June 1969

Computer-based Information Centre on Organizations and Entities

Criteria, Coding, Processing Requiements and System Implications

-- / --

Annex VII to the report entitled: *Need for a world management information network -- to assist initiation and coordination of global development programmes*. This Annex was also distributed as *Criteria, Coding, Processing Requiements and System Implications for a Computer-based Information Centre on Organizations and Entities*

Design criteria

The following factors guided decisions on the design of the file:

1. The file structure should not stress unnecessarily the difference between types of organization (or link between organizations) since, whatever definitions are used, different types blend into one another on some dimensions whilst being distinct on others. Similarities between types may be greater than diferences. Accepted and conventional distinctions should be possible but should not distort the file structure. This is the only possible means of making the file useful to *a* wide variety of researchers and decision-makers interested in the functions performed by overlapping classes of organization.

2. A sequential file of data on organizations is completely insufficient in terms of resent and expected future demands for information. The file musttherefore provide means of showing the links between organizations. This form of cross-referencing within the file is the first step towards representing a variety of 'flows' between organizations.

3. A network file structure can therefore be conceived as made up of nodes and links. The nodes can be organizational entities of any kind, programmes independent of any particular organization, treaties, meetings, etc. The links, whether input or output, are the channels along which the node receives (or transmits) information, funds, non-financial aid, recommendations, etc. Such links may also represent the memberships relationship of 'members' of the node. Links in this general sense can also represent consultative, collaborative, informal and other relationships as necessary.

4. The network file structure should facilitate use of an adaptation of the network and input/output analysis techniques employed in operational research and analysis of electrical networks. Since these techniques have not yet been adapted to this use, the consequences for the file design are simply to separate, to the extent possible, coding relating to node characteristics (static) from those relating to link performance (frequency, volume, type). Provision should be made for the inclusion of coding which world reflect the maximum number of dimensions along which communication and collaboration can break down. The objective of this type of approach is to maximize the possibility of constricting models which would be partly quantitative and predictive as suggested by Karl Leutsch (Nerves of

Government, p.126-7):

> 'A part of this development would be the application of cybernetic concepts to the system, making larger and more explicit use of time variables as well as of probabilistic and statistical considerations. This would mean, among other things, the measurement or estimation of the extent and probable distribution of imbalances in the transaction flows, of the corresponding loads upon the equilibrating or adjusting mechanisms in the subsystems; of the lags, gains, and leads in their responses; and hence of the probable stability and future states of the entire system and its parts.'

5. Associated with the long-term requirement of systematic network analysis is the *simpler*requirement that the file structure should facilitate detection of weaknesses (as defined by the user) in coordination or communication between organizations concerned with the same or related problem areas, in order that such bodiescould be notified of each others activities.

6. Aside from the problem of distinctions between organizations based on conventional definitions of formal organization types, similar problems arise in attempting to distinguish between permanent bodies and temporary bodies, and between independent and dependent or internal bodies (within an organizational structure

A temporary structure such as an independent meeting or a programme may be considered to have an important integrative effect starting from the time it is proposed (and papers are called for) to the time the report or recommendations are finally available as a stimulus to further effort. The complete cycle may in seme cases be up to 10 years or more. This exceeds the life of many formal. constituted 'permanent bodies'. In addition, the borderline between a meeting and an organization, particularly if the meeting forms part of a series and has an informal continuing committee, can only be arbitrarily established.

In the case of independent and dependent bodies, it was again decided that, whatever the degree of autonomy, the file structure should permit, if necessary, treatment of the entity in question as a node in the network. This avoids the unsatisfactory procedure of preestablishing the sub-system boundaries and thus predetermining what is system-external and what is system-internal. The location of sub-system boundaries may itself be an important research objective. In addition, this draws attention to the fact that although communication and coordination between an outside organization and some subsidiary body may be eminently satisfactory, there is no guarantee that the relationship between the central body and the subsidiary body is satisfactory. A sub-sub-system of sub-system A may be affected by a sub-system Bwithout sub-system A as a whole being significantly affected. This has many important consequences.

