#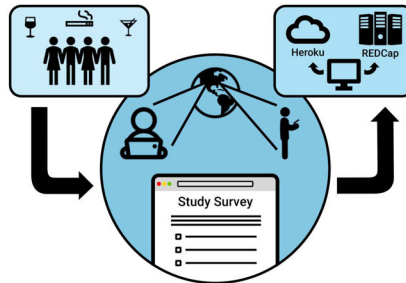 The Transformation of a Gold Standard In-Person Substance Use Assessment to a Web-Based, REDCap Integrated Data Capture Tool

**R Martin-Willett**[1], **Z McCormick**[2], **W Newman**[1], **L Larsen**[1], **MA Ortiz Torres**[1], **LC Bidwell**[1]

[1]The University of Colorado Boulder, Boulder, CO

[2]Trailblazing Technology, New York, NY;

## Graphical Abstract



## 1   Introduction

It is well known that perhaps the largest and most influential development in the research and delivery of health care has been the adoption of computer systems for gathering, managing, and analyzing health data. For example, the number of hospitals switching from paper to electronic health records increased nearly nine-folds from 2008 to 2015 [1]. Further, in health-related research, the data-management system *Research Electronic Data Capture* (REDCap) is used worldwide at nearly 3000 institutions with 745 thousand users since its creation in 2004 [2]. Overall, as these trends accelerate, the usage of pen-and-paper methods for collecting data and managing health records will increasingly be replaced by computerized methods.

One classic "pen-and-paper" assessment in health and substance use research is the *Timeline Follow-Back* (TLFB). The TLFB is the gold standard in self-reported substance use, developed in 1996 by Sobell et al. to assess alcohol consumption patterns and later other

Corresponding Author: Renée Martin-Willett, University of Colorado Boulder, Institute of Cognitive Science, CUChange Lab, 1777 Exposition Drive Ste. 222, Boulder, CO 80301, renee.martinwillett@colorado.edu.

substances such as marijuana or tobacco over discrete timeframes [3–9]. Traditionally this assessment is in the form of a structured in-person interview and utilizes a blank retrospective calendar, which the participant populates with events unique to their life, acting as a visual trigger for reporting on substance use.

As computerized health data management strategies advance, the TLFB has been modified by some research groups for use as a web-based assessment [10–12]. Online versions are thought to increase accessibility and privacy for participants and eliminate the need for participant travel [10], as well as reduce data collection time [11], given that the traditional administration of the TLFB is time-intensive for study staff. Despite these efforts towards digitization, many online TLFB assessments have limitations. Research teams have yet to develop an online TLFB that could simultaneously be leveraged across multiple studies with differing substance use measures and timeframes. There is also an increasing need to more widely assess substance use across specific domains as modes of administration shifts over time. For example, in states where cannabis can be legally purchased, there is a rapidly expanding market in alternative forms of administration beyond smoked flower or bud, including edibles, tinctures and oils, or concentrates. This trend is demonstrated by market statistics from the state of Colorado. Though general marijuana sales have increased 51.6 percent from 2015 to 2017 ($996 million to $1.5 billion), concentrated cannabis product sales increased by 114 percent and infused edible sales increased by 67 percent over the same period [13], demonstrating the increasingly diverse ways that cannabis is administered to the public, increasing the need to more accurately characterize cannabis use in legalized markets.

It is also the case that existing web-based TLFBs and other e-research tools do not utilize the latest developments in accessibility in web design, including features such as descriptive headers, a consistent navigation scheme, and informative error messages for incorrect text inputs [14]. Web-based TLFBs could furthermore be optimized to leverage the application program interface (API) capabilities of data capture tools such as REDCap, eliminating the need for secondary data entry by research staff. This optimization would utilize data flow automation practices, increasing the security and integrity of health-related data [15,16].

While the validity of the in-person format of the TLFB has been demonstrated [17], there has been considerably less research towards developing a reliable online TLFB assessment addressing the aforementioned limitations. As such, this paper describes the team-oriented process by which a new online TLFB (O-TLFB) was conceptualized, built, and implemented, including the technical details of integration with REDCap via API.

## 2  Method

The O-TLFB was developed to provide a platform for the capture of substance use data that could simultaneously and dynamically adjust to the parameters of multiple study protocols, while being optimized for web accessibility and streamlined data capture. Given the criticality of the online tool's ability to replicate the benefits of the validated in-person structured interview as closely as possible, we took a team-oriented approach to its development. The technical development team was fully integrated with the data collection

team from the onset of the project, to ensure that the insights from those who were trained and experienced in the administration of the in-person TLFB were assimilated into the design and implementation of the online version. Project initiation was divided into Conceptualization, Development, Implementation, Testing and Monitoring (Figure 1).

## 2.1 Conceptualization

The conceptualization process for the O-TLFB was guided by a transdisciplinary working group that determined the elements needed to make the O-TLFB function well both on the "front end" for participants, and on the "back end" for the research staff. The working group consisted of a clinical psychologist with expertise in the pharmacology and neuropsychology of marijuana use and who was the lead investigator for the project, a postdoctoral associate with expertise in behavioral neuroscience, a senior research team member with extensive experience in the administration of substance use assessments–including the TLFB–in a research context, three members of the study team that would be piloting the O-TLFB, two software developers respectively responsible for the development of the web application and the backend data flow pipeline, and a data management specialist. During the final development and implementation stages, two additional software developers devoted to the backend data flow pipeline joined the team.

