Department of the Environment:
Contaminated Land Research Report

INFORMATION SYSTEMS FOR LAND CONTAMINATION

Prepared by Meta_Generics Ltd

CLR Report No 5

1994

Department of the Environment:
Contaminated Land Research Report

INFORMATION SYSTEMS FOR LAND CONTAMINATION

Prepared by Meta_Generics Ltd

CLR Report No 5

1994

Printed and published by the
Department of the Environment 1994

DEPARTMENT OF THE ENVIRONMENT

CONTAMINATED LAND: Identification, assessment and control

INFORMATION SYSTEMS FOR LAND CONTAMINATION

META_GENERICS LTD

This report is one of a series of reports financed under the contaminated land research programme of the Department of the Environment. The current series deals with: information needed to assess risks; procedures for categorising and assessing risks; and evaluation and selection of remedial methods.

The purpose of the reports is to provide regulators, developers and other interested parties with authoritative and researched advice on how best to identify and assess the problems contamination can pose and what can be done to tackle them. They cannot, however, address the specific circumstances of each site. Every site is unique. Anyone using the information in a report must, therefore, make appropriate and specific assessments of any particular site or group of sites. Neither the Department nor the authors can accept liability for the use or interpretation of the contents of any report.

General guidance on assessing contaminated land and developing remedial solutions which is complementary to the series is provided by the Construction Industry Research and Information Association (CIRIA).

The Department wishes to acknowledge the work of the late Michael Biscoe-Taylor of Meta_Generics Ltd in preparing the report, and the help of the Local Government Management Board.

CLR 5 prepared by:
Meta_Generics Ltd
The Jeffreys Building
St John's Innovation Park
Cowley Road
Cambridge CB4 4WS

CONTENTS

1. INTRODUCTION

1.1 This report provides guidance on the development and management of systems for holding information on land where contamination is known, or suspected, to be present, or which is affected in other ways by contamination.

1.2 The guidance is intended for use by any organisations which wish to hold information on land affected by contamination. These may include industrial or corporate landowners, developers, local planning and environmental health authorities, and any other bodies with either interests in, or responsibilities for, such land [1].

1.3 The guidance covers:

- consideration of the types of organisation which may need to compile or make use of information on land contamination, and the ways in which they may use it;
- the types of information which may be useful for such organisations to hold; and
- the management of information, including recommendations on quality control.

1.4 The level of sophistication of the information system needed depends on the users' requirements. The general principles set out in this report apply to both paper-based systems and computer-based systems. However, as many users are likely to adopt computer-based systems the emphasis is on this method of holding information. Guidance on specific information technology and on data specification is given in Annexes A and B.

1.5 This report does not cover either the collection of information on land which may be affected by contamination or any aspect of its assessment or treatment. These matters are the subject of other reports published by the Department of the Environment or in preparation for publication.

---

[1] **Readers of this report should note that the Government has withdrawn its proposals, under s143 of the Environmental Protection Act 1990, for a statutory register of land subject to contaminative use.**

2. USERS AND USES OF INFORMATION ON LAND CONTAMINATION

USERS

2.1 Individuals and organisations who may find information on land contamination helpful include:

- Local authority departments
- Utility companies
- Regulatory authorities
- Emergency services
- Insurers
- Financial institutions
- Property developers
- Industrial land owners
- Other corporate land holders
- Development agencies
- Estate agents
- Surveyors
- Solicitors and other legal advisers
- Private individuals
- Environmental groups
- Scientific, historical, and archaeological societies

USES

2.2 Information on land contamination may be used both to determine long-term policies and to deal with day-to-day problems.

2.3 The main uses are:

- Identifying and assessing possible effects on health and the environment
- Planning remedial action
- Planning future uses of land
- Assessing land value
- Assessing liabilities

2.4 Holding the information required in a structured way will assist organisations to carry out their activities in an efficient and effective manner, and may in some cases be the basis for competitive or strategic advantage.

3. RELEVANT INFORMATION

3.1 This section indicates the broad categories of information on land contamination that is likely to be useful to many organisations. Both the extent and nature of the information held and the sophistication of the information systems required will depend on the particular organisation's strategic and operational requirements and its resources.

LOCATION OF LAND AFFECTED BY CONTAMINATION

3.2 The following information will usually be needed to identify the location of a site:

- a unique reference, preferably indicating the organisation as well as an alphanumeric systematic numbering of the record;
- the name of the site;
- the address of the site;
- a map reference for the centroid (centre point) of the site;
- a map showing the site and surrounding development; and
- any qualifying statement regarding the accuracy of the information on location.

Note: Large areas may be simply shown on maps or given map references for estimated site boundaries. Sometimes it may not be possible to plot the site boundary accurately, for example an old waste disposal site. In such cases it may be possible to estimate the site boundaries by reference to current or historical features. Where historical features are used, the site might be plotted on an historical map and the references transferred to a current map. Any assumptions made in defining boundaries should be recorded and the possible inaccuracy of the plotting noted.

BS7666: Spatial Data-sets for Geographic Referencing, parts 1, 2 and 3, specifies an alternative mechanism for identifying land and property uniquely.

3.3 It may also be appropriate to record other information relevant to the location of the site, eg:

- the district or town;
- the local authority;
- the region; and
- details of neighbouring or other nearby sites.

HISTORY OF THE SITE

3.4 The following information may be needed as a basis for assessing the possible and actual contamination on a particular site and planning any further action:

- past and current uses of the site;
- dates and nature of any investigations;
- dates and nature of any remedial action; and
- details of any other action relating to the site, eg development proposals, legal actions or applications for funding.

INTERESTS IN THE SITE

3.5 Certain organisations holding data on individual sites may need to include:

- the name and address of the owner of the site;
- the names and addresses of others with interests in the site; and
- relevant local regulatory bodies.

3.6 A record of individual users of the information, showing the nature of their enquiries and which items of information they have used, will enable them to be contacted if information is subsequently changed. Such a record may also identify people or organisations who may be interested in information about new sites.

CONTAMINATION AND ITS EFFECTS

3.7 It may not be practicable to record all available scientific and technical data on individual sites or on a group of sites in a uniform way. However a standard summary sheet could contain references to more detailed reports held on individual files. The following information may be appropriate:

- data on known effects of contamination, eg reported pollution incident;
- results of any investigations, eg copies of checklists or tables of results;
- geotechnical data;
- the groundwater regime;
- NRA groundwater protection zone or river quality classification;
- other environmental information, eg location of SSSI;
- cross-reference to relevant source of information on potential problems or to checklists or other guidelines to ensure uniformity of assessment; and
- information on any assessment of the site or categorisation of problem.

3.8 Where there is no information on a particular substance or condition, it should be made clear in the summary whether this is because no information is available or whether the substance or condition has been investigated and found to be absent.

PROPOSED ACTION

3.9 The nature of action on sites is of course dependent on the nature of the organisation holding the information. Information which may be relevant in dealing with contamination may include:

- proposals for any further investigation;
- proposals for any remedial work;
- applications for funding; or
- proposals for any other action, including any notification requirements.

IDENTIFYING DATA SOURCES

3.10 Organisations collect data for different reasons, and data are always to a greater or lesser degree coloured by the reason they are collected. Data may also be incomplete. Original sources of information, therefore, should be clearly referenced, whether or not the actual information is reproduced in the system. Any interpretations put on data should also be recorded. So also should any contradictions between one piece of data and another.

3.11 Careful recording of the sources of data and its limitations may suggest what level of confidence may be placed in the data. It may also indicate whether further and more detailed studies are necessary.

3.12 Supporting information about a particular site or other relevant information may be held in other locations or on other systems. Where this is the case, information should be included with the site details to show users how to obtain those further data.

CONFIDENTIALITY OF INFORMATION

3.13 It may be necessary to identify some information which cannot normally be released, for example to protect commercial confidentiality.

LEGAL REQUIREMENTS

3.14 Organisations holding information on land contamination should be aware of, and comply with, legal requirements relating to the collection and provision of any such information which is relevant to their function. The following provisions may be relevant:

- Data Protection Act 1984;
- Environmental Information Regulations 1992;
- Copyright law; and
- Current case law and precedents relating to use of electronic data in evidence. (As of July 1994, this area of law is under consideration by the Law Commission. Laws which currently affect this area include the Civil Evidence Act 1968, the Police and Criminal Evidence Act 1984, and the Criminal Justice Act 1988.)

3.15 Specialist legal advice may also be needed on the storage, use, sale and dissemination of information owned or obtained from other people or organisations.

4. ORGANISATION AND MANAGEMENT OF INFORMATION

STORING AND PRESENTING THE INFORMATION

4.1 The resources used in collecting and storing information, whether in 'real' or electronic form, will be wasted if the information cannot be used properly. Would-be users need to know about the information, to be able to use it with the minimum of time and cost, and to be able to trust the material. "Users" should be interpreted widely; they could be people in other parts of the organisation which holds the information and possibly people outside it.

4.2 This section sets out a general approach to holding information so that it is useful, understandable and reliable.

4.3 The information may comprise a large number of individual records and elements of data. These should be stored simply and logically, preferably using standard formats throughout an organisation.

4.4 If a standardised system is to be used it is essential that all users' requirements are assessed at an early stage, both in information and operational terms, to ensure that the system meets the relevant needs. Specialist advice may be required to ensure that any new system is the most suitable for a particular organisation and delivers what is expected.

ACCESSING THE INFORMATION

4.5 Users need to know how to access the information. They may also need to have guidance on how to interpret it. This requires effective communication between the people who compile and hold the information and those who use it. Adoption of data standards (see para 4.16) will help both the compiler and the user.

4.6 The existence of information systems and where they are should be made clear to those who may need to use them. Users should be trained or given instruction in the use of the systems and, where necessary, in the interpretation of the data. They should also be made aware of any limitations in the data and any factors which may affect its interpretation.

4.7 Access to information may be facilitated by indexing the information. Information on sites may, for example, be indexed using any of the following data:

- the unique reference for the site;
- the address of the site;
- a general geographical area (eg district);
- current or past site uses; and
- contaminants known or suspected to be present.