7. A consequence of the decision not to restrict attention to particular types of organization is that arbitrary definitions of 'international', 'national', 'regional', 'local' or 'governmental', 'non-governmental', 'commercial', etc. are avoided. This permits a researcher to establish his own definitions of such sub-systems with a maximum amount of flexibility.

This is in line with the conclusions of Andrew M.Scott (The Functioning of the International Political System) that the nation-stages are no longer the only significant actors on he international political scene. The file design should facilitate the systems approach suggested by him which would 'help overcome the sharp separation between domestic affairs and international politics, because it operates equally well at either level and can move between the two.'

8. Most information systems are designed as means of speeding up the processing, storage and retrieval of documents. Because of the high volumes involved, such systems are very costly and where they are less costly, this is only achieved by a considerable degree of specialization in order to reduce the volume. To avoid this dilemma and yet optimize information on the world system as a whole, it was decided to concentrate on the producers of information rather than the, information produced in document form.

The information producing and processing points in the world system are organiz ations of one kind or another. These represent the points at which decisions and control activity regarding the production of information occurs. A focus on suchpoints therefore maximizes the possibility of obtaining a clear, overall picture of the world system. Such a picture is an essential basis for management type decisions concerning the allocation of resources.

A management information system requires information on bodies controlling, evaluating, formulating, and implementing programmes, and coordinating memberships (in thebroadest sense), relationships and information networks linking them to problem areas. It is therefore focussed on the coordination achieved and necessary for current and planned or proposed activities. A documentation information system concentrates on on the information produced when it eventually appears in published form,

The first is focuesed on the initiating points for present and future activity *whilst* the second is focussed on the published record, if any, of past activity. The fact that one organisation can coordinate the production of many documents in the context of one programme, isan indication of the volume of information in each case, the scale of the problem in each case, and the cost of each type of system. Most important though, is that it illustrates the relatively much higher value of information on the current programmes of organizations.

Intermediate between these extremes, is Information on sources of information produced in document form (e.g. bibliographies of bibliographies, directories ofperiodicals, etc.) which can be incorporated in a management information system, since it reprerents the key to information collecting points and systems in a particiular problem area. Such ch information is of relatively much higher value if it is produced regularly within a series rather than as a one-off publication.

The file structure is therefore deliberately not orientated toward the solution of the documentation problem and the associated 'information explosion'. Such solutions imply the retrievability within a 'reasonable' period of time of an optimum number *of*past relevant documents on a subject. A management information system implies - he immediate availability of information on all currently active bodies, programmes and information networks with in the world system. It can, to some extent, predict future document production.

The emphasis on the management approach is based on the view that even if document information system can provide an optimum selection of relevant material on a problem, this does not facilitate the solution to many important subsequent problems. Specifically:

- decision-makers are increasingly in a position in which they can no longer afford the time to wait for libraries and information centres to complete the documentation retrieval cycle. Having received a pile of documents, they are no longer in a position to read and assimilate all the information supplied. Not only does the time factor come into play, but also the problem for the decision-maker of determining the relevance of analytical results based on the techniques and assumptions of disciplines with which he is not familiar. If the are 'foreign' to him, his inclination to use them, will be low, even if he studies the results in detail. This is a major problem in the utilization of research implications for policy formulation.
- a request for documents is specific. The documents received are the answer to the request, They do not automatically supply an operational context for the problem area in question, particularly where it may cross specialist or jurisdictional boundaries. The documentation information system is 'blind' to this approach, particularly when set up within a specialist organization with an accession profile designed to minimize acquisition of material from other fields. The more general the request, the more material supplied which *must* be interpreted, restructured and assimilated.
- the response of a documentation information system is a response from the past and cannot take into account current developments (even the lag between production and publication of a journal article may be several years).
- a documentation system is not dynamic. It cannot permit analyses which could signal probable problem areas. The decision-maker is therefore dependent on historical reports to detect a problem, unless it hasreached crisis level and been reported through not documcntary channels across the accepted jurisdictional boundaries.

