A CRM Practitioner's Guide to

Data Migration Survival

Anosh Wadia
CRM Project and Technical Manager
XTIVIA Inc.

A Publication of
XTIVIA

TABLE OF CONTENTS

About the Author **04**

Introduction **05**

1. ISSUES AT THE SOURCE **07**
A. Data is not easily accessible or exportable 08
B. Data/Database format is non-standard 08
C. Data has issues/needs cleansing 09
D. Lack of knowledge of the source system 09

2. ISSUES AT THE DESTINATION **10**
A. Data may not match the destination system and needs further massaging 11
B. Validation rules may not be met in the new system 11
C. Data could be truncated at the destination 11
D. Lack of knowledge of the destination system 12

3. DATA TRANSFER ISSUES **13**
A. Data has to be mapped accurately to the destination 14
B. Breath of the data increases integration completely 14
C. Depth/Volume of data increases integration complexity 14
D. ETL Tool limitations/capabilities may require multiple tools and/or import procedures 15

TABLE OF CONTENTS

4. TESTING ISSUES **16**
A. Good sample test cases must be selected..... 17
B. UAT test cases have to be outlined..... 17
C. Depth/Volume of the data can make testing challenging 17
D. Lack of knowledge of the source system 18
E. Scope Creep..... 18

5. ONGOING MAINTENANCE **19**
Ongoing Maintenance 20

SUMMARY..... **21**

THANKS FOR READING..... **22**

About the Author

Anosh Wadia is a Project Manager at XTIVIA and has been working with Customer Relationship Management (CRM) systems for over 12 years. He was involved with over 100 CRM implementation/integration projects and holds developer certifications in both InforCRM/SalesLogix as well as Salesforce.

Anosh has a Master's degree in Management Information Systems, a Graduate Certificate in Telecommunications Management as well Logistics & Supply Chain Management. In his spare time, he enjoys listening to classical music and long rides on his motorcycle.

Introduction

Jim was excited! It was his first day working at the new startup firm downtown his friend Chris had been telling him all about for weeks—and finally convinced him to join the team. They had played college football together, and Jim was looking forward to the new opportunity and working with his friend.

“Jim, let's dive in and get you acquainted with our systems. There’s so much I have to show you!” said Chris. “All of our clients are in our custom-built CRM application, which you can access to via this icon on your desktop. We’ve got all the billing handled by our accounting system on our shared network drive. If you want sales numbers, we can pull those out of our transaction database on our ERP system. We do our time reporting in our HR/Payroll system—which is not in-house, but is cloud-based and hassle-free. So there you have it—all this information is at your fingertips!”

“Great!” Jim exclaimed. “That’s wonderful! Now, the first thing I’d like to do is run an analysis of who our top 10 clients are by sales volume—as well as dollars—and figure out how much time our sales team spends converting those opportunities into closed business. Is there a way I can pull that up?” Chris thought to himself, “Hmmm… well, not really. Accounting does provide us with some annual sales numbers by the customer in a report, so maybe you can start there?”

Introduction (continued)

Jim replied, “Well, yes I could, but I doubt that will tell me how long it’s taking our sales team to convert those opportunities into business. Wouldn’t it be ideal if we had a single, centralized system with all this data integrated so we could run our analysis in one place instead of having to look up multiple sources of information?”

If the story above sounds familiar, it’s because you’ve probably asked the same question Jim did at least once in your career. All organizations, large or small, run into this issue. New applications are constantly added to leverage core benefits and improve business processes, but making them talk cohesively to the rest of the enterprise data and systems is the challenge.

At XTIVIA, we pride ourselves in getting two disparate systems to talk to each other successfully. With numerous successful data migrations under our belt, our team has overcome major pain points and common issues encountered in such endeavors. This e-book outlines these challenges and will provide some helpful suggestions to make your next integration/migration project a great success.

1

Issues at the Source

A. Data is not easily accessible or exportable

Often, the data is on a legacy system which cannot be exported easily to a standardized format like CSV or a delimited file. An API may not exist to extract the data. Finally, due to legacy systems or security reasons, the data may not be accessible at all to a standardized ETL tool.