MAINTAINING RECORDS

4.8 It is obviously important to ensure that the information held is correct and is kept up-to-date. The date of compilation of the record should be recorded and

additions, amendments and deletions should be carried out systematically. The deletion of records should be strictly controlled. As of August 1994, the Local Government Management Board, LGMB, (see section 5.2) was due to publish a paper entitled "Standards for digital geographic information".

4.9 A system should be established to ensure that changes to records can only be made by authorised individuals. A key safeguard is the maintenance of a log of all entries and changes to the records, showing who carried out each change and the date. Further advice on managing security is given in paragraphs 4.22-4.25.

NOTIFYING USERS OF CHANGES

4.10 Information on individual sites will change as new data become available, inaccuracies are corrected and new interpretations are made. There will be a knock-on effect on data held by other organisations, and there may also be implications for decisions made on the basis of previous information. Maintaining a log of users who have accessed or copied particular information, or who have an interest in it, will allow them to be notified of changes. This process can be automated if computer systems are used.

4.11 Where users hold copies of data, there is a risk that they will keep and use out-of-date material. There could be confusion with both superseded and new information in circulation. A system of "version control" can be used in such cases to enable the current data sets to be identified easily. Versions or updates may, for example, be identified by numbers, or by including the date on which the information was created or last amended.

4.12 In some cases it may be appropriate to inform users that information is only valid for a specific period of time.

INTEGRATING INFORMATION

4.13 Many different organisations will have information on contaminated land, or will perform activities which require such data. Often, several systems within individual organisations will be based on similar underlying data structures; for example Local Authorities with environmental health, planning, building control, estates, housing and works departments.

4.14 Benefits to be gained from integrating data sets within organisations and between organisations include:

- reducing duplication;
- improving consistency;
- improving currency;
- providing fewer points of access;
- providing more complete information; and
- providing faster access.

4.15 Manual integration of data can be achieved through paper-based indexing systems. This can be further improved by providing the indexes on simple computer systems. Integrated computer systems provide automatic searches of different databases to provide information about a common site or identification of sites with common features.

4.16 The integration of electronic information requires conventions to be agreed and followed. Standards are required to define data structures, minimum data sets, naming conventions and data transfer methods. The Land and Property Gazetteer Working Party run by the LGMB is one group which defines data standards for land-based information holdings. The Working Party has recently developed a British Standard for recording information about land and property, which is described further in Annex B.

4.17 As mentioned in paragraph 3.10, different organisations often have varying reasons for collecting or holding data. Factors which might affect the interpretation of the information must be taken into account when accessing or integrating data from external organisations.

CONTROLLING QUALITY

4.18 The quality of information is principally determined by the accuracy and completeness of the underlying data. It is impossible to guarantee that data are error-free. Searching for errors and having effective arrangements for correcting them are essential.

4.19 A formal quality control method will help to reduce inaccuracies, errors of input and omissions. As a minimum, the quality control process should comprise:

- data verification procedures built in to the system, eg at point of entry of computerised records;
- a definition of the responsibilities of maintainers of the system;
- procedures for maintaining and using the system, including data checking and reporting errors; and
- periodic audits of the system.

4.20 Indicators of the success of the quality control process include the reduction or elimination of:

- typographical errors;
- incomplete records of sites and associated details;
- occasions on which information is misinterpreted; or
- complaints received.

4.21 Safeguards against corruption of IT-based data from viruses should also be built in to the system management procedures.

MANAGING SECURITY

4.22 Information on contamination can affect land valuations or development decisions. It is important, therefore, that the information available is correct and not susceptible to unauthorised changes.

4.23 Unrestricted access to systems or information is not normally desirable. The normal method of restricting access in computer systems is to provide password protection to the system as a whole and to grant individual users access

privileges to particular areas of information. This might take the form of identifying different groups of users, some of whom have the ability to modify data, with others restricted to viewing or copying information.

4.24 It may be necessary to maintain master copies of information, to which access is restricted. Working copies of the information can be released to users, if required. This is more difficult to achieve with paper-based systems where it is not practical to maintain several copies of substantial quantities of underlying data.

4.25 Responsibility for the security of the information should be formally placed with a specified post, and incoming holders of the post must be made aware of their duties at the outset.

5 USE OF INFORMATION TECHNOLOGY

5.1 Many of the considerations relating to the use of information technology for information on contamination are similar to the considerations which apply in other areas. The decisions on which method to use and the extent of computerisation are likely to be taken in the context of wider organisational issues.

5.2 Advice on the selection of technology is available from a number of organisations. Local Authorities may like to contact the Local Government Management Board (Arndale House, The Arndale Centre, Luton, Bedfordshire, LU1 2TS). Other organisations may have access to similar central advice, for example to "Information Management, a Director's Guide" produced by the Institute of Directors for its members, or can contact the Association for Information Management (20-24 Old Street, London, EC1V 9AP.)

5.3 This section sets out the general considerations which affect the decision to use an IT-based system and the form it might take. Annex A reviews the key hardware choices for electronic data storage and input, and Annex B provides details of standard format and terms which can be used in databases of information.

ADVANTAGES

5.4 Computerisation offers scope for efficiency and flexibility in storing, accessing, manipulating and presenting data on contamination. The following represent the major advantages possible, many of which can be realised with low levels of computerisation:

- rapid scanning, retrieval and sorting of information, including the import of data from paper or electronic sources;
- transfer of information to other applications such as word processing and spreadsheets;
- provision of information in a variety of presentational formats, including display of textual and graphical information together and highlighting or screening of different elements for individual users and types of user;
- rapid distribution of information in high quality format;
- direct and quick transfer of information between organisations, allowing data to be accessed more quickly, improving the currency of data, decreasing the amount of duplication and hence possible errors;
- provision of management information (such as usage of data) as a by-product of the system operation rather than having to be collected separately;
- reduction in the volume, and requirements for access to, hardcopy information; and

- improvements in accuracy and security of data.

COSTS

5.5 In considering the types of system to be employed, it is advisable to perform a cost-benefit assessment. Costs include:

- equipment and software;
- installation;
- implementation, including data acquisition and entry;
- changes to operational procedures;
- training;
- system management (access; amendment; security; publicity); and
- on-going maintenance and system development.

5.6 Purchase of systems which are based on modular components or otherwise allow for expansion can provide a means of spreading the cost of installation and implementation of a new system or tailoring it to match demand.

SPECIFICATION OF REQUIREMENTS

5.7 As a minimum, a computer system to support information holdings on land contamination should allow:

- recording of data in a structured way;
- amendment to existing records and creating of new records;
- deletion of erroneous or superfluous records; and
- viewing of information held.

5.8 To reduce errors, the system should validate information as it is entered. Examples of this include making sure that descriptions correspond with allowed terms and that dates recorded are in the correct sequence.

5.9 The system should ensure that a record is created of changes, identifying the user performing the action and the nature of the action.

5.10 Additionally, the need to do the following should be considered:

- view map and other graphical information;
- view ancillary data (such as detailed information under a particular category or other related information);
- print data in a variety of formats;
- import and export data to and from other systems;
- produce summary statistics on the information; and
- produce summary statistics on the use of the information.

5.11 The system should be capable of restricting access to data or built-in operations, by individual users. An audit trail should be included to ensure that users are making the intended use of the system. The following points will need to be considered in developing the specification of the system:

- the breadth of information to be made accessible - whether it is to be restricted to information on contamination, or extended to include

additional related operational information;

- whether or not users will be allowed to take copies of the information holdings for themselves;
- whether access is to be provided to a paper or electronic copy of the information;
- the methods of accessing information to be used (for example on-line, remote access, telephone support);
- if access is to be through an electronic system, what functions will be provided and what restrictions there will be, particularly on modifying information; and
- the geographical locations of users who will be given access to the system.

SYSTEM MANAGEMENT

5.12 The success of an information system depends on its design, implementation and on-going development. The chances of success are greatly increased if a single post within an organisation has delegated authority and responsibility for the whole system.

5.13 After developing the system the following points require particular attention:

- Implementation, including debugging and problem-solving and changes to working practices. A method of recording problems with the system, its use and the data should be defined. In some organisations a central advice point may be appropriate.
- Training, including use of system, changes in working practices and presentation and use of information.
- Security, by means of periodic checks.
- Data maintenance. Accuracy of data should be reviewed periodically.
- System use. Ideally the system and practices should be reviewed periodically to ensure that the maximum benefit is being obtained.

ANNEX A: COMPUTER TECHNOLOGIES

PCS AND WORKSTATIONS COMPARED WITH MAINFRAMES

A1 This annex provides introductory information which may be relevant in selecting a system. It is not intended to provide an exhaustive review of the available technologies and software available, since this would rapidly become out of date.

PCS AND WORKSTATIONS COMPARED WITH MAINFRAMES

A2 The costs and timescales for developing PC and workstation software are lower than those for mainframes, and the computer industry is tending to concentrate more on PCs and workstations and less on mainframes. There is therefore greater competition between suppliers of "open systems" than between suppliers of mainframes.

PCS COMPARED WITH WORKSTATIONS

A3 The distinction between PCs and workstations is becoming blurred as the power of PCs increases. Workstations have higher screen resolution, speed of screen updating, processing power, storage capacity and cost. It is likely that, increasingly, systems will include both, using PCs for general text and graphics work and workstations where the extra performance and display capabilities are required.

COMPUTING SOFTWARE FOR PCS AND WORKSTATIONS

A4 Commercial databases specifically designed to store information about land contamination are being developed. These would need to be checked on a case by case basis by any organisation considering their purchase. Factors to consider include:

- relevance to the current and likely future requirements of the particular organisation;
- format and completeness of data fields;
- flexibility for adaptation or extension;
- accuracy, efficiency and underlying scientific or technical basis of any built-in operations;
- compatibility with existing or external information bases;
- hardware requirements; and
- training requirements.

A5 Many systems have the facility for several views of the same information to be shown concurrently. These can also allow different applications (e.g. spreadsheet, database and word-processor), to run simultaneously and to transfer data.

