

Journal Title: Business Information Review

Article Number: 616235

## SUBMITTED MANUSCRIPT

N.B. When citing this work, cite the original published paper

Morton, J., Beckford, J. & Cooke, L., 2015. Towards a Rosetta Stone for translating data between information systems. *Business Information Review*, 32(4), pp.220–230.

## **Towards a Rosetta Stone for translating data between information systems**

### **Abstract**

Information systems are an important organisational asset and offer numerous benefits. However, organisations face continued challenges when upgrading aging information systems, and the data contained within, to newer platforms. This article explores, through conversations with information systems professionals in four organisations, the potential development of a 'Rosetta Stone', which can translate data between systems, and be used to help overcome various challenges associated with their modernisation. Despite mixed feedback regarding the Rosetta Stone concept from interviewees, solutions highlighted in literature combined with participant feedback presented theories for its development, primarily as a tool to enable meaningful interpretation of data, rather than direct translation. The conclusion reflects on data collected to recommend a framework for how the tool might be developed, and has potential to be of significant interest to practitioners, open-source communities and organisations.

### **Keywords**

Asset management, Data migration, Data translation, Information systems, Information system modernisation, Legacy systems, Open-source, Rosetta stone.

## **Introduction**

Information systems (ISs) are used by organisations for collection, storage and dissemination of information and data. Additionally, they are utilised for planning, decision making, maintaining competitive advantage and providing ongoing feedback to management, as noted by Issa et al. (2013). To Aloini et al. (2007) ISs have been adopted as a way of managing operations, with businesses focused on their potential benefits. These benefits are valued as components that help organisations achieve key goals, as emphasised in work by Stair and Reynolds (2010), with many investing heavily in IS technology over an extended period of time. However, organisations often fail to recognise the challenges surrounding ISs and their management. Sumner (2000) considered potential issues for organisations implementing new, cost effective systems and raised a series of potential challenges and risks, many of which continue to be prevalent in today's information landscape, as stressed by Howard (2015a). These challenges are explored further in this article, and are particularly prominent in large organisations, where there is a presence of ISs which are outdated and in need of replacement, often referred to as 'legacy information systems'. Legacy ISs can be further defined as any system based on older technologies, which continue to provide core services to an organisation (Paradauskas and Laurikaitis, 2006).

The original Rosetta Stone was discovered in the eighteenth century by French scholars. Its inscriptions displayed the same text written in three languages: Egyptian Hieroglyphics, Demotic and Greek. This made the Rosetta Stone a key to deciphering

Hieroglyphics, enabling modern understanding of the writings and history of Egypt (Andrews, 1981). The term has since been used to represent problem solving and language connection, including in the field of IS (e.g. Heminger and Robertson, 2000; O'Grady, 2010).

The aim of this article is to explore the potential development of a 'Rosetta Stone for information systems', meaning it can translate data between legacy and contemporary systems, and help overcome a series of challenges now present with ISs in organisations. This innovation should maintain data integrity whilst avoiding typical costly and value negative processes of data migration and data cleansing. These terms and processes are introduced in more detail later in the article.

### **Research design**

The research was designed around two main components. First was an in-depth literature review, a shortened version of which is presented in this article, which highlighted a number of key areas including IS use, challenges, existing data migration solutions and potential ways of creating the proposed Rosetta Stone data translation tool. Second were a series of interviews with IS professionals. The interviews focused on the following three areas, related to the research aim:

- The challenges associated with information systems and their modernisation.
- The potential for the proposed Rosetta Stone tool.

- The suggested methods for creating and implementing the Rosetta Stone tool, and perceived difficulties in its creation.

In addition to these main data collection methods, exploratory conversations were used to discuss the general scope and purpose of the research with participants, and offered the chance for any initial comments, questions or ideas to be discussed.

Additionally, there was the opportunity to use ISs from a 'users' point of view' in one of the participating organisations, enabling further insight to the workings and interlinking of their numerous systems.

Four organisations participated in the empirical investigation; a University IT Department, University Library, Public Utility Company and Fortune 500 Hardware and Software Manufacturer. There were six participants with varying backgrounds, selected to enable sampling from professionals who had recent or ongoing experience of IS change.

### **Information systems: Use in organisations and possible challenges**

Over time the need for innovative ISs has increased, and the strategies behind their implementation within businesses have seen considerable change (Avison and Elliot, 2005). Such advances have not only impacted how organisations operate, but have been a catalyst in changing how society as a whole communicates and shares knowledge and information (Rainer and Cegielski, 2010). The advent of the Internet, and transition into an Information Economy (Tapscott, 1997), coupled with a wealth of

new technologies and ideas have created new possibilities. This is recognised as being responsible for transforming economic activity (OECD, 2001), and for many businesses, ISs have helped streamline operations and business functions by providing a common interface. The benefits provided by ISs are abundant and this is evident in a number of business areas, as noted by Kornkaew (2012). However, organisations are under significant pressure to have more effective strategies surrounding IS implementation, with Kornkaew indicating a series of issues or challenges in doing so that require considerable attention. Four key challenges have been drawn from literature, and are highlighted as being of particular importance in the context of this article:

1. The existence of legacy ISs, and the issues that persist when developing a strategy for replacing these often disparate systems, and implementing newer alternatives. Rice (2009) emphasises that the need for legacy replacement is, in part, due to the growing competitiveness of commercial markets and the global economy. Peslak (2012) explains that larger organisations are viewing upgrade of legacy systems to newer alternatives as an increasingly critical issue. Schneider (2013) highlights that whilst legacy systems were state of the art when introduced decades ago; they are now significantly outdated, but still have great relevance to organisational operations, and have been modified into bespoke systems over many years making replication or replacement problematic. This has incited significant concerns, particularly with longer established organisations, where numerous systems of varying age and sophistication exist. Reddy and Reddy (2002) note that organisations could suffer from a lack of

competitive flexibility in effectively utilising new technologies, and without significant investment, many are left burdened with legacy systems for the foreseeable future.

