Digital Floor Plan Database: A New Method for Analyzing Architecture

White Paper

National Endowment for the Humanities
Office of Digital Humanities
HAA-258763-18
Elise King (Project Director), King-Ip (David) Lin (Co Project Director)
Baylor University
Due: 4/30/20
Submitted: 4/30/20
Digital Floor Plan Database: A New Method for Analyzing Architecture

This report describes the project activities, accomplishments, audiences, evaluation, continuation of the project, long-term impact, and award products.

Background

Architectural history acts as a lens through which to interpret historical, social, cultural, and ideological components of a given era and society. As such, the analysis of built environments, both past and present, is essential for enriching our understanding of the human condition. The logistics of examining large samples of buildings, however, remain challenging. Floor plans, specifically, can provide a snapshot of family relationships and roles, building culture, and the human need to define one’s cultural identity in time and space. Floor plans are information rich and, among architectural drawings, offer the most comprehensive overview of a building, informing the reader about elements such as rooms (location and name), doors and windows, built-ins, plumbing fixtures, and stairs. In the past, examining and comparing large numbers of floor plans was difficult. Most scholars, therefore, were limited to analyzing floor plans manually (digitally or by hand). While others had to employ a combination of tools requiring specialized technical knowledge. Recognizing floor plans as an underutilized trove of data, we began developing the Building Database & Analytics System (BuDAS) to bring together manual data entry and image recognition techniques to partially automate the process of floor plan analysis.

In 2016 we developed the BuDAS prototype, which allowed users to manually enter information about a floor plan (e.g., house name, square footage, construction dates, and architect) and store it in a relational database. Through the database, we also implemented an application where users can retrieve basic information about floor plans (e.g., find all houses designed by Frank Lloyd Wright) and analyze graphically results of simple queries. The prototype proved to be a useful system for organizing and analyzing plan information, but the process of collecting and entering the information was cumbersome, time consuming, and often inaccurate. Calculating and then inputting plan information into the prototype, which includes room names, square footage, number of doors, etc., took several hours for a single house. We applied for a Level II Digital Humanities Advancement Grant to create BuDAS 1.0, adding image recognition software for semiautomatic floor plan detection, along with an expanded database, and advanced query/analysis tools.

Project Activities

Our major project activities centered on adding an automated floor plan detection component to the BuDAS system. To accomplish this goal, we developed a floor plan image extractor and annotator.

The first part of automated detection is the image extractor. After the user uploads an image (jpeg) of a floor plan, the extractor detects rooms and room names (see appendix A). We studied and tested different methods and ultimately decided to approach floor plan detection as a contour detection problem. Essentially, with contour detection the aim is to extract the outline of a shape or object. For floor plans, ideally, each contour represents a room. Contour detection offers a number of advantages: it is generic (can be applied to a variety of plans and plan symbols) and allows for multi-level contours which can be useful for more complex plans. We found, however, that there were several challenges to this approach, including image noise (reduces effectiveness of detection) and even windows and doors ( Breaks the contour). To
minimize detection issues, we developed a five-step process that includes text detection and removal, noise removal and wall closing, external boundary detection, contour detection, and room area checking. These processes were designed to balance accuracy and efficiency. Our semi-automatic detection approach is not intended to achieve perfect detection results, which no system has been able to accomplish. Instead, we want to achieve a baseline level of accuracy through automatic detection and then allow the user to make revisions using the annotator.

After the extractor runs, the annotator interface displays the detection results (see appendix A). The annotator shows the original floor plan image with detected rooms overlaid in a colored fill. The annotator also allows users to make revisions or additions to the extractor results using a suite of tools. We developed more than ten tools that provide the ability to add, delete, or split a room; combine two rooms; add a door or opening; add room labels; and set a scale. There is also a reset tool so that users can go back to the original detection results if they make an error. Once satisfied with the modifications, the user can save the results. After running the extractor and saving the annotator results, a file of room information is automatically created. Users can import this information into the database using a standalone python program that is included in the package. This connectivity package is currently in the beta phase and available through GitHub. The extractor, annotator, and database are available through the project website (http://sites.baylor.edu/budas/) or Github (https://github.com/linki0/BUDAS-extractor and https://github.com/linki0/BUDAS/). They are implemented in Python 3.7 using publicly available modules. We also included an executable option, but it's only available for Windows users at this time.

