Salesforce Data Integrity Plan

Executive Summary

Data Integrity is a measure of perceived trust and confidence in the veracity, quality, credibility, accuracy, prompt delivery, and plausibility of information JobTrain stores in Salesforce.

Automated Salesforce tools used to maintain and measure data integrity are input data validation and duplicate control rules, record types, page layouts, and pick lists; automated generation of unique identifiers in data records, data integrity reporting and analysis software packages; and Salesforce monitor and optimizer reports, and logging tools

Manual processes to maintain and quantify data integrity are documented inspection and reconciliation procedures; inspection of Salesforce system monitoring and data quality reports; writing user documentation, training users, and using test environments

The quantitative baseline measure of data integrity is figures published in the JobTrain annual report over the last ten years. If figures in these reports can be verified quickly using data from Salesforce then data integrity is good

The qualitative baseline measure of data integrity is perceived trust and confidence. Information JobTrain publishes, based on what is in Salesforce must be perceived with trust and confidence from critical constituencies: senior executives, teachers and counselors, clients, donors, auditors, directors, funders, local employers, and the people who live in this area.

Operational measures of data integrity are perceptions of user satisfaction, data completeness, quantities of missing information, data validity and reconciliation, ease of use, how promptly information requests are serviced, size of the back log, and the number of completed and pending service requests

Data security is a strategic mission-critical element of data integrity. A publicized breach of confidential client information would cause long-term damage to our credibility and sterling reputation.

JobTrain does well with user support and capitalizing on Salesforce automation to measure and improve data integrity

JobTrain must improve data security, protecting the personally identifiable information, service request turnaround time, hygienic data management, and use more automation

Continuous improvement. Improving data integrity is a continuous processes. Information in Salesforce changes every day. Data integrity is not static: it is either getting better or worse. *The "Plans" section of this document on page 33 enumerates initiatives planned and under way to continuously improve data integrity at JobTrain*

Salesforce Data Integrity Plan

Introduction and Document Purpose

This document is the JobTrain Salesforce Data Integrity plan. It explains what data integrity is and what JobTrain does to make it better.

There are definitions in this document of what data integrity is and how data integrity is measured. There is scope and limit to the definition. There are high-level descriptions in this plan of automated and manual procedures in place and planned to quantify and improve data integrity. Human factors are critically important to data integrity planning. Data security is critical to data integrity strategy. Measurement, and improvement of data integrity are tightly coupled processes and use many of the same automation tools and operating procedures.

There are established business practices at JobTrain that are effective in improving and measuring data integrity, such as effective use of automation and well-designed manual inspection procedures. There are areas where data integrity process and procedure is deficient or absent, ands need to be fixed, such as data security, training, user documentation, self-service, and service turnaround time. Automation is under-used.

Information in Salesforce at JobTrain changes every business day. Improving and measuring data integrity is a continuous practice. Plans for improvement are explained in the context of JobTrain's strategy for maintenance and continuous improvement of data integrity.

There are automated and manual measures in place to ensure and measure data integrity. Too many processes are manual, not enough are automated. In some cases automated procedures are redundant or overkill. Automated and manual processes to ensure and quantity data integrity are explained in this plan. In some cases, as illustrated in the Salesforce Optimizer report, too many data integrity controls are in place. This impedes performance and is inconsistent with best practice.

The summary and conclusions enumerate what JobTrain does well, and not well, with regards to data integrity; and outlines plans and projects to improve data integrity.

Audience: This plan is written because the CEO asked for it.

Scope. This plan is an executive level description of data integrity practices, measurements, deficiencies, continuous improvement strategy, and future plans.

There are limits to the technical detail of this plan, and there are definitions of critical constituencies whose perception of data integrity is important.

These limits are:

This is a management plan and not a technical document. Automated and manual processes to promote and measure data integrity are enumerated and described in this document. They are not explained in technical or operational detail. There are software user guides, HowTOs, self-documenting features in Salesforce, training videos, optimizer reports, data quality dashboards, and other materials that document technical details of Salesforce data integrity management controls that are active.