A major requirement for a management information is that it be highly structured, eliminate non-significant data in order to highlight problem areas and areas requiring decisions. It should also relate a problem area to associated problem areas across discipline and jurisdictional boundaries. It should indicate the location of resources and the channels through which they could advantageously be moved. An attempt should therefore be made in designing the file structure to facilitate the development of techniques of this kind.

9. Another approach to the analysis of the world system is through the use of political and social indicators based on statistical analysis of the relationship between key variables in a manner analogous to that used for economic indicators. Major difficulties associated wit) this approach are cost, comparability of data collected in different countries and ensuring regular updating.

This approach provides indications of conditions of clearly defined classes either by national, regional or local averages. It does not tie these conditions directly to the organizational structures within society by which they can be modified and tends to gloss over the structure of sub-systems and communication within and between them. Thus although primary problems can be detected, the detection of secondary problems is not facilitated e.g. the structural weaknesses which obstruct the effective recognition of, or implementation of solutions to primary problems (nor does it facilitate the detection of structural strengths by which solutions can be speeded up).

The greater the emphasis placed on structural elements within the world system and dynamic relationships between them, the greater should be the practical value of the file when set up. The incorporation of general political and social indicators was therefore envisaged but only as a part of the node or link description coding.

10. The file design should not be an attempt at model building but should rather provide the elements from which a wide variety of partial or general models could be built. It should, be left to the researcher to define the classes into which he wishes to group entities for model building purposes.

The advantage of this approach is that an attempt is made to include as many different types of entity as can be detected. The researcher is therefore forced to explicitly exclude certain types of entity when building partial models, rather than merely neglect certain types of entity because their significance has not been brought to his attention

11. Additional factors governing the design arise because of the Tactical problems of implementing and maintaining the system. These are:

- **flexibility of development.** It would be impractical to introduce a large amount of data before making use of the system. The file should therefore make provision for build-up
    - (a) in number of entities included over time
    - (b) in detail included about entities
    - (c) of new types of detail not envisaged at the time when the file structure was designed (This permits the file to be extended in response to demand and as funds become available without any need to follow a predetermined order of development. Stored information should be of optimum utility of each stage in order that it should immediately justify funds allocated to the project)

- **initial focus on the international system**. Since the network of international organizations and related entities supplies a basic structure for the world system, the file should be developed down from international organizations, through their national members and then includeother national entities and local bodies of great significance. In this way the file would be focussed on the most 'coordinative' entities of the world system, at each stage.

- **low priority for commercial bodies.** Since commercial organizations are very well documented and have already been incorporated into many sophisticated information systems, it should not be necessary to include them initially. Exceptions to this would be multinational enterprises and their national subsidiaries, togethter with research institutes set up for commercial purposes. The file organization should not however preclude incorporation of profit-making bodies as such, in t!.ose cases where they are considered to be of interest.

- **mailing list preparation.** To provide a source of funds, as well as to facilitate file maintenance, it was considered necessary to design the file in such a way that names and addresses of organizations could be conveniently listed in a flexible manner for

mailing, survey questionnaires and directory preparation purposes. Unless the system is used a great deal in this way, insufficient mail returns are received to feed back corrections and keep e system up to date, and therefore of continuing value. It is only by being in a position to supply mailing list information that he system can make practical use of research techniques developed to detect unnecessary communication and coordination gaps and their effects on programme implementation. Following on from this, the greater the extent to which the mailing list use of the system can be facilitated, the greater should be the value of it to those organizations included which are faced with communication and coordination problems.

- **receptiveness to data in a wide variety of formats.** In order to maximize the value of the system to different research groups and to increase the detail included on entities, the file should be able to incorporate survey data on entities and links from many sources without any need to completely restructure and recode the data.