In the project we proposal identified that the extractor would detect the following information from plan images: general building/project information (location, builder, client(s), scale, orientation), information about individual rooms (size, number of windows/doors, room names), and relationship between rooms. Based on our evaluations of the extraction system and discussions with the advisory board, we slightly altered what information is extracted automatically versus entered using the annotator. This was done to ensure the widest possible use and access to BuDAS. The goal and aims, however, remain unchanged. As we expressed in the proposal, BuDAS is designed to be accessible to as many users as possible. It is free and open-source and does not require the use of expensive add-on software. But as we began to implement the automated plan detection, our team discovered that many of the existing, open-source solutions required users to ground truth plans, essentially training the system to detect a particular group of plans by labeling and defining objects/symbols. This typically requires the user to pre-process plans, which is time-intensive and in most cases requires specialized software (e.g., Photoshop). Also, the ground-truthing process would likely have to be repeated each time a different group of plans is introduced to the system, depending on the degree to which the graphic standards/symbols changed. The barrier for entry to use such a system appeared quite high, so we went another route, applying contour detection to floor plans (described in accomplishments/activities). As a result, our focus shifted to room detection (which is the part of the process that is most time-consuming and error-prone), and we moved the identification of doors and windows to the annotator. Users simply select the door tool and click on the location of doors on the plan. The process is quick and easy, and more accessible to users than ground-truthing plans. The end result is the same, the process has just been adjusted, moving openings from automatic detection to the annotator.

We accomplished our goal and aims of adding automated floor plan recognition capabilities to BuDAS, but we had to adjust our timeline and make several accommodations. In spring 2019 a key team member (not the PD or co-PD), had to take an unplanned medical leave of absence, withdrawing from the university. As a result, we requested, and were granted, a six-month no-

---

1 Identifying information has been omitted because of our small team size. We provided additional details about the team member and their role to NEH when requesting the extension.
cost extension. To continue development, during the summer of 2019 we recruited another student to work on the project until the original team member returned in the fall. Unfortunately, after returning in the fall the original team member was only able to work on the project for 2/3 of the semester before having to withdraw again for medical reasons. The co-PD took on the team member’s work and ensured the completion of that portion the project.

**Accomplishments**

We were able to accomplish our main goal as stated in our project proposal: “to expand our prototype and add a third component—a floor plan image extractor and annotator—to enable the automatic detection of plan information.” This was achieved through the development of the annotator and extractor. Additionally, we completed our two main aims. First, we aimed to make floor plan detection quicker than completing the process manually. On average, for simple floor plans (most floor plans) the BuDAS extractor completes floor plan detection in several seconds. For plans with a lot of complexity (e.g., plan grids, landscaping, rooflines, etc.), the process can take up to a minute or two. This compares to calculating the same information by hand, which can take 45 minutes or more per floor. Another aim was to make the system accessible to a variety of users (e.g., humanities researchers, archives, libraries, historians). We accomplished this by creating a system that is free, open-source, and does not require additional software to run. Although unexpected delays necessitated a one-time, no-cost extension, our team achieved the goals in our proposal.

In several instances, we were able to go beyond what we had originally proposed in our work plan. For example, we originally proposed three project evaluations, but we completed five. Also, we originally intended to present at one conference (pending acceptance), and we presented a paper and poster prior to the close of the project cycle and another paper just after. We presented a poster examining challenges to floor plan detection at the Digital Frontiers Conference in Austin, TX in September 2019. Then, in October 2019, we presented on the application of BuDAS to interior design at the Interior Design Educators Council (IDEC) Southwest Regional Conference in Waco, TX. And just following the close of the project cycle, we presented our initial evaluations of Wright’s floor plans from the Storrer collection (as described in our proposal) at the IDEC national conference in Tulsa, OK. We also have a manuscript that will be ready to submit late spring or early summer 2020.