The constituencies whose perception of data integrity is important are:

- The five most senior JobTrain executives

- The eleven instructors and counselors at JobTrain

- Board of directors

- Private donors

- Government agencies that give us money

- The general public in San Mateo and northern Santa Clara counties.

While JobTrain is a large and diverse social services business, we are primarily and educational institution. The people whose service delivery is most important, and whose perception of data integrity matter, are those who *teach*.

JobTrain has a large service footprint but for purposes of measuring perceptions of data integrity it is important to know the attitudes of people who live in immediate proximity to our primary service area, which is in San Mateo and northern Santa Clara counties

Data Security risks, risk estimates, mitigations, and plans in this document are restricted to Salesforce data security, and not to a more generalized practice of corporate data security. This document

enumerates data security risks and mitigations as they apply exclusively to Salesforce.

Definition

Data Integrity is a measure of perceived trust and confidence in the veracity, quality, credibility, accuracy, prompt delivery, and plausibility of information JobTrain has in Salesforce.

Data integrity is based on people's perception. JobTrain enjoys a good reputation in the community. It is based on trust. Figures published in the JobTrain annual report are believed. Broken Pathways received significant favorable publicity in the print and broadcast media. When JobTrain quotes numbers, people believe us.

The JobTrain Salesforce Administration team contributes to that sustained reputation, serving as custodian of information that is believed true and verifiable.

Human Factors

Who? Data integrity is measured by perception and attitudes of people. Perceptions that are important are those of senior executives at JobTrain, counselors and teachers, directors, professional auditors, clients, private funders, government entities that give us money, and the general public in San Mateo and north Santa Clara counties.

Human factors are relevant to data integrity planning because data integrity is a perception, human error causes flaws in data integrity, data integrity controls and measures are designed and interpreted by people, and people manage and use data in Salesforce. *Perception (of data integrity) is reality.*

People cause errors. Errors in data integrity are caused by people. Computers do not make mistakes. Mistakes, by people who are educated, well-intentioned, dedicated, and hard-working; in data entry, system design, execution, data security, interpretation, and decision making; are what causes error and flaws in data integrity.

Some human errors are direct, such as mistakes or omissions keying information into Salesforce, either through direct data entry or programmatic loading of data in quantity.

Sone human errors are indirect. Flaws in design of objects, data entry screens, and reports lead to inaccurate information. Incorrect analysis of reports can lead to wrong conclusions and bad decisions. Accurate and correct information on a report can be inappropriately interpreted leading to bad decision making. Accurate and correct reports can be accurately and correctly interpreted, and flawed decisions can still be made, either through bad decision making, or the absence of decision making because people are too overloaded with work to make thoughtful decisions. Even automated and manual processes intended to reinforce data integrity can have flaws, produce misleading results, or lead to erroneous interpretation.

Bias. There are many ostensibly objective and factual ways to measure and verify data integrity, such as consistency testing, data validation, completeness measures, statistical results, and audits. The problem is that objective measures are designed and selected by people. This means that all measures of data integrity, directly or indirectly, are subjective and reflect the bias

and perspective of people participating in or directing the process or making decisions.

Mitigation. The best method of minimizing bias from measures of data integrity is to employ and engage the participation and input of multiple independent reviewers who employ their own independently designed measures. For this reason, measuring perceived data integrity relies on input from multiple important constituencies: senior executives, teachers and counselors, outside auditors, clients, directors, government auditors, and salesforce administrators.

Using independent teams is the strategy for mitigating bias. The best way to remove harmful bias from the evaluation of data integrity at JobTrain is employ and engage multiple independent teams of knowledgeable, educated, and responsible people who can articulate a candid and unvarnished opinion.

People are the source of errors and perception is reality when assessing data integrity, people maintaining data integrity, and improvements in data integrity mean improvements to what people do. This is implemented using independent auditors to compare the annual report to information in salesforce, privately and independently

survey senior executives, teachers and counselors, directors, funders, government agencies that give money, and the public living in San Mateo and north Santa Clara counties.

Strategy

This section explains key elements of JobTrain's data integrity strategy.

Key strategic elements in this plan are

- Reliance on automation

- Use both subjective and quantitative measures

- Measure perceptions of data integrity from multiple independent teams

- Statistical facts in the JobTrain annual report are the baseline for quantitative measurement of data integrity

- Conformance with Salesforce recommended best practice

- Best practices in data security that protect the confidentiality of information in Salesforce and that protect the reputation of JobTrain.