Word-Processor based systems

A6 With the exception of graphic display of maps, much information on contaminated land could be held in word-processor systems. Most organisations have standard word-processing packages which are familiar to their employees. Word-processor packages offer fewer facilities for manipulating data than database systems as the information is held as "text". However, word processors require less training and only basic computer hardware.

Databases

A7 Databases allow the information held to be structured and inter-related. This enables the data to be manipulated and interrogated easily. Combining the database with user-interface applications allows the information to be presented on screen and in ways tailored to users' requirements, thus reducing the need for paper-based reports.

A8 Database systems offer advantages in ensuring the accuracy and consistency of data. Information can be checked at the point of entry; key words can be confirmed and the information verified in the context in which it exists. The relationships between data elements also allow changes to be made throughout the database in a consistent manner; for example, criteria for estimating zones of potential influence of contamination can be changed universally and the new boundaries automatically recalculated.

A9 There are many proprietary packages suitable for constructing and managing databases. These require varying levels of programming ability and offer different degrees of sophistication. Systems can be built on large computer networks or on single-user PCs.

Distributed systems

A10 Distributed information systems are becoming increasingly common. These allow many users to access a single copy of the information. Users may be connected directly to the system (eg via terminals to a mainframe) or may access the data via computer networks, or the connection may be a combination of the two. The information system may be required to operate concurrently with other systems (eg a PC running two applications simultaneously), or to allow data to be passed directly between it and other applications.

A11 Distributed information systems only require one set of data to be maintained and result in better version control. However the structure of a distributed system is more complex than that of a stand-alone system and is more difficult to implement. As a result they are often initially more expensive than other types of system. However, they offer greater potential for expansion and integration.

Geographical Information Systems (GIS)

A12 Commercially available GIS are principally electronic map systems. They are capable of handling vector and, in some cases, raster based "electronic maps" (see section A17), storing information on the type, location and geographical coverage of features, displaying maps on screen and presenting analyses of the data stored.

A13 Extra information about features (for instance road numbers and town names) may be stored in the database and retrieved, and analyses of the information made. Facilities are incorporated for editing maps, adding features to maps and defining new features. A GIS may provide the benefits of "electronic paper": easy editing and viewing, automated searching and automated analysis of the contents.

A14 A GIS normally requires a powerful workstation for processing large amounts of map data and features. GIS packages can be costly (of the order of £100,000 at 1993/94 prices). There are smaller GISs which will run on PCs, handling smaller amounts of information. These are suitable for limited, single-user operation.

A15 The costs of maintaining a GIS database can be high. Many existing GISs require detailed expertise to operate the system and are often difficult to integrate with other systems. The underlying data structures and functionality are often geared towards a particular application and type of user. These may prove difficult to adapt to the needs of contaminated land information.

A16 The term GIS is also used to describe more general information systems. These often have a graphical user-interface which shows information overlayed on a display map. They may allow users to access further data by selecting map coordinates. Such systems are designed to meet much broader information handling requirements than commercial GISs, and therefore to meet the operational requirements of many different types of user.

MAPS

A17 Paper maps are very desirable as visual aids in positioning sites and in understanding issues such as the potential for migration of pollution to or from other locations. Electronic maps are now available for use on computer systems. There are two choices of electronic maps; raster and vector as described below.

Raster maps

A18 Raster maps have the following characteristics:

- They are essentially an electronic photograph of a paper map. The computer is not supplied with any information to interpret the features on the map and therefore, for example, cannot count the number of houses shown.

- They are difficult to edit. Typically, old electronic images would be replaced wholesale by new sheets rather than edited piecemeal.

- The data files can be large and slow to display or move across a network. However, compression techniques can allow files to be reduced to more manageable sizes.

- They do not scale well; being similar to photographs they become increasingly grainy on enlargement, and on reduction they lose important detail - text quickly becomes illegible. Typically, in place of scaling, maps of several different scales have to be recorded individually.

- As with paper copies of a map, problems can occur when trying to align the information on neighbouring map images. Scrolling may be restricted to areas within a given map to avoid these problems.

- There is generally no control over the density of detail that is displayed. Limited control can be achieved if maps are separated into several overlays, some containing only major features, such as large towns and major roads, and others containing other, more detailed features. The techniques used for storing and overlaying features selected by the user are essentially vector map-based techniques. However, this may be sufficient in small systems.

- Print-outs, especially large ones, can have a jagged appearance because they are restricted by the resolution of the image file (and by the printer).

Vector maps

A19 Vector maps have the following characteristics:

- They are composed of individual features which are stored separately.

- They can be re-scaled easily, while retaining display quality. For example, text can always be printed legibly and symbols changed to suit the scale.

- The type, extent and level of information displayed may be controlled. This allows maps to be redrawn to show only selected features and types of features, avoiding displaying unwanted information.

- It is possible to correct any mismatches between page edges when building vector maps. This allows maps to be scrolled freely in any direction without discontinuities appearing in the display.

- It is possible for the appearance of the print-out to be as good as the printer can deliver. Features can be described in sufficient detail to make lines and edges smooth.

A20 Vector maps are a versatile tool, with strong presentation capabilities. The computer power required depends on the size of the application. Users' features are coded in much the same way as base map features and are incorporated within the map.

A21 Vector maps offer significant features beyond display capabilities. Since the

computer identifies each feature, it can make use of this for other applications, for example calculating the density of residential housing around a site, or determining the numbers of planning applications in a given area.

A22 However, vectorising maps for Geographical Information Systems can be expensive, requiring considerable manual input.

Availability of electronic maps

A23 Several major projects have developed raster mapping on laser disk, for example the BBC's Doomsday project and Shell's travellers' direction finder. However, if organisations wish to construct raster maps for storage on computer hard disks (rather than laser disks) they may have to commission the production of maps.

A24 The Ordnance Survey (OS) has a programme for vector digitisation of all their major mapping for the UK. This will be completed by December 1995. It is not likely to be economic for most organisations to vectorise maps for only their own purposes. However, where the OS is not able to meet an organisation's requirements, other users of geographical information, such as the public and private utilities, may be able to help. Local Authorities have access to the digital OS data free of charge under a Service Level Agreement. Non-digitised, large-scale maps may still be required for detailed survey work.

A25 For both rasterisation and vectorisation of map data, bureau services will be available, to which organisations could subcontract work. The total cost of converting all information required must be appreciated in advance - otherwise conversion costs may be seen as escalating to the extent where they cannot be justified, stopping the conversion process before completion and resulting in useless partial information.

SCANNERS

A26 Scanners provide a means of transferring map or other graphical data into electronic form. The type of scanner required will depend on the method of map representation chosen.

A27 For vectorisation, scanners will need:

- big sheet handling capability;
- wide contrast range;
- good linearity; and
- good accuracy.

A28 They are however expensive, (of the order of £40,000 at 1993/94 prices). A good computer support system is a necessary part of the package, providing automated vectorisation with verification, and the ability to edit features to correct mistakes and resolve uncertainties.

A29 The specification for scanning raster maps depend on the display resolution required. The technology is relatively straightforward and cheap (of the order of £5,000 at 1993/94 prices). The equipment required is a high resolution colour camera, a good lens, and a PC with a "frame grabber" and processing software. Ideally, a single camera image will correspond with one image on the

target computer screen.

DIGITISATION TABLETS

A30 A digitisation tablet is an electronic plotting table and provides a method of inputting small amounts of graphical data manually. Maps are laid on the table surface and points identified using a special pointing device. The point location is detected and translated into co-ordinates offset from an origin on the tablet, which can be passed on to a computer for calculating grid co-ordinates.

A31 A large size tablet would be needed for typical historical map sheets, in which case a tablet with good accuracy will be expensive. Using digitisation tablets can be helpful, speeding up the process of scaling between different maps, calculating changes in coordinate systems and transferring site positions and boundaries to current maps. However, it is possible to build an information holding without using one. Some errors in historic maps may be corrected by checking site positions in relation to adjacent features which are recorded accurately on more recent maps.

ANNEX B: FORMAT AND SPECIFICATION OF DATA

B1 This Annex sets out general criteria and specific examples of the way in which data can be recorded. It covers:

- general format;
- creation of records;
- specification of data; and
- a typical specification

B2 The information on contamination may be maintained in a variety of ways, according to the facilities available to the individual organisation. The format and general specifications in this Annex have been developed to be appropriate for all physical forms of data holding, including but not limited to typed documents, word-processed documents, databases and geographical information systems.

B3 Specifying a common format for all data holdings ensures consistent appearance of information and will also allow free interchange of information. For data available in electronic form, the use of the common format will allow automated loading of entries into a database or geographical information system. For information only available in paper form, the common layout will reduce the manual intervention required for the electronic capture of entries by means of optical character recognition.

B4 A specification for electronic data is given which conforms to BS7567: 1992 "Electronic transfer of geographic information".

GENERAL FORMAT

B5 For large holdings of data on land contamination the following basic approach is suggested:

- a map index: this could be a scale map indicating the locations (centroids) and boundaries of land, for land for which information is held, and a reference to the number of the record or file where the information is held; and

- one record or file per site or other land area, for example a basic land or property unit, for which information is available, which contains the information about the site or land area.

B6 Useful optional sections include:

- a summary listing the names of each site or other land area for which data is held, and other key data relevant to the holding organisations interests; and

- an index, by key information feature, for example ownership or stage of development, to all sites or land areas for which data is held.

B7 An introductory section is recommended to provide guidance for those who have access to the information. This could include:

- a summary of the purpose of the information and the method of collection;
- a description of the structure of the information;
- instructions for operating any computer systems required to access the information;
- any indexes to maps that would assist users of the information;
- the procedure for accessing source data and contact personnel;
- the procedure for obtaining copies of data; and
- any disclaimers or limitations of responsibility for correctness and completeness of information.

CREATION OF RECORDS

B8 The record or file is the key unit of information. This section sets out issues to be considered in creating a record or file.

Entry criteria

B9 It is expected that the basic criteria for entering information will be related to the actual or possible condition of an area of land, for example its current status, or its specific condition established by investigation and/or assessment. These will therefore refer to a particular land area which should be specified in the record or file.

