting point is to use simple form-based data collection which can already use simple data validation to employ specific adaptations that Chen et al. (2010) tested such as setting likely defaults, re-ordering of options, automatic warnings for invalid entries.

Database applications are well beyond activity or event records; but our focus here is on activity records because of their importance in informing and facilitating logistical and strategic decisions. In some applications, these data will provide the fuller contextual metadata needed to drive artificial intelligence and machine learning algorithms, run

simulation models (Antle et al., 2017), or support economic analyses. With the overall goal of providing a simple primer to databases and a logical entry point for farmers and advisers to digitize records effectively, our specific objectives for this work were to:

1. Explain basic database features so Extension Educators and Trainers can extend this knowledge.
2. Illustrate the value of well-structured, tidy, and open data (i.e., the simplicity and interoperability with spreadsheets or other software).
3. Introduce prepared Airtable templates and related resources.
4. Cite examples of private databases for events and activity type data that were developed by others.

Farm and Business Examples of Data Needs

Budgeting

Suppose you were evaluating the tradeoffs of re-establishing an alfalfa field or the cost/benefit of a previous fertilization program on your alfalfa. A partial budget analysis would require a tally of all inputs. The inputs used or dispensed each pass over the field, including information about the machinery, labor, fuel, etc. must be known. Memory would not suffice to have a complete tally. But if there was a record for each time the field was covered, including date, time, equipment used, material applied, or yield removed, all information would be compiled and in one place.

Scheduling

A specialty crop farm operating a community-supported agriculture operation with various vegetables has complex management tasks regarding planting dates that lead to an anticipated crop availability schedule. Tracking plant status and projecting dates of harvest and yields could be facilitated with simple data recorded at key events and occasional scouting reports.

Nutrient management

Whether required by law or not, it seems fitting to have manure application records across all fields as part of a nutrient management plan. This is information that the

fertilizer advising agronomist would want to know. A tally sheet in the cab would suffice, but a digital record that is preserved and secured would be better.

Evaluation

If a farm manager desires to perform true profitability analysis on a per-management zone basis or for a couple of similar fields that were managed differently (in terms of tillage, variety, planting and harvest dates, seed rates, fertilizer date/rate, weed control methods, etc.), they would need to have complete records. Having complete records on hand at analysis time could improve insights; for instance, were the differences observed due to dates and weather or to technology or management choice?

Driving AI

Tech-savvy agronomy consultants want to apply AI to UAV imagery of fields, crop yields, and soil data to provide improved recommendations. This AI needs information about planting dates, fertilizer rates and dates, irrigation, and products used for weed, disease, and pest control. Most farms have yet to capture all this information in a FAIR format.

Vehicle records

A farm or firm with a fleet of vehicles and/or implements will have repair and maintenance events with associated expenses. Optimizing the decisions to refurbish, replace, or rent requires better data. That data is also useful for tracking/scheduling maintenance, documenting equipment or repairs under warranty, and completing tax forms. A simple database that tracks these events for each vehicle will yield multiple benefits.

Livestock treatments

Whether administered by a farm employee or veterinarian, livestock treatments should be recorded. A simple database could track deworming, growth implants, antibiotic injections, etc., based on individual animals or groups. Such records are important for withholding periods, analyzing efficacy, and tracking expenses.

Database Basics

Spreadsheets such as Microsoft Excel or Google Sheets are organized as workbooks with sheets that can be renamed, cross-referenced, etc. Individual cells (the matrix of intersected rows and columns) are filled with numbers, text, and insightful charts. Airtable looks and feels like a spreadsheet, but it is a base with interactive tables. It can have formulas, but we introduce it here as a holder of data. Because of the similarity to spreadsheets, it is a good entry point for novices. It is platform-agnostic and can be used in-browser or as a native app on Android or iOS devices. Airtable has a free tier that should suffice in most farm and small business situations. If a fee-based tier is needed, it is still very reasonable when you consider the relative costs of business operations.

Airtable is cloud-based. It can be used offline; however, online usage brings practical benefits of real-time saving and sharing. “Sharing” has assorted levels of security with permission levels of *owner/creator*, *editor*, *commenter*, or *read-only*. Using the form view of a database table, you can configure the database to be populated or informed by workers or affiliates who do not have access to it. This separation of the ability to input data (i.e., record an event or activity) from the ability to access the data could be very important when dealing with financial items. This separation also provides a degree of data security. Not everyone contributing information or populating records needs to understand databases; they need to know how to answer questions and fill out a form.