- **new computer input/output techniques.** Since the system would be developed over a period during which remote and/or visual display terminals will become increasingly accessible and low in cost, it is necessary to minimize the difficulties in making use of these devices for retrieval and display of information. Use of visual display devices in particular, should considerably facilitate attempts to represent the operation of the world system, both in general and at a detailed level, from the static (structural) point of view and from the dynamic aspect (inter-entity flows, proposed structural modifications).

12. Finally, the file organization had to be kept reasonably simple to facilitate input, and updating.

The most important factor implicit in many of the points above is the generality of the required file design, because of its generality, the system should be of value to a wide variety of users, but it is only the generality which facilitates response to a wide variety of cross-category queries and permits the construction of models of the world system as a whole.The difficulty inherent in optimizing a general design (aside from that of locating financial support) is illustrated by a quote from Bertram M. Gross (The State of the Nation, p. 138) on the preparation of general social indicators:

> 'Most proponents of new indicators, however, are mainly interested in some special category of data - say, educators in educational indicators, psychiatrists in mental health data, sociologists in information on strat- ification and mobility, political scientists in voting behavior and political attitudes. Activists in all fields are interested in new information that will help to vindicate their position or indict the opposition... Only a small minority of proponenets -- whether on the producing or the using side -- are interested in enough new indicators to provide comprehensive social systems accounting.'

Entities included

The file will permit the inclusion of the following types of organizational entity. It **is** however highly probable that the different groups will be given different levels of prioity, approximately that of the order here. Individual entities from low priority groups could of course be included at any time if necessary. The groups are based on conventional categories, but the file organization will of course permit much more flexibility in selecting categories.

- **international organizations:** international governmental organizations; organizations of international non-governmental non-profit organizations; international non-governmental non-profit organizations; regional international organizations; international meeting series; multinational business enterprises

- **commissions and sub-commissions of international organizations** (particularly where they may have independent fields of activity, names which may create the impression that they are unconnected with the parent body; also cases where the links between the secondary body and its parent may be of significance to an understanding of the operation of the parent body or the mechanism by which a particular problem is dealt with)

- **organizations of national non-governmental, non-profit organizations** as the major coordinative bodies for non-profit activity

- **libraries and information centres**

- **national organizations** (governmental and non-governmental) with international programmes or interests (significant state or local organizations with international programmes or interests particularly where such organizations are important to the implementation of international programmes and where they are the only ones of their type in the country (or he world) and may therefore be considered of international significance. Such organizations may also represent the major source of potential membership of international non-governmental organizations, or in the case of governmental bodies, for the implement- ation of international recommendations)

- **bilateral international organizations**

- **international programmes, projects, 'days', etc.** (particularly where these are independent of any individual organization or have names which create the impression that they are organizations or independent; also cases where collaboration of organizations through the programme is of importance to an understanding of the mechanism by which a particular problem is dealt with)

- **international treaties and agreements** (particularly where these take over the normative function s of organisations or are the

principal reason for the existence of an organization)

- **international journals, directories, abstracting or bibliographical services** particularly where these in effect take over the information processing and disseminating function of international organizations or are the principal reason for the existence of a particular organization or are in effect the most important 'coordinative structure in that field

- **individuals:** holding positions international organizations; international roles or position (particularly where the positions held by one individual are such that he himself performs an important integrating function in linking organizations (e.g. cross-linking directorships in business enterprises, or individuals holding positions in government and in non-governnental organisations)

Clearly there are many similar types of entity at the national level which could be included if this was considered justified. The emphasis above has been placed on the geographical coordinating function of entities. Equal emphasis could be placed on cross-disciplinary or cross-jurisdictional coordinating functions, and priorities could be allocated accordingly.

The concept of entity is sufficiently general to permit inclusion of other types of entity if necessary. Possibilities are considered in a later section.

Comments on other possible entities

The purpose of considering other possible entities is to arrive at greater facility in identifying and describing parts of the world system.