BuDAS also received recognition and publicity from several different sources. In fall 2019, we were awarded “Best Paper” for our presentation at the Interior Design Educators Council (IDEC) Southwest Regional Conference. BuDAS was also promoted in two university press releases, highlighted on university and college-level social media channels, and mentioned in posts on Architizer's and the Society of Architectural Historians’ websites. We are committed to continuing to develop BuDAS as well as sharing and publicizing our findings and progress.

**Audiences**

BuDAS is free, open source, and available through our project website and Github. Prior to developing BuDAS we identified several potential user groups. They include humanities researchers, archives and libraries, and others who research the built environment (e.g., designers and architects). Our advisory board and user test group include representatives from each group. Additionally, we presented on BuDAS at three conferences, two focused on interior design and the other on digital humanities. These opportunities allowed us to engage in formal evaluation and casual dialogue with our targeted users. Additionally, collected contact information and business cards from educators and researchers from around the United States
who are interested in using BuDAS in their class or providing additional archival materials to test the detection system.

**Evaluation**

**Summary of Evaluation Methods**

Evaluation was an important component of our project cycle. Our project team completed two internal evaluations of BuDAS's detection accuracy, and we requested feedback from different audiences and user groups over the course of the grant. This feedback helped to improve and shape BuDAS, and it has also provided us with future development directions. In accordance with our original work plan, we solicited feedback from our advisory board on two occasions, students from PD King's architectural history course, and other audiences during conferences and presentations. Additionally, we completed several rounds of internal evaluation to test the accuracy of our plan detection.

**Advisory Board**

Our advisory board consisted of four members, with representatives from architectural history, libraries and archives, and digital humanities. During the grant period, two of our board members moved to a different institution. They remained available to provide feedback, but we transitioned our in-person advisory board meeting to a digital format as a result. One of our advisory board members also shared BuDAS with two of her colleagues, and they were able to provide collective feedback, which was very helpful to the project. During the spring of 2018, prior to adding automatic plan detection (i.e., the extractor and annotator), we sent members of our advisory board a link to access the BuDAS prototype. They manually entered houses into the database and used the analysis tools. Following their testing of the prototype, we asked them to complete a questionnaire and System Usability Scale (SUS) assessment. Our project team created a nine-question questionnaire in Qualtrics, focusing on three areas: readability and usefulness of graphs/visualizations, future tools/features, and potential research questions that BuDAS could help answer. After reviewing the advisory board’s feedback, our project team identified several opportunities for improvement, including adding instructions to the graphs/visualizations in the analyzer and correcting a glitch where the entry field reset to “0” when users entered the cost and date information. Regarding future tools, the advisory board favored adding graph mining for room relationships, a geographic mapping tool, and an inflation adjustment tool for cost information. Before completing the extractor and annotator, we addressed these concerns by adding instructions to the charts and visualizations and correcting the issue with the date and cost entry tool.

Then, at the end of December 2019, we commenced our final evaluation with the advisory board. We provided board members with a copy of the annotator and extractor, along with a sample floor plan for detection, and instructions for use. In addition, we sent them a document containing mockups of proposed changes to the annotator graphic user interface. To ensure that all board members could assess the annotator and extractor, we also provided a screen capture of the system in use. Advisory board members were asked to answer five questions about the extractor, annotator, and proposed changes to the interface. Board members suggested that some colors might need to be revised to meet accessibility standards: “Be mindful of accessibility issues, especially with regards to color choices and contrast and the usability for color blind and low vision users.” Additionally, they suggested implementing a way for users to modify not only the detection results but also to correct any inaccuracies on the plan itself. We also asked the advisory board to review a list of tools/features and to select the one that would be most useful. The options included an undo button (currently, there’s an option to restore to the original but not undo), drawing guides/assistance for room divisions, and a zoom tool. The response was unanimous in favor of an undo button. Our advisory board
members provided helpful feedback, and also were favorable in their assessment of the project and its future: “All in all, you all have a great project and have made a lot of progress! Most of our initial feedback was addressed in your proposed revisions,” and “Congratulations to you and the team on all the amazing work here! Everything is looking really nice and I’m excited about the launch of the site, I think it will be a great resource!”