The conceptualization process included regular in-person development meetings and conference calls over the course of eight weeks. De-bugging and refinement of the O-TLFB was managed using a Google Drive document in which the study team reported errors and queries, which were subsequently assigned to the developers as action items. Using these methods, the collaborative process by which the project was completed was well documented. These materials were subsequently utilized in the development of the 'readme' and video manual for the O-TLFB.

The research team identified nine core requirements for the O-TLFB that included the following:

- Secure and encrypted data collection and transfer

- Capability to capture a variety of substances, including alcohol, nicotine/tobacco, cannabis, prescription medications, and illicit drugs.

- Compatibility across multiple platforms including desktop computers, tablets, and smartphones.

- Intuitive for research participants to complete without the help of the research staff.

- The ability to track longitudinal data/multiple entries from each participant.

- Simultaneous use of the tool by participants from multiple studies, who would need to access different forms of the tool [i.e., including different substances (prescription drugs assessed in one study and not another) or different retrospective recall periods (one study needs to assess substance use over 14 days vs. 30 day recall for another study)].

- A backend interface enabling manual download, review, or modification of the raw data way that mirrored the format of the data in REDCap.

- Documentation of the technical specifications for the O-TLFB web application, enabling troubleshooting or modification of the O-TLFB in the future and/or for other research teams.

- Elimination of manual data entry on the researcher end through the automation of data transfer.

## 2.2 Development

In order to meet the nine core requirements from the working group the O-TLFB was built using a series of independent components wired together, otherwise known as a component-wise architectural approach (Figure 2). The first component is a form framework, designed for quick and easy iteration of participant facing survey content as well as allowing for variously complex hierarchical structure of forms. The second component is a secure storage engine, which allows for temporary storage/buffering of survey results on a secure database before transfer to an analytics tool. The third component is an extract-transform-load (ETL) interface, allowing export to REDCap (the study team's preferred data management/pre-analytic tool).

All of the components were written primarily using Python, in addition to HTML, JavaScript, and CSS. Python was chosen as the primary language for the project given that REDCap's API functionality is based on existing open-source Python code, as well as its increasing use in the biological sciences, specifically in neuroscience. The O-TLFB also relies heavily on the open-source Django web application framework. Alternative frameworks such as Flask and Pyramid were considered, however Django was ultimately selected due to its all-tools-in-one approach to web application development, as well as the rich ecosystem of open-source components for form rendering and manipulation.

**2.2.1    Form Framework—**The form framework refers to the participant facing content menus which allow them to endorse and select the various substances they have used. As many of the project goals revolved around the flexibility of various forms to accurately reflect and capture detailed substance use content, two types of forms were utilized: drill-down forms and response forms. Drill-down forms are used to organize multi-dimensional data and are displayed most often as a series of simple checkboxes or drop-down menus. Participant selections are processed by the server to determine which additional forms should be shown – either more drill-down forms or response forms. Response forms record the final data as determined by previous drill-down entries; these forms can be radio buttons, checkboxes, or free text. The combination of these two types of forms allow for diverse participant workflows with little modification to the software (Figure 3).

An example workflow for O-TLFB might be as follows: after first entering his/her unique code to initiate the assessment (described in 2.2.3), a study participant will encounter the general instructions for entry and a blank retrospective calendar. The participant makes entries into the calendar with "marker dates" and then proceeds to the next component involving entry of specific substances used on a given day. The participant selects their first

date in the calendar by clicking that rectangle, and the web application queries the study participant first on whether or not they used any substances within a selection of broad categories (in drill-down form) based on the protocol they are involved in. In this hypothetical example, the participant endorses prescription drug use on this day. Subsequently, the web application queries which sub-categories of drugs they endorse in a second drill-down form. If they chose a category like "sleep aid" for example, they would be presented with a final response form asking about the name and dosage of the particular medicine. Once this day's entry has been submitted, the corresponding calendar day is marked with a green check mark, and the participant can proceed to the next day's entry. A session can only be submitted upon completion of every date on the retrospective calendar, including entries for non-use. This example participant workflow is easily achieved using two drill-down forms and a response form – allowing the end user to skip response forms for irrelevant responses.

**2.2.2    Storage Engine—**The storage engine for the O-TLFB leans heavily on Django – an open-source web framework allowing for rapid development of web applications that render HTML, process forms, and store data in a commodity database such as MySQL or PostgreSQL. By using Django and adhering to its general usage guidelines, the O-TLFB can be hosted easily and inexpensively with Heroku, which fully supports deploying and maintaining web applications using Python, Django, and PostgreSQL. Ultimately this provides a number of best-in-class storage features – such as at-rest encryption, regular backups, and strong access control – with no additional cost or effort. As mentioned earlier, alternatives were considered, but the benefit of simplicity by using an all-tools-in-one solution like Django with a managed hosting provider like Heroku outweighed any benefits of using alternative frameworks like Flask or Pyramid, or alternative hosting providers, such as AWS or DigitalOcean.