B10 Organisations will need to decide for themselves whether they wish to maintain individual records based on information on a site's history, or on current criteria, for example ownership or development areas. It is however recommended that sites or land areas with distinct, separate histories or other distinguishing features are kept separate in records as far as possible. As a guide, records could be combined if they meet one or more of the following conditions:

- Each site should individually account for at least 75% of the area of the aggregated site, ie most of the ground covered is common.
- There is a high density of individual sites, with related common characteristics, within an area. They may be combined to form a single site if there is, for example, reason to believe that they were operated in such a way that the ground between them is possibly contaminated or if their centroids are less than 50m apart. This avoids the unnecessary creation of repetitive records.

B11 Where sites are combined in this way, the boundary on the index map would be the boundary containing the land covered by all the sites, with the centroid calculated based upon this boundary.

B12 It is important to note that the boundary of land which is or may be affected

by contamination does not necessarily correspond to legal ownership - the site may cover more than one basic land and property unit, or less than a single unit.

Uncertainty in boundaries

B13 There will be occasions when the exact boundary of a site or land area cannot be known, particularly when information is transferred from previous records or maps. Typical reasons for the boundary not being definitive in relation to land contamination are:

- the only boundary known relates to land ownership, not necessarily to the specific history or extent of contamination on the site or land area.

- even where history and ownership are broadly known, parts or the whole boundary of a site, or part of a site, are not clear. Best estimates should be made based on understanding of historic practices and clues from other nearby features, for instance rivers, roads and buildings used for other purposes.

- a historic map clearly marks a site boundary, but because of inaccuracies in mapping and a lack of nearby features, the site cannot be positioned on modern maps to perhaps closer than 25m. In this case, a larger area which encompasses the possible variations in boundary may need to be recorded.

B14 In the event of any uncertainty in site position or boundary, a note to this effect should be made on the record/file. Organisations are advised to keep a record of all assumptions and estimates made in setting the boundary. The uncertainty should be shown on the boundaries drawn on the map. Different symbols could be used to indicate the degree of uncertainty.

Centroid

B15 The centroid (centre of area) of a site can be derived from the site boundary. Generally, it is recommended that the centroid be calculated by taking the centre of the minimum bounding rectangle (i.e. the smallest rectangle that encloses the site boundary). This is illustrated in Figure 1a. However, there are some occasions when a different approach is needed:

- For sites or sections of sites with nearly triangular boundaries, it is recommended that the centroid be taken from the intersection of the lines that join the midpoint of each side with the opposite vertex of the triangle, as illustrated in Figure 1b.

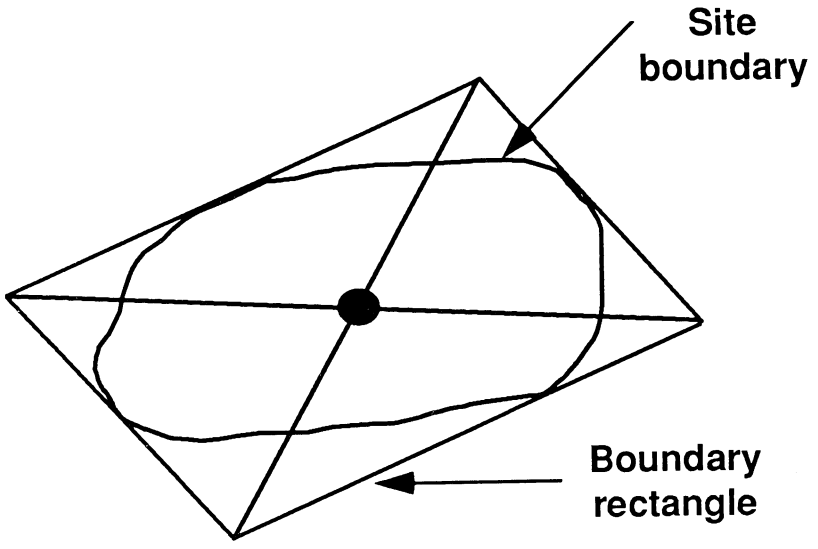

Figure 1a

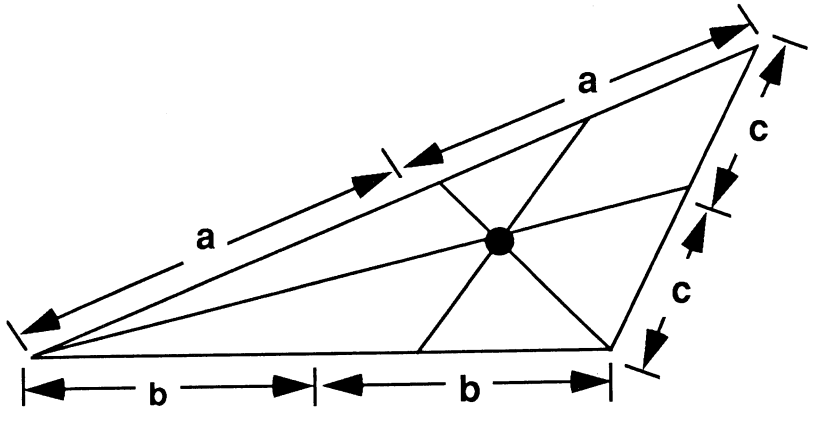

Figure 1b

Figure 1 : Calculation of Site Centroids

- There are cases where it may be necessary to split a site into several sections in order to obtain useful centroids. Examples would be a long, filled water course, a large "L" and other irregular shaped sites. Sites should be split into sections, each with their own centroid, if:

  either

  i the ratio of length to width of the bounding rectangle is greater than 4:1 (for example a section of canal);

  or

  ii the site is long, thin and bends (eg a bending river) but if the site was straightened out, then the ratio of the sides of its bounding rectangle is greater than 4:1;

  or

  iii the centroid does not lie within the boundary;

  and

  iv the length of the site is at least 300m, when marking a 1:10,000 scale map, or at least 100m, when marking a larger scale map.

- Alternatively, boundaries can be determined in accordance with BS7666: Spatial Data-sets for Geographic Referencing. Centroids can be shown as the visual centre of the area which avoids the overheads involved in determining the accurate centroid.

B16 For these cases, the site should be broken into sections and the centroid calculated and recorded for each of the sections. In such cases, each of the centroids should be given the same RECORD IDENTIFIER (see section B39) and, following the symbol "-", a series of centroid coordinate numbers should be appended (e.g. 35-1, 35-2, 35-3). In all cases, the site should be divided so that the bounding rectangle for each of the sections is only a little larger than the section itself.

Marking Sites

B17 The following suggestions are for those organisations who wish to develop standard methods for recording site information on map based systems which could be transferred to electronic format.

a) Mark the centroid of each site, or site section, on the index map using a 3mm diameter brown filled circle.

When sites are marked on an insert map, mark their centroids on the index map without their corresponding identifiers, unless the density of centroids obliterates too much detail on the index map, in which case the region of dense sites should be shaded [brown with 0.5mm diagonal brown lines at 5mm spacing]. Where a larger scale map has been attached to the 1:10,000 index map, mark all sites with centroids lying within the boundary of the insert map in full on the insert map rather than on the 1:10,000 map.

b) Show the record identifier for a site to the right of the centroid marker and , if possible, adjacent to it. Print the record identifier on a plain white background at a font size corresponding to minimum 12 points. If it is necessary to offset the identifier(s), use a [0.3mm solid black] line to link them with the marker circle.

c) Designate site boundaries with a broken line of width 0.5mm and colour brown, containing alternating dashes of [6mm and 3mm] at [3mm] intervals.

- If there is more than one centroid marker within the boundary, then the boundary should be linked to its corresponding centroid marker with a [0.3mm solid black] line.

- Where the boundary is not clear, a larger area, containing the boundary should be indicated - a note should be made in the record of the uncertainty.

- If the site boundary is not wholly contained on a single map, or is greater than 5cm in its largest dimension when plotted, then the interior of the boundary can be marked with inward pointing arrows.

B18 Figure 2 illustrates the marking of sites.

