IDAI.FIELD: DEVELOPING SOFTWARE
FOR THE DOCUMENTATION OF ARCHAEOLOGICAL FIELDWORK

1. Introduction

The German Archaeological Institute (Deutsches Archäologisches Institut, DAI) conducts a variety of different types of field research, each with its own unique documentation requirements: excavations, surveys and architectural surveys. The resulting differences are reflected in the workflows, the recording methods and the documentation. In addition, the DAI's international work has to comply with the guidelines of the respective heritage agencies in the host countries. This results in a whole range of requirements that software for field data recording must fulfil:

– It must provide a sufficiently flexible, adaptable data model.
– At the same time, a standardised core data schema should allow for data comparison across geographical, temporal and thematic boundaries.
– Multilingual data entry must be possible.
– It must work both offline and online and allow robust methods of (delayed) data synchronisation.
– It should also be possible to publish primary research data, together with the research results, with minimal additional effort and in accordance with the requirements of the international research community and the funding providers.

An initial version of a unified field recording system for the DAI was based on the proprietary software Filemaker. This was used and modified by numerous projects at the DAI. However, systematic evaluation revealed problems in guaranteeing data quality and comparability as well as integration into existing workflows. Furthermore, the use of proprietary software and its closed data formats caused problems with upgrades that broke backward compatibility. For these reasons, the decision was made to design the new iDAI.field from scratch, relying exclusively on open source software and current web technologies. The focus of the redesign was on enabling complete documentation of archaeological fieldwork and its automatic synchronisation/integration between peers working on site. This distinguishes iDAI.field from more generic platforms, such as GIS or CAD.

The individual roles of the software were split into separate components: a desktop application for on-site data entry, a server component for synchronisation via the Internet (or a local network) and an online platform for publishing the primary data. So far, the focus has been mainly on transferring already known, analogue documentation workflows into the digital

application. In the future, it should also be evaluated whether and how the possibilities of using these digital technologies can improve those workflows in order to open up new knowledge discovery processes.

2. Development history

Previously, the DAI used a whole range of individually developed, project-specific systems for the digital documentation of field research. Maintaining these systems was resource intensive and there was no uniform data model. Information stored in these systems could not be easily compared and was only published in the context of traditional publications. In order to overcome this, the development of a modular system was started, which should both meet the requirements of the individual projects and offer sufficient data conformity to be able to compare the collected data with one another later on.

The first version of iDAI.field was developed from summer 2006 (SCHÄFER 2011; CUY *et al.* 2019) by the IT department of the DAI and the Cologne Digital Archaeology Laboratory (formerly the Research Archive for Ancient Sculpture) at the University of Cologne (Fig. 1). This version was developed in close cooperation with the Pergamon excavation of the DAI Istanbul and initially adapted to their requirements. The system was also used there for the first time. In 2008, a so-called master clone of the database was
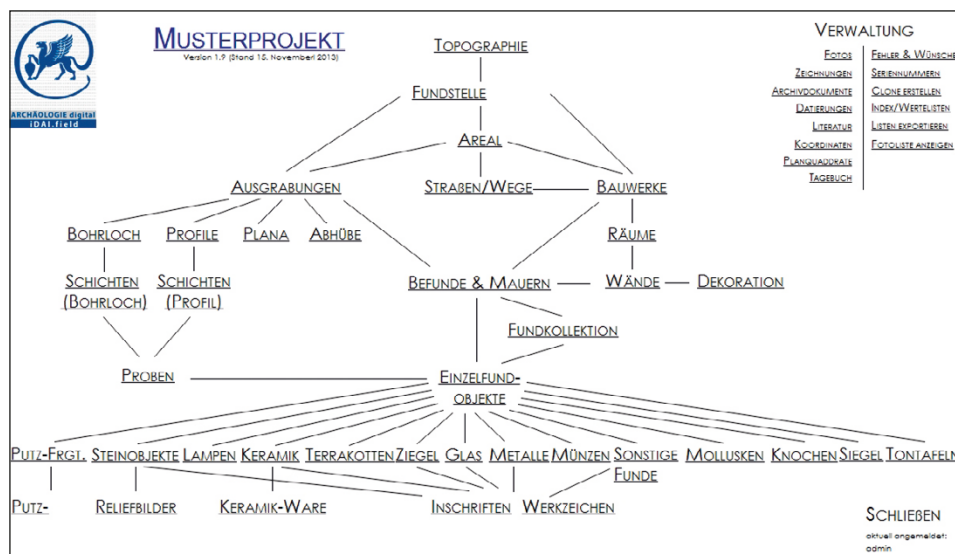