History of Interiors II Class

Additionally, our work plan included an evaluation of BuDAS by students in PD King’s History of Interiors course. The students were working on a project using Frank Lloyd Wright’s floor plans, and they were asked to complete a survey about automatic floor plan recognition. Twenty-one students completed the survey, and we have included a summary of their responses here. When students were asked about the biggest benefit of automated plan recognition (as compared to manual), seven discussed visualizing plan information, six wrote about improving comparisons between plans, and four responded about the time saved calculating the square footage of rooms. We also asked students about additional sorting or filtering categories that could be added to the database to help them sift through information. Examples of their suggestions include orientation, style of residence, number of floors, number of rooms, specific room types, and construction types. The students also offered suggestions for new analysis features. Their responses included adjacency matrices and bubble diagrams for room relationships, mapping features, and an option to sort by materials. Finally, twenty out of twenty-one students stated that automated floor plan recognition would be beneficial for analyzing floor plans.

Internal Evaluations

Finally, we completed two internal evaluations of BuDAS’s automatic detection capability (detection without user input/revision). While BuDAS employs a semi-automatic detection method, detection accuracy is still important because the greater the accuracy of automatic detection, the less user input (time) is required post-detection. During summer 2019 we evaluated the effectiveness of BuDAS’s detection on a set of unedited floor plans. Pre-processing or cleaning plans (e.g., removing noise, increasing contrast, erasing landscaping, or other potentially “distracting” elements) is commonly done prior to detection to increase accuracy. But this practice requires additional time and in many cases, specialized software, so we elected to run our detection without any pre-processing. Our evaluation metric is a slight variation of the Heras et. al (2014) protocol, which compares how the algorithm can reduce the work of annotation following automatic detection. We found that, on average, half the rooms on the floor plan are matched in some way, reducing the time for the user to enter information by nearly 50%. We used two sets of plans for our evaluation: an “easy” and a “difficult” set. The difficult set included a selection of Frank Lloyd Wright floor plans from the William Allin Storrer Collection. Evaluating Wright’s work using BuDAS was one of our aims in the project proposal. The easy set included plans from publicly available mid-century house plan books. Wright’s plans were more difficult to detect because of their added complexity: grids, non-orthogonal layouts, landscaping elements, rooflines, etc. Interestingly, the overall match rate is better for more difficult plans. One plausible explanation is that fewer lines are removed during noise-removal, so the algorithm detects a greater number of rooms. With more rooms detected, the chances increase that more actual rooms will be matched, at least partially. We have written a manuscript about our findings from this evaluation, which we plan to submit late spring or early summer 2020.

Following the summer 2019 internal evaluation, we completed a second assessment to identify plan characteristics that were contributing to differences in detection rates. We used the extractor to test 80 variations of the same floor plan: drawing conventions (single and multiple line weights); page setups (plan only, plan/border, plan/drawing title, and plan/border/drawing title); and plan content (e.g., blank plan, plan/wardrobes, plan/wardrobes/cabinets, plan/
wardrobes/cabinets/furniture, etc.). Our findings indicate that single line weight plans (as opposed to plans drawn with multiple line weights—a common graphic standard) had the highest detection rate, and the full match rate was lower for plans with grids, landscaping, or materials. We presented a poster on these findings at a regional digital humanities conference in September 2019. We will use this information as we continue to develop BuDAS and refine the detection capabilities.

Continuation of the Project

We plan to continue developing the BuDAS following the conclusion of the grant cycle. We are currently exploring broadening our team to include individuals with expertise in artificial intelligence and human computer interaction. Artificial intelligence could improve the detection accuracy rate and require fewer user revisions. Additionally, we are planning to work with experts in human computer interaction to develop a new interface that considers the unique challenges of floor plan data, which involves representing three-dimensional forms in two-dimensions.

Moving forward, we will also expand the analysis features to provide additional tools for comparing buildings, specifically including methods for more objective and statical analysis. Our team is exploring the fields of space syntax analysis and graph mining as areas of consideration. Drawing on the feedback we received from users, we also plan to incorporate tools to improve the ease-of-use, including an undo button and more editing tools. To support the continued development and expansion of BuDAS, we will explore additional funding opportunities from internal and external sources, including NEH Level III funding.