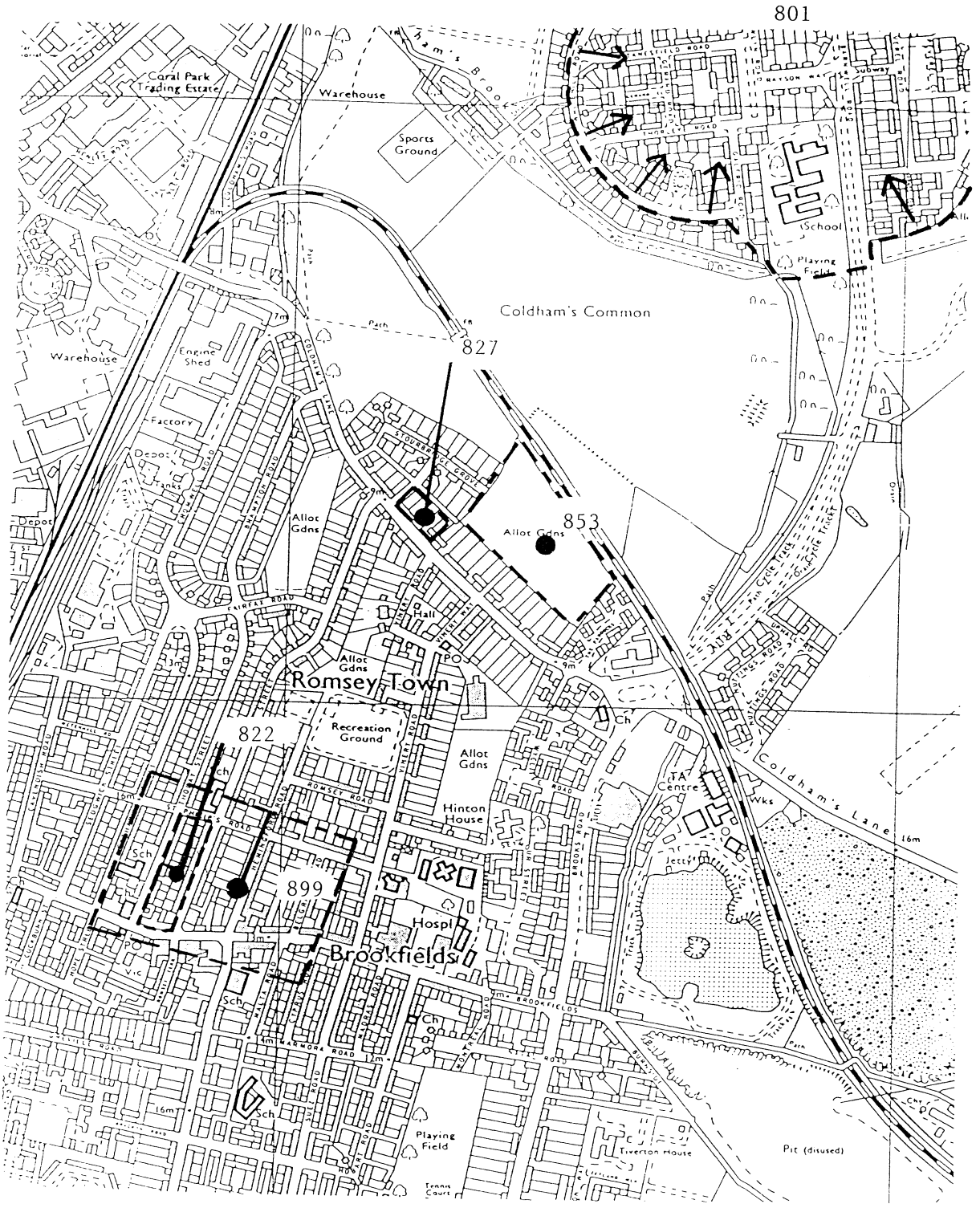
801
827
853
822
899
Coldham's Common
Romsey Town
Brookfields
Coldham's Lane
Sports Ground
Warehouse
Coral Park Trading Estate
Recreation Ground
Allot Gdns
Hinton House
TA Centre
Hospl
Playing Field
Pit (disused)
Triverton House
Engine Shed
Factory
Depot
Ch
Sch
PO
Hall
Wks

Figure 2

SPECIFICATION OF DATA

B19 Detailed specifications are likely to be prepared by the organisations holding information to identify the data to be included, the way in which it is recorded and the format in which it will be presented. Such specifications should include:

- layout control instructions;
- field names; and
- data entries.

Layout control instructions

B20 General layout control instructions, for example pagination, fonts, page headers, page footers and margins, can be used in the specification where these help organisations who wish to ensure the common appearance of data or facilitate data transfer.

B21 Layout control instructions can also include BLANK LINE (at least one line left blank - extra blank lines may be used to enhance readability); [INDENT] (text indented by at least 3 blank character spaces - the **minimum** number of spaces represented by [INDENT] is variable, but should be fixed within any individual set of records to facilitate electronic exchange); and { and } for repetition (items within the brackets are repeated as many times as necessary). Additional spaces may be added to improve the legibility as required provided that the structure of the information is still clear.

B22 Information is structured in terms of "heading fields" and "data fields". Heading fields are identified by a heading title, and contain one or more fields within the heading. The fields within may be subsidiary heading fields, or data fields, or a mixture of the two. Data fields are identified by a data field title and specify entry of an item of data.

Example specification.

In this specification the heading field is **"Distribution"**; the data field is "Copy:"; and the data type is "DISTRIBUTION". The data can be text (naming a person, with department and organisation, who holds a copy of the data).

"Distribution"
BLANK LINE
{
[INDENT]"Copy:"[INDENT]DISTRIBUTION
}

Specification: The distribution field lists all previous recipients of copies of this record who should be sent a new copy of any updated record.

This would result in:

Distribution
Copy: Sam Weston, Recreation and Leisure Department, Tongleton B.C.
Copy: Jean Ketley, Finance Department, Tongleton B.C.
Copy: Water Protection Manager, Bluhme Regional N.R.A.
Copy: Miss M. Bicknell, Head Librarian, The West Regional Collective Library of Local Archaeological and Historical Artifacts.

B23 In the specifications that follow, "N" is used to indicate any decimal digit (i.e. 0 to 9), and "A" represents any uppercase alphabetic character (i.e. A to Z). An appropriate character must be placed in the data item for each repetition of either of the above. For example AAAN specifies that three alphabetic characters must be entered, followed by a numeric character. Any text appearing between quotation marks (" ") in a specification should be reproduced character by character, excluding the quotation marks.

Specification of Data

B24 Items of data which may be relevant to information holdings are listed below in alphabetical order together with their specifications. As a general requirement it is suggested that all fields appearing in the printed form contain only printable characters, as defined in the ASCII (American Standard Code for Information Interchange) character set.

B25 Some data items may have comments enclosed in parenthesis "(" and ")" appended to them. These should be used to record information sources and notes on conflicting or incomplete information, using the general specifications above.

COUNTY NAME

B26 Standard County names are specified in BS6879: 1987 "British Standard Specification for Codes for the representation of names of counties and similar areas".

DATE AND DATE RANGE

B27 Although the dates should be as accurate as the source data permits, it is inevitable that the exact dates of events may not be known. Sometimes the information will be limited to knowing the date before which something had not happened and after which it had happened, for example when a change of use of a site does not appear on the map published in a particular year, but does appear on a later edition.

B28 The table below gives examples of different date formats for electronic data entry.[2]

Format	Example	Meaning
YYYY	1981	The year 1981
"Q"N YYYY	Q3 1981	The third quarter (July to September) 1981
MM YYYY	October 1981	October 1981
OD MM YYYY	18 October 1981	18 October 1981
YY"00s"	1800s	19th Century
"D"ON YY"00s"	D2 1800s	The second decade of the 1800s, i.e. the period 1810-1819 inclusive.
<YYYY	<1894	Before 1 January 1894
>"Q"N YYYY	>Q2 1945	After 30 June 1945.
YYYY-YYYY	1988-1991	The range from 1 January 1988 to 31 December 1991
"Q"N-"Q"N YYYY	Q1-Q3 1942	The range from 1 January to 30 September 1942
"Q"N YYYY-"Q"N YYYY	Q3 1932-Q2 1933	The range from 1 July 1932 to 30 June 1933

---

[2] Notes:

- *Y* is a numeric digit of a valid year
- *MM* represents a valid month name, which should be written in full.
- *D* is a numeric digit of a valid day of the month
- *O* is an optional extra numeric digit in the event that the number is in the range 10-99, rather than 0-9.

INVESTIGATION TYPE

B29 Consideration could be given to the use of BSI and other recognised terms for the stages of investigation such as:

- desk study;
- site walkover;
- preliminary or exploratory survey; and
- detailed site investigation.

and, for types of investigation:

- topographical;
- geotechnical/engineering;
- hydrogeological;
- biological/ecological;
- chemical; and
- gas/methane.

LOCAL AUTHORITY NAME

B30 Local authority names are specified in "Code For Local Authorities And Other Administrative Bodies, Department of The Environment, 1986".

ORGANISATION CODE

B31 A unique identifier or code for an organisation. Local authority codes are specified in "Code For Local Authorities And Other Administrative Bodies, Department of The Environment, 1986".

OS GRID COORDINATES (SPATIAL REFERENCE)

B32 Grid coordinates should be taken from current OS 1:10,000 or larger scale maps. The standard form specified by the Ordnance Survey should be used, to 10 or 12 alphanumeric characters. Currently, the form is two letters uniquely defining the 100 km square within the UK, followed by a total of either eight or ten decimal digits in two equal groups: the first group specifying the "easting" and the second group specifying the "northing". The format is therefore AA NNNNN NNNNN (eg GK 51274 39187) or AA NNNN"X" NNNN"X" (eg GK 5127X 3919X). Note the "X" inserted in place of the "missing" digit.

B33 Ten digit numerics (giving a resolution of 1m) should be used when working from 1:2,500 or 1:1,250 scale maps. Eight digit numerics may be used when working from 1:10,000 scale maps (giving a resolution of 10m) if the accuracy is not required to locate a site, or not possible.

Note: The purpose of using this level of precision in recording is not to suppose that this accuracy of positioning on the ground is attained. Rather, it is to identify small sites accurately on the map from which the grid reference of the site is taken. (Inaccuracies in mapping and copying can mean the grid reference does not correctly identify the site on a different edition or version of a map). For example, 10 character precision (10m accuracy) is not sufficient

to locate many small sites (e.g. those which are about 5m square) on a 1:2,500 map. (These small sites are generally not depicted individually on a 1:10,000 scale map.)

OS REVISION

B34 This is the year of a revision to an OS map, the year to be in DATE form.

OS SCALE

B35 This specifies the scale of the Ordnance Survey map and for current maps is expected to be one of the values 1:10,000, 1:2,500 or 1:1,250.

OS SHEET

B36 This is the sheet reference in the form specified by the Ordnance Survey. (For example, for current 1:10,000 scale maps, the sheet number is of the form: AA NN AA e.g. TL 45 NW.)

POSTCODE

B37 A code comprising two sets of alphanumeric characters determined by the Royal Mail.

RECORD CODE

B38 This would normally consist of the ORGANISATION CODE for the organisation holding the data, a colon (":") and the RECORD IDENTIFIER, eg DOE:CLL123.

RECORD IDENTIFIER (REFERENCE NUMBER)

B39 This should be an alphanumeric code identifying the record or file uniquely and systematically.

REFERENCE

B40 This is used to define an information source and should contain as many of the following as available: the publisher or commissioning organisation, the full title of the publication with reference and edition or version number, the author or editor, the date of publication, the ISBN number, the page number.

SITE NAME (PROPERTY NAME)

B41 This is usually the customary name of the site, but is rarely a single name.

TEXT

B42 Unlimited number of characters.

TEXT TERM

B43 Text string taken from a list of options, specified for the particular data item.