Fig. 1 – The first version of iDAI.field: 1.0. Screenshot of the welcome page with an overview of existing categories and the project management options.

created, which could be copied and thereby made available to other projects for use. In 2009, a multilingual prototype was added. FileMaker was chosen as the technological basis for the system, as this application was sufficiently functional to be able to map the chosen object-relational data model. In addition, the application could be installed on all desktop computers without any problems. Local synchronisation of the data within the excavation network was possible and FileMaker also offered a server component for online operation and synchronisation via the Internet.

The modular structure chosen was supposed to make it possible to offer a uniform data model at the core, which could be supplemented with project-specific extensions. In practice, however, these adjustments turned out to be very time-consuming, as changes made during the project phase always had to be transferred back to the master clone and then to the other project copies. The use of proprietary software also posed a problem. For example, software changes were not always backwards compatible. Therefore, the switch to a newer version also required a migration of the data. Additionally, desktop and server licences for numerous excavation computers represented a considerable financial expense. Finally, the use of proprietary software contradicted the open source idea to which the DAI-IT is committed.

Despite these problems, the Filemaker version of iDAI.field had been used relatively successfully by more than 35 projects at DAI. This showed the demand for an excavation software that could meet the requirements of the DAI. At the end of 2009 it was therefore decided to address the problems that had been identified and to develop a new database solution based on common web technologies. For this purpose, the DAI cooperated with the Brandenburg University of Technology Cottbus-Senftenberg, which had developed the Cottbus Information System for Archaeology and Building Research (CISAR), also an application for documenting archaeological field research (Henze *et al.* 2013; Schäfer *et al.* 2013).

The medium-term goal was to combine the respective strengths of CISAR and iDAI.field in one system and to make this available as a free, web-based successor solution (working title: OpenInfRA). Overall, the scope of the application to be developed was very ambitious and covered almost all possible application areas of archaeological fieldwork including GIS, 3D and much more. The project was pursued until 2016, while at the same time the first version of iDAI.field continued to be used. Unfortunately, the results were almost exclusively theoretical. A functional application has not been developed. Therefore, in 2016, it was decided at DAI to start a new project and develop the application from scratch (Fig. 2). The specifications were:

– a generic data model that can deal with most documentation requirements of archaeological fieldwork while also trying to maintain standardisation where possible;
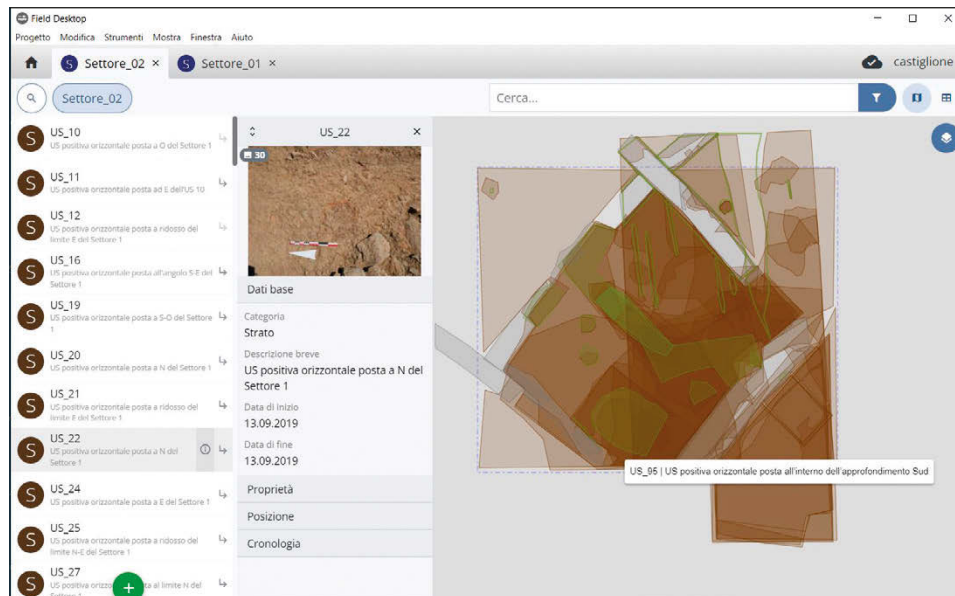
Fig. 2 – The current version of iDAI.field: 3.0. Screenshot of the application with data from the excavation project in Vetulonia/Castiglione della Pescaia of the Rome Department of the German Archaeological Institute.