**2.2.3    Extract-Transform-Load Interface—**The extract-transform-load (ETL) interface allows for the easy movement of the stored responses from the storage engine to REDCap. Once a participant workflow is completed, all of the raw data is uploaded to the Django database (it can also be extracted manually at this point). The next step is converting that raw data into a comma-separated-value (.csv) format, assigning labels for each individual data point, calculating aggregate values, and then making an API call to a REDCap cohort key table project to assign the data to a study group and participant ID. After the transform phase is completed, the data is uploaded to REDCap using a second API call (Figure 4).

While REDCap is the tool of choice for this research group, the ETL interface was designed to transfer data to another system, such as SAS or Tableau, requiring only trivial code changes. By following a component-wise architecture with regard to this interface, the lifetime cost of reusing the code for multiple studies should be reduced. More plainly, the only part of the code that would need to be rewritten in order to use the O-TLFB with SAS or Tableau would be the ETL interface component – the other components would require no modifications and would work in exactly the same way they work in the current implementation.

### 2.3 Security

Because of the sensitivity of the data being collected and according to standard best practices and the rules and regulations of the Institutional Review Board (IRB) for human participant research, participant's names are never used in the O-TLFB. Instead, users are asked to enter the last five digits of their phone number and the first letter of their first name. These inputs are immediately hashed using a format-preserving encryption algorithm that is based on the secure block-cipher method TDEA [18,19], and thus the phone number digits and the first letter of the participant's first name are never explicitly saved.

These inputs were selected because they are easy for participants to remember and are statistically unlikely to yield duplicates based off sample sizes typical of large biopsychosocial research studies (Equation 1). For example, if 1000 participants were using the O-TLFB simultaneously, there is a 0.01% chance that any two participants would have duplicate inputs (Figure 5). When uploading to REDCap, this hashed identifier is modified to include the date the test was submitted and a random 4-digit value, to ensure that if a user completed the assessment multiple times on the same day the PID would still be unique. (The PID is formatted XXXXXXXX-YYYY-MM-DD--RRRR where XXXXXXXX is the hashed identifying record ID, and RRRR is the random number assigned to them upon participant workflow completion).

Equation 1:
Probability equation for duplicate O-TLFB sign-in credentials

$$P(X_n = X_m) = P(X_n \neq X_m)^c \quad (n \neq m)$$
$$P(X_n \neq X_m) = \left(1 - \frac{1}{T}\right)$$
$$c = \frac{s(s-1)}{2}$$
$$P(X_n = X_m)|_{n \neq m} = \left(1 - \frac{1}{T}\right)^{\frac{s(s-1)}{2}}$$

Along with anonymization of identifiable data, the hosting provider (Heroku) seamlessly supports HTTPS with automatically generated and deployed security certificates. HTTPS is a standard practice adopted to protect confidential information and all submissions have mathematically guaranteed confidentiality and integrity when sent over the public Internet. Moreover, HTTPS is supported by REDCap at the University of Colorado Clinical and Translational Sciences Institute, thus data sent between the O-TLFB and REDCap, and the API tokens in for the REDCap O-TLFB and Cohort Key Table projects are only assigned to the Principal Investigator's REDCap user account.

The last level of security is on the hosting site. Although the code is hosted on a Heroku server, it is initially stored on a private GitHub repository. Any changes committed to the code are documented and stored in the private repository, where only collaborators can make changes and notations indicate which user modified the code. This helps to ensure an API token will not be stolen and changes cannot be made to the code without review. Moreover, any changes to API tokens or the code itself will be recorded in a built-in audit log.

### 2.4 Implementation

The implementation of the O-TLFB followed a standard software development lifecycle known as "agile", where the software requirements are defined as "user stories" composed of statements such as "as a user, I can select *prescription drugs* in the first form so that I can provide details about prescription drugs on the next page of the survey" and are iterated on to completion [20]. Moreover, an "agile" approach involves rapid feedback from the "product owner," in this case the research team conducting the study, so that aberrations are caught early and changes can be introduced before the product is fully complete. Notable requirements addressed during the agile testing and implementation process were multi-platform support (i.e. support for mobile, iPad, and desktop usage), the ability to inspect the raw data and re-export to REDCap, and a reliable error-management solution for confirming proper operation of the software when participants were using it in an unsupervised setting.

Multi-platform support was achieved by using a multi-page web application paradigm (the default paradigm for the Django web framework) combined with the Bootstrap 4 framework originally authored by Twitter. These tools are regarded as having extremely high compatibility with nearly all combinations of browsers, devices, and operating systems. Moreover, when Django is used as intended, an administrative interface is automatically generated, making inspection of raw data trivial to enable. Finally, Sentry was chosen as the tool/service for reliable error management (http://www.sentry.io), as it integrates easily with Django and provides a variety of useful features, such as real-time email alerts and roll-up of similar errors into concise reports, enabling fast analysis and deployment of bug fixes.

### 2.5 Testing and Monitoring

Testing was done via rapid feedback from the study administrators before deployment. After the initial deployment, testing used a "staging instance" strategy, or the use of a separate instance of the application hosted on an entirely different server and database than the currently-in-use application. This way, the team was able to roll out updates without any negative consequences to any current study participants using the tool. Some changes that were developed because of this testing environment include adding the capability to remove prescription medication as an option for clinical studies and modifying the PID to prevent multiple consecutive entries from a single user to be regarded as duplicates, and thus not transferred to REDCap.