TREATMENT TYPE

B44 There are no standard terms for types of treatment but broad categories commonly used include:

- containment on site ;
- in ground venting (eg for landfill gas);
- continuous extraction (eg for landfill gas);
- biological treatment;
- physical treatment;
- chemical treatment;
- thermal treatment;
- solidification;
- hydraulic methods;
- removal of source of contamination by excavation;
- off site disposal; and
- other methods.

USE

B45 Previous or current use of the site as described in the source information.

USE CLASSIFICATION AND CLASSIFICATION CODE

B46 The Department of the Environment provides standard classifications and classification codes in for example, surveys of land use (Department of the Environment Statistical Bulletin (90) 5, chart 1.)

B47 Uses may also be classified in more detail, for example into the headings used for DoE industry profiles.

VALUE and VALUE RANGE

B48 All numeric values should be given as decimal numbers. The decimal point, represented by the symbol "." should only be used for non-integral numbers when it is followed by one or more decimal digits.

B49 In some cases it is necessary to give a range of numbers. This may be specified by the inclusion of a hyphen ("-") between two numbers. An optional qualifier, either ": odd numbers" or ": even numbers" may be added after the value range to restrict the numbers included; this is used on occasions for addresses.

TYPICAL SPECIFICATION

B50 In the specification that follows the following conventions are used:

- Main heading field names are shown in upper case and bold and subsidiary headings are shown in bold for clarity. Field names and other terms and characters, that would be reproduced as written in the specification, are enclosed in quotation marks. Text in italics is provided to specify circumstances in which information is inserted. These elements in italics do not, themselves, form part of the record.

- In general there is no more than one data item on each line of the record, unless multiple items are explicitly shown on a single line in the specification (e.g. Summary Of Use).

- Repetition is denoted by curly brackets "{" and "}" surrounding the field to be repeated. The curly brackets themselves do not form part of the record. For example, {USE} indicates that as many items as needed, of type USE should be entered.

- For clarity of presentation, the repetition of [INDENT]s preceding data fields is not shown, although data fields are always indented. In these cases, each data item should occupy a separate line of the record, with each continuation line indented one [INDENT] with respect to the field name.

Example

Specification:

[INDENT] "County(ies):" [INDENT] { COUNTY NAME }

This can be translated as:

```
County(ies):    Buckinghamshire
                Hertfordshire
```

Alternatively lines could be broken after a data field name and before data field entries. In this case, the [INDENT] before the data item is offset from the position of the first character of the data field name. For example, another translation of the above specification is:

```
County(ies):
   Buckinghamshire
   Hertfordshire
```

- If a repeated field includes a numbered heading or field name, this is shown in the specification by an "N". For the first use of the heading or name, "N" takes the value "1"; for each subsequent use it is incremented by one. For example "Amendment " N is translated to "Amendment 1", "Amendment 2" etc.

- Some data fields are inappropriate under certain circumstances. For instance, addresses are variable and any of the property names, street numbers and districts can be present or absent. These are omitted when appropriate, rather than inserting the title but leaving the data entry section blank. The specification and guidance notes for the field or fields in question describe the relevant circumstances in which they may be omitted.

B51 The specification below describes fields and data entries which may be relevant to information on land contamination.

"AMENDMENT RECORD"

BLANK LINE

[INDENT] "**Initial Entry**"
[INDENT] [INDENT] "Entry made:" [INDENT] DATE
[INDENT] [INDENT] "Entry made by:" [INDENT] TEXT

Specification: Give date at which an entry was first made for this site and name of the person making the initial entry.

{
BLANK LINE
[INDENT] "**Amendment " N**
[INDENT] [INDENT] "Amendment made:" [INDENT] DATE
[INDENT] [INDENT] "Amendment made by:" [INDENT] TEXT
[INDENT] [INDENT] "Description:" [INDENT] { TEXT }
}

Specification: Give date at which the entry was updated for this site, the name of the person making the update to the entry, and a description of the items changed with reasons.

Repeat amendment field, once for each time the record has been amended, although several amendments made at the same time may be made within a single entry; list the amendments in chronological order, oldest first, with the heading field "Amendment 1", "Amendment 2" etc.

Incorporate the most recent [ten] amendment entries, or amendment entries made over the [five] previous years, which ever is the fewer amendments, within the information.

"IDENTIFYING INFORMATION"

BLANK LINE

[INDENT] "Record Code:" [INDENT] RECORD CODE

Specification: code allocated to this record.

[INDENT] "Organisation:" [INDENT] ORGANISATION NAME

Specification: name of the organisation compiling the data (one entry only for this field).

[INDENT] "Customary name of site:" [INDENT] SITE NAME

Specification: name in common usage that describes the site.

"LOCATION AND AREA"

BLANK LINE

[INDENT] "Region(s):" [INDENT] { REGION }

Specification: The names of all regions in which the site lies (one or more values for this field).

[INDENT] "Local authority:" [INDENT] { LOCAL AUTHORITY NAME}

Specification: The names of all local authorities in which the site lies (one or more values for this field).

{
[INDENT] "**Current Address " N**
[INDENT] [INDENT] "Property name:" [INDENT] TEXT
[INDENT] [INDENT] "Property number:" [INDENT] VALUE RANGE
[INDENT] [INDENT] "Street:" [INDENT] TEXT
[INDENT] [INDENT] "Village or district:" [INDENT] TEXT
[INDENT] [INDENT] "Town:" [INDENT] TEXT
[INDENT] [INDENT] "Postcode:" [INDENT] POSTCODE
[INDENT] [INDENT] "Source:" [INDENT] { REFERENCE }
}

Specification: Information on current address of the site. If there is more than one address, the whole block between the curly brackets "{" and "}" is repeated.

{
[INDENT] "**Historic Address " N**
[INDENT] [INDENT] "Property name:" [INDENT] TEXT
[INDENT] [INDENT] "Property number:" [INDENT] VALUE RANGE
[INDENT] [INDENT] "Street:" [INDENT] TEXT
[INDENT] [INDENT] "Village or district:" [INDENT] TEXT
[INDENT] [INDENT] "Town:" [INDENT] TEXT
[INDENT] [INDENT] "Postcode:" [INDENT] POSTCODE
[INDENT] [INDENT] "Source:" [INDENT] { REFERENCE }
}

Specification: Information on addresses of the site from historic information, eg old maps or trade directories, recorded in chronological order, starting with the earliest. If there is more than one address, the whole block between the curly brackets "{" and "}" is repeated.

Historic trade directories may not give the address of the actual site of operation of a business. For example, sometimes trade directories list the home address of the tradesman rather than the address of the site where the business was carried out. References should be recorded to all texts giving the address and also to any maps used in order to locate the site from its address.

It may be necessary to add a text field to comment on the relationship between historic addresses and present day addresses, eg renaming of road, district or site.

Guidance: The purpose of including the historical address(es) information is to provide an audit trail. Historical addresses should be included where they are available, the most likely sources being trade directories and other textual material.

[INDENT] "OS Map sheet number(s):" [INDENT] OS SHEET

Specification: Ordnance Survey sheet, current at the time the record was made, that was used for locating the site and reading off its centroid coordinate.

Guidance: The purpose of this and the two subsequent fields is to identify uniquely the current (at the time the record was made) map used to locate the site and centroid. All map references are given in the context of the specific sheet from which they have been measured. This should skirt around any problems which might occur from any survey inaccuracies in OS maps, or changes in coordinate alignment.

Small sites, i.e. those under half a hectare, should preferably be located on 1:1,250 or 1:2,500 scale maps rather than 1:10,000 scale.

[INDENT] "OS Map Revision:" [INDENT] OS REVISION

Specification: The revision of the Ordnance Survey map used for location of the site centroid.

[INDENT] "OS Map Scale:" [INDENT] OS SCALE

Specification: The scale of the Ordnance Survey map used for location of the site centroid.

{[INDENT] "Centroid Coordinate " N ":" [INDENT] OS GRID COORDINATE
}

Specification: The location of the centroid of the site. When there is just a single centroid, then N is omitted.

If a site is split into sections, then one coordinate is given for the centroid of each section; these are named sequentially, "Grid Coordinate 1:", "Grid Coordinate 2:" etc.

[INDENT] "Sites grouped" [INDENT] TEXT

Specification: This field is included only if sites with identifiably different boundaries have been grouped according to the rules set out in B10. The text could be standardised, as a minimum to YES or NO or with reasons for grouping.

[INDENT] "Area (hectares):" [INDENT] VALUE

Specification: The area calculated where possible or unless otherwise specified to an accuracy of:

i 0.01 hectares if the total site area is less than 1 hectare

ii 0.1 hectares if the total site area is less than 10 hectares

iii 1 hectare otherwise

[INDENT] "Boundary expanded, site area (hectares):"
[INDENT] [INDENT] VALUE
[INDENT] [INDENT] TEXT

Specification: This field is included only if the position of the site was not clear, and therefore the area has been expanded to be sure of encompassing the site. The text field can be used for comment on the expanded boundary.

[INDENT] "Correspondence between source and current maps:" [INDENT] TEXT TERM

Specification: Qualitative description of the accuracy with which it has been possible to match site location from historic maps to current OS 1:10,000, 1:2,500 or 1:1,250 maps. The TEXT TERM should be standardised, eg one of "Same map", "Accurate", "Approximate", or "Poor".

[INDENT] "Reference(s):" [INDENT] { REFERENCE }

Specification: The references should list all maps and other sources of information used to determine the position and boundary of the site. This includes any maps which do not show the site itself, but which were used to locate the site on a modern map because they record nearby features also on previous or later maps.

"CURRENT USE"

BLANK LINE

[INDENT] "Classification:" [INDENT] CURRENT USE CLASSIFICATION

Guidance: The information on current use is most likely to be needed for statistical analysis and to assist in any required "action analysis". The information may be obtained from that given on the Ordnance Survey 1:10,000 maps for the area, or local authority records.

[INDENT] "Classification Code:" [INDENT] CURRENT USE CLASSIFICATION CODE

Specification: This is the code associated with the classification given above.