2. Data integrity; the maintaining of data to avoid problems including duplication, incomplete or corrupted data (Boritz, 2005) is another issue, and one that often exists with the presence of legacy ISs. These impurities create particular difficulty when data is repeated or duplicated across one or a number of different systems. This also raises problems when transferring data, with maintaining of data integrity being paramount. Data cleansing is used as a technique for removing undesirable incomplete or duplicate data; described by Muller and Freytag (2003) as integral to data integration processes. However they also outline various challenges with data cleansing, highlighting it as an undesirable process which can be time-consuming and expensive.

3. The migration of data from legacy ISs to more contemporary platforms is a significant challenge, which due to technological advances are likely to be different in data architecture to previous legacy platforms. Tools continue to be developed to ease this process, known as data migration. Data migration is required anytime data needs to be moved between platforms, and are relevant to any organisations who use ISs (Howard, 2015a). Laszewski and Nauduri (2012) emphasise the importance of finding ways to automate data migration, as it becomes increasingly challenging for organisations with large amounts of data held across legacy systems. To Bertolucci (2012) data migration is high risk, and selecting the correct strategy and thorough risk assessment is pivotal to its success, especially considering challenges such as

transferring data between platforms and the recognition of specific business needs.

Morris and Fosker (2007) similarly express that data migration is not a straight forward task; it is expensive and prone to failure. For example, Howard (2015a) highlights how recent research suggests that over fifty percent of data migration projects experience significant problems, whilst almost twenty percent are cancelled. Nahar et al. (2012) consider that many modern organisations are seeking affordable and reliable solutions to the migration challenge, with increasing desire to pursue solutions which keep efficient storage infrastructures with zero-downtime; meaning minimal effect on everyday operations. They explain that demands for such solutions make data migration one of the most prominent challenges in IT.

4. Finally, recognition of the human elements and processes involved with ISs is important, rather than ISs being viewed entirely as the obtaining of hardware and software (Butt and Zaman, 2012). Noffsinger et al. (1998) express that human elements including staffing and knowledge management and retention are relevant, with diminishing numbers of staff with direct knowledge of original systems being a significant issue. More recently, Liem et al. (2006) outline that there must be clear key success factors for system implementation; a combination of human and technical. Kroenke (2014) agrees, demonstrated by the view that ISs consist of Hardware, Software, Data, Procedures and People.

Finding an appropriate solution for these four issues presents a challenge in itself; therefore it is important to explore what solutions are possible, particularly regarding



realistic, more time and cost efficient ways of moving data from legacy to modernised systems. Highlighted in table 1 are some of the favoured modern methods for the transformation of ISs, and these ‘legacy modernisation’ techniques are defined by Norfolk (2014a) as being the cost effective re-use or replacement of existing technologies.

Solution	Description
System rewriting and incremental upgrade	A decomposition of legacy systems so that the components can be selectively replaced and reproduced in a new format, reducing the costs and risks of moving all assets together (Cimitile et al., 1998).
Legacy conversion	Various widely used conversion methods exist, revolving around converting data from old to new platforms all at once, or in on-going stages (Mallach, 2009).
Web and cloud migration	The use of cloud platforms for ‘cloud migration’ (Khajeh-Hosseini et al, 2010; Pant and Thakur, 2013). Pant and Thakur explain that cloud migration involves moving a business’s data, applications or other elements from an organisation’s computers to the cloud. It is being used increasingly by large organisations such as IBM (Howard, 2015b).

**Table 1.** Data migration and legacy modernisation solutions

### **Legacy modernisation: potential for a data Rosetta Stone**

As the need for alternatives to legacy ISs has grown, so has the market for solutions.

Vendors are offering ‘all-in-one’ packages, promising to counter increasing risks

associated with data migration whilst providing innovative system architectures and interfaces (Radding, 2011). Norfolk (2006) highlights a series of system development approaches which have come to fruition, including an increase in constraint-driven approaches, allowing new systems to learn from legacy ISs and data to be incrementally translated between the two (Norfolk, 2014b). Many of these approaches derive from work on intelligent systems, which have an ability to understand, think and act instead of being automated or operated by instruction (Annadurai, 2013). Though the definition of what makes an 'intelligent system' is subject to debate, in the perspective of computer systems the intelligence is characterised by flexibility, learning, memory capability and ability to manage imprecise data (Rudas and Fodor, 2008), and have been popularised through seminal works such as Cybernetics and Management (Beer, 1959) and Decision and Control (Beer, 1966). One example is the Erudine Behaviour Engine (Rice and Rice, 2006) which uses mathematical modelling of behaviour as a means of systems development, and is designed to utilise automated features to target critical legacy system issues. Erudine also operates to develop the new system whilst running parallel with the old, meaning no disruption of business activity. Another example is C2EIF (Trudel, 2013), an automated method of modernising legacy computer languages by translating them into contemporary formats. Trudel emphasises the need for tool automation, which is inexpensive and more time-efficient than manual methods.

In addition to ready-made or marketed solutions, open-source is increasingly popular for developing software and systems, and is defined by its accessibility, meaning it can be developed by anyone under free licencing agreements (Lakhani and Von Hippel, 2003). Many current open-source translation tools focus on translating human languages. For example Forcada et al. (2011) mention the development of Apertium, which matches numerous languages using rule-based translation techniques. Another example is GREAT, which Gonzalez and Casacuberta (2011) describe as a statistic driven translation tool. By comparison there is less evidence of open-source tools which offer a solution for translation of computer languages. One existing example is Constraint-Driven Development (CDD, 2014), which takes a similar approach to Erudine, with more focus towards migration to web and cloud formats. Norfolk (2014a) believes that organisations should be investigating the open-source route for cost reduction purposes, and to solve the legacy problem, stating that it offers useful ways forward.