Salesforce is an innovative and powerful tool for managing information. Most of its capabilities have not been put to use, and precious little of it is understood by people who work here. There are automated, self-documenting, and analytic features of salesforce that will be capitalized on further in calendar year 2018.

Key elements of the JobTrain data integrity strategy are:

Reliance on automation. A key element of this data integrity strategy is to fully exploit capabilities in Salesforce to prevent errors, enforce completeness and consistency, and deliver information instantly. Salesforce has automated capabilities to prevent bad information from getting into the system, schedule the production of critical daily, weekly, monthly, and annual reports; trigger backups, and automate consistency checking that is done manually.

Automation of all periodic reporting for users. There is a conspicuous periodicity to the reporting requirements of users and senior executives at JobTrain. Information requirements are easily predicted. There are monthly operational reports needed by teaching and instruction teams. There are annual performance figures that go into

the JobTrain annual report that is one of the benchmarks of data integrity. If information requirements can be anticipated then automation can and will be used to deliver access to that information to the managers who need it.

Self-service. Users are self-sufficient in accessing information that is needed to get their job done. People do *not* wait for someone else to complete a service request and then furnish spreadsheets or hyperlinks that require further analysis.

Instant service. Users have immediate access to information needed to do their job and to answer pressing questions from senior executives and grant writers. The "old" days of making a request and waiting for announcement of completed service are no longer part of how JobTrain does business. Salesforce has the capabilities to deliver information immediately and will be implemented to that end. Dashboards presenting company performance, operational status, and data quality are implemented and being improved.

Go get it, don't receive it. The strategy for delivering information services is to help people *go into salesforce* and use information, *rather than pull it out* to do

analysis, further transformation, and formatting for appearance.

Reduce handling and transformation. Human error is the cause of errors and flaws in data integrity. The strategic objectives of automation, self-service, and using Salesforce for reporting and analysis are designed to significantly reduce opportunities human error caused by people when people transform, manipulate, or otherwise handle information.

Critical constituencies. Data integrity is a perception. The people whose opinions matter are a discrete and finite list: Five senior executives, eleven teachers and counselors, donors, auditors, directors, funders, local companies that hire graduates, and the general public in south San Mateo and north Santa Clara counties

The authoritative quantitative benchmark of data integrity at JobTrain is the last ten annual reports. Figures in these reports are compared to what is immediately accessible from Salesforce. If the information can be fetched promptly and the numbers match, data integrity is good enough. Otherwise not.

Customer satisfaction. Having internal customers who are satisfied with technical support, user service, and the ability of Salesforce to immediately deliver information is a strategic objective in this data integrity plan. The software has the capability of immediate delivery. Customers of any product or service are more likely to trust it, believe in it, look for ways to fully capitalize on it, and support changes and improvements.

Coupling of measurement and enforcement. Automation and manual processes to enforce and measure data integrity are in most case the same tool and the same process. Manual consistency checks are designed to enforce integrity, and results of these efforts are a measure. Validation rules and work flows are designed to prevent entry of incorrect information enforce data integrity, and enumeration of execution instances is a measure. The tight coupling if measurement and enforcement is not an element of strategy, but rather it is recognition and acknowledgement of technical fact: Data integrity enforcement and measurement are coterminous.

User education. Data integrity is maintained and improved if users are well educated in the use, administration, and data security of Salesforce at JobTrain. Internal user training is developed and

delivered consistently and periodically. A variety of educational media and knowledge transfer mechanisms are used, such as formal stand-up classroom instruction, written user documentation, training and instructional videos, context-sensitive help, instant messaging, and more. There is explicit evidence that user knowledge and education levels need to be elevated, evidenced by how painfully long some things take to get done, lack of immediate turnaround on reporting requests, a sizable backlog of service requests that are not started, and common user complaints about limitations of Salesforce.