"SUMMARY OF PREVIOUS USE"

BLANK LINE

{[INDENT] USE CATEGORY "," USE "," DATE RANGE }
{[INDENT] SURVEY TYPE "," DATE }
{[INDENT] TREATMENT TYPE "," DATE }

Specification: The summary of previous use should contain one entry for each

use recorded in the history of previous use, investigation and treatment sections of the record. The entries under each three types should be merged and listed in chronological order, from earliest to most recent.

Guidance: The purpose of this summary is to allow the user to establish whether the entry is of interest.

"HISTORY OF PREVIOUS USE"

BLANK LINE

{
[INDENT] "**Use " N**
[INDENT] [INDENT] "Use:" [INDENT] { USAGE }
[INDENT] [INDENT] "Category:" [INDENT] { USE CATEGORY }
[INDENT] [INDENT] "Operator:" [INDENT] TEXT
[INDENT] [INDENT] "Owner:" [INDENT] TEXT
[INDENT] [INDENT] "Start date:" [INDENT] DATE
[INDENT] [INDENT] "End date:" [INDENT] DATE
[INDENT] [INDENT] "Sources:" [INDENT] { REFERENCE }
}

Specification: There may be more than one use for a site, at different periods, in which case the whole block between the curly brackets "{" and "}" is repeated, with the oldest use being given first. If there is only a single use, then N should be omitted, otherwise the blocks should be numbered from 1, e.g. "Use 1", Use 2" etc.

All source material should be referenced.

A site record may need to include more than one concurrent use, either because more than one activity was actually carried out within its boundary or because different source materials give contradictory information. In the latter case, all contradictory uses should be recorded. A note acknowledging the concurrent uses, indicating contradiction if present, should be appended to the record, prefixed with the title "Comment:".

Guidance: Where precise commencement and termination dates of previous use cannot be established, the period of use should be indicated by narrowing the start and end dates down as much as possible. For example there may be a site which does not appear in an 1847 map sheet, but does on an 1863 sheet and also an 1857 directory. Thus the start date is recorded as between 1847 and 1857.

The source from which each date recorded is derived should be referenced so that the information can be easily reconstructed and verified.

"INVESTIGATION AND TREATMENT"

{
BLANK LINE
[INDENT] "**Site investigation " N**
[INDENT] [INDENT] "Investigation:" [INDENT] INVESTIGATION TYPE
[INDENT] [INDENT] "Date:" [INDENT] DATE

[INDENT] [INDENT] "Source: [INDENT] REFERENCE
}

Specification: There may have been more than one investigation for a site, in which case the whole block between the curly brackets "{" and "}" is repeated, with the oldest investigation being given first. The subsequent blocks should be labelled "Investigation 2", "Investigation 3" etc. If there is no record of any investigation relating to the site the survey type data item associated with "Investigation" should contain "No investigation recorded for this site" and the "Date:" and "Source:" fields should be omitted.

Guidance: Investigations may include both site surveys and desk studies. These are identified by the INVESTIGATION TYPE.

{
BLANK LINE
[INDENT] **"Site treatment " N**
[INDENT] [INDENT] "Treatment:" [INDENT] TEXT
[INDENT] [INDENT] "Date:" [INDENT] DATE
[INDENT] [INDENT] "Source:" [INDENT] REFERENCE
}

Specification: There may have been more than one treatment applied to a site, in which case the whole block between the curly brackets "{" and "}" is repeated, with the oldest treatment being given first. The subsequent blocks should be labelled "Treatment 2", "Treatment 3" etc. If there is no record of any treatment relating to the site the text data item associated with "Treatment" should contain "No treatment recorded for this site" and the "Date:" field should be omitted. Any interpretation of the treatment efficacy should be fully referenced, prefixed with the title "Comment:" (and see below on assessment studies).

{
BLANK LINE
[INDENT] **"Assessment study " N**
[INDENT] [INDENT] "Date:" [INDENT] DATE
[INDENT] [INDENT] "Reference:" [INDENT] REFERENCE
[INDENT] [INDENT] "Comment:" [INDENT] TEXT
}

Specification: Any studies done, including any of migration from the site, should be referenced if possible. The comment field could be standardised if in-house or other classification systems are used.

"SUMMARIES"

Summaries may be useful for individual organisation's requirements and for passing information to other users. The particular specification will depend on the nature of the summary data, for example:

"Summary of site histories"

{

DOUBLE BLANK LINE
[INDENT] "Site location: " OS GRID COORDINATE ", Area(hectares):"
VALUE
{[INDENT] "Use:" USE CATEGORY "," USE "," DATE RANGE }
{[INDENT] "Investigation:" REFERENCE }
{[INDENT] "Treatment:" REFERENCE }
}

Specification: The summary of site histories should contain the site centroid coordinate, area and one entry for each use, investigation and treatment recorded in the record. Approximate start and end dates of use should be entered in the data range. In doing so, the "decade" or "quarter century" means of date specification should be used if the dates are not known more accurately.

The corresponding "use", "investigation" and "treatment" data fields should be displayed together. Where there is more than one such set of information for a site, they should be displayed in chronological order.
This record is repeated for every site.

ELECTRONIC STORAGE AND TRANSFER

B52 The preceding section has described the basic format of information. It is anticipated that many organisations will wish to store their information electronically in databases or geographical information systems. The National Transfer Format (NTF) (BS7567: 1992 "Electronic transfer of geographic information (NTF)") provides media independent file and data record descriptions for the exchange of geographic information.

B53 NTF is sub-divided into five levels, denoted from 1 to 5 in order of increasing complexity. Level 3 is the least complex level capable of transferring the information and is the recommended standard for electronic transfer of data on land contamination from database and GIS systems. The remainder of this section specifies how the NTF standard can be used for information on land contamination and should be read in conjunction with part 1 of BS7567.

Correspondence between records and NTF structures

B54 Each record on land contamination corresponds to a NTF level 3 "complex feature", each of which comprises a "collection" of "simple features". At NTF level 3, "simple features" include polygons, points and text objects that are to be shown on the electronic map. A polygon feature is used to express the site boundary, a point feature the site centroid, and text features the text that appears on the index map itself. Figure 3 shows an annotated and simplified version of the topological data model for NTF level 3 as applied to records of land contamination.

Implementation of NTF Records

B55 This section complements the NTF standard by specifying particular implementations of the NTF records that should be used for transfers of records on land contamination. The name for each NTF record type is taken from the BS7567 as is the logical structure of the collection of records comprising a single transfer known as the "transfer set" as shown in Figure 4.

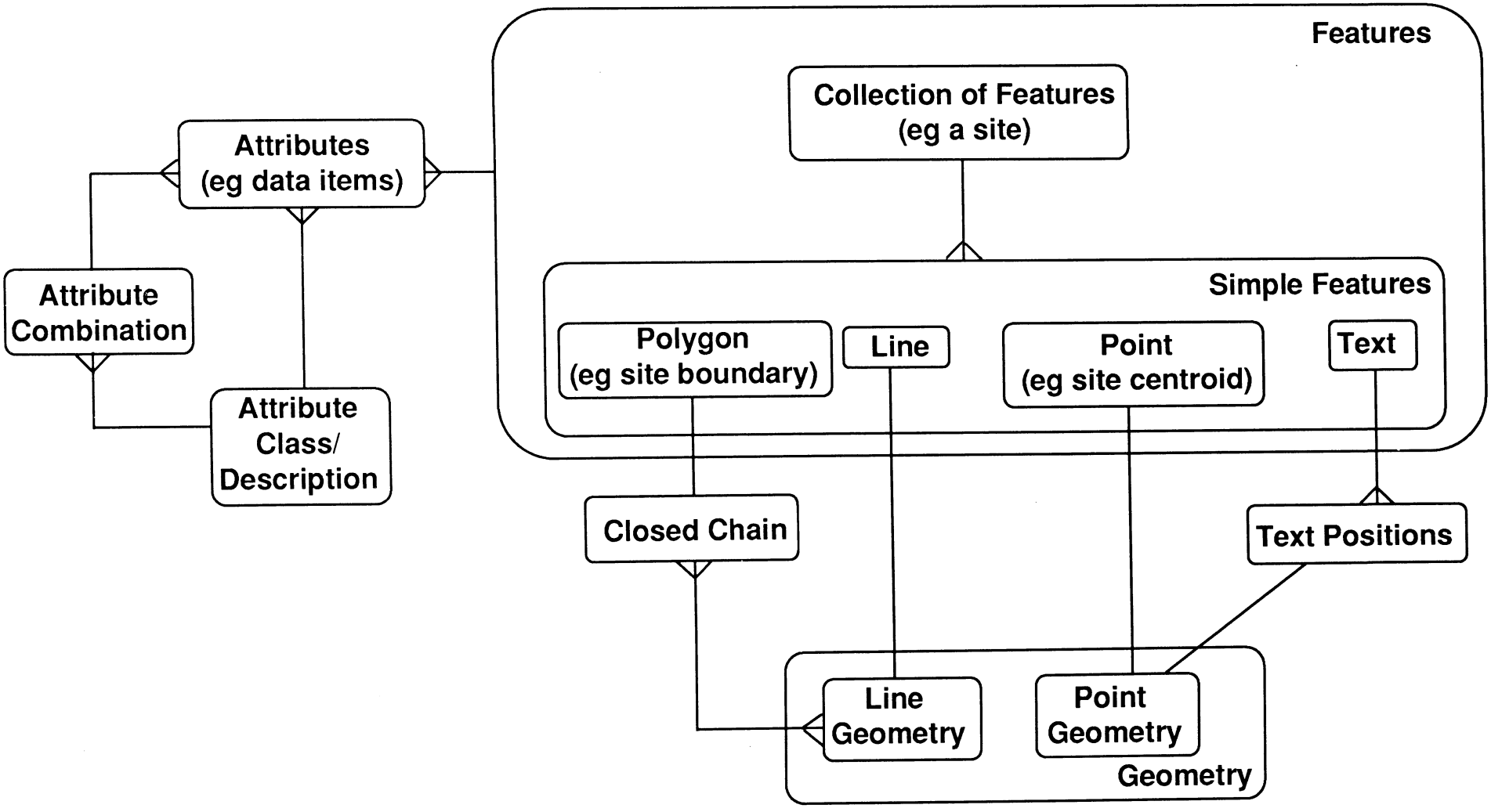

Figure 3 : NTF Level 3 "Complex Spaghetti Model"