Continuous monitoring was achieved via integration with Sentry for immediate feedback on errors, as well as with periodic monitoring of REDCap to ensure that valid O-TLFB data was uploading. To this end, Sentry was able to help the team quickly discover errors related to browser and device-specific behaviors that were not caught in initial testing. The quick discovery and high level of detail provided by Sentry led to very short response times for fixing such errors, as opposed to human-centric feedback forms, which require the user to remember exactly what he/she did to cause the incorrect behavior to happen.

After completing testing and deployment, the development team devised a standard operating procedure for the ongoing monitoring of the O-TLFB. It was determined that a two-step functional testing protocol would be executed on the system monthly to ensure the

quality of the data being uploaded. In the first step, a study team member completes the O-TLFB with a dummy 5-digit-number (i.e. 11111, 12345) and confirms that the results uploaded to REDCap match the entries in the Django "backend." This dummy data is immediately logged to document the audit and deleted from the REDCap project and Django "backend" once accuracy is confirmed.

The second step is a comparison of the number of completed records between the Django backend and the REDCap project for a given audit period. If the values do not match, then data may have been uploaded to REDCap multiple times or not uploaded at all. The raw data should always upload to the Django backend when a participant makes a complete submission, thus the study team can quickly ascertain if an expected entry (for instance, at a pre-defined protocol time point) has not been recorded in Django due to a user error (such as forgetting to select "submit" after entries are complete), or if the error lies in the interface between Django and REDCap. If the error is from the latter, Django is configured so that any uploading errors can be remedied by manually pushing data to REDCap by selecting the participant's missing data and selecting 'Resubmit to REDCap'.

In addition to systematic testing for technical errors, the study team conducted a validation study among undergraduates at the University of Colorado Boulder of the OTLFB as compared to in-person administration of the TLFB, and also in comparison to other established measures of substance use such as the Alcohol Use Disorders Identification test (AUDIT) [21] and the Marijuana Dependency Scale (MDS) [22]. The results of the validation study are described elsewhere [23].

## 3 Results

Since development and validation of the O-TLFB, the tool was implemented in the context of two large-scale, federally funded studies of *Cannabis* use on anxiety or chronic pain respectively.

The first time a participant sees the O-TLFB in the context of these studies, it is administered through a web browser on an iPad. A research assistant loads the OTLFB, hands the iPad to the participant, and they are asked to double enter in the last five digits of their phone number and first letter of their first name to create a new record and verify participant identity. The research assistant reviews the instruction page with the participant, explaining that the O-TLFB is a tool used to document substance use over a particular timeline, and explains the concept of marker dates. Once the participant has entered in any applicable marker days, the research assistant is there for any questions while they otherwise make the remainder of their entry independently. All subsequent entries are initiated through follow up REDCap surveys sent directly to participant email addresses that redirect to the web address for the O-TLFB, which the participant can then complete on their own time and without the help of a research assistant. For one study utilizing the O-TLFB, participants have been 87.8% compliant at a one-month follow-up in completing their entries within the designated window.

Following implementation of the O-TLFB, Google Analytics were initiated to assist the study team in identifying usage patterns. As of the development of this manuscript, there are approximately 350 entries in the O-TLFB, entered over the course of approximately 189 days. The O-TLFB is used on average between 10–11 times per week, and between 1–2 times per day. During this same period of time, the O-TLFB has not had any duplicate participant inputs (five digits with first letter of first name), security breaches, or REDCap upload errors.

## 4    Discussion

This new O-TLFB stands to benefit health-related research in numerous ways, as well as be a model for interdisciplinary development of future tools in other fields. This new assessment is easy to use for research study participants and for research staff; it eliminates the need for manual data entry, leveraging free or inexpensive and widely available resources, lending itself well to eventual open-source sharing.

The "front end" of the O-TLFB follows best practices for user accessibility across multiple platforms such as phones, tablets, and computers, thus reducing the burden on participants making self-reports to the study team and hopefully contributing to participant retention. The visual layout of the assessment is attractive and intuitive, combining best practices in web-based accessibility with flexible drill-down and response forms in a way that is familiar to most users.

The O-TLFB also eliminated the need for manual data entry by research staff while simultaneously increasing data quality and security. Prior to its implementation, the administration and subsequent coding of in-person TLFB data was among the most labor-intensive processes across studies in the research center. While it is a limitation of all online tools that respondent errors can almost never be tracked and/or corrected, the use of the O-TLFB does eliminate the possibility of transcription errors by research staff as well as increases the security of the data through multiple layers of encryption and common-sense security practices (such as annual password updates and automatic log-off from REDCap or Heroku).

Finally, the O-TLFB leverages free or inexpensive and widely-utilized resources such as Heroku, GitHub, and REDCap. The use of GitHub and Heroku is immensely useful for maintaining the O-TLFB. With proper branching and distributed workflow, the development team is able to push changes that can be reviewed before merging with the master branch. With such a workflow, a better understanding of each change is documented and reversible. Furthermore, in cases when the code does not run on Heroku but on a local machine, Heroku will automatically redeploy the application after a push to master and any resulting errors can be correlated to the changes pushed to the GitHub. Given this exhaustive documentation process, it is significantly easier to trace errors and deploy bug fixes.