User satisfaction is a key element of data integrity. Complaints and indications of dissatisfied users typically are legitimate and reflect deficiencies in Salesforce, or its implementation, or the ability of support engineers to provide prompt and adequate service. Satisfaction and education go together. Knowledgeable users are typically more satisfied and productive; and data quality is improved. Disgruntled users do not care as much about data quality.

Conformance with Salesforce Best Practice. Salesforce has written white papers documenting best practices in a wide variety of areas in system management. The recommendations and best practices in these white papers all seem to be well thought out and reflect the

accumulated knowledge and expertise of a small army of software engineers at Salesforce. It is the strategy of JobTrain to rely on and mimic the Salesforce documented best practices in system management data quality.

Data security is critical to perceived data integrity because a publicized breach can damage our reputation for credibility. Best practices in data security are to know all the risks, estimate probabilities and consequences, and have mitigation and contingency plans in place.

Automation

This section explains Salesforce automation tools in effect to improve and measure data integrity.

Automation forms a small, growing, and not-big-enough part of enforcing and measuring data integrity. Salesforce offers a substantial arsenal of powerful data integrity enforcement and measurement tools. *The primary and very effective strategy for using automation to enforce data integrity is preventing bad data from getting into the system in the first place*. Automation is also used to improve data integrity by: 1- Remind people to keep records current as conditions change, and 2- Flag inconsistencies with "in-situ" data.

Salesforce automation tools in use today at JobTrain that ensure and measure data integrity are:

"Enemy at The Gates"* There is substantial deployment of Salesforce automation tools that stop incorrect and incomplete information from getting into the system in the first place. In Salesforce parlance these tools are called validation rules, work flows, visual work flows, duplicate record control rules, Salesforce processes, automated error notification, and triggers. They are all mechanisms that alert people to information being entered that is wrong, and to impose a hard stop.

∗ 1973 novel by William Craig about the 1942-43 siege of Stalingrad

These various obstacles to entry of invalid or incomplete information are applied to all the critical information entered into Salesforce:

- Clients
- Donors
- Class schedules
- Students in classes

- Students who graduate
- Students who get work
- Social services
- Youth services

Duplicate record prevention. Most of the Salesforce duplicate controls work autonomously and automatically, and are not the result of an intentional deployment. Presence of duplicate records in some cases, such as donor related tasks, funding opportunities, student attendance records, and classroom enrollments simply cause the system to immediately lock up, freeze, and grind to a screeching halt. Issue notification therefore is not much of a problem. These facts tend to make most problems with duplicate records self-evident with some alacrity, and diagnosis of root cause, based on experience to date, is easily traced. It could be argued that this is an example of a data integrity flaw induced by human error. However, regardless of definition, it is an effective method of flagging and causing the removal of certain duplicate records.

White papers published by Salesforce and various consulting companies suggest duplicate error rates below two percent of population is a healthy indicator. While the duplicate record rates at JobTrain are a fraction of

that (less than 50), the duplicate records that remain have proven to be an intractable problem that significantly interferes with users getting work done such as client intakes. There are about 45 duplicate contacts and 45 duplicate accounts. Attempts to merge and then purge these duplicates fail because there are about 160 anonymous account records that are criteria matches for Salesforce standard duplicate control Rules.

Anonymous accounts cannot be arbitrarily deleted because they have donor opportunities associated with them. Before the anonymous accounts are deleted the donor opportunities must be reconciled, deleted, or reassigned to active accounts. After this is done the remaining duplicate accounts and contacts can be merged and purged.

Pick Lists. Salesforce pick lists are used whenever possible in data entry screens. This minimizes human error by presenting pre-defined values for fields with limited allowable values.

Page layouts. Salesforce page layouts and record types present data entry screens to users that are restricted to fields that need to be entered. Default values are used on data entry screens when possible. The layouts support

context-sensitive help with explanations of users about when needs to be entered.

Data typing. Salesforce field definition tools help enforce entry of correct information. Salesforce supports a large collection of data types to define fields, such as date, date and time, formula, index, short or long text, system-generated unique identification numbers, automated summaries, binary flags, telephone number, latitude and longitude, encrypted text, and web page address. Using data types helps to automatically minimize and mitigate human error when entering formation.

Automated notification Errors are made known to users through immediate messages posted to a screen. Email messages of certain perceived serious errors, such as process failures, are sent to a team of system administrators.