Our NEH grant provided the opportunity to share BuDAS with different audiences, and as a result we were able to strengthen existing partnerships and make connections that will allow us to expand into new directions. As described in our proposal, we partnered with the Alexander Architectural Archives at the University of Texas to use the William A. Storrer papers as the sample for this project. The Storrer papers include one of the most comprehensive and detailed collections of floor plans by renowned American architect Frank Lloyd Wright. Over the grant period we have strengthened our partnership with the Alexander Architectural Archives (AAA). Also, PI Elise King connected with the Head of Architectural Collections, Katie Pearce Meyer, at the Digital Frontiers Conference at UT Austin. King and Pearce Meyer were both presenting posters at the conference. King and the other representatives from UT Austin are planning to meet in the coming months to discuss BuDAS further. Presenting on BuDAS at conferences created opportunities to foster new collaborative partnerships. We have an invitation from another university to visit their archives and potentially use BuDAS to analyze their collection materials. And following our presentations at the Interior Design Educators Council (IDEC) conference, several educators expressed interest in using BuDAS with their students and inquired about becoming beta testers.

Long Term Impact

As we continue to develop BuDAS we plan to build on potential partnerships and collaborations discussed in the continuation section. This includes partnering with classes (at Baylor and other institutions), archives (at the University of Texas and other institutions), and researchers. The ultimate impact of BuDAS will be making floor plan documents more accessible and useful and unlocking a data source for scholars that was previously difficult to access. Our NEH grant has provided a significant step forward on that path.
By improving the ease of collecting and analyzing information from floor plans, BuDAS will continue to open up new ways of studying the built environment. BuDAS systematizes the storage of plan information (database) and includes tools (extractor and annotator) to improve the accuracy and speed of data collection. This has the potential to transform the way humanities researchers (e.g., historians) use floor plan information, allowing them to complete more comprehensive analyses using larger numbers of plans and more data points. BuDAS helps researchers study buildings objectively, removing style/formalism. As we continue to expand BuDAS and share the tool, it will allow others to form new research questions and address current questions through a different lens. Additionally, BuDAS has the potential to make archival collections more accessible. Through our partnership with the University of Texas Alexander Architectural Archives and the William Allin Storrer Collection, we have already made an impact in this area. In accordance with our third project outcome, we also posted our initial analyses of Storrer’s Frank Lloyd Wright floor plans to the Texas Data Repository (TDR) (https://demo.dataverse.org/dataverse/BuDAS#). In the future, we plan to expand our presence on TDR, which at present is minimal, but the initial posting provides an additional venue to share about BuDAS and to link to the Storrer Collection. These connections are critical for building recognition and developing a community. Our project team plans to expand the use of BuDAS to other archival collections, and through academic conferences, we have been able to secure contacts with other interested universities. Since BuDAS is available without cost and open-source, we also plan for other archives and libraries to independently begin making the system available for use with their collections.

**Award Products**

BuDAS project website: https://sites.baylor.edu/budas/
Extractor and Annotator: https://github.com/linki0/BUDAS-extractor
Database and Analyzer: https://github.com/linki0/BUDAS/

Citations:

King, E., Wu, Q & Lin, D. (2020, March). Exploring Frank Lloyd Wright through the lens of 'big data.' Presentation at Interior Design Educators Annual Conference, Tulsa, OK.


Appendix A

BuDAS Work-flow

1. Work-flow overview
2. Extractor process
3. Annotator process
4. Database
5. Room relationship graphs
6. Analyzer
BuDAS Work-flow Expanded

BuDAS contains four main components:

1. **Extractor**: component that extracts information from floor plans (applies a five-step contour detection process to remove noise and detect rooms)
2. **Annotator**: graphic user interface that allows users to revise or add to extractor results
3. **Database**: SQL relational database that stores floor plan information and annotator results
4. **Analyzer**: analysis tools, including charts and visualizations, that are connected to the database
Extractor Process

Plan Source: William Allin Storrer papers, Alexander Architectural Archives, University of Texas Libraries, The University of Texas at Austin.

Floor plan selected from possible versions by extractor.