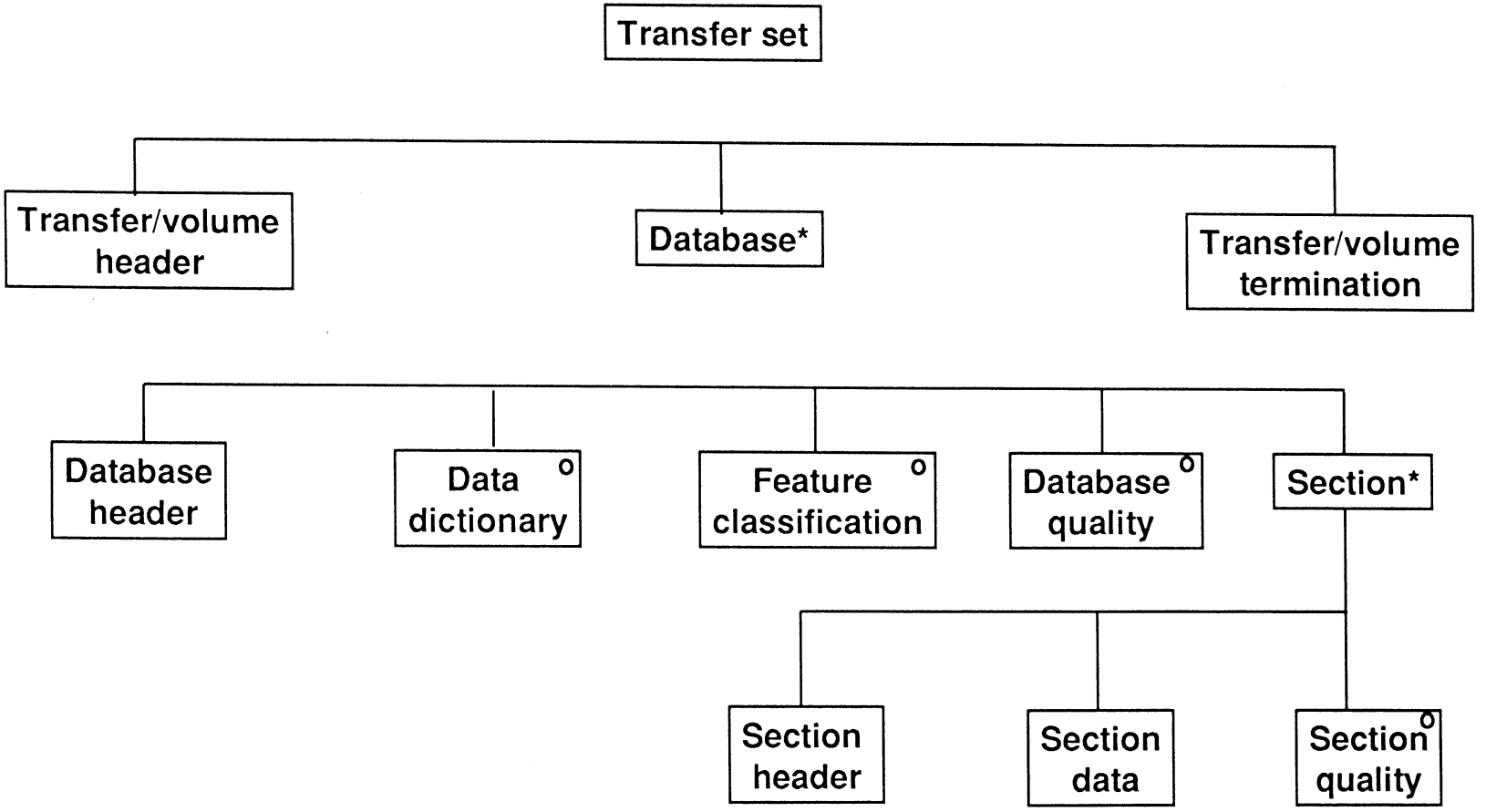

Figure 4 : Logical Structure of the NTF Structure Set

This diagram should be read from top left to bottom right. An asterisk '' indicates that the entity may be repeated several times; an 'o' indicates that the entity is optional within a Transfer Set.

Transfer Volume Records

VOLHDREC There are no special requirements for this record.

Database Description Records

DBHREC The following requirements are placed on specific fields in this record.

DDNAME = 01.00_DEFAULT if the data dictionary described in this report is used. If the organisation wishes to enlarge the data dictionary to allow additional information to be included in the transfer, then an appropriate DDNAME should be given.

DATA_MODEL = 3 (Complex Spaghetti).

ATTDESC This record is used to define the linkage between the field names and the attributes referenced by the features. Each ATTDESC record must contain the following fields:

VAL_TYPE	Two character mnemonic identifying the attribute.
FINTER	Interpretation of field.
ATT_NAME	Attribute name.
FDESC	A textual description of the attribute. This is used to specify the data type of the data field.

An example of the attribute description for typical records is given in Table A.

ATTCOM This record allows common combinations of attributes to be defined. Each record consists of a value-type (VAL_TYPE) as the record's unique identifier, followed by one or more value-types identifying the other ATTCOM and/or ATTDESC records in the grouping.

FEATCLAS This record is not required.

CODELIST This record is required to enumerate codes used in the transfer set, as specified in BS7567. It may also be used by the donor to assign codes to particular values used in the information base, thereby indicating the presence of structured data.

DBQUAL This record is optional.

Section Header Records

SECHREC There are no special requirements for this record.

SECQUAL There are no special requirements for this record.

QUALREC There are no special requirements for this record.

Section Data Records

COLLECT — There should be one COLLECT record for each record or file. This should specify the IDs of all the simple features associated with the Record. It should also list all ATT_IDs associated with the Record.

LINEREC — Line features are used to represent additional artwork on the map, such as the lines used to join the text identifying a site to the corresponding centroid.

POLYGON — Polygon features are used to describe boundaries and so the only attributes referenced by them should be properties of the boundary, rather than the site, which is linked to the COLLECTion record (as shown in Figure 3). Each POLYGON record should reference an attribute describing the type of boundary (see the enumeration of the VAL_TYPE field for the ATTDESC record above).

POINT — Point features are used to describe the site centroids and the only attributes referenced by them should be properties of the centroid, rather than the site, which is linked to the COLLECTion record (as shown in Figure 3). Each POINT record should reference an attribute describing the type of point (see the enumeration of VAL_TYPE for the ATTDESC record above).

TEXT — There are no special requirements for this record. It could be used to represent any fixed text appearing on an index map.

TEXTPOS — There are no special requirements for this record.

CHAIN — There are no special requirements for this record.

GEOMETRY1 — This record is used to specify the points that form the vertices of the polygon representing the closed boundary of the site.

ATTREC — This is used to assign values to the attributes and so contains the non-geometrical information associated with each record.

Table A

Attribute Description Records

Identifier (VAL_TYPE)	Interpretation (FINTER)	Field Name (ATT_NAME)	Data Type (FDESC)	Combine (VAL-TYPE)
01	NO_DATA	Record	-	-
02	NO_DATA	Amendment Record	-	01
03	NO_DATA	Initial Entry	-	02
04	A*	Entry made	DATE	03
05	A*	Entry made by	TEXT	03
06	I3	Amendment N	-	02
07	A*	Amendment made	DATE	06
08	A*	Amendment made by	TEXT	06
09	A*	Description	TEXT	06
10	NO_DATA	Identifying information	-	01
11	A*	Entry Code	ENTRY CODE	10
12	A*	Customary name of site	TEXT	10
13	A*	Organisation	TEXT	10
14	A*	County(ies)	TEXT	10
15	A*	District(s)	TEXT	10
16	A*	Civil parish(es) and/or Ward(s)	TEXT	10
17	NO_DATA	Location and Area	-	01
18	A*	OS Map Sheet Number(s)	OS SHEET	17
19	A*	OS Map Revision	OS REVISION	17

20	A*	OS Map Scale	OS SCALE	17
21	A*	Grid Coordinate N	OS GRID COORDINATE	17
22	A*	Area (hectares)	VALUE	17
23	A*	Reference(s)	REFERENCE	17
24	A*	Correspondence between source and contemporary maps	TEXT	17
25	I3	Address N	-	17
26	A*	Name	TEXT	25
27	A*	Number	VALUE	25
28	A*	Street	TEXT	25
29	A*	Village or district	TEXT	25
30	A*	Town	TEXT	25
31	A*	Source	REFERENCE	25
32	NO_DATA	Current Use	-	01
33	A*	Classification	CURRENT USE CLASSIFICATION	32
34	A*	Classification Code	CURRENT USE CLASSIFICATION CODE	32
35	NO_DATA	Previous Use	-	01
36	I3	Use N	-	35
37	A*	Use	USE	37
38	A*	Category	USE CATEGORY	37
39	A*	Operator	TEXT	37
40	A*	Start date	DATE	37
41	A*	End date	DATE	37

42	NO_DATA	Sources	-	37
43	NO_DATA	Investigation and treatment	-	01
44	I3	Site Investigation N	-	43
45	A*	Investigation	INVESTIGATION TYPE	44
46	A*	Date	DATE	44
47	A*	Source	REFERENCE	44
48	I3	Site treatment N	-	43
49	A*	Treatment	TREATMENT TYPE	48
50	A*	Date	DATE	48
51	A*	Source	REFERENCE	48
52	NO_DATA	Distribution	-	01
53	A*	Copy	DISTRIBUTION	52

Numbers and dates should be treated as ASCII strings, not encoded.

Copies may be purchased from the
Publications Sales Unit,
Block 3, Spur 7, Government Buildings,
Lime Grove, Ruislip, HA4 8SF.

Price £7