

## **Case Studies**

**Digitisation Projects of**



## **CASE STUDY 4: The Qatar Heritage Rare Book Project : Digitisation and Online Publication of Treasures from the Persian Gulf**

### **Background:**

Qatar and the Middle East region have a unique heritage, dating to thousands of years, when Arab civilizations have made enormous contributions to art, craft, religion and science.

Qatar has been ruled by the Al-Thani family ever since gaining its independence from Great Britain in 1971. The present ruler is Sheikh Hamad bin Khalifa al-Thani. Though closely regulated the Emir of Qatar has brought about many reforms to diversify and empower the regions economy and people.

A commitment towards preserving Qatar's Cultural Heritage, make Qatar a centre of excellence and research, in addition to an important contributor to world knowledge His Highness the Emir of Qatar, Sheikh Hamad Bin Khalifa Al-Thani and Her Highness Shieka Mozah have established 'The Qatar Foundation, for Education, Science and Community Development'.

The primary aim of the Qatar Foundation is to develop the people of Qatar through world-class education and research facilities to meet the challenges of an ever changing world and to make it a leader in innovation. A program is underway to promote awareness of Qatar's cultural heritage, and to stimulate new creative projects.

Through her occupation at Qatar Foundation, Her Highness helps preserve Qatar's heritage and promotes cultural richness. Museums and libraries, magnificent collections of Arab writings, calligraphy and art will be available for research and public appreciation in order to implement the objectives. The creation of new institutions for the performing arts, and a revival of creative disciplines in education has made Qatar a dynamic focus for the arts in the region.

The Heritage Book Collection, located in Doha, contains approximately 120,000 items; 10,000 of which are extremely rare or unique. The original focus of the collection centered on western views of the Middle East.

The collection consists of books, manuscripts, newspapers, journals and maps. The materials are written in several languages with English and Arabic being the most prevalent, but also including German, French, Farsi and Hebrew.

In 2004, The Qatar Foundation supported CMU to open a branch campus in Qatar offering undergraduate programs in computer science and business. Dr. Raj Reddy, who helped establish the campus in Qatar was also with the Million Book project, a universal digital library that brings together content from United States, India and China. (Thrinaina is an active participant in this endeavor as Indian partner – On how this project has achieved its goals successfully can be read in the Case Study for 'The Million Book Project').

### **Requirement:**

In October, 2006 the Qatar Foundation in Collaboration with Carnegie Mellon University (CMU) began a Pilot project to digitise the Qatar Heritage Book Collection and make the collection available to the world. The Pilot makes available online 300 rare materials and 5000 books from the general collection of the Heritage library by August 2009.

In November 2005 members of CMU Libraries visited Qatar and were introduced to the

Heritage Library collection. It immediately became clear that this magnificent collection should be made available to researchers around the world through the creation of a Digital Library of Qatar. An agreement was reached between the Qatar Foundation and CMU to undertake a Pilot project to digitise and make available a portion of the Heritage Collection.

The Pilot project will illuminate problems on a small scale; however, these problems may become larger when the full-blown project to digitise the rest of the Heritage Book Collection is conducted.

The Digitisation projects goals encompass - the study of and address the practical problems surrounding intellectual control of the Collection, digitisation workflow, OCR and management of the Arabic character set and the configuration of the information management system for world-wide access to the digital library.

This case study addresses only on how the book2net alliance was able to stand up to the challenge of providing a comprehensive Technology solution through setting up of standards, supply of scanners (Microbox), workflows (Thrinaina), Trained manpower (Thrinaina) and supervision of the digitisation process (Thrinaina).

The Pilot project will be operational from August 2008 to July 2009.

#### **Solution:**

In order to ensure that the Partner it chose would bring out the best results, CMU selected book2net, who is presently a partner for their ongoing Million Book project, through Thrinaina Informatics Limited in India. Thrinaina was engaged by CMU for its proven track record in handling similar projects and ability to provide face-to-face assistance at all stages.

Book2net and Thrinaina with their vast pool of resources and existing workflows were up to the challenge.

The client's needs and specifications were captured through the initial interactions. Carnegie Mellon University, shared the requirements with book2net and Thrinaina Informatics limited. The results of such interactions between the client (CMU) and the solution provider (book2net) resulted in formulating the requirement and standards for successful digitisation :

1. Digitisation Equipment – The Scanners.
2. A comprehensive Workflow configurable to manage and achieve the volumes and diverse requirements of such kind of project.
3. Specifications for Computer Hardware and Networking required to provide a reliable backbone and platform to manage and achieve the required targets.
4. Trained staff to supervise and man the Scanning stations with all necessary skill sets.

A specifications capture questionnaire was provided to the client which allowed the capture of all process standards. The answers to the same resulted in capturing the requisite standards for :

1. Physical document details
2. Digitisation specifications

for Eg.:

- i) Rare materials have been digitised at 300 DPI, 24 bit color, using lossless JPEG 2000 compression.
- ii) The General Collection have been digitised at 600 DPI, bitonal scan with a CCITT G4 compression, etc.,).

3. Indexing and Tagging specifications
4. Rendering Specifications
5. File and Folder structures/nomenclatures

into a specifications sheet.

To scan the Heritage books without damaging the original materials and to make high quality digital images Book2net supplied its 'Booky A2 RGB' book2net scanners which have been specially designed for converting books and historical documents into digital form. These scanners were selected as they were able to meet the exact needs of the project along with many features such as unrivalled resolution, ease of operation among others. The book2net scanners also come with a fully featured image enhancement software.

To meet the workflow monitoring and process adherence requirements of the Project, Thrinaina's proposed its Digiflo Software Suite.

Thrinaina was also entrusted the duty of training the required staff with the skill set required to man the scanning stations. The staffs were given comprehensive training on how to handle historical books, use book2net scanners to achieve high quality images and various modules of the Digiflo Software Suite. This task of training the staff was ably completed by Thrinaina due to ready availability of proven training manuals and technologies.

At the now set up state-of-the-art digitisation centre equipped with book2net scanners and computer systems Thrinaina had installed its Digiflo Suite and requisite Data Capture tools. The process is then initiated by setting up the operations to follow criteria as per the specifications for finished document delivery (Scanned, Cleaned, OCR'd, converted to PDF and Text formats at original size and ready for print). The Digiflo Suite with its various modules is set up to ensure that the document being processed is at all times adhering to the delivery specifications through its multiple quality check points and at final delivery too.

To summarise, book2net can confidently say that it was able to provide The Qatar Foundation a viable and dependable model to achieve its goal of digitising its Heritage and Rare books.

#### **Learning:**

1. The setting up digitisation centre with assets transferred from different parts of the globe was prone to many logistical problems. But many of these have been overcome through diligent working and planning of the book2net team with the client.
2. Timelines for delivery of scanners, making availability of the fully configured Digiflo software suite, training of staff were defined with enough provisions for fallbacks.
3. These timelines were monitored closely by all concerned to ensure that each member/organisation was made to feel is part of a team that need to successfully deliver a quality end product.
4. Due to these concentrated efforts of the whole team the Pilot project is now on its way to meet all the deadlines.