Paper to be presented at the 7th SA Online Information Meeting to be held at Misty Hills on 3-5 June 2003

Creating an Expert Knowledge Base

Presented by Chris Cromhout

Madam chair, ladies and gentlemen,

The reason for making this submission is mainly because I have had numerous enquiries from information professionals about the knowledge base used at Business Partners. I can only but trust that someone may learn from our experience and the lessons that we have learnt.

A. BACKGROUND

We started in 1996 – at a time when the term “knowledge management” was still in its infancy. It was driven not by what industry and technology could offer, but by what was a necessity for the survival of a newly ‘designed’ company.

To put the development of an expert data base into perspective a short historical view is necessary. The presentation will thus start with a short history of the Small Business Development Corporation (SBDC) and the transition to Business Partners Ltd. It will show the role of both state and private sector before and after the transition.

B. HISTORY

The Small Business Development Corporation Limited (SBDC) started in 1981 as a result of the Carlton Conference, which was an initiative between the private sector and the state to enhance the development of small businesses in South Africa. The SBDC commenced business in 1981 fulfilling a vital role at the time in the development of small, medium and micro sized businesses. This took place in what was at that time the boundary of the RSA i.e. not in the independent and self governing areas.

Up to realignment we gave information and advice to approximately 2,5 million entrepreneurs and potential entrepreneurs and assisted financially by granting 96 000 loans to the value of R2,8 billion to SMME's.

In 1996 the state and private sector shareholders entered into negotiations and the result was that the private sector shareholders bought 30% of the government's stake resulting in 80% of the company now being owned by the private sector and the balance owned by the state. In addition the negotiations resulted in a new market focus set for the company with new product lines. What was a company focused on providing term loans before, now became a company investing in equity and venture capital. The necessity for the improved distribution of knowledge to all offices countrywide by using easier and faster methods was placed high on the agenda.

The change of name in 1998 from the SBDC to Business Partners further strengthened the role and the focus of the company. This new focus also meant that the information requirements needed a revamp with a more focused approach. An assessment of the needs and the means of fulfilling those needs were done and we found the following:

1) There was a wealth of data available with the focus on financial and statutory reporting, but the availability of operational and industry information was limited.

2) Information that accumulated over many years, particularly in the library, was focused on functional areas and on the SME industry in general. Information on specific industries was limited

3) Much of the knowledge available, was centralized in the minds of individual members of staff.

4) Some regions was grouped based on an industry focus. Restructuring in the company resulted in this industry focus structure dissolving.

5) There was no proper directory. The existing directory of contacts, products and other sources of information were insufficient, outdated and costly to maintain. In an effort to recoup some of the costs, the publication was sold. However, the focus on making this directory available to the SME sector differed from the focus we had internally, resulting in a conflict of interest. For this reason it was difficult to ensure the continued production of the directory

6) Making use of the postal system was at best a very lengthy process to disseminate information from head office to operational staff in outlying branches

After putting all these facts together the decision was made to develop an electronic support system to enhance the collection and dissemination of information and knowledge. The concept of an expert data base was born. Only after its development did we realise that it was in fact an Industry Knowledge base.

Its purpose is to provide front line staff with the knowledge necessary to enable them to make a fully informed decision of whether the business under investigation is a viable investment proposition or not.

C. TYPICAL REQUESTS

Obviously we all work in our own specialised environment. For this reason the content of the information with which we work differs. Some may need information about medicine or illnesses or physics. Some may be searching for legal issues while others may be involved in specific product lines such as diamonds, gold etc. Our world is a business world where business terms and jargon is at the order of the day. The questions that we are confronted with on a daily basis amongst others are the following:

1) How big is the market for a specific product in a specific geographical area?
2) What are the critical success factors for this type of business?

3) What are things that will hinder this kind of business to achieve success?
4) Is this the right location?
5) If the pricing policy is X, is it correct and in line with the general norm in the industry?
6) How should the product be promoted?
7) What is the typical industry 'jargon'?
8) What are the benchmarks for gross profit margins, cost ratios, etc. that could be used in evaluating the business?
9) What traits, personality, skills or other abilities does an entrepreneur require to successfully start, manage and improve the business?

D. PROCESS

Some time will be spent on the process that followed.

1) **Obtaining national inputs**

National inputs, commitment, buy-in and specific requirements were obtained from the relevant senior members of staff at a think tank. This paved the way to put together a development team.

2) **Market Scan**

As a first step an in-depth investigation was done in order to identify what information sources of a business nature were available in the marketplace, but none were found – not for SME the sector. Practitioners in the small and medium enterprise field will be quick to comment that what is experienced in this field is different from what is experienced in the so-called big business arena. Many big businesses are highly diversified. Making use of their results to analyse and give advice to entrepreneurs in the SME sector is not only misleading, but at best very risky.