– to support distributed work and data entry in the field even in places with no Internet coverage (need for a very robust synchronisation between the clients and the server);
– the integration of spatial data and descriptive documentation;
– to use only widely used open source technologies;
– to offer a low technical barrier and general ease-of-use (no high barrier of entry due to the technical complexities involved in setting up web servers and databases and configuring the respective software stack);
– and as a goal with lower priority: to later extend the application further to also feature a meta-search that will enable cross-project queries.

3. Technology

From the very beginning, the application has been based entirely on open source web technologies. The main application is written in TypeScript and uses established open source frameworks such as Angular and Leaflet. One advantage of this approach is that the functionality and layout of the application only have to be implemented once and are then available identically (with minor adjustments) on all supported operating systems (Windows,

macOS and Linux). Another advantage is that parts of the code could be more easily reused in the supplementary web publishing platform iDAI.field web.

By using the electron framework (https://www.electronjs.org/), the application can be delivered as a single installable desktop application without additional dependencies. To enable robust support for offline work with later syncing between researchers, the Javascript database PouchDB (https://pouchdb.com/) is used. PouchDB is a browser variant of the NoSQL database Apache CouchDB. The central feature of both CouchDB (https://couchdb.apache.org/) and PouchDB is the ability to replicate: the state of one database can be transferred to another one. This replication can also be automated continuously between databases (synchronisation), which is the core functionality used for iDAI.field. The desktop application can be used online and offline – alone or connected to installations of the software on other computers.

Another argument for using PouchDB was the fact that a NoSQL database works great with user-configurable data models. iDAI.field's data model should, on the one hand, be based on the foundation of a fixed core data model common to all projects to enable comparability of data across project boundaries. On the other hand, it should also allow for project-specific customization that would have been more difficult to support with an SQL database. By using PouchDB, documents can be easily modelled and stored as JSON objects with arbitrary fields. The structure of these JSON objects is stored within the project configuration, which also exists as a document in the database.

4. Data model

In iDAI.field, each entity is a resource belonging to a specific category (e.g. "Stratigraphic unit", "Find", "Image"). Resources are described by a set of fields and linked with each other via relations (e.g. hierarchical, temporal or spatial relations). The project configuration determines which categories are available for selection in a project and which fields can be filled in depending on the category. The configuration is composed of three different parts, each fulfilling a specific role between the poles of comparability and adaptability.

The core configuration defines the framework within which all iDAI.field projects operate. Here, the main categories (e.g. "Find") as well as their possible linkings via relations among each other are defined, furthermore the most important fields that are directly referenced in the application code (e.g. "identifier", "geometry"). The library contains a number of subcategories (e.g. "Pottery" as a subcategory of "Find") and fields for a further description. The fields were chosen by evaluating projects created with iDAI.field 1 and identifying the most commonly used fields for each category. The use of the library is not mandatory but should motivate users to reuse existing categories and fields instead of always creating new ones. Therefore, it is

planned to gradually extend the library in the future using the configuration data of existing projects. Finally, the project-specific subcategories and fields can be created to meet the unique requirements of a project. Data entered in these can therefore only be compared with those of other projects to a limited extent. However, it is also planned to incorporate a curated selection of project-specific configuration elements to the library and thus make them available to all users.