### **Conversations with information systems professionals**

The findings that resulted from the empirical data collection are discussed here. Organisations and participants have been anonymised for this article, as shown in table 2.

Organisations	Participants
<p><b>A-</b> A large Water Utility company, based in the North of England.</p>	<p><b>Participant A1</b> - Operations Solutions Manager, responsible for ISs and IS staff.</p>
<p><b>B-</b> The library of a UK Public Research University in Yorkshire.</p>	<p><b>Participant B1</b> - Head of Library Digital Services and Systems, responsible for learning resource teams and associated systems.</p> <p><b>Participant B2</b> - Library Systems Manager, responsible for the management of all University library systems.</p> <p><b>Participant B3</b> - Learning Resources Coordinator, responsible for maintaining and updating library systems and associated learning resource materials.</p>
<p><b>C-</b> The UK branch of an American hardware and software manufacturer, based in Oxfordshire.</p>	<p><b>Participant C1</b> - Director of Location Intelligence, responsible for Geographical ISs.</p>
<p><b>D-</b> The dedicated IT department of a Public Research University in Yorkshire.</p>	<p><b>Participant D1</b>- Senior Analyst and Programmer, responsible for University SAP systems.</p>

**Table 2.** Details of organisations and participants

Exploratory conversations helped explain the project and explore initial views with participants, whilst acting as a tool to aid the structure of interview questions. Two days were also spent observing staff actions and use of systems in organisation B. This involved experiencing systems from a user's perspective, with the primary goal of gaining an understanding of the systems in the organisation, including insight into potential issues and challenges. Additionally, organisation B provided notes from their recent library reading list migration project (an example of which can be seen in figure 1); one of the main projects being performed by the IS team during this period was modification to the reading list system structure. Re-structuring the system was explained as being essential for organising datasets, and making information retrievable for users. Seeing the reading list system structure first-hand emphasised a primary observation regarding extensiveness of the data held on the library ISs, and how efficient organisation of this data is imperative to its successful operation. Another observation was the level of maintenance required for the different library systems. For example figure 1 indicates the scale of maintenance required for entering data when University courses and corresponding modules change; ongoing changes such as this are happening to the hundreds of modules across the University. A final observation relates to the quantity of systems used; the main Library Management System comprised different sub systems which require extensive input, and the learning resource systems comprised of multiple sub systems often containing instances of the same data.

## Representative data structures:

### 1. Simple Module with lists [vast majority of courses]

1. Module (course\_id)
2. List (list\_id) <-top-level list
3. Sub-List ▲
4. Sub-Sub-List ▲

### 2. Module with Sub-course

1. Module (course\_id)
2. List (list\_id) <-top-level list
3. Sub-List ▲
4. Sub-course (course\_id)
5. List (list\_id) <-top-level list
6. Sub-List ▲
7. Sub-Sub-List ▲

### 3. Module with Sub-course and Topic

1. Module (course\_id)
2. List (list\_id) <-top-level list
3. Sub-List ▲
4. Sub-course (course\_id)
5. List (list\_id) <-top-level list
6. Sub-List ▲
7. Sub-Sub-List ▲
8. Topic (course\_id)
9. List (list\_id) <-top-level list
10. Sub-List ▲

**Figure 1.** Representation of scale for maintaining University information systems

The semi-structured interviews, which are the main focus of this article, involved five of the six participants, with each lasting between forty five and ninety minutes. The sixth participant was unable to take part in the main interviews, but contributed through the exploratory interviews and by providing information sheets about IS use in

organisation C. The analysis is ordered by themes derived from a focused coding process, rather than by interview question.

### *Background and experience*

The interviewees had varying backgrounds and experience with ISs. Participant A1 noted management of Geographical, SCADA and Telemetry Systems and emphasised the large number of asset systems within their organisation. Participant D1 explained that his role as systems specialist involves maintaining SAP systems, which contain all University asset repositories. Participants B1, B2 and B3 had roles related to maintaining library systems, including those hosting resources for both staff and students, in addition to working with metadata.

It was important that the interviewees had experience of data migration processes. For example, Participant A1 noted his involvement with various migration projects, and explained that organisation A was planning a large project relating to their numerous asset systems. Participant B1 had managed a data migration project moving from a legacy Library Management System to a newer alternative. Participants B2 and B3 both identified numerous migration projects, concentrating primarily on a recent migration of a library reading list system. Participant B2 was part of this project before moving to a new role as Library Systems Manager. Participant D1 outlined that he had managed data migration initiatives in a number of organisations, including at the University and at a large British food manufacturer.