Documentation. Details of the implemented data integrity controls are documented in Salesforce and in internally written user documentation. Automated processes, validation rules, duplicate controls, work flows, and error notifications are self-documenting to some extent. Readers of this plan with further interest in

the details and workings these processes they can view technical details from hyperlinks in the appendix of this data integrity plan.

Reminders. Information in Salesforce changes every day, and if the reality of our students and clients is inconsistent with what is recorded in Salesforce then this is a data integrity flaw. Salesforce automatically generates reminders, mostly to counselors and teachers, to keep critical client information updated.

Examples of continuously changing information that must be kept current to maintain data integrity are students, who have:

- Graduated, whose graduation is not recorded

- Attended a class on a certain date, whose attendance is not recorded

- Found a job, whose employment is not recorded

A critical data integrity flaw in this area is documenting student outcomes. JobTrain does not keep adequate records of what happens to students after graduation. Reports from Salesforce are now automatically generated monthly to remind people to get this information and keep it current in Salesforce.

Reports of internal consistency and completeness. A recent consolidation of information from disparate systems into salesforce has resulted in exposing redundancies, inconsistencies, incompleteness, and other errors not previously detectable when information is housed in different systems. A series of reports, some completed as of this writing and others in design, exist in order to flag inconsistencies and incomplete information.

Data Integrity Dashboards are deployed to help quantify and measure the data integrity state and condition of in-situ data. These graphs and charts present statistics about numbers of validation rules, populations and percentages of empty fields, enumeration of missing fields, missing and incomplete data, record ownership.

Salesforce Optimizer report is used to help measure and quantify the data integrity state of the database. It contains information that alerts system administrators to conditions that may trigger outages or instances of information becoming inaccessible to users because of error conditions, license expirations, and exceeding licensed limits on resource use.

The Salesforce Optimizer further lists issues with data integrity:

- Deployed and unused resources such as validation rules and work flows that are intended to improve data integrity that are not used or never executed.

- Unused fields, page layouts, reports, dashboards, permission sets, and page layouts.

- Applications making calls to outdated and soon to be phased versions of the Salesforce API

- Excessive numbers of deployed validation rules that should be consolidated or eliminated to help improve system performance

- Resources that use features, such as notes and attachments, that Salesforce will shortly phase out

- Non-conformance with Salesforce recommended best practices

- Number of custom fields on standard objects, which are very high in some cases

- There are three files with more than 20 validation rules. Salesforce considers this excessive

- How many people have privileged user profiles, which in the case of JobTrain is way too many

Automated logging. Most logging features in Salesforce relate to date security and are explained in the Security section of this plan. There are two logging features in Salesforce used for data analysis and troubleshooting and form useful automated tools to enforce and measure data security.

- **Duplicate records** detected by Salesforce at data entry are logged. While most of the implemented duplicate control rules obstruct data entry, there are cases where contacts of similar names or addresses are admitted into the system, following prompts and messages to a user. In these cases the duplicates are automatically logged and available for reporting, analysis, and correction

- **Debug logging** is activated normally for purposes of troubleshooting and error diagnosis. This is a useful data integrity enforcement tool allowing prompt detection and correction of errors.

Data Security

Salesforce data security is a self-evident mission-critical element of data integrity strategy. JobTrain stores significant amounts of personally identifiable information (PII in data security parlance) about clients and employees, including social security number, income, address, offender status, etc. A publicized breach of client information would damage JobTrain's reputation for credibility, and it would take years to restore that reputation.

Data security practices at JobTrain are trusting and not strict. There are modest internal restrictions to donor information. There is room for improvement. The probability of a data security breach is low but the consequences would be severe and long-lasting.