5. FEATURES

iDAI.field currently consists of three different software components: the desktop application Field Desktop for entering data, the server application FieldHub – providing a centralised synchronisation target – and the web platform iDAI.field web for publishing the data. Field Desktop is the main component of iDAI.field, where most of the development work has gone into. The main functionalities of the application include:

– Database and file synchronisation between individual desktop clients via HTTP, mainly used within local WiFi or LAN on excavation.
– A comprehensive resource management with a hierarchical view as well as basic search capabilities (full-text search for specific core fields, field specific search).
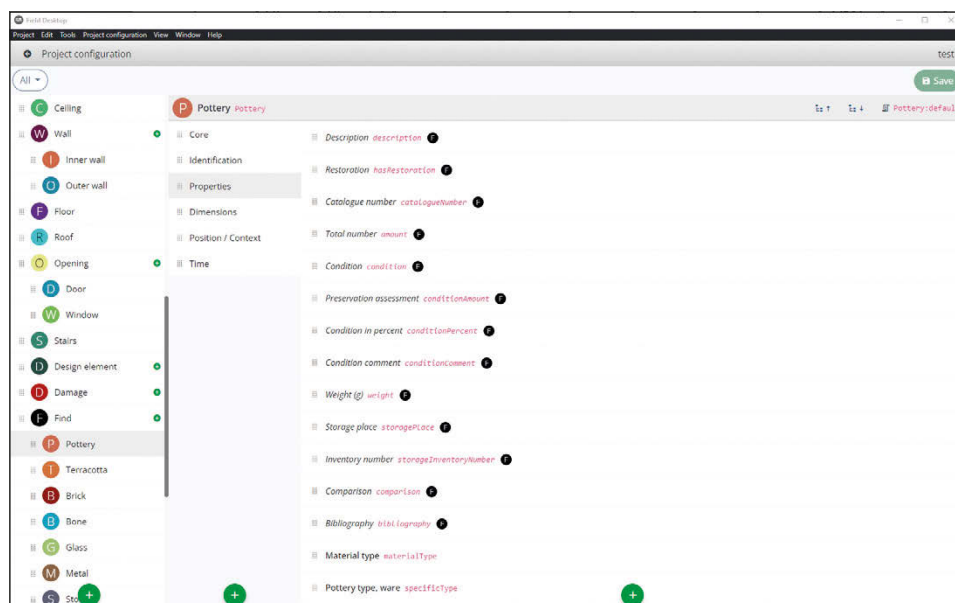

Fig. 3 – Screenshot of the iDAI.field 3.0 configuration editor.

– Integration of geodata into the application and the joint display of geodata with the respective resource data.

– Image management: Images can be imported into the application, linked to resources and displayed as map layers after adding georeferencing information by importing a world file.

– Type management: Types can be sorted hierarchically in catalogues and linked to finds.

– A configuration editor for an easy customization of the project configuration, including a value list editor (Fig. 3).

– A simple matrix view, which is automatically generated from the project data depending on the temporal or spatial relations set for the resources (the matrix view is implemented in a simplistic way and needs to be revised in the future.).

– Import and export of project data in the formats CSV, GeoJSON and Shapefile; furthermore there is the possibility to export or import type catalogues including all images.

– Creation and recovery of project backup files.

– Multi language support: with version 3.1, data can be entered in different languages per field; the user interface is available in English, German, Italian and Ukrainian.
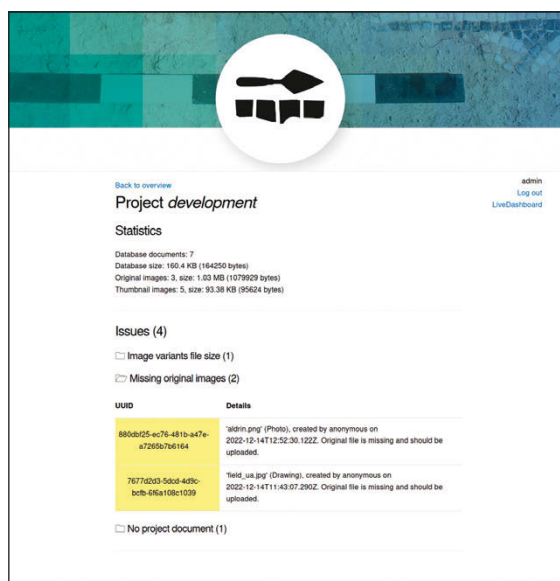