### *Information systems: Challenges experienced*

The interview questions were designed to complement findings from the literature review, with interviewees giving feedback surrounding IS challenges. The first significant issue, raised by participant A1, related to disparity of ISs and existence of bespoke systems which had been updated over long periods. After this an additional question was added, asking if the problem was a result of legacy systems, to which A1 responded:

*“Yes, legacy systems, I guess you could say it is a lack of governance really, long term lack of information management. But it’s come through historically developing a system and then needing extra functionality, and a decision being made at the time to develop a new system just to do that task, when that might have been the point to take something off the shelf that did everything that we wanted to do. So there’s lots of bolt-ons.”*

Participant D1 also highlighted the challenges of working with systems which had been modified. The issues with integration and disparity of systems were also existent in organisation B, and two of the interviewees mentioned that integration was a huge consideration, as was integration of the new Library Management System with the many other library systems. For example, B1 stated:

*“One big area was integration, as the library management system doesn’t stand alone; it also integrates with other systems. For example, linking with vendor systems for ordering, fulfilment data, exporting of catalogue data to other databases, links with*



*University systems for user data. All sorts of things that are vital for running the library”.*

Participant B2 highlighted existence of legacy systems and the disparity of data, specifically in terms of making it difficult to extract data ready for migration, and Participant D1 expressed that before any data migration can commence they must plan how to access data held in these systems. Another theme arising from this conversation around disparate systems was the presence of unstructured data, and challenges this presents when extracting data and upgrading platforms. This was noted by participants A1 and B2. A1 provided an example relating to storage of council permission contracts when excavating roads:

*“One of the stages in this data migration that we're going to have to do will be looking at that unstructured information and saying is there stuff on this unstructured information that we want to extract at this point and put into a field so that we can use it more intelligently”.*

Participant B2 also highlighted challenges of unstructured information, especially during extraction. This led to one of the prevalent challenges conveyed during the interviews; the vast amounts of data held in legacy systems, and deciding how and what to migrate. Organisation A were deciding how this would be done for their upcoming project, compared to organisations B and D where the process being discussed had recently taken place. For example, B1 stressed that migration projects

have been a learning experience, and the library had learnt to be vigilant about what data to transfer:

*“One of the key things we agreed very early on was we would migrate as little data as possible, partly because of time pressure, and secondly because data migration is an expensive business. It’s not just getting the data over, but you also have to test it. If you don’t need the data don’t migrate it”.*

Participant B3 emphasised refining the data and making it extractable as the biggest challenge in data migration, whereas Participant D1 further highlighted how difficult these processes can be, noting that data often needs ‘cleansing’ before being input to a new platform. Data inconsistencies and redundancy were mentioned frequently by the interviewees; with all underlining it as a challenge related to IS management. D1 stated that quality of data after extraction was often variable, and could cause compatibility issues with the new platform. He also mentioned data redundancies:

*“Data is often duplicated, some parts will be missing. You have to fill in the gaps and create a standard way of formatting so data can be consistent”.*

Participant A1 highlighted that past experiences of data migration, and experiences of what organisation A were learning in the build up to their migration project, reflected a need to make sure the whole business is committed:

*“I think lessons we are already learning at this point, and from other experiences, is to make sure that we engage the business before we start, for them to see this as very much a business project and not an IT project”.*

This prompted the other interviewees being asked what they thought of data migration being a business versus an IT problem. This features heavily in literature and participants generally agreed.

Links between knowledge and legacy systems was also discussed with A1, and if challenges existed in relation to those who developed systems no longer being in employment. A1 agreed it did, stating that although it can be worked around, it involves hiring experts making it costly and time-consuming. D1 also mentioned concerns for organisational knowledge retention:

*“Legacy systems have to have in-house people with the knowledge to work them, if you look around our department most of the people are approaching retirement, so that’s something the University considers; it’s a business risk”.*

Interviewees B2 and B3 both recognised this as a challenge, but noted they were fortunate that B1 knew their legacy formats, and was able to assist. Participant B1 expressed emphasis on change management and ensuring staff receive sufficient training, so change to new systems did not have an adverse effect.

### *Overcoming the challenges*

Issues highlighted and participants’ experiences of data migration led to insight into how they handled the process and overcame challenges. Library participants, and D1, indicated that the migrations they have been involved with were successful. In preparing for their migration project, A1 mentioned that organisations A would want

to get their systems down to a base of as few platforms as possible, to help minimise future issues.

Both of the library migration projects were carried out using an incremental approach; utilising migration scripts which B1 described as being customisable and allowing extraction of data from the legacy Library Management System. Participant D1 explained his typical step-by-step approach for data migration. He noted that during data migration for a previous employer, bespoke automated tools were developed that assisted with the process. Participants B1, B2 and B3 stated that, in their experience, automating processes for data migration projects was a considerable success factor. B1 explained:

*“Try to automate as many procedures as possible for migration. It makes it easier to maintain the data extract process. And it also helps with consistency compared with running processes manually, as everything manual is a potential source of error; you want to run things automatically for consistency”.*

The library staff explained that they benefitted from vendor migration support. It was also apparent that the new Library Management System was a cloud platform, meaning hosting was external. B2 expressed that cloud and web platforms must now be a consideration for organisations undertaking data migration projects. A1 noted consideration for cloud technologies, based on a system-by-system evaluation. A1 commented that this could be beneficial for increasing business value and overcoming problems related to having disparate platforms:

*“Where we are at the moment is developing a cloud strategy. It looks at what we should do when we assess each new system. It's not the case where we can move everything to the cloud, there is no business case for it, but if you've got a business case to implement a new asset management system naturally you have to ask if it should be in the cloud”.*

#### *Potential for the Rosetta Stone*

This led to asking participants what they thought of the proposed Rosetta Stone, and what, if anything might be a better alternative. The reaction was mixed, with most responses relating to the difficulties of creating such a tool when considering currently available technologies and solutions. However, despite this there were positive comments about how it might be achieved.