The process of managing data security is:

- Enumerate risks and potential exposures

- Estimate probability of failure

- Estimate severity of failure

- Document plans for mitigation and contingency action

- Data security education and training for all employees

- Deploy logging tools to record system activity and audit user actions

The enumerated data security risks for Salesforce at JobTrain are:

- An unauthorized user gets a user's password, logs into the system, and downloads client personally identifiable information (PII)

- An unauthorized user is able to see client data

- An unauthorized user is able to see a client social security number

- An employee laptop, that has cached Salesforce access credentials, is lost or stolen

- Student interns who have access to the production instance of Salesforce

- The Eventbrite and Mobile Cause integrations allow entry and creation of duplicate client and account records

- An unauthorized employee gains access to donor records

Data security initiatives in place and planned, with the objective of protecting our information asset and improving data integrity, are as follows:

Access. JobTrain enforces common and typical industry standard requirements concerning presentation of access credentials; password length, re-use, complexity, and expiration; session timeout, and login failures. There are three individuals who are consultants and not JobTrain employees, who log in remotely, and whose activity is closely watched. Accounts of people who leave the company are immediately suspended.

Separation of Donor and Client pools. The only real separation in side Salesforce is protecting access to information about donors and contributions. Teachers and counselors have limited access to any information about donors and donations. Donor Management people at JobTrain do not have acc

Single sign-on is not used, and should be. Single sign-on is generally considered in the industry as a safer form of privileged data access and security compared to using separate credentials for accessing different resources.

Ease of use. If security is easy to do, then users do it; if not, they may not. Single sign-on would require users to present credentials once to access all company resources, including Salesforce. This would significantly alleviate a constant user frustration and complaint. Currently users present credentials for access to the JobTrain network, email, Salesforce, and other internal

and external resources. With engineering and system changes a single sign-on to all protected resources can be implemented.

Logging. Several forms of logging exist in Salesforce that are used to improve and measure data integrity

- **Field history tracking.** Changes to critical information about clients and accounts is recorded. Salesforce calls this "Field history recording"

- **Session logging.** All user logins and sessions are recorded. Excessive logins and off hour logins are considered a data security red flag

- **System changes** are logged

- **Message logging.** Email messages generated by Salesforce, such as error notices or messages generated by automated processes are logged

- **System security health check.** Salesforce system security "Health Check" reports identify risks, exposures, deviations from Salesforce best practices

- **Login forensics.** Salesforce login forensics is an advanced analytical tool in Salesforce that does forensic analysis of user login and data use patterns. While this tool is interesting, the small

and intimate size of JobTrain's user base suggests this tool may at some date be helpful but right now it is worth knowing about but not using.

Encryption Salesforce encrypts data in transit using Transport Layer Security (TLS) versions 1.1 and 1.2. This is considered strong encryption and an industry standard in cloud computing. The Salesforce database is hosted at a Salesforce data center and is not under JobTrain's internal control. Security practices at Salesforce data centers are well documented, and are considered strong and reliable.

Test Environment. The integrity of user data in Salesforce at JobTrain is protected with the use of test environments for developing changes and improvements that affect uptime or data. Salesforce calls these environments a "sandbox" and are precise copies of the live environment with limited subsets of data. Sandboxes have been used to develop major system changes such as introduction of contact notes, purging duplicate accounts and contacts, changes to entry of Services data, implementing the Salesforce State and Country Pick List feature, merging the morning and afternoon sessions of the ASE and HSE classes, deleting most of the attendance history data, and testing packages downloaded and installed from the Salesforce App Exchange

Installed packages. There are about 17 add-on packages downloaded from the Salesforce app exchange installed into Salesforce. This represents a security risk and a data integrity exposure. There is also an external third-party package installed that accesses, updates, changes, and deletes user data for purposes of merging and purging duplicate records. Installed packages represent a security risk, especially those, about eight, not written by Salesforce. This risk is mitigated by installation and testing in a test environment before deployment, and careful monitoring.

Event Brite customizations admit duplicate records into the system. Salesforce was customized to allow enrollment and contact updates from the Event Brite web site. This customization includes coding of programmatic bypasses to standard duplicate control records. The result is periodic introduction of duplicate client records. The volume of records is small and there are reports and messages that make this known. The workaround and mitigation is to manually inspect and correct problem records.

Mobile Cause customizations do not collect enough client information to create an adequately populated

client record. The result is periodic automation failures when attempting to create new contact records. The error messages and failure notifications product intimidating and nasty messages that look like critical system failures but in fact are merely notices that an automated attempt to create a contact record failed. While this is not a direct data integrity failure, the messages are cause of constant concern and interest, and the Mobile Cause code needs to be improved in order to capture more information about contacts.

