



DCI Solutions for Business Surveys¹

1. Collection by Paper - Scanning and Intelligent Character Recognition (ICR)

1.1. The current Integrated Data Capture (IDC) system for handling paper forms needs upgrading and replacing. Alternative methods of scanning and ICR have been researched, and the future software will utilise question recognition rather than the current form recognition principles using templates. This step is essential to meet the anticipated requirements of bespoke forms using a question library, and will have the effect of reducing the constraints and lead times required for producing paper forms. Dependency on the 16-bit software components of the current software will be eliminated. The new methods will allow the procedures used for handling paper forms to share common systems and standards with other media, via the Business Collection Database.

1.2. Benefits of this approach include:

- the move to 32-bit software, essential for the proposed introduction of future versions of the Windows operating system;
- reduced lead time needed to make changes to inquiry forms;
- savings in the DVU resulting from faxed forms being integrated into the ICR process, rather than being keyed from image by the DVU (except on recognised 'golden days');
- this is one of the essential foundation stones needed before a Bespoke Forms system can be implemented; i.e. increased flexibility to change inquiry questions or to vary questions within inquiries.
- further integration of processes within the Integrated Data Capture area;
- extending the range of inquiries for which it is cost effective to introduce scanning and ICR, thus reducing DVU effort;
- introducing further standardisation across different modes of collection.

2. Fax

2.1. A considerable number of completed forms are returned to ONS by fax. Currently faxed forms for the majority of inquiries (excluding those inquiries with the larger more complex forms e.g. ABI) are now returned to the central fax server (FacSys). The images are indexed by the Integrated Data Capture team, and are then passed to the DVU for online data take-on. The current scanning and recognition software means that it is not possible to integrate incoming faxes

¹ This is an internal Office for National Statistics (United Kingdom) document provided for specific use in the OECD *STES timeliness framework*.

into the standard ICR process, as the quality of the incoming faxes is too poor. This depends on the sending fax machine, over which ONS has no control.

2.2. However, the introduction of the new question-based recognition software will enable incoming faxes to be integrated with the mainstream processing. This will reduce the need for online data take-on by the DVU, although this may still be needed in certain circumstances, for example, on the day prior to results being run, when batch updating will be too slow.

3. Telephone Data Entry

3.1. A new, digital, Telephone Data Entry (TDE) system has been procured and installed, and all existing inquiries have now migrated to the new system. Currently approximately 20% of forms are collected using this method. We believe there is scope to increase this percentage; firstly by implementing TDE on a number of additional inquiries that are suitable; and secondly by increasing the take-up of TDE by businesses.

3.2. MG is currently starting work to investigate potential modal bias of TDE, and to recommend standards for dialogues to ensure bias is eliminated or minimised. All inquiries (existing and new) using TDE will implement these recommendations.

4. Internet

4.1. Two pilot projects have been completed, on the Prodcum and Research & Development inquiries. These have given valuable information and have guided the future strategic direction of Internet collection. The proposal for the collection of business statistics over the Internet has been redesigned to use the infrastructure supported by the Office of the e-Envoy (OeE) for external communication and associated standards (e.g. e-GIF), and to adopt technical solutions within ONS which will fit with the strategic direction of the Information Management Programme. This will also enable the potential use of developments achieved by other departments (e.g. Inland Revenue, Customs & Excise).

4.2. As a public-facing organisation, ONS engages in two-way transactions with the public and with other parts of Government. The integrity of our processes and the confidentiality of the data collected is paramount. The Government Gateway will be used for the collection of business data because it provides:

- secure transfer between businesses and the ONS secure data repositories;
- an authorised and approved public interface;
- an authentication management service to enable Public Key Infrastructure (PKI) and Level 2 services for data collection;
- access to the authentication agreements of other departments that can be applied by businesses dealing with ONS for data collection activities;
- robust and resilient infrastructure for hosting ONS communication services.

4.3. Detailed proposals can be found at [. In summary, the services would fall into four main areas:](#)

Secure email for unstructured data

Certain business surveys require a few contributors to supply large quantities of data. Over the years informal arrangements have been made for return by spreadsheet or text file, but not to any consistent or recognised standards. These are known as "Special Arrangements". These contributors want to replace diskettes with some kind of electronic transfer. In addition, currently we actively try to discourage and prevent unsecured email, but without any alternative we are beginning to lose that battle.

This service is an interim measure to meet the demand for electronic transfer designed to avoid what would otherwise be an increasingly probable security incident. Having registered, the business would connect to the secure email system which would allow any file to be uploaded to the ONS website, converted to text and transferred by XML message through the Government Gateway. The only development required within ONS, additional to registration, is a system to route the data to the correct place in the office which is straightforward to implement, within the existing DCI workflow applications.

Interactive Web collection

This is the core service which needs to be made available for all business surveys by 2005 in order to meet Information Age commitments. It allows interactive validation at the point of capture, therefore improving data quality and reducing compliance by minimising recontact with the contributor.

Automatic transfer of XML documents from business and accountancy applications

This would take accountancy and other business data direct from the software applications used by the businesses, as described in the section below. This is still at the discussion stage with Customs & Excise, Inland Revenue and Companies House. However, it is a nice feature of the Gateway that once we have published the XML standards and some software companies have implemented them, the infrastructure will be ready for immediate implementation. Realistically, this is going to take at least another two years but has major implications for reducing the compliance costs for businesses completing business inquiry forms when successfully implemented.

Bulk capture of administrative data

The Government Gateway and/or GSI could also be used for the capture of data from bulk administrative sources; for example, from the Customs & Excise's VAT system for input to the Inter Departmental Business Register, or from Inland Revenue's PAYE system for the New Earnings Survey.

4.4. MG will be involved in advising on the design of Internet questionnaires and other issues to ensure modal bias is eliminated or minimised.