Participant A1 was the first to be asked about the Rosetta Stone. He commented that whilst liking the concept, it was difficult to see how it would be possible, and stressed that it would need to work with bespoke systems and that the way forward was wide adoption of a standard data format:

*“No. Perhaps in the future when, and if there is less bespoke systems in organisations. The spanner in the works is bespoke systems. The Rosetta Stone is a great ideal, but practically speaking I don't really see it coming off. What you end up doing instead is trying to create a layer, and have all of the various systems feed into this layer. You're not talking about one piece of software that just plugs in and does it all, you're talking*

*about hard graft and lots of analysis and bespoke development on each of those systems around the outside to put it all into that layer. There's no silver bullet". I like the idea of the Rosetta Stone, but personally I think it's information and data migration Nirvana. I saw people 15 years ago talking about tools that promised to do this, it's an old problem. It's almost too good to be true. The only surprise is that we haven't adopted standards that would have allowed us, by today, to all be working with standard data formats".*

B1 expressed that the Rosetta Stone would be difficult due to the amount of effort it would likely take to create. B1 used human language translators to illustrate how it could be achieved, but believed translating system languages in this way would also be challenging. He expressed that creating a new standard to aid creation of the Rosetta Stone idea might be possible, but again would be difficult. Instead participant B1 proposed that a 'Rosetta Framework' would be more achievable, and could act as a communication protocol between systems:

*"What you really want is a Rosetta framework, which works in a similar way to the SOAP and REST protocols, and can tell you how to integrate with other systems rather than map it all across automatically".*

The interviews with participants B2, B3 and D1 concluded with equally mixed reviews, with further reservations about how the Rosetta Stone could be achieved. B2 expressed that it would depend on scale, and would likely have to be a combination of a tool and project management platform, highlighting project management as crucial:

*“It depends on scale; if it’s a massive system then it’s probably not feasible and would be problematic. A lot of it would come down to project management. Nonetheless there might be a tool that could do so for certain situations”.*

Additionally B2 noted that the implementation of the Rosetta Stone would be especially demanding in larger organisations. Utilisation of open-source platforms was mentioned by B2 as having potential in allowing communities to collaborate and contribute code for such a project. B2 felt the amount of work required to build a Rosetta Stone would then be more manageable, and knowledge from a range of professionals could be congregated. B1, B2 and D1 again stressed that automation was imperative, and that current methods of migration such as extract scripts and tools with increased automation were still sufficient for organisations in most cases.

B3 felt creation of the Rosetta Stone was doubtful, mainly due to lack of awareness of anything that would make it possible, stating the following:

*“Artificial intelligence would have to go somewhere very different to where it is now to make it practical. I’m sure there is somebody who could do a decent job of something like that, but the best option is probably trying to aid automation with a good framework”.*

Participant D1 was also unsure of this being possible, suggesting open-source as the only way he could imagine the Rosetta Stone being manageable. D1 concluded that human knowledge of business processes and what data is useful would be

irreplaceable, so the Rosetta Stone would have to work alongside these considerations:

*“Looking at the feasibility of a Rosetta Stone, perhaps it would have to be an open-source type collaboration. In the open-source community you get people who learn from each other. Mozilla is an example, it’s all people in their spare time making things that are useful for others. It would be an awfully big process but that might be the best way”.*

### **Summary of findings**

This article aims to address three main areas regarding: the challenges associated with ISs and their modernisation; the potential for the proposed Rosetta Stone tool and potential methods of creating this; how it might be implemented; and perceived difficulties in its creation. These are summarised in this section.

#### *The challenges associated with information systems and their modernisation*

The main challenges highlighted revolved around legacy ISs and the need to migrate data to newer platforms. It was interesting to see that participants mentioned all of the challenges identified in the literature, as well as highlighting additional issues from experience, including integrating data to new systems, what data to keep, what type of system to move to and managing change. Participants were thorough in explaining what made data migration challenging for organisations, whilst sharing the view



expressed in literature that data migration can be wrongly seen as an IT problem, with human and business aspects often overlooked. Organisation A in particular highlighted the issues also explored in the literature surrounding the disparity of ISs in modern organisations, and participant A1 emphasised this as being significant when implementing new systems. Overall, findings from literature and the participants illustrated numerous issues present with ISs today, and which solutions might prevent similar challenges prevailing in the future.

#### *The potential for the proposed Rosetta Stone tool*

There were interesting findings in the literature, particularly regarding solutions which claim to translate between ISs. Although participants were somewhat negative about the practicalities of the Rosetta Stone, responses to its ideology were generally positive, with suggestions about how it might be achieved. Most of these suggestions focused on the Rosetta Stone having more potential on a small scale; within smaller organisations. Positive suggestions came from participants B1, B2 and D1. B1 mentioned the protocols SOAP and REST which act as a framework to link web based standards, and believed something could be created in a similar way, thus making a 'Rosetta framework' which could aid the linking of systems and meaningful interpretation of data. This was viewed as being more realistic than a tool which could automate IS translation. Participants B2 and D1 implied the Rosetta Stone might be possible using an open-source platform. B1 mentioned examples of networks for

computer programmers, which provide a base for collaboration. D1 also expressed that, due to the workload the Rosetta Stone would require, facilitating an environment where it could be community built would be beneficial. Participants' mention of open-source created a positive and encouraging link to open-source possibilities highlighted in the literature. Overall, although the responses to the Rosetta Stone did not present detailed ideas of how it might be created, the combination of the primary and secondary research helped generate concepts about how it might be achieved in theory.