Recommendation: Determine the accessibility and export capabilities of the source data prior to estimating the scope and effort for the integration project as well as delivery timelines.

B. Data/Database format is non-standard

Sometimes the data can be exported when it is not in a usable format and therefore has to undergo an interim process. For instance, the export may comprise of multiple CSV files which have to be joined and related together which tends to be very slow, particularly for large data sets. In such a case, it is more efficient to import them into a database format prior to running an integration. One may also find the data export uses a legacy format that requires additional conversion and correction prior to import, adding another layer to the process.

C. Data has issues/needs cleansing

After the information is extracted, only part of the battle is won. The data usually requires some form of cleaning and transformation. The source data may contain duplicates and may need to be de-duplicated prior to import. The data might not be in a normalized format and may need to undergo an interim normalization process prior to final import. Additionally, the data may need to be transformed further to match the destination system (for example, values in a single field may need to be transformed to a 1 to many tables in the destination system). These efforts add significant time to a project.

D. Lack of knowledge of the source system

An often-overlooked pitfall in the integration process is a lack of knowledge of the source system. It is useful to have as much insight regarding the source system and its data as possible to ease its migration to another system. Without this, one could spend hours trying to locate and figure out where the correct data resides and in what format the data is in—when a person familiar with the source system could point this out in a matter of minutes. However, the reality in many organizations is legacy systems exist with little to no internal knowledge about their workings due to employee turnover and lack of documentation. This reality is often why the organization opts to migrate to a new system.

2

Issues at the Destination

A. Data may not match the destination system and needs further massaging

Data usually has to be transformed from the old system to the new. For instance, pick lists or selection lists might exist on the old system which will not match the new system and will need to be matched up in the mapping document. The destination system may also have the same data in a completely different format and will need conversion and/or transformation. In some cases, entirely new tables and data objects will need to be created in the destination system to house the data. Finally, the issue of duplicates could surface at the destination as well. The data may already exist at the destination and the incoming data may need to be merged/de-duplicated. Sometimes, due to a change in the data structure between the source and the destination, a duplicate could surface which did not exist in the source data to begin with.

B. Validation rules may not be met in the new system

Depending on the methodology used to add information to the destination system, the data may or may not meet the validation criteria of the destination system which could result in errors in that system subsequent to the import. Hence, this should be a vital aspect of the QA testing process.

C. Data could be truncated at the destination

The field lengths and data types in both systems may not match and that data can possibly be truncated. Again, the mapping document is the key to avoiding some of these pitfalls in advance.

D. Lack of knowledge of the destination system

Just like with the source system, having good knowledge of the destination system and its data objects is critical. When sufficient information regarding the destination system is lacking, invalid or incorrect data imports, or even worse, data corruption and errors on the destination system can result. Hence, it is absolutely vital that the mapping document and subsequent data be reviewed by people knowledgeable with the target system.

3

Data Transfer Issues

A. Data has to be mapped accurately to the destination

An elementary mistake is to jump in and start working on the integration without clearly defining a mapping document the stakeholders agree upon. The mapping document is the bible the integration must follow. A good mapping document has tables and fields (including data type and length information) from both the source and destination listed as well as any transformation instructions for each field. It also dictates the frequency of each integration run for each dataset.

B. Breath of the data increases integration complexity

“Breath” implies the number of columns of data in each integration. While the number of columns itself does not automatically increase complexity, it does, however, increase the time taken to correctly map each column as well as test the data integration. Thus, this is a significant factor in estimating the cost of an integration.

C. Depth/Volume of data can cause issues

“Depth" refers to the rows or number of records of data. In most cases, the depth does not affect integration estimates and costs because once the mapping is done, the number of rows run is not as much of a factor. However, depth *does* factor in when the dataset is sizeable. Large datasets can lead to integration time-outs, data connector timeouts, or just cause the integration to run for excessive periods of time. In such cases, strategies need to be implemented to piece-meal the data import in stages or smaller chunks, or more aggressively, segregate the data so that only delta values are updated instead of entire records.