5. Groupware / workflow

5.1. ONS's current policy is to use Lotus Notes as the core support tool for internal communications and business management. PBG has a number of workflow applications to assist the processing and validation of returned forms. With the advent of Notes 5 further improvements in integration are possible. Improvements in communication between different systems are also being considered, using Lotus Notes as a front-end. These initiatives will ensure that PBG

processes business inquiry data in the most efficient way possible, minimising duplication of effort and ensuring that quality control procedures are built into the process.

6. The Business Collection Database

6.1. All of the above systems (for paper, fax, TDE and Internet) can be developed and implemented as separate systems, but this implies large overhead, maintenance and data integration problems. We propose to solve these problems with the introduction of the Business Collection Database.

6.2. The Business Collection Database will be used to drive all the processes connected with the capture of data for inquiries, integrating the various collection media. It will underpin the collection of data, irrespective of the medium by which that data is collected - it will cater for data returned to the office on paper, by fax, via the Telephone Data Entry (TDE) system, and via the Internet. Eventually the database will also drive the production of the collection instrument; ultimately it could potentially be used by the DVU when validating and editing returned data.

6.3. Benefits of this approach include:

- a single integrated system supporting and driving data capture via all types of media, and managing increasing mixed mode data collection in a standard way;
- increased commonality and standardisation of processes, and therefore reduced maintenance costs;
- the provision of considerable data for the management of the collection processes, including for audit and monitoring processes;
- it is one of the essential foundations that must be in place before Bespoke Forms can be implemented for the fast and flexible production of forms; which leads to
- increased flexibility for customers to change inquiry questions and the ability to vary questions within inquiries;
- therefore giving a reduction in compliance costs;
- a number of technical solutions to dealing with the transition to the Euro become available, in contrast to the current scanning and recognition system which offers limited scope for a solution and would be at risk of being unable to cope with the required increase of formtypes.

6.4. The Business Collection Database will be designed and implemented in such a way that it can be used to drive the collection of any data required by ONS or other Government Departments, not just the collection of business data.

6.5. Further details of the Collection Database, and how it fits into the proposed corporate vision, can be found in Appendix C.

7. Business systems software

7.1. The use of accountancy software to gather information from businesses without the need for them to complete business inquiry forms has significant potential to reduce compliance costs for

businesses. This is particularly the case for short-term business inquiries where the relatively aggregate data collected is more readily mapped to business accounts which do not have any standard framework or definitions. BDD is working with Customs and Excise (C&E) and the Business Applications Software Developers Association (BASDA) to build up a common data requirement to be used by the commercial software developers. There are also potential quality and congruence benefits of accessing company data via accountancy software.

7.2. A number of companies currently market business applications software, including accountancy software, to help businesses manage their company accounts. Much of the information that ONS collects via business inquiries is held within software. However, it is not held in a way in which a business that uses such software can readily extract it for passing to ONS as an alternative to completing an inquiry form. There are two reasons why the information cannot be readily extracted. Firstly within the United Kingdom there is no standard accounting framework. This means that the information stored by businesses will vary in detail and will be coded differently. Secondly the level of detail requested in ONS business inquiries, particularly the detailed annual inquiries, will not match the data held in the accountancy software.

7.3. Firms that market accountancy software have shown some interest in the past in working with ONS, to develop modules that extract the information from the accounts data held within the software, in a format that readily provides the information needed by our business inquiries. However, this interest did not lead to any positive outcome in the delivery of any modules. One reason for this was thought to be the case was that ONS believed the software companies would develop the modules from their own resources, as they would see a marketing benefit. This proved not to be the case. BDD is now taking an alternative approach, working closely with C&E.

7.4. There are clear benefits from businesses being able to provide their data directly from accountancy software:

- data quality should be high as it would be taken directly from the systems that the businesses use to monitor the company finances;
- businesses should have less difficulty in understanding what data is being asked for than on the paper inquiry forms, as the software would link to the correct data;
- businesses would have a much lower compliance burden as the software would extract the required data automatically, thus avoiding the transcription needed for other data collection mediums, and also eliminate associated transcription errors;
- there is scope to take more data from a business than is strictly needed by an individual inquiry, as this would be as easy for a business to provide as a more restricted set;
- the software would ensure that the data sent to ONS was in an ONS defined format.

8. The Comprehensive Business Directory

8.1. Over recent years the ONS has been leading a project to develop a Comprehensive Business Directory (CBD). Although not an integral component of the DCI Strategy the CBD is closely linked, particularly with regards to the work being undertaken on Internet collection.

8.2. The main thrust for the CBD project has been to:

- work towards achieving the joined-up government principles set out in the Modernising Government agenda, and
- work towards the 2005 electronic government targets

8.3. The project has been undertaken in collaboration with HM Customs & Excise, Inland Revenue and Companies House. Its primary aim is create a CBD as a central facility, which can provide a range of services by acting as a hub containing information to link data relating to the business community held in administrative data sources.

8.4. To address this the ONS CBD team has commenced development of a prototype virtual data warehouse model, whereby electronic access by ONS to HM Customs & Excise and Inland Revenue data can be promoted and strictly controlled. The model also includes improved access to public domain data held by Companies House and Yell. Better access to data will result in improved quality of business statistics. There is also the potential for achieving more consistency between departments, especially in the case of the classification of businesses.

8.5. Data transfers from the administration departments will be via direct and secure electronic means, providing greater resilience and more streamlined and timely flows. Security of the data transfers is key to the concept, and the project will make use of the government standards for data exchange and authentication that are provided within the Government Gateway strategy.

8.6. The development of a CBD model is based on matching of data, in the absence of a single business registration system for government. The quality of the matching tools will determine the extent to which the model can be implemented more widely within government.