Sub-Systems and Classes of Entities: Sub-systems may conventionally identified by name (e.g. international NGOs, the American banking system, etc.). Descriptive coding can be supplied, as can keyword coding. The actual entities which make up (i.e. are 'members') of the sub-system can be clearly defined, individually or as classes and cuttin down access time. Such cards would be a useful means of avoiding analysis. A library of sub-system cards could be built up as a result of each analysis of the file as a whole. Each sub-system would be defined according to the special definitions used by the invesigator. The result might *be* that of a series of overlapping classes which had together employed definitions which effectively excluded some specific entities registered within the file. This in itself would be useful.

In some cases the sub-systems would in fact represent a non-existant umbrella or- ganization.

Depending upon how the systems were defined, it could be useful to include 'black box' system cards known to be important parts of the system with known inputs and outputs, but about which it was impossible to provide any description with certain ty.

Religions, Armies, Tribes and Clans: The manner in which the system is conceived does not preclude treatment of these as entities. Their heirarchical structure and cross-links to toher entities could easily be indicated.

Movements of Opinion and Informal Organizations: Since a structure can be identified for informal organizations and, using classes of entities, movements of opinion, there is no reason why these important features of the world system should not be included, if this was considered necessary.

Information and Communication Systems or Networks: Information networks may be independent of any particular organization and may therefore be considered to be important integrating factors in their own right. They possess a well-defined structure and may therefore be included if necessary.

Decisions: Where a decision is taken as the result of the deliberations and activities of a wide range of organizations not necessarily formally linked, it would be an advan- tage to treat the decision as a type of entity in its own right. The organizations which participate in the decision-making process may then be treated as 'members' of this entity.

As a detail of an organizational structure, cards of this type could be used to indicate the inputs and outputs to decision centres.

Financial Equilibrium: As one possible means of evaluating an organization and detecting classes of orga- nizations, elements of its financial statements could be expressed in ratio form. This could perhaps be considered as a supplementary description card.

There is a case for coding absolute amounts only and choosing the ratios to be calculated according to the research goal. This avoids the initial manual calcula- tions prior to input.

Propositions: At some stage, it would be an advantage to store propositions concerning the functioning of the world system and its sub-systems. They could be filed at any stage of verification, so that apparently contradictory propositions could exist together. Each would have its status changed as it moved towards acceptance or rejection.

The value of including propositions once the file is used for simulation and deci- sion-making is to offer the user a choice of relationships governing a field in which he is interested, *plus* all the necessary qualifications. A proposition veri- fied for a limited set of cases could be drawn to the users attention as a possible guide for a decision in his unexplored area.

It would be useful to express propositions concerning flows or restriction on flows between entities or classes of entities as simple mathematical functions. In fact it is probably only propositions which can be so expressed which could be usefully included.

Criticisms: Inclusion of data from a variety of sources will clearly lead to *a* situation where two or more sources of different standpoint will disagree. This disagreement is itself *3* feature of the system and important to an understanding of its Operation.

Provided a critic card follows the same format as the card of data criticized, either may be chosen, or the two compared to establish the degree of dissonance. An example would be a comparison between stated objectives and some evaluation of the 'real' objectives, or of what is really being achieved.

Problem Areas: Consideration has been given to means of coding problem areas, as distinct from subject or field of interest areas. An organization can be concerned with a field of interest selected from some sort of representation of the totality of possible fields of interest and ordered into classes and sub-classes. It would be useful to develop a structure of problems with which entities can be concerned. In effect this is an ordered collection of ways in which any entity and in particular (by extension) the world system, can malfunction.

This problem thesaurus could be used as a qualifier on field of interest coding to indicate in what way the field is of interest or is a matter of concern, thus clarifying the objectives and activities of the organization. Alternatively, problem areas could be treated as entities with a 'membership' corresponding to those bodies concerned with them.