There are numerous databases available in the market – many of which have an industry focus – but all have a focus on big business. Market research at the time proved this to be true and if I may add - the same holds true for today. We are not aware of an information source in the market with a focus on industry related information in the SME field.

This fact was a big driving force in the decision to establish our own industry base that could cover industry information from both internal and external sources.

3) **Deciding on a standard Industry code**

First and foremost in the development process was to agree on which standard should be used when identifying an industry. The choice came down to designing our own codes and making use of existing codes in the market.

a) Designing own codes

1. At first we considered designing our own codes. The most acceptable manner in which industries are classified is either through a classification of occupations, commodities or economic activities. Each of these methods of classifications has its own advantages and disadvantages.
2. The classification based on occupations would not have supplied us with a proper basis for analysis as we are more focused on the business operation rather than the personnel involved.
3. Commodities are commonly associated with particular industries, e.g. clothing industry or rubber industry, but bear in mind that each industry generally turns out a variety of commodities which, although related, could nevertheless be accommodated separately in a commodity classification.
4. The classification of economic activities of industries is based on establishments engaged in the same or a closely related kind of economic activity based mainly on the principle class of goods produced or services rendered.

Mainly because of the shortcomings inherent in existing classification mechanisms there was a strong bias towards designing our own coding system at the time, but we soon realised that to undertake such a task would be a project completely on its own. In addition, the development and particularly the maintenance thereof will consume literally hundreds of man hours – a luxury that we could ill afford.

b) Using existing codes

1. SIC codes from Stats SA

    a. Although the code developed by Stats SA is based on the International industry classification it was not completely in touch with all our requirements. We nevertheless decided to use this as basis. The 5th version of the National Standard Industrial Classification of all Economic Activities (SIC) had then just been released.

    b. The added advantage we got from deciding to make use of the existing SIC codes was that all existing data we had captured on our systems already had a SIC code reference even though it was based on the 3rd version of the codes. In order to have that data available in the new version, we converted the data form the 3rd via the 4th to the 5th version. The reason for doing it in this way, was because there were conversion tables available from versions 3 to 4 and from 4 to 5, but not from version 3 to version 5.

    c. As was done in the past, we decided that each piece of data or information that would be captured in future must contain a SIC code. At the time an article appeared in Fortune magazine quoting a very well known British banking group which had just decided to capture an industry code on their database in all future transactions so that meaningful industry analysis could be performed. We pride ourselves on the fact that we had designed our legacy systems in such a way that that information was captured since inception in 1981.
    Although it was decided to use the industry classification of Stats SA, this by no means solved all our classification problems. Let me quote a typical example or two:

        i. According to the industry codes the code 63500A represents the "Retail sale of automotive fuel", but excludes the convenience store, the wash bay, a service centre and vehicle sales, each one having a different code. Constant change also takes place.

ii. The SIC code for "hotels, motels, botels and Inns registered with the SA tourism board" is 64101A, but there's a completely different rating system today.

iii. The question could also be asked what the definition is of a Guest House (64103C) or Guest Farm (64103B) and how it differs from a Hotel and Motel.

iv. What about a 'Bed & Breakfast' establishment for which there is as yet no code available?

2. Modifications

The following modifications were made to the codes and existing data:

a. The list of codes published by Stats SA were analyzed, some slightly modified in their description and about 10% were added to tailor make it for our needs bringing the total number to 943 codes.

b. The data that was captured over the years on our database was originally based on Version 3 of the SIC codes published in January 1981. This was reclassified from version 3 to version 5 programmatically. The correspondence tables supplied by Stats SA in their publication were put to good use in achieving this.

c) Identifying sources of information

1. Internal

a. Existing Investments

We realised that the most important source of information came from those investments that we did in the past. The most important task was important was to make the information from these deals available to every portfolio manager throughout the country. The deals mainly consist of a word processing document and a spreadsheet. We created a directory on each file server throughout the country that was accessible to everyone and then arranged that all deals be copied to this directory once the deal was approved.

As part of the knowledge base we created a link to our legacy systems enabling the user to do enquiries to the data base thus locating a deal which contains the relevant information.

b. Knowledge of portfolio Managers

   There's a saying that the mind contains the most information. Realising this, we made sure that the enquiry that the user does also contains the name of the staff member and geographical area where the deal was done.

c. Profiles

   Getting information on paper is good, but we realised that there is a lot more information available than that which is generally produced in a report to an investment committee. We then went through a process of obtaining a profile of a business from each one of our staff members and also made that available for use. We provided a structure within which they should work and they in turn forwarded the profile using a word processor.

Having focussed on the sources of information we had at our disposal in house, we then started focusing on external sources of information.