Backup and restore. All user data is backed up automatically once a week, and more frequently as conditions warrant, such as introduction of major changes or planned deletion of unused data. The accessibility of this archived data is tested by restoring it to a sand box and making sure it is accessible and intact.

Manual Processes

There are manual operating procedures in place to ensure, improve, and measure data integrity. This includes inspections and check lists, running reports, verifying changes to data that reflect changes in the student population

Periodic inspection of data for inconsistencies and anomalies includes the following checks:

- At the beginning of each fiscal year (in July and August, after year-end reporting), check for new placements having Position Start in the previous fiscal year; they will not be counted as placements for the new fiscal year. Check with job developers: might they be able to add newer placements that can be counted for the new fiscal year (since those client's placements missed being counted for the previous fiscal year)? If so, the vocational training enrollment references in Graduated From should be moved to the newer placements.
- Check that a JD record exists with a Graduated From value corresponding to each enrollment having the JD/Position checkbox checked.
- Check enrollment for change in completion status (no longer a graduate)
- Watch out for placements with no wage that are really just internships.
- Check for JD Outcome "None planned" where the client is eligible for work and Contact Notes fail to give a good explanation.
- It may be necessary for JD Outcome to be changed to "Interim outcome" due to a subsequent vocational training enrollment within one year of graduation.

- If JD Outcome starts with "Placed" and Graduation From exists, check for:
  - Occupation missing
  - Hourly Wage missing
  - Hours Per Week missing
- Check for training related placements for which a training related occupation is not selected. These must be explained in Contact Notes.
- Check for placements for which Position Start and Position End are the same. These must be explained in Contact Notes.
- Check for unusual Hourly Wage.
- Check for Job Retention data that is needed.

System reports. Daily inspection and monitoring of:

- Error notification email messages
- System overview page
- Security health check report
- Data quality dashboard
- User logins
- Salesforce optimizer report
- Field update history
- System update history

Backup archival. After a scheduled automated backup is complete it is a manual process to move the archive data to a secure folder on an internal file server

Training. Development and delivery of periodic user training is a manual process for improving data integrity. The relevance and importance of user training is written in the Strategy section of this plan. The most effective user training seems to be the daily one-on-one help that people get when they call for help with a specific issue.

Documentation. Giving internal customers information about the operation and use of Salesforce is a manual process for improving data integrity. The chronic problem with documentation in the application software industry is that it is static and people do not read it. Documentation that exists today at JobTrain is a written user guide that nobody reads and training videos that nobody watches. Context-sensitive on-line help visible from screens that people are using gets occasional viewing.

Easter eggs. Planting easter eggs in user documentation and training videos is a tool to find out if anyone reads the documents or watches the videos. User documentation that is read and training videos that are viewed would reflect positively on the state of data integrity at JobTrain. So far, easter eggs planted in user documentation in August 2017 have apparently not been

found, nor have any planted in training videos been discovered.

Plans

Plans and actions to improve data integrity are listed below in order of priority. The most immediate tasks are using Salesforce to prove the last ten annual reports are accurate, and survey the attitudes of critical constituencies regarding data integrity at JobTrain. There are also plans to improve data security through user education, implement additional Salesforce capabilities, and continued user training and education.

The plans are:

What	Why	When

1. Run reports from Salesforce that produce figures on the JobTrain annual reports for the last ten years	Most critical quantitative data integrity validation	10 days
2. Survey attitudes and perceptions of key constituencies regarding perceptions of data integrity	Most critical qualitative data integrity validation	3 months Negotiating volunteer work with a Stanford undergrad class
3. Hire outside party to audit data security and data integrity	Key qualitative and quantitative measure, best practice	? Costs real money unless we find volunteers
4. Write **Data Security** HowTO for all internal users	Critical data integrity improvement	10 Days

5. Single Sign On	Data security, ease of use, and customer satisfaction	3 months
6. Implement **web to lead**	Improve of use, and customer satisfaction	1 month
7. Activate Salesforce State and Country pick list feature	Data integrity: Standardize representation of state values, ease of use, customer satisfaction	3 months
8. Install package to auto-complete city and state based on zip code.	Data integrity, ease of use, customer satisfaction	6 months