The templates introduced below employ data validation and are exemplars or reference implementations that illustrate good practices to keep your data tidy and structured. The preset lists of options (such as names of employees or equipment) enable users to choose from drop-down lists rather than typing of freeform values. Benefits are:

- fewer errors (only valid responses and no misspellings),
- improved consistency (for instance, not having different users typing in “Bill” or “William” to indicate the same person), and
- less tedium (being able to choose “24-row planter” from the list rather than typing it out).

With automatic sorting, the most recently used option can appear at the top of the list the next time the form is completed. For example, if you were spraying yesterday and you go to a field to perform another activity today, it is highly likely that you are spraying again. The templates also apply conditional data requests to simplify the forms by requesting only the data pertinent to the operation. For example, seed rate and variety are only requested if planting is the operation.

Using Well-Structured Tidy Data

Airtable data will be organized as tabulated tidy data (Neo, 2020). “Tidy” means each column is a different variable, each row is an observation, and each cell is an individual value. Table 1 illustrates tidy data from the Horticultural Crop Activity Records (Buckmaster, 2023a). Values can be of several types, such as integers, real numbers, text, hyperlinks, or complete documents (e.g., photo or other file types). Tidy data facilitates sorting, filtering, and interpretation. It also facilitates export and usage in other software (such as exporting to a spreadsheet to do pivot tables).

Table 1. Sample subset of tabular data (grid view) exported from the horticultural crop activity records database template as CSV into Microsoft Excel.

Who	Where	What	Duration	Notes	created time	Power Unit	Implement(s)	Seeds planted	Seeding Rate (seeds/ac	Products applie	Fertilizers applie	Fertilizer Rate (lb/ac
Purdue Pete	Bed 72	Tillage	40	left disc needs adjustment	12/20/2022 11:35am	Tractor 2 JD X120	bed shaper					
Suzie Jones	Bed 72	Plant/Transplant			12/20/2022 11:37am	Utility tractor	water wheel transplanter	onions - candy				
Suzie Jones	Bed 72	Harvest	30	bundles of rhubarb	12/20/2022 11:50am	human powered						
Purdue Pete	Field 1	Spread/Spray	120		12/20/2022 11:39am	Gator	150 gal sprayer			Glyphosate		
Suzie Jones	Field 1	Plant/Transplant		burn down looked effective	12/20/2022 11:41am	Tractor 2 JD X120	seed planter	corn - sweet - 82 day	30000		9-18-9 starter	50
Purdue Pete	Field 1	Scout		popcorn is near ready	12/20/2022 11:42am							
Purdue Pete	Zone D	Scout		all looks great	12/20/2022 12:30pm							

Templates or Reference Implementations – a Good Place to Start

Wiginton (2022) identified excellent Airtable examples and used cases as diverse as accounting, customer relationship management, marketing, fitness tracking, timesheets, and live dashboards.

We have released templates or reference implementations of Airtable databases specific to production agriculture. ***Digital Field Records*** (Buckmaster, 2022) is an

Airtable template that enables farmers or researchers to collect digital records of “what happened here?” in wide-acre fields and forage crop operations. As the backstory to give a frame of reference for interpreting yields, labor or machinery efficiencies, or other production aspects, the sequence of events leading to later circumstances is critical.

The digital field records template is released with video documentation and brief tutorials that fully explain how the database works, how to duplicate and customize it, and how to control access. This and the other templates enable data validation which simplifies data entry (Figure 1) but options for entering free-form text or uploading photos remain.