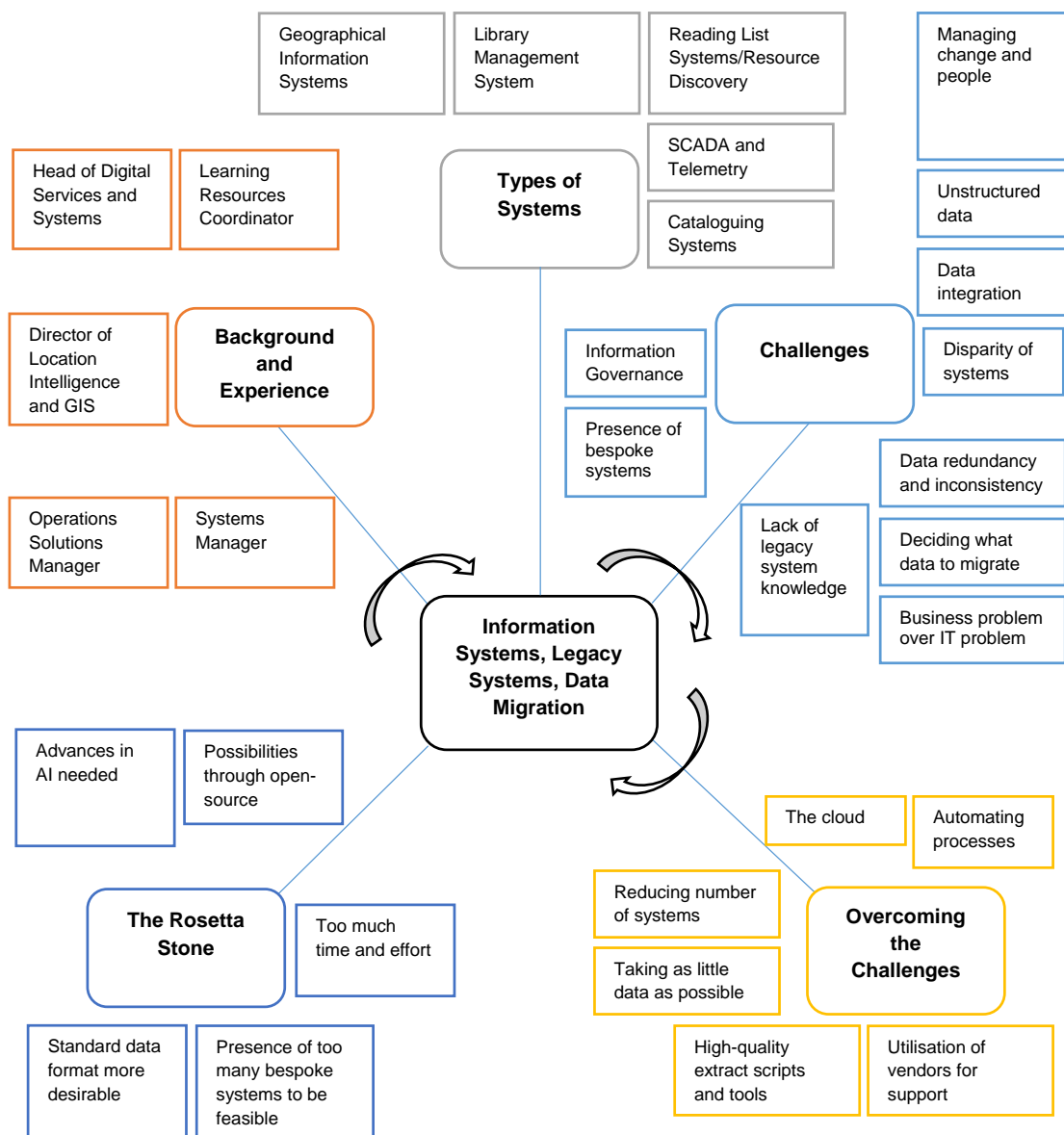
*The suggested methods for creating and implementing the Rosetta Stone tool, and perceived difficulties in its creation*

As the participants alluded, the implementation of the Rosetta Stone would be more achievable on a small scale, and with the right level of support. Potential creation links back to two suggestions mentioned throughout the article. The first regards intelligence and automation in systems, building on existing theories for machine learning to explore potential creation of a tool that can actively learn and translate computer languages. Examples mentioned indicate potential methods of translating data between ISs. The second, perhaps more realistic option would be the creation of a Rosetta framework, to provide improved understanding of the meaning of data held in legacy systems, rather than direct translation. As scale and time considerations have been widely highlighted as problematic, the chosen solution could benefit from

utilisation of an open-source platform, allowing developers and open communities to collaborate towards its creation.

The findings from the literature review and the interviews with the five IS professionals are conceptualised in a 'thematic network' (Attride-Stirling, 2001), shown in figure 2.

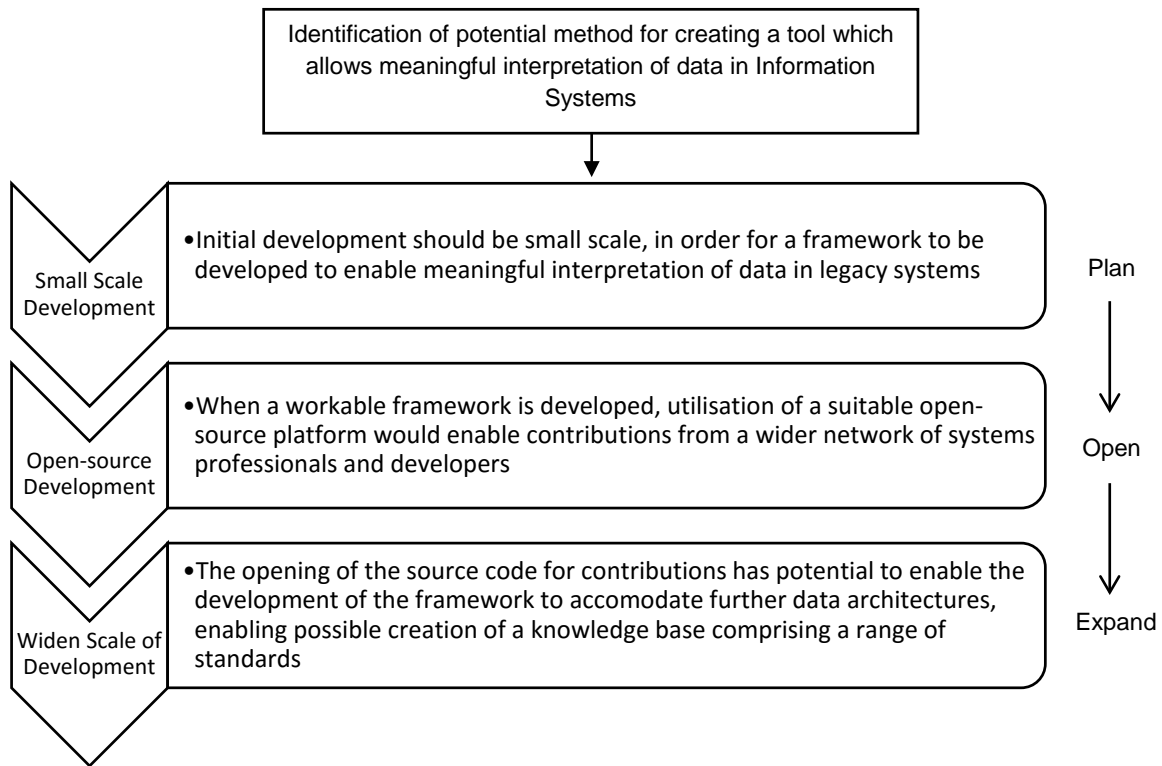
This helps provide an overview of the main themes discussed.