The development process described here included limitations. While a traditional agile development process enabled rapid change to the O-TLFB software during development, the need to rapidly provide features sometimes came at the expense of structure and

extensibility. This is classically labelled as "tech debt" within the software development community and is akin to monetary debt in the sense that future changes may require more effort than if the features had been developed without shortcuts (i.e. "paying interest" on tech debt). Thus, the current iteration of the tool does have some limitations that would be beneficial to address in the future.

To begin with, the current infrastructure of the O-TLFB relies on hard-coded questions and language to support the different variants for hiding or displaying data entry sections or for customizing the number of days in the submission. In other words, a developer must update the code of the O-TLFB overall in order to make minor changes. A more thorough implementation of this feature utilizing the "Admin Site" functionality of the Django framework and the migration from hard-coded field variables to the PostgreSQL could provide increased flexibility for the research team to customize the O-TLFB directly for specific study purposes, without need for developer assistance. Additionally, the current method for re-uploading data to REDCap after a modification of the assessment infrastructure could be improved upon to allow data to be easily uploaded to REDCap projects. A short-term goal for the study team is to pursue continued improvement of the O-TLFB server as a separate model from the rough data, where the update of REDCap data after a modification could be streamlined.

Overall, as health-related research continues to shift towards increasing reliance on technological tools, it is important that new methods of assessments are developed in a collaborative and interdisciplinary way that accounts for not just the perspectives of software developers and engineers, but of health researchers and study participants. There is great potential for the O-TLFB described here to further expand substance use categories or even improve on other aspects of the O-TLFB functionality, benefitting the field of behavioral and health research overall through broad application. Furthermore, though clinical applications were not considered in the development of the O-TLFB, there is nonetheless a potential use for this application towards improving mental health care into integrated health care settings. While the backend of the O-TLFB is currently optimized for the export of data for research analysis, it could potentially be modified for easy data viewing by clinicians seeking to learn more about their patient's substance use patterns. Towards this end, longer term goals include open-sourcing the code developed by the team for broad application in health-related research and even potentially clinical applications.

## Acknowledgments

## References

[1]. Henry J, Pylypchuk Y, Searcy T, Patel V, Data Brief 35: Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008–2015, 2016 https://

www.healthit.gov/sites/default/files/briefs/2015_hospital_adoption_db_v17.pdf (accessed August 2, 2018).

[2]. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG, Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support., J. Biomed. Inform 42 (2009) 377–81. doi:10.1016/j.jbi.2008.08.010. [PubMed: 18929686]

[3]. Harris KJ, Golbeck AL, Cronk NJ, Catley D, Conway K, Williams KB, Timeline follow-back versus global self-reports of tobacco smoking: a comparison of findings with nondaily smokers., Psychol. Addict. Behav 23 (2009) 368–72. doi:10.1037/a0015270. [PubMed: 19586155]

[4]. Hoeppner BB, Stout RL, Jackson KM, Barnett NP, How good is fine-grained Timeline Follow-back data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level, Addict. Behav 35 (2010) 1138–1143. doi:10.1016/j.addbeh.2010.08.013. [PubMed: 20822852]

[5]. Robinson SM, Sobell LC, Sobell MB, Leo GI, Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use., Psychol. Addict. Behav 28 (2014) 154–162. doi:10.1037/a0030992. [PubMed: 23276315]

[6]. Phan O, Henderson CE, Angelidis T, Weil P, van Toorn M, Rigter R, Soria C, Rigter H, European youth care sites serve different populations of adolescents with cannabis use disorder. Baseline and referral data from the INCANT trial, BMC Psychiatry. 11 (2011) 110. doi:10.1186/1471-244X-11-110. [PubMed: 21749677]

[7]. Aalto M, Alho H, Halme JT, Seppä K, The alcohol use disorders identification test (AUDIT) and its derivatives in screening for heavy drinking among the elderly, Int. J. Geriatr. Psychiatry 26 (2011) 881–885. doi:10.1002/gps.2498. [PubMed: 20661878]

[8]. Magee SR, Bublitz MH, Orazine C, Brush B, Salisbury A, Niaura R, Stroud LR, The Relationship Between Maternal–Fetal Attachment and Cigarette Smoking Over Pregnancy, Matern. Child Health J 18 (2014) 1017–1022. doi:10.1007/s10995-013-1330-x. [PubMed: 23892790]

[9]. Ouimette P, Read JP, Wade M, Tirone V, Modeling associations between posttraumatic stress symptoms and substance use, Addict. Behav 35 (2010) 64–67. doi:10.1016/j.addbeh.2009.08.009. [PubMed: 19729250]

[10]. Rueger SY, Trela CJ, Palmeri M, King AC, Self-administered web-based timeline followback procedure for drinking and smoking behaviors in young adults., J. Stud. Alcohol Drugs 73 (2012) 829–33. [PubMed: 22846247]

[11]. Pedersen ER, Grow J, Duncan S, Neighbors C, Larimer ME, Concurrent Validity of an Online Version of the Timeline Followback Assessment, Psychol. Addict. Behav 26 (2012). doi:10.1037/a0027945.

[12]. Wilks CR, Lungu A, Ang SY, Matsumiya B, Yin Q, Linehan MM, A randomized controlled trial of an Internet delivered dialectical behavior therapy skills training for suicidal and heavy episodic drinkers, J. Affect. Disord 232 (2018) 219–228. doi:10.1016/j.jad.2018.02.053. [PubMed: 29499504]

[13]. Orens A, Light M, Lewandowski B, Rowberry J, Saloga C, 2017 Market Update, Denver, Colorado, 2018.