2. External

   a. Individual Contacts

      There are many individuals that could be classified as knowledgeable - people that have accumulated a lot of practical knowledge over many years. In many instances it just needs a phone call to one such individual to ensure a successful application. The functionality was added to the system to enable a staff member to add individuals whom they considered to be experts. Any staff member thereafter may then have access to the name and contact details of the individuals they need to get in touch with.

b. External Associations

In addition to individuals there are many organisations, institutes, associations etc. in the marketplace that have accumulated a wealth of information and who are able to share it. These organisations can also be added to the database for use by other members of staff.

To start the process, a database of industrial organisations was acquired and added. This resulted in approximately 1 000 records added to the database.

c. Press

A third element of information that was identified are articles that appeared in a variety of newspapers and trade magazines. At the time our front line staff did not have access to the Internet to do searches on the available Internet sites for information required when necessary. To make these articles available, as is, would obviously result in an infringement of the Copyright Act. To ensure that this would not be the case, we went through the following process:

i. Identify relevant newspapers and magazines.

The first step was to identify those newspapers and magazines that from time to time would contain relevant information about industries that may look attractive from an industry point of view.

ii. Obtained authorisation from each publisher.

A letter was addressed to each editor requesting authorisation to copy selected articles and make these available in-house to front line staff only.

We obtained authorisation from the majority of editors with the proviso that it be used only in-house for our own information purposes.

iii. Maintaining the same level of purchasing.

In the case of those publications that we scan articles from, we ensured that we maintained the same subscription level that we had before we started with the process. In many instances the subscription actually increased from what it was before.

iv. Making a selection of the articles

From the printed matter we then make a selection of the articles needed. This selection is to a great extent driven by two principles:

1. The articles must refer to SME-type operations and not big business.
2. It must be an 'industry' related article. By this it is meant that articles covering the functional areas of business are eliminated except if such an article has a specific industry focus.
3. At present we scan some 15 000 pages per month. Of this, approximately 400 articles are added to the Industry Information Base.

v. Scanning the articles

The selected articles are passed through a scanner and indexed

vi. Indexing the articles

An index is added to each article comprising of:

1. The source
2. The date of publication
3. A title
4. Sic code/s
5. Any franchise/s that may be mentioned.

vii. OCR Process

Each article then goes through an OCR process using Adobe Capture

viii. Store

The article, together with the index that was captured is then saved for future use.

E. TECHNOLOGY

1) Internal information.
    a) Our legacy system for investments resides on an Adabas database. A consolidated version thereof is exported monthly to SQL and made available using VisualBasic.
    b) Each portfolio manager was targeted and required to provide a profile of a business. A SIC code was added to each profile and each profile was then made available in HTML-format.
    c) SQL/VB programmes were written to accommodate individual contacts and external associations. We make use of Internet Explorer to reference ASP and HTML pages that retrieve data from a SQL database.
2) External
    a) Press articles are scanned using Adobe Acrobat Capture and captured with an index. They go through an optical character recognition (OCR) process using Adobe Capture. Microsoft Internet Information Server ensures full text search capability.

F. Internal Marketing

Shortly after the introduction of the database it was found that very few members of staff actually visit the Intranet site. In an attempt to enhance the use of the database (at least initially) two steps were taken

1) The submission
    a) At first a requirement was added that when a staff member submits an investment to management for approval, that staff member must quote information that was obtained from the database.

2) Newsletter
    a) We started with a monthly Expert Database Newsletter that was a brief summary of some of the information that was added onto the database during the past month. Initially this was in hard copy format, but for the past 18 months has been

done electronically and includes links to the relevant pieces of information on the database.

The requirement to add information in the submission phase of the investigation process has since been removed. We found that, having made the information more easily available by means of an electronic link, clearly enhanced the use of this source.

G. Sad moments

There were many successes, but also some sad moments: The saddest moment probably was when we were forced to get rid of most of the traditional library we had. Before the transition as a company we were internationally known for our library resource that was built up over many years. This resource focussed on information for the SMME sector across the globe. We were suddenly forced to reduce our library space from 350m5 to 40m5 and later even further to 24m5. The result was that we had to discard more than 80% of all the books, magazines and other publications that were accumulated over many years. May I add that we contacted numerous organisations that we thought would be interested in obtaining these sources, all of them thanking us for thinking of them, but no one showing any real interest in adding this resource to what they had at that point in time.

Madam chair, ladies and gentlemen:
Let me close by thanking the SA Online User Group for allowing me the opportunity of sharing some of our experiences. Thank you for listening to me and I trust that you gained something from my talk that you could put to good use in your own company.

Chris Cromhout
Business Partners Ltd
Tel: (011) 480 8707
Fax (011) 642 2791
E-Mail: CCromhout@BusinessPartners.Co.Za

Website http://www.BusinessPartners.Co.Za