$600/year license |
| **9. Implement USPS API** call to retrieve city and state from zip code | Data integrity: Standardize representation of state values, ease of use, customer satisfaction | 12 months |

10. Improve follow-up reminder automation so that counselors keep information about graduates current and updated	Data integrity: automation, faster turnaround, ease of use, customer satisfaction	1 month
11. Improve documentation with the intended result that people use it	Ease of use, customer satisfaction, employee education	6 months Ambitious plan
12. Write index of all manual operating procedures that are used for data integrity improvement.	Data integrity	3 months

13. Explore Salesforce Knowledge and CRM Content as a possible repository for more suitable and usable user documentation	Data integrity, ease of use, employee education	3 months
14. More education and training for users	Employee education, satisfaction; greater autonomy, faster turnaround	6 and 12 month milestones
15. More education and training for users and administrators	Faster turnaround, customer satisfaction	6 and 12 month milestones
16. Eliminate unused resources documented in the Salesforce Optimizer report	Salesforce best practice	1 month
17. Reduce number of validation rules in effect	Salesforce best practice	1 month

18. Continue studying Salesforce best practice documentation	Salesforce best practice	3 months
19. All report requests are serviced immediately	Salesforce best practice, data integrity, customer satisfaction	1 year
20. Automate manual processes for reconciling inconsistences in in-situ data	Data integrity, more automation, less manual procedure	1, 3, 6, and 12 month milestone tasks
21. Find more data quality reporting tools		3 months
22. Log all service requests, measure volume and turnaround time	Best practice, system management	1 month

23. Reconcile and delete the 160 anonymous accounts	Data integrity, duplicate management	3 months; ongoing
24. Merge and purge remaining duplicate account and contact records. There are about 85 of these records.	Data integrity, duplicate management	3 months; ongoing

Summary

Data integrity is a measure of perceived data quality.

There are qualitative measures of data integrity. Data integrity is a measure of trust and confidence in the accuracy and truthfulness of information JobTrain stores in Salesforce. There is a discrete identifiable constituency of people whose perceptions of this confidence are critical.

There are quantitative measures of data integrity. The strategic quantitative measure of data integrity at JobTrain is whether figures in the last ten JobTrain annual reports can be accurately and immediately substantiated with information extracted from Salesforce.

Improving and measuring data integrity is a continuous process because information in Salesforce changes every day.

The most immediate threats to Salesforce data integrity at JobTrain are:

- Automatic generation of duplicate contact and account records by Eventbrite and Mobile Cause

- Lack of completeness. Client contact, demographic, and job development information is limited and not complete

- Limits to the knowledge and capability of administrators, impeding prompt completion of support tasks and service requests

- Limits to the knowledge and capability of internal users, impeding prompt completion of work

The strategy for improving and measuring data integrity is reliance on automation tools in Salesforce software, education and training, and best practice white papers written by Salesforce (the company)

Successful data integrity management is accomplished using a huge arsenal of automated and manual practices, designed with the intent of controlling human error.

Direct and indirect flaws in data integrity are caused by human error.

There are data integrity initiatives planned and underway intended to improve user education and satisfaction, use more automation, deliver information to users more rapidly, and enhance security. These planned improvements are listed in the previous ("Plans") section of this plan.

Reference Information

Most of the reference information in this plan came from Salesforce, IBM, college text books on Security and Quality Assurance, Arthur Anderson, McKinsey, and

National Bureau of Standards publications about data security.

Salesforce publishes a generous amount of documentation, references, tutorials, and white papers. There are numerous white papers from Salesforce presenting best practices in system implementation, data quality, and how to deal with duplicate records.

http://www.salesforce.com/assets/pdf/misc/BP_Implementation.pdf

Achieve Outstanding CRM Administration from Salesforce
https://help.salesforce.com/articleView?id=000007548&type=1

Links to the JobTrain Salesforce User's Guide and training videos are visible on the Useful Links list visible from the Home tab in Salesforce. https://jobtrain.my.salesforce.com > Home Tab > Expose "Custom Links"