[14]. Kirkpatrick A, O Connor J, Campbell A, Cooper M, Web Content Accessibility Guidelines (WCAG) 2.1, W3C. (2018).

[15]. Dunn WD, Cobb J, Levey AI, Gutman DA, REDLetr: Workflow and tools to support the migration of legacy clinical data capture systems to REDCap, Int. J. Med. Inform 93 (2016) 103–110. doi:10.1016/j.ijmedinf.2016.06.015. [PubMed: 27396629]

[16]. Yamamoto K, Ota K, Akiya I, Shintani A, A pragmatic method for transforming clinical research data from the research electronic data capture "REDCap" to Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM): Development and evaluation of REDCap2SDTM, J. Biomed. Inform 70 (2017) 65–76. doi:10.1016/j.jbi.2017.05.003. [PubMed: 28487263]

[17]. Robinson SM, Sobell LC, Sobell MB, Leo GI, Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use., Psychol. Addict. Behav 28 (2014) 154–162. doi:10.1037/a0030992. [PubMed: 23276315]

[18]. Barker E, Mouha N, Archived NIST Technical Series Publication Archived Publication Series/ Number: Title: Publication Date(s): Withdrawal Date: Superseding Publication(s) Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher Recommendation for the Tripl, (2012). doi:10.6028/NIST.SP.800-67r2.

[19]. Black J, Rogaway P, Ciphers with Arbitrary Finite Domains, n.d

[20]. Beck K, Beedle M, Van Bennekum A, Cockburn A, Cunningham W, Fowler M, Grenning J, Highsmith J, Hunt A, Jeffries R, Kern J, Marick B, Martin RC, Mellor S, Schwaber K, Sutherland J, Thomas D, Manifesto for Agile Software Development Twelve Principles of Agile Software, (2002). doi:10.1177/0149206308326772.

[21]. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG, The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care, World Heal. Organ (2001) pp1–40. doi:10.1177/0269881110393051.

[22]. Stephens RS, Roffman RA, Curtin L, Comparison of extended versus brief treatments for marijuana use., J. Consult. Clin. Psychol 68 (2000) 898–908. doi:10.1037/0022-006X.68.5.898. [PubMed: 11068976]

[23]. Martin-Willett R, Helmuth TB, Abraha M, Hitchcock LN, Lee K, Bryan AD, Bidwell LC, Validation of a Multi-Substance, Dynamic, and Accessible Online Timeline Followback Assessment, Under Rev. (2018).

The Timeline Follow-Back assesses self-reported patterns of substance use

As substance use patterns shift over time the need for an adaptive assessment grows