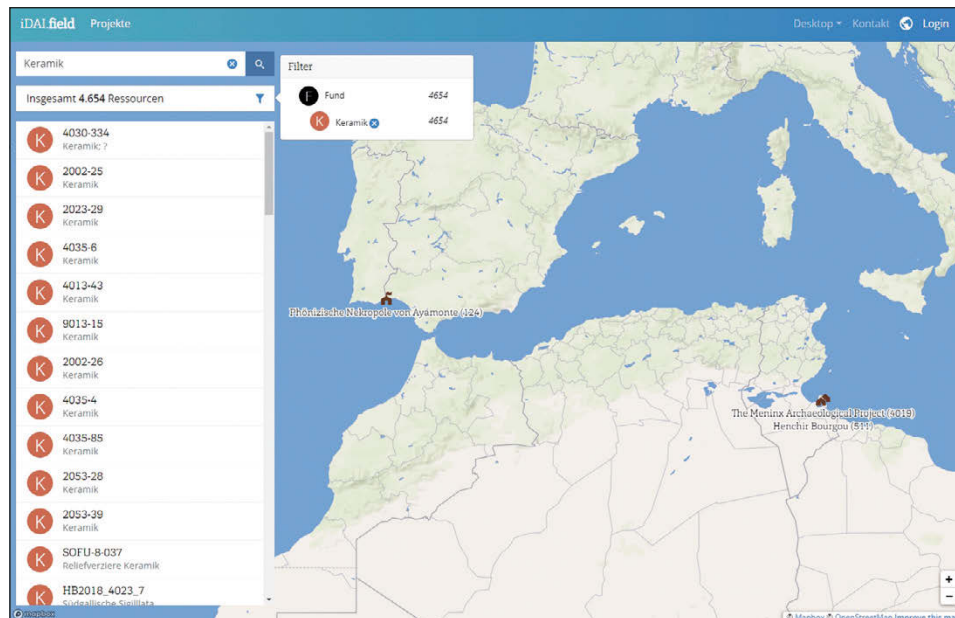Fig. 4 – Screenshot of the FieldHub data monitoring interface.

Fig. 5 – Screenshot of iDAI.field web with cross-project search for ceramic finds.

FieldHub is a small server application, written in Elixir, that institutions may provide for their researchers as a permanent syncing target and to collect all data by all running projects centrally (Fig. 4). FieldHub can be deployed using pre-built Docker images we host on our GitHub container registry (https://github.com/dainst/idai-field/pkgs/container/field_hub). The server provides some monitoring that enables project members to track issues (for example if someone did not upload all their high resolution images to the server, which would be essential for later publications).

iDAI.field Web is the DAI publication platform written in Elixir (backend) and React (frontend) for projects created with Field Desktop (Fig. 5). It provides an overview of all published projects on a world map and allows submitting search requests across these projects. The possibilities of the common core data model are already being used to some extent in the form of a cross-project search with the option of filtering by category. These search options are to be further expanded in the future. Currently iDAI.field Web still contains some dependencies on the DAI's infrastructure, which will have to be refactored in order to make it usable for an audience outside the DAI.

6. Development methodology

In the development of the Field software, an agile approach with two-week sprints has been implemented. Already at a very early stage in development, a pre-release version of the application was used by archaeological excavation projects (such as the Meninx Archaeological Project: https://field.idai.world/project/meninx-project); their feedback was then used for the further development of the application in the next iteration. This approach ensured that the application would not be developed past the needs of its users and could be used productively in the field.

In order to cope with the foreseeable increasing complexity of the application, it has been written according to Clean Code principles based on the guidelines of Robert C. Martin (2009) to keep it as intuitively understandable as possible. This includes frequent refactorings and code reviews. In addition, an extensive test suite with currently about 900 unit tests and about 140 end-to-end tests ensures that no functionalities are lost. While the end-to-end tests in particular involve a great deal of maintenance work because they have to be adapted constantly, they have also proven very effective in preventing errors.