Note: the steps on this page occur automatically after starting the extractor (BuDAS.py). The user can change parameters, however, including the number of versions run by the extractor (default is 100).
Annotator Process

Step 1: Raw results from the extractor open in the annotator GUI

Step 2: Revisions to results using the “delete” and “draw rectangle room” tools

Step 3 (optional): Add room names (if removed by extractor), set scale, add doors and openings.

Typically, this takes from several seconds to a few minutes per plan.
Database

Importing Annotator Data into the Database

As a semi-automatic system, BuDAS relies on a combination of information entered automatically by the extractor and manually by the user. The most time-consuming and error-prone aspect of collecting floor plan information involves the collection of room information, which we automated with the extractor and annotator tools. The database brings the manual and automated processes together, storing both as well as providing analysis tools through the analyzer.

The user begins by entering basic information about the project manually to create a house profile. After running the extractor and saving the annotator results, a file of room information is automatically created. Users can import this information into the database using a standalone python program that is included in the package.

Relational database for the storage of plan information.
Database and analyzer: https://github.com/linki0/BUDAS/
Room Relationship Graphs

Room Relationships

A room relationship graph is automatically created for each house in the database. Room relationship information can be entered after the extraction results have been imported or done completely manually. The graph provides a visual overview of plan information and layout, using nodes (rooms) and edges (relationship between rooms). Nodes are colored according to function and sized based on relative room size.

Room relationships graph example.
Enter the tool by clicking on a completed house in the database, which will bring up the house information dialog box as shown here. Click on the room relation graph tab at the far right to bring up the graph.

Original room relationship tool (pre-grant)

Revised room relationship tool
Rooms (nodes) based on room size, stylized room relationships (edges), zoom tool, colored key
Analyzer Tools

Analyzer

Like the room relationship graphs, the analyzer tools are automatically connected to information in the database. We currently have three different analysis options: room space usage, room relationship comparison, and attributes.

Room space usage.
Select a room function (e.g., bathe, eat, cook, sleep, entertain) and a starting year to see how space is allocated over time.

Room relationship comparison.
Select two room functions and a starting year to compare over time.

Attributes.
Select an attribute (cost, size, etc.) and a starting year to compare over time.
Appendix B

Grant Products

1. Project website
2. Github
3. Texas Data Repository
4. Citations
BuDAS

Project Information

The Building Database and Analytics System (BuDAS) is an all-in-one floor plan recognition and analysis system. Upload scans or images of floor plans (jpeg) and use BuDAS to detect rooms, doors, room dimensions, and other plan information.

We employ a semi-automatic detection approach whereby users can revise or add to automatic detection results. This allows BuDAS to work on a wide variety of floor plans, regardless of symbols or drawing conventions. Currently, however, BuDAS is designed primarily for residential floor plans. The database is populated with information related to heaven, but the extractor (plan recognizer) should detect rooms for any building type.

How does it work?

1. Upload an image of a floor plan (jpeg)
2. Run the BuDAS Python script or use the executable option – extractor
3. Revise the extractor results in the annotator
4. Save the annotator results and then upload to the BuDAS database or a program of your choosing
5. If using the BuDAS database, use built-in analysis (energy) and search functions

Download:

Database and analyzer: https://github.com/roknBU/DUAS
Extractor: https://github.com/roknBU/DUAS-extractor

Funding:

BuDAS is made possible by the generous funding of the National Endowment for the Humanities and Baylor University’s Vice Provost for Research.

Project website homepage (https://sites.baylor.edu/budas/)
Github Links:

**Extractor and Annotator:** https://github.com/linki0/BUDAS-extractor  
**Database and Analyzer:** https://github.com/linki0/BUDAS/

Documentation, python packages, instructions, and other information needed to use BuDAS can be found through the above Github links. Need to download both in order to utilize all the features described in this document.
In accordance with our third project outcome, we also posted our initial analyses of Storrer’s Frank Lloyd Wright floor plans to the Texas Data Repository (TDR) (https://demo.dataverse.org/dataverse/BuDAS#). In the future, we plan to expand our presence on TDR, which at present is very minimal, but the initial posting provides an additional venue to share about BuDAS and to link to the Storrer Collection. These connections are critical for building recognition and developing a community.