D. ETL Tool limitations/capabilities may require multiple tools and/or import procedures

Not all ETL tools are built the same and they have varying abilities and features. Know the capabilities of the tool being used and what its limitations are. Sometimes, it may be easier to transform the data outside of the ETL tool rather than having the ETL tool itself handle the transformation. In other cases, post-processing data after it has been handled by the ETL tool, rather than having the ETL tool handle certain changes, may be more efficient.

4

Testing Issues

A. Good sample test cases must be selected

After all the data migration is complete, select good test cases to validate the data has come over successfully. This is vital. A good test case is one with complete data in all fields and related objects/data entities. Other good test cases to select are records that involve all the possible business process variances the data encompasses. For instance, when integrating Sales Orders, have a test case of a Completed Order as well as a Cancelled or Pending Order.

B. UAT Test cases have to be outlined

Aside from actual data, pick actual business processes to be tested when it comes to an integration. Merely transferring data from one system to another and validating said data may not always be sufficient. Using the example above, if a Sales Order is processed as completed and then later canceled by the customer, the entire business process use case has to be considered from the aspect of the integration so the two systems accurately reflect the same information.

C. Depth/Volume of the data can make testing challenging

With large data sets, trying to review most of the data to determine accuracy is impossible. In such cases, statistical probability approaches should be taken to determine how many actual records ought to be tested. This will determine the degree of confidence the data is migrating across successfully.

D. Serious issues found during testing can cause lengthy re-imports

The QA testing of data is often the most overlooked time factor when estimating the time and cost of an integration. A data migration could take hours or even days to complete, and once QA tested, issues could be uncovered, causing the data to be purged and reimported again. Also, this is often the first time users can see their data in the new system and are likely to bring up issues and concerns after their first glance at the data. So it is critical to accommodate sufficient re-runs of data during the testing and QA process.