We developed and validated a new web-based version of the Timeline Follow-Back

This new tool adapts to changing use patterns and is optimized for data security

This new tool can be leveraged broadly across research and clinical domains
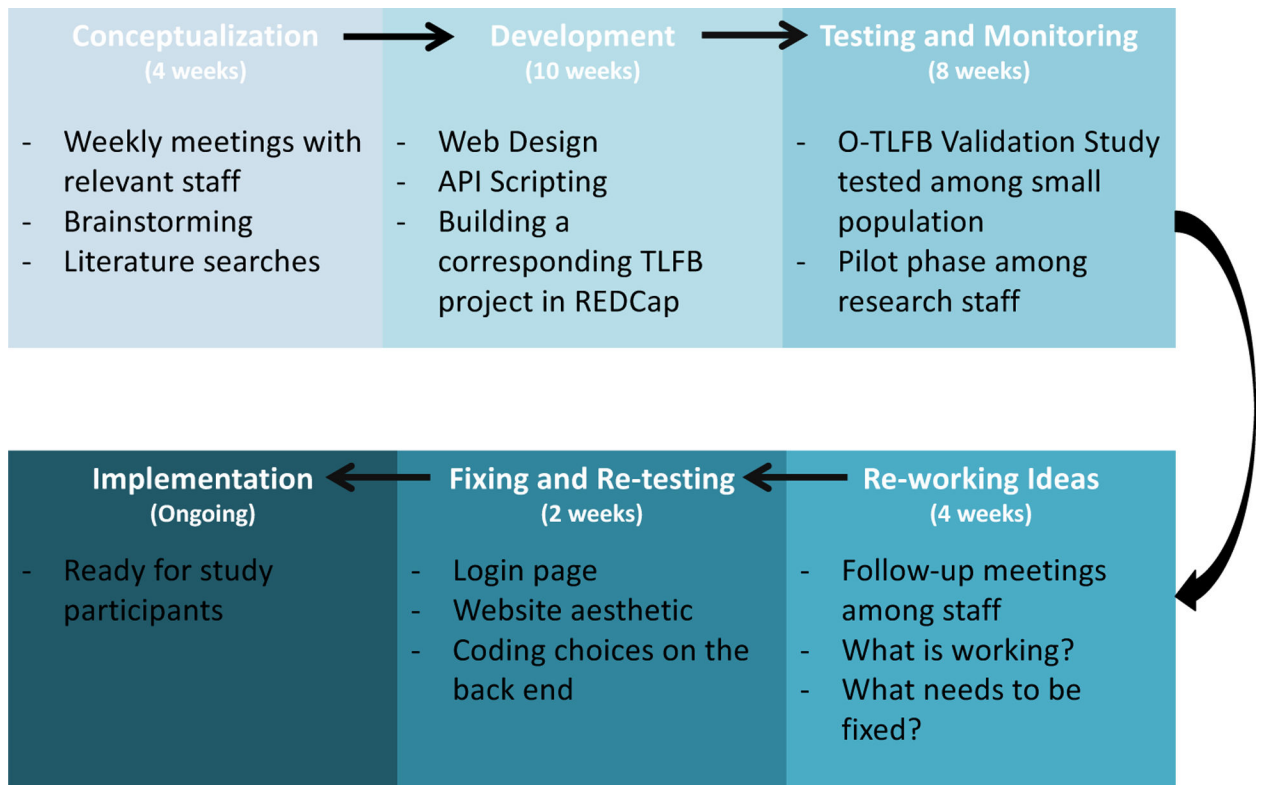
**Figure 1:**
Project Development Flow
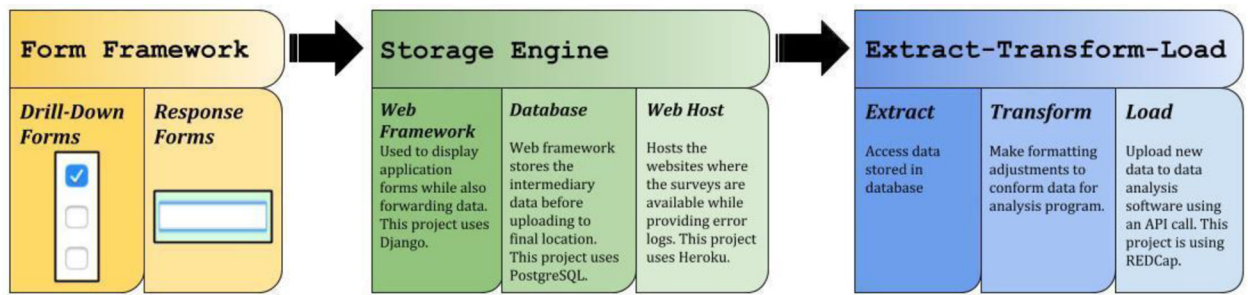
**Figure 2:**
Component-wise architectural approach for the optimized O-TLFB
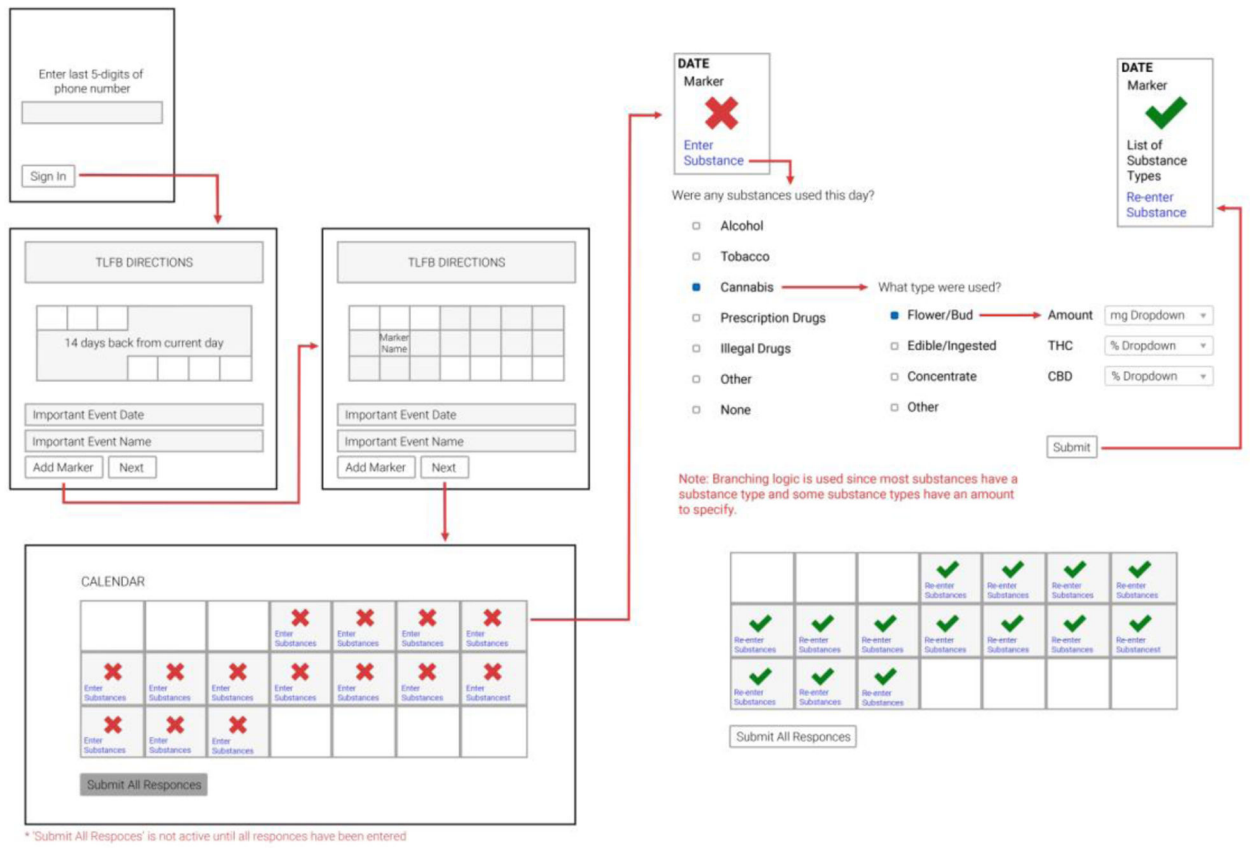
**Figure 3:**
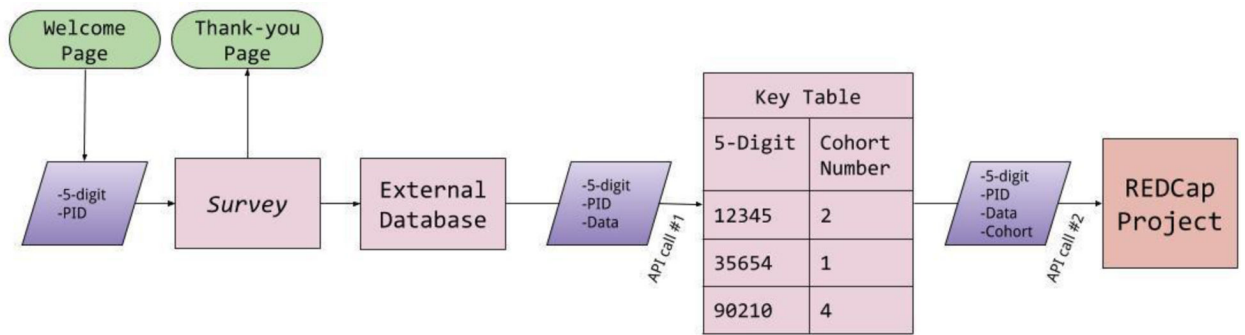Form framework wireframe for the O-TLFB