In the first case a valuable predictive tool would be created. For if analysis shows that a number of organizations are concerned with a limited number of problems within a particular problem area, it will bring out those aspects with which no organization is concerned and concerning which data should be obtained, even if only as a check. In this way a systematic picture could be built up of what might go wrong in the future, or might be wrong, but be undetectable because no body employs the conceptual categories necessary to detect the problem, possibly because it is interdisciplanry.

Presumably such problem hierarchy would at its more abstract end include the vague concepts included in organization objectives, about which it is possible to enthuse e.g. cooperation, well-being, etc. At its more detailed level, it would include statistics of the problem as measured. It is the intermediate levels which would prove of value as a guide to decision-making.

The disadvantage of this approach is that no generally accepted and highly develo- ped problem thesaurus exists. Insuffiecient is known about system malfunction in the most general sense. Even 'problem'does not seem to be very well defined in the system sense.

The second approach is simpler but more closely related to the field of interest coding. The problem area could be treated as an entity with a membership. Related problem areas could be linked using the inte-entity link cards. Because this is an associative type of coding, no predictive feature is available, but it does increase the ability to evaluate the degree of coordinated response to a problem area.

It fields of interest were coded as entities, then problem area coding would blend into field of interest coding. A field of interest could then be considered as a problem area in a broader sense.

The link between an organization (say) and a problem area is then that the organi- zation is set up because the problem there is considered critical. Organizations could then be considered as society's response to a problem area. An organization may express concern (general interest) about certain symptoms, but consider that an indirect approach was necessary. The problem area attached may therefore not be identical with the symptoms of concern.

An advantage is that an attempt is made to distinguish between ways in which an organization is concerned about a subject area and how that subject area is defined as a problem and how it is proposed to attack that problem. This sort of qualifica- tion on field of interest coding would avoid superficial analysis identifying dupli- cation when the two bodies were concerned about the same area in different ways. It would also highlight those cases where an organization is apparently the autho- rity in a certain field of interest but in fact is only responsible for certain aspects of that field of interest.

This sort of problem area approach would help to take the emphasis off documentation about a field of interest and place it on the way in which that field of interest constitutes a problem and what needs to be done about it.

Entity description and card types

An important design criterion is flexibility in the amount of information associated with each entity at a particular time, in the way in which the information is built up over time, and in the addition of new types of information after the program is operational.

This is achieved by having a number of different input card groups for each entity by which a number of types of description of the entity can be stored. Each group contains cards or series of cards having the same format for similar types of information. In the initial design several options thus created have been taken up, but many remain for new types of data on the entity or on its links with other entities.

As the file is developed the number of card groups used to describe the entity and its relationships with other entities may progress from three (minimum) to the limits Imposed by the card coding system and the tape format. In addition, within each group the number of card types used may also be built up. Needless to say, large amounts of information would usually only be available or desirable for enti- ties which are considered to be of great significance to an understanding of the operation of the world system.

The range of card groups which it is possible to have associated with an entity is as follows :

1. Name cards: one card per entity name (with one English version wherever possible or necessary) and including the abbreviation (initials) normally used ; pro- vision will be made for a single overflow card.

2. Address cards: one set of two cards per address. In general each address will be consi- dered as indicating a different entity (even where the address is not the main address of the parent entity), but in some cases an entity may have two principal and equal addresses.

3. Description cards: single cards containing descriptive coding. The first card in this group will contain the minimum information necessary to distinguish between different type of entity, plus a field in which more detailed summary coding can be entered.

4. Class link cards: one card per type of link. Used where the entity cannot for practical purposes be shown as specifically linked to many single entities (e.g. information or recommendations sent to '10,000 agriculturalists'). The main use envisaged for this initially is as a field of interest key- word/keycode card.

5. Geographical link cards: one 1 to 4 card series per type of link. This permits coding of types of link with up to 220 individual countries (or a similar number of geographical areas within a country, etc.). Initially the main series envisaged are

- countries in which a given entity has 'members' (in any sense spe- cified)
- countries in which a given entity has activities

This series will be the principal means by which information on 'flows' of any type and frequency between entities and geographical areas will be stored.