E. Scope Creep

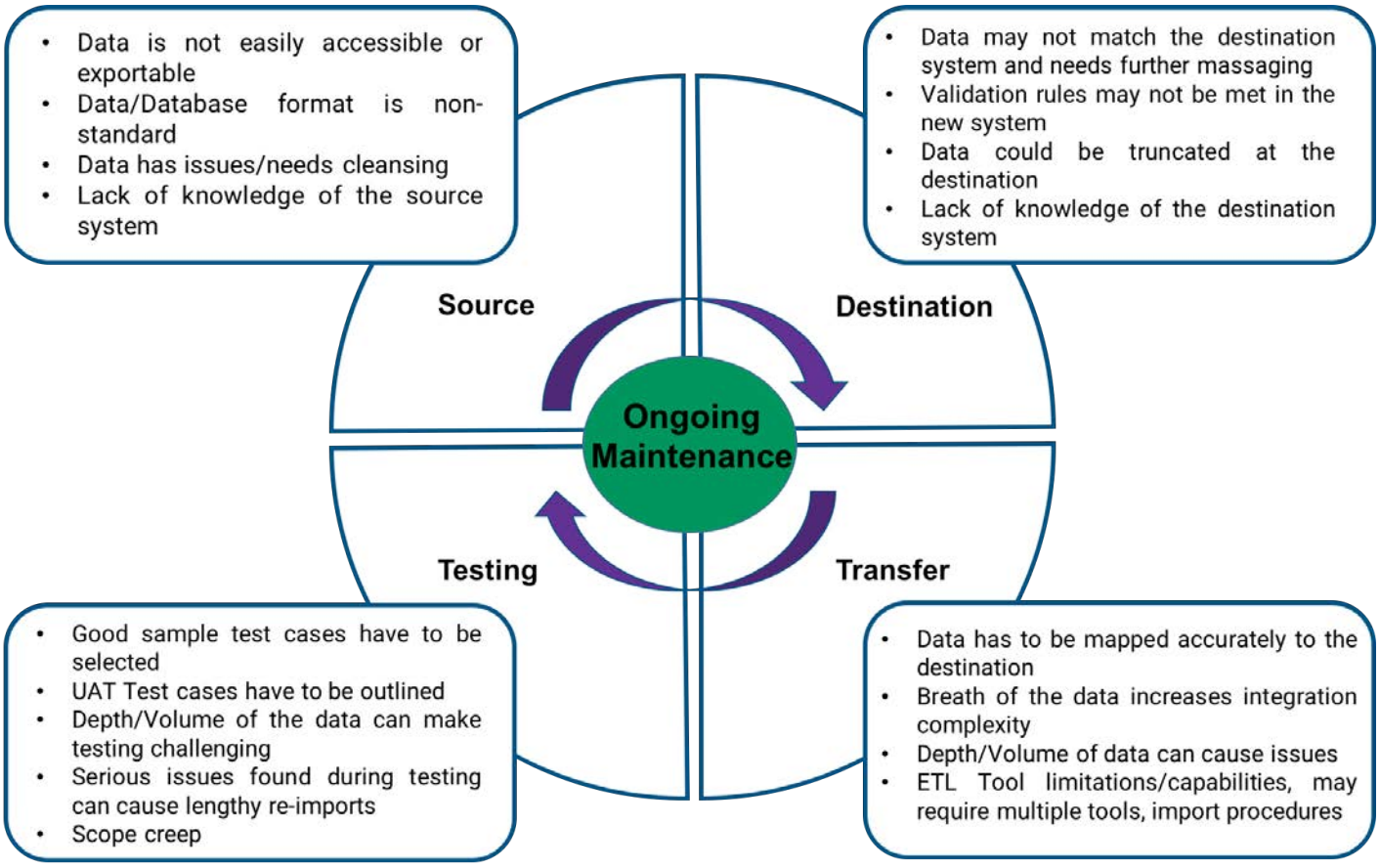
As mentioned in the point above, the QA and UAT process is often the first time users get to see the data in the target system, which they may or may not have used before. Therefore, they may have not had a chance to envision what the data would look like during the initial discussion and data mapping process. Scope creep is very likely to come in during this process since users may bring up additional requirements for the data integration. Factor in a certain degree of scope creep, particularly for these situations.

5

Ongoing Maintenance

Ongoing Maintenance

Once the testing process is complete and the integration is running in a production environment, continuous monitoring is crucial. Make sure to tweak and improve upon this production environment periodically. Reference the chart below:



Summary

Let's get back to our friends Jim and Chris. Their organization decided integrating their CRM and accounting systems provided a wonderful insight into their business revenue and forecasting. They could now track sales life cycles accurately, project their expected annual sales, reward their frequent customers with discounts, pinpoint products with poor turnover and also track products with low return rates and gauge customer satisfaction. Getting here wasn't easy, but it started with discussing the key objectives, planning and implementing a comprehensive mapping document for the stakeholders to agree upon, and then proceeding with developing and testing an integration for their CRM. As a result, it had the most successful return on investment for their startup to date! Riding on this success, next on their organizational roadmap is the integration of their CRM and HR/Payroll systems which hope to provide them with even more insight into their salesmen's work day and product specific sales cycles.

Data migrations and integrations can be challenging! They require great care in planning, execution, testing, and ongoing monitoring. In most cases, project managers underestimate the time and effort necessary to import data. One can easily fall into the trap of assuming that data migrations are a matter of a simple export and import and everything should work fine. In reality, it is rarely that simple and complexities surface only once the mapping document is comprehensively outlined and completed. Without this step being done, XTIVIA often sees actual import times 2-3 times longer than the original estimate.

Armed with the knowledge above, I hope you can avoid some of the common pitfalls regarding data migration, and most importantly, XTIVIA hopes your next data integration project is a victory!

Thanks for reading

A CRM Practitioner's Guide to Data Migration Survival

Need Data Migration Guidance?
Call 1-888-685-3101 ext. 2
www.xtivia.com