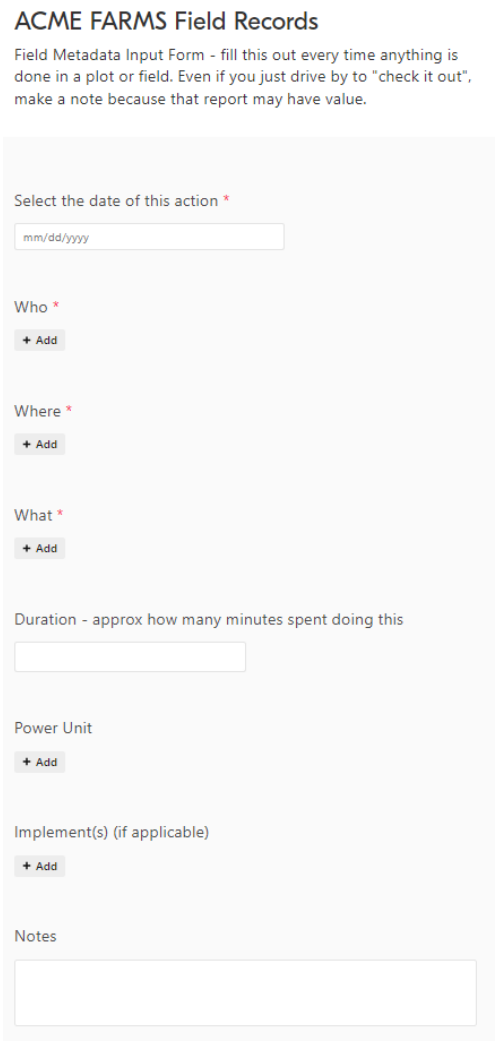

Figure 1. Sample Airtable form from the digital field records template. Some conditional data element questions to be posed based on the operation are not shown.

Users can quickly duplicate this free template and customize it (including permissions) for their operation. The records from this private database would be valuable in tracking season progress, documenting machinery usage and labor efficiency, supporting claims about the crops produced (e.g., organic, GMO-free), and informing enterprise budgets. These records can also round out information from FMIS or OEM platforms with necessary data regarding climate and carbon balance (Illinois Soybean Association, 2022).

A derivative of the digital field records, customized for ***horticultural or specialty crop operations***, was released with tutorial videos (Buckmaster, 2023a). It accommodates specialty crop operations of all types. It can accommodate areas as large as fields or in diverse production areas such as high tunnels, benches, zones, blocks, or beds. The resulting data could inform community-supported agriculture (CSA) schedules and inventories and improve seeding and transplanting management in the coming years.

The Food Safety Modernization Act (FDA Food Safety Modernization Act, 2011) requires that certain records be kept by firms that produce and process food. Those records relate to the preventative controls and hazards in growing, harvesting, packing, and storing produce. using The Airtable toolkit (Buckmaster et al., 2022) makes ***FSMA records*** easier and more complete for those unfamiliar with FSMA requirements or for whom it is simply tedious or confusing. Using this toolkit will ensure compliance with FSMA.

Grain marketing decisions throughout the growing season and after harvest can be complex. The evaluation of current marketing options to expectations in coming months must also consider previous commitments. Then, later, as deliveries toward contracts occur, tracking can be a challenge. The ***Digital Marketing and Delivery records*** template (Buckmaster and Soonthornsima, 2022) can track contracts and deliveries of commodities in simple tables; those tables can facilitate reconciliation, and if this data is always current, it can quickly inform decision-making. With such records completed as each contract or delivery is made, users will have fingertip access to information about

unmet commitments. They can evaluate marketing decisions about futures price risk and basis risk separately.

These templates and associated video series offer immediate value to users and can spawn ideas for digitizing other important records of events, activities, or decisions. In each case, the templates may be duplicated and customized.

Beyond Templates – Making Your Own

As noted above, activity or event databases have many applications for farms, firms, and families. Steps to generate these private databases are straightforward once one understands the structure of the data, understands concepts of data validation, conditional requests, and the specific operation of software to create and modify forms. For this reason, a set of five brief instructional videos that explain how to generate and use a customized activity record database in Airtable was recorded (Buckmaster, 2023b). While completing that tutorial series, users will generate a simple income and expense system. That web resource also points to a template anyone could duplicate to keep vehicle operating, maintenance, and repair expense records.

Practical Private Databases for Event and Activity Records

With very limited training as is contained in the aforementioned video series (Buckmaster 2023b), many farmers and farm workers have generated useful private databases for record keeping and event or activity tracking. Table 2 lists a few examples (of the authors' knowledge) with hopes that these are motivating to others who might want similar records. Some of these were also synced with automation to feed Google Sheets workbooks that contained pivot tables and charts to facilitate interpretation.

Table 2. Sampling of private databases generated and used by others.

Topic	Database Records Included
Livestock records	Feeding, vaccinations, medication, date
Farm income and expenses	With conditional data requests, was complete with date, fuel, feed, hardware, parts, tools, livestock purchases, hay, livestock sales, labor
Grain transport and inventory	Date, driver name, truck ID, commodity, field ID, bin ID, weight, moisture, destination
Spring planting records	Date, operator, tractor, implement, field, field size, corn/bean hybrids
Vegetable records and scouting	Expenses, income, yields, pest and disease log
Farm records	Date, chemical application, seeds, fertilizer, field, operator, crop, event, harvest, transport, prices, profit/loss by field and total
Field records	Operator, operation, field, equipment, implement, fertilizer, seed, micronutrients, manure, date, irrigation
Specialty crop deliveries	Date, bin count, crop, destination
Crop yield and yield calibration	Field, crop, calculation mode (yield or calibration), area, weight, test weight
UAV logbook	Flights, dates, batteries, weather conditions, purpose
Business service record	Customer ID, service done, parts used, time spent, follow-up notes
High tunnel and vegetable irrigation records	Date, equipment used, location, crop, fertilizer, duration