**Figure 2.** Significant research findings presented as a thematic network, summarising key data from the literature review and empirical data collection

## **Conclusion and potential avenues for future work**

This article concludes by offering a theoretical, rather than technical stance as to how the Rosetta Stone might be created, based on reviewed literature and analysis of expert opinion of IS professionals. It is recommended, at this stage, that the tool be developed as a framework to enable meaningful interpretation of data, rather than direct translation. This will enable the operator, whether human or machine, to understand the meaning of the data, and allow insight into how the data might need to be transformed to meet the constraints of target systems. It is also recommended that the tool be implemented on a small scale, between a nominal number of ISs. This could then be tested and potentially allow accommodation of data from a further number of ISs. Findings indicate that a logical method for development would be through open-source platforms and communities, presenting cost and time effective development for businesses to tailor legacy modernisation to specific needs. Open-source development might also provide a solution which offers zero-downtime implementation, indicated as being increasingly desirable. Based on these conclusions, a model for developing the Rosetta Stone is proposed in figure 3.



**Figure 3.** Four-point model for developing the Rosetta Stone as a tool for meaningful interpretation of data, based on the project findings

The next progression would be testing any potential tool in practical settings, to provide insight as to whether the method would be effective and beneficial. This article has opened up avenues for future research in two ways, first is to follow the theoretical design of figure 3, and create a workable Rosetta framework. Second is to follow the theme of this article and further explore how a Rosetta Stone for directly translating data between ISs might be possible, with an increased focus on the potential of machine-translation and meta-systems for computer languages. Future

research of this nature should pursue increased input from participants with technical viewpoints, and extend to include input from those in open source communities.

## References

Aloini D, Dulmin R and Mininno V (2007) Risk Management in ERP Project Introduction: Review of the Literature. *Information and Management* 44(6): 547-567.

Andrews C (1981) *The Rosetta Stone*. London: British Museum Press.

Annadurai A (2013) Intelligent Systems: A Concise Description for Pedagogy. *International Journal of Advanced Technology & Engineering Research* 3(5): 70-73.

Attride-Stirling J (2001) Thematic networks: an analytic tool for qualitative research. *Qualitative Research* 1(3): 385-405.

Avison D and Elliot S (2005) Scoping the Discipline of Information Systems. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.7159&rep=rep1&type=pdf> (accessed 28 September 2015).

Beer S (1959) *Cybernetics and Management*. London: English Universities Press.

Beer S (1966) *Decision and Control: The Meaning of Operational Research and Management Cybernetics*. London: John Wiley & Sons.

Bertolucci J (2012) 10 Big Data Migration Mistakes. Available at: <http://www.informationweek.com/big-data/software-platforms/10-big-data-migration-mistakes/d/d-id/1105724?> (accessed 18 September 2015).

Boritz JE (2005) IS Practitioners' views on core concepts of information integrity. *International Journal of Accounting Information Systems* 6(1): 260-279.

Butt MA and Zaman M (2012) Legacy Information System Development: Problems and Issues. *International Journal of Computers and Distributed Systems* 2(1): 114-117.

CDD (2014) CDD: Constraint Driven Development Open-source. Available at: <http://www.constraintdrivendevelopment.org> (accessed 21 September 2015).

Cimitile A, De Carlini U and De Lucia A (1998) Incremental migration strategies: data flow analysis for wrapping. In: *Fifth Working Conference on Reverse Engineering*, Honolulu Hawaii, USA, 12-14 October 1998, pp. 59-68. Honolulu: IEEE.

Forcada ML, Ginesti-Rosell M, Nordfalk J, O'Regan J, Ortiz-Rojas S, Perez-Ortiz JA, Sanchez-Martinez F, Ramirez-Sanchez G and Tyers FM (2011) Apertium: a free/open-

source platform for rule-based machine translation. *Machine Translation* 25(2): 127-144.

Gonzalez J and Casacuberta F (2011) GREAT: open-source software for statistical machine translation. *Machine Translation* 25(2): 145-160.

Heminger AR and Robertson S (2000) The digital Rosetta Stone: a model for maintaining long-term access to static digital documents. *Communications of the AIS* 3 (1): 1-28.

Howard P (2015a) Bloor Research Article: Data Migration. Available at: <http://www.bloorresearch.com/technology/data-migration> (accessed 5 October 2015).