**Figure 4:**
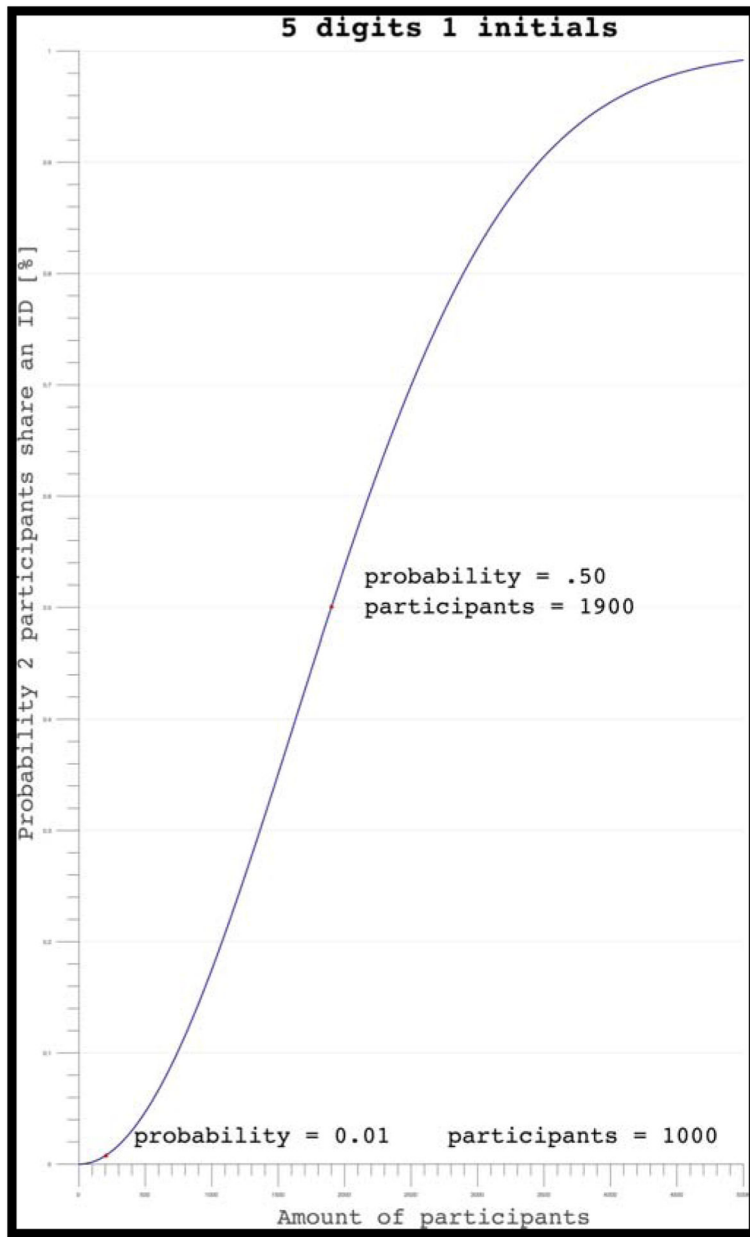Extract-Transform-Load (ETL) interface for the O-TLFB

**Figure 5:**
Probability graph for O-TLFB sign in credentials