Conclusion

Private databases can be simple and useful introductory tools for farmers and others wanting to digitize activity or event-type records. This approach uses simple forms completed with data validation which simplifies data input and makes data more consistent. Data is updated in real time. Airtable database examples were introduced with pointers to online resources that fully document how to duplicate and customize each reference implementation. These templates are practical for production and research farms and serve as an introduction to more complex databases and analysis for applications well beyond production agriculture.

Literature Cited

AgGateway. 2021. *ADAPT Code and Application Data Model*. Accessed 1/24/2024. https://github.com/ADAPT/ADAPT.

Airtable. 2022. What is Airtable and why should you use it? *Airtable Guides*. Airtable. Accessed 1/24/2024. https://www.airtable.com/guides/start/what-is-airtable.

Antle, J.M., B. Basso, R.T. Conant, H.C.J. Godfray, J.W. Jones, M. Herrero, R.E. Howitt, B.A. Keating, R. Munoz-Carpena, C. Rosenzweig, P. Tittonell, and T.R. Wheeler. 2017. Towards a new generation of agricultural system data, models and knowledge products: Design and improvement. *Agricultural Systems* 155: 255-268. https://doi.org/10.1016/j.agsy.2016.10.002

Buckmaster, D.R. 2022. Digital field records: an easy way. Accessed 1/24/2024. https://ag.purdue.edu/news/2022/04/digital-field-records.html.

Buckmaster, D.R., A. Deering, and S. Monroe. 2022. Digital records for FSMA – a free toolkit. Accessed 1/24/2024. https://ag.purdue.edu/news/2022/02/digital-records-for-fsma.html.

Buckmaster, D.R., and K. Soonthornsima. 2022. Digital marketing and grain delivery records. Accessed 1/24/2024. https://ag.purdue.edu/news/2022/09/digital-marketing-delivery-records.html.

Buckmaster, D.R. 2023a. Digital horticultural crop activity records. Accessed 1/24/2024. https://ag.purdue.edu/news/2023/01/digital-horticulture-crop-activities-records.html.

Buckmaster, D.R. 2023b. *Simple Personal Databases – Make Your Records Digital Simply.* Accessed 1/24/2024. https://ag.purdue.edu/news/2023/06/simple-personal-databases-make-your-records-digital-simply.html.

Chen, K., J.M. Hellerstein, and T.S. Parikh. 2010. Designing adaptive feedback for improving data entry accuracy, pp. 239-248. *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, New York, NY. https://doi.org/10.1145/1866029.1866068

FDA Food Safety Modernization Act. 2011. 21 U.S.C. §§ 2201-2251. https://www.fda.gov/food/food-safety-modernization-act-fsma/full-text-food-safety-modernization-act-fsma

Hickson, N. 2020. Using Google forms for crop recording. *Community Supported Agriculture*. Accessed 1/24/2024. https://communitysupportedagriculture.org.uk/resources/using-google-forms-for-crop-recording/.

Illinois Soybean Association. 2022. Carbon and data guidebook. Accessed 1/24/2024. https://www.ilsoyadvisor.com/carbon-data-guidebook/.

Joly, N. 2011. Shaping records on the farm: agricultural record keeping in France from the nineteenth century to the Liberation. *Agricultural History Review* 59(1): 61-80. Accessed 1/24/2024.

https://www.researchgate.net/publication/233565051_Shaping_records_on_the_farm_agricultural_record_keeping_in_France_from_the_nineteenth_century_to_the_Liberation.

Meijer, L. 2022. What is Airtable and what makes it great? *Softr*. Accessed 1/24/2024. https://www.softr.io/learn/airtable/what-is-airtable.

Neo, B. 2020. *What is tidy data*? Accessed 1/24/2024. https://towardsdatascience.com/what-is-tidy-data-d58bb9ad2458.

Pagano, P., L. Candela, and D. Castelli. 2013. Data interoperability. *Data Science Journal* 12 GRDI19–GRDI25. https://doi.org/10.2481/dsj.GRDI-004

Wiginton, H. 2022. *47+ best Airtable examples and tutorials*. Accessed 1/24/2024. https://hannahwiginton.com/blog/software/airtable/curated-list-of-airtable-articles/.