Howard P (2015b) Bloor Research Blog: IBM and the Cloud. Available at: <https://www.bloorresearch.com/blog/im-blog/ibm-and-the-cloud> (accessed 1 October 2015).

Issa T, Isaias P and Kommers P (2013) *Information Systems and Technology for Organizations in a Networked Society*. Hershey: IGI Global.

Khajeh-Hosseini A, Greenwood D and Sommerville I (2010) Cloud Migration: A Case Study of Migrating an Enterprise IT System to IaaS. In: *Cloud Computing (CLOUD) IEEE 3<sup>rd</sup> International Conference*, St Andrews, Scotland, 5-10 July 2010, pp. 450-457. St Andrews: IEEE.

Kornkaew A (2012) *Management Information System Implementation Challenges, Success Key Issues, Effects and Consequences: A Case Study of Fenix System*. Master's Thesis, Jonkoping University, Sweden.

Kroenke D (2014) *Using MIS*. Upper Saddle River: Prentice Hall.

Lakhani KR and Von Hippel E (2003) How open-source software works: 'free' user-to-user assistance. *Research Policy* 32(1): 923-943.

Laszewski T and Nauduri P (2012) *Migrating to the Cloud: Oracle Client/Server Modernization*. Waltham: Elsevier.

Liem I, Schatten A and Wahyudin D (2006) Data Integration: An Experience of Information System Migration. In: *Proceedings of the 8<sup>th</sup> International Conference Information Integration, Web Applications and Services*, Yogyakarta, Indonesia, 4-6 December 2006, pp. 1-10. Yogyakarta: iiWAS.



Mallach EG (2009) Information System Conversion Strategies: A Unified View. *International Journal of Enterprise Information Systems* 5(1): 1-4.

Morris J and Fosker T (2007) Practical Data Migration: The Challenge at British Telecom. Available at: <http://www.practicaldatamigration.com/white-papers> (accessed 21 September 2015).

Muller H and Freytag JC (2003) *Problems, Methods and Challenges in Comprehensive Data Cleansing*. Berlin: Humboldt-Universität zu Berlin.

Nahar N, Huda N and Tepandi J (2012) Critical risk factors in business model and IS innovations of a cloud-based gaming company: Case evidence from Scandinavia. In: *Technology Management for Emerging Technologies proceedings of PICMET*, Vancouver, Canada, 29 July-2 August, pp. 3674-3680. Vancouver: PICMET.

Noffsinger WB, Niedbalski R, Blanks M and Emmart N (1998) Legacy Object Modelling: Speeds Software Integration. *Communications of the ACM* 41(12): 80-89.

Norfolk D (2006) Firing up the Erudine engine. Available at: [http://www.theregister.co.uk/2006/11/16/erudine\\_behaviour\\_engine](http://www.theregister.co.uk/2006/11/16/erudine_behaviour_engine) (accessed 21 September 2015).

Norfolk D (2014a) Bloor Research Article: Legacy Modernisation. Available at: <http://www.bloorresearch.com/issue/legacy-modernisation> (accessed 22 September 2015).

Norfolk D (2014b) Bloor Research Article: Simulating behaviour to replace legacy. Available at: <http://www.bloorresearch.com/analysis/simulating-behaviour-to-replace-legacy> (accessed 16 September 2015).

OECD (2001) The Internet and Business Performance. Available at: [http://www.oecd-ilibrary.org/science-and-technology/the-internet-and-business-performance\\_233751540382](http://www.oecd-ilibrary.org/science-and-technology/the-internet-and-business-performance_233751540382) (accessed 18 September 2015).

O'Grady J (2010) *System Synthesis: Product and Process Design*. Boca Raton: CRC Press.

Pant P and Thakur S (2013) Data Migration Across The Clouds. *International Journal of Soft Computing and Engineering* 3(2): 14-21.

Paradauskas B and Laurikaitis A (2006) Business Knowledge Extraction from Legacy Information Systems. *Information Technology and Control* 35(3): 214-221.

- Peslak AR (2012) An analysis of critical information technology issues facing organizations. *Industrial Management & Data Systems* 112(5): 808-827.
- Radding A (2011) Five Big Risk Areas of Technology Migration. *Business Finance*, 2 June, 1.
- Rainer KR and Cegielski CG (2010) *Introduction to Information Systems: Enabling and Transforming Business*. Hoboken: John Wiley & Sons.
- Reddy SB and Reddy R (2002) Competitive agility and the challenge of legacy information systems. *Industrial Management & Data Systems* 102(1): 5-16.
- Rice M (2009) Erudine and NCC publish new guidelines on evaluating legacy. Available at: <http://www.theinformationdaily.com/2009/08/17/erudine-and-ncc-publish-new-guidelines-on-evaluating-legacy> (accessed 12 September 2015).
- Rudas IJ and Fodor J (2008) Intelligent Systems. *International Journal of Computers, Communications and Control* 3(1): 132-138.
- Schneider, A. (2013), "When Companies Become Prisoners of Legacy Systems", *The Wall Street Journal*, 1 October, pp. 5.
- Stair R and Reynolds G (2010) *Principles of Information Systems*. Stamford: Cengage Learning.
- Sumner M (2000) Risk Factors in Enterprise Wide Information Management System Projects. In: *Conference on Computer Personnel Research*, Chicago Illinois, USA, 6-8 April 2000, pp. 180-187. Chicago: SIGCPR.
- Tapscott D (1997) *The Digital Economy: Promise and Peril In the Age of Networked Intelligence*. New York: McGraw-Hill.
- Trudel M (2013) *Automatic Translation and Object-Oriented Reengineering of Legacy Code*. PhD Thesis, ETH Zurich, Switzerland.