Although the different software components also differ technologically, some parts of the code are currently being reused by multiple components. For example, one of the most important core components of the application, which is responsible for loading the project configuration, is currently used in Field Desktop as well as in the backend of iDAI.field Web and in the prototype of the mobile application. To simplify this process, all components reside in a single code repository.

7. Open source development

Since its redevelopment with version 2, iDAI.field has relied exclusively on open source libraries and has been published on GitHub as an open source project under the Apache License 2.0. However, the application was initially only used by excavation projects of the DAI and other institutions that are in close contact with the DAI. To work with a customised project configuration, a JSON file had to be created and submitted to the developers, which was then deposited in the GitHub repository and shipped with a new version of the desktop application.

With version 3.0, released in spring 2022, several steps have now been taken towards opening up iDAI.field. Project configurations are stored in the database itself and can be created and edited in a configuration editor integrated into the desktop application, so that consultation with the DAI is no longer necessary to use the software. Nevertheless, certain dependencies still remain. The library refers to DAI systems such as iDAI.gazetteer and iDAI.bibliography, and the publication platform iDAI.field Web, as part of the

iDAI.world, cannot easily be set up in other contexts (unlike Field Desktop and FieldHub). The goal is to further resolve these dependencies in the future and turn Field into a "true" open source project that is jointly developed and used by a range of different institutions and individuals.

Simon Hohl, Thomas Kleinke, Fabian Riebschläger, Juliane Watson

Deutsches Archäologisches Institut

simon.hohl@dainst.de, thomas.kleinke@dainst.de, fabian.riebschlaeger@dainst.de, juliane.watson@dainst.de

REFERENCES

Cuy S., Oliveira D.M., Gerth P., Kleinke T., Schierenbeck J., Watson J. 2019, *iDAI. field 2: A modern approach to distributed fieldwork documentation*, in W. Börner, S. Uhlirz (eds.), *Proceedings of the 22ⁿᵈ International Conference on Cultural Heritage and New Technologies (Wien 2017)*, 1-8 (https://archiv.chnt.at/wp-content/uploads/eBook_CHNT22_Cuy_etal.pdf).

Henze F., Magdalinski N., Schwarzbach F., Schulze A., Gerth Ph., Schäfer F. 2013, *Concepts and technologies for a comprehensive information system for historical research and heritage documentation*, «International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences», XL-5/W2, 325-330 (https://doi.org/10.5194/isprsarchives-XL-5-W2-325-2013).

Martin R.C. 2009, *Clean Code. A Handbook of Agile Software Craftsmanship*, Upper Saddle River, Prentice Hall.

Schäfer F. 2011, *iDAI.field. Ein modulares Dokumentationssystem für archäologische Feldforschungsprojekte*, «Kölner und Bonner Archaeologica», 1, 193-200.

Schäfer F., Schulze A., Schwarzbach F., Gerth Ph., Henze F. 2013, *OpenInfRA. Storing and retrieving information in a heterogeneous documentation system*, in G. Earl, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska, D. Wheatly (eds.), *Archaeology in the Digital Era. Volume II. Papers from the 40ᵗʰ Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA) (Southampton 2012)*, Amsterdam, University Press, 422-431.

ABSTRACT

The German Archaeological Institute (Deutsches Archäologisches Institut, DAI) conducts a variety of different types of field research, each with its own unique documentation requirements: excavations, surveys and architectural surveys. The resulting differences are reflected in the workflows, the recording methods and the documentation. In addition, the DAI's international work has to comply with the guidelines of the respective heritage agencies in the host countries. iDAI.field is the system for documenting archaeological fieldwork at the DAI. From the very beginning it was developed to meet these very different requirements. The development spanned several years, major versions and associated, far-reaching technology changes. The latest iteration of the application relies exclusively on open source technologies and is published on GitHub under the Apache License 2.0 in accordance with DAI-IT's open source policy. In order to open up the application to other interested researchers and/or developers, the focus of the last year has been the implementation of an extended configuration interface and the removal of dependencies from the DAI infrastructure. This article outlines the development history, introduces the currently available functionalities, and briefly discusses the data model, followed by an overview of the technologies used. It also describes the development into a real open source product and gives a short outlook on the future plans.