6. Inter-entity link card: one n-card series per type of link. This permits coding of the frequency and importance of the link between the entity and each of those with which it is linked (as indicated by the identification number of each). Initially the main series envisaged are :

- 'members' of the entity (in any sense specified)
- entities of which the entity is a 'member' (in any sense specified) This series will be the principal means by which information on 'flows' of any type and frequency between entities will be stored.

7. Link detail cards: one card per link. This permits detailed descriptive coding of a particular link between two entities.

No coding is planned and clearly this type of card would only be used for links of considerable significance (an important example being the Washington-Moscow hot-line).

8. Sub-entity cards: one card per sub-entity. These cards are coded in a similar way to the Name cards. They are used to identify either dependent or internal bodies for which it is not considered necessary to prepare complete coding as separate entities, or temporarily applicable details concerning the entity (e.g. where the continuing entity is a meeting series, each new meeting may have a special title in addition to the series title. This information is stored here).

9. Other card groups: A number of other card groups are under consideration.

Relationship between dependent, independent, continuing and temporary entities

The file structure has been designed on the assumption that it isvery difficult to distinguish in a generally acceptable manner between continuing and temporary entities, and between independent and dependent (possibly internal) entities. Provision is however made for conventional distinctions of this type to facilitate some mailing list uses of the file.

To enable the entities to be grouped in as many ways as possible, temporary entities **are** coded in the same way as continuing entities, and dependent (or internal) entities in the same way as independent.

A complication arises because:

> a) a dependent entity may use the same address as the main body b) a temporary body may use the same address as the main body(e.g. one meeting in a meeting series) c) an organization may be responsible for a meeting series for which information on one meeting (temporary) must be stored d) it may not be considered necessary to store more than the title of the dependent or temporary body (where the latter is also dependent).

These difficulties are resolved as follows:

- a description card carries coding to distinguish between dependent and independent (The user does not have to rely on this somewhat arbitrary coding), and possibly, uncertain
- a description card also carries coding to distinguish, between continuing and temporary
- where the entity is
    - temporary but not dependent: it is treated as an independent entity and no problem arises
    - temporary,dependent, significant, dependent address: it is coded separately, the address being replaced by a cross- link to the main body
    - temporary, dependent, significant, independent address: it is coded separately with its own address
    - dependent, insignificant: it is coded by name only as a sub-entity card of the main entity

This solution does imply that not all dependent entities are tacked onto the main entity in a manner convenient for listing without a preliminary sort. This feature world be required for periodical checks on the file.

Entity identification numbers

Each entity coded separately, whether independent, dependent, permanent or temporary, provided it is considered significant, will be allocated a unique 6 digit identification number. The numbers will be allocated sequentially and will have no significance in themselves, in order to avoid the many difficulties when entities change over time. When entities cease to function, the identification number will become free for allocation to a new entity.

This numbering system gives a maximum of under one million addresses which *is* certainly sufficient for medium-term goals. It is possible to consider using alphabetic characters in the code. If this was done, six positions might not be necessary. Alphabetic coding might however prove unsatisfactory if much sorting is expected on the identification number, as is planned. In addition, since it would be an advantage to facilitate use of the file on direct access devices, numerical coding might be more efficient.

Card coding system

Depending on whether alphanumeric coding is permissable one or two columns of code will be required to distinguish between data cards, movement cards, processing request cards, etc. for system purposes. The first two columns have been reserved for this purpose.

Within each card group, two further columns may be required to permit a sufficiently wide variety of additional cards. Needless to say, some of these codes will only be allocated to certain entity classes, so that the number of cards, in a particular group of a particular entity will generally be small, unless many surveys have been conducted on that type of entity. In general, each survey of entities will be allocated a code in these two columns, except in those cases where the number of columns of data is small. In such cases, the data would b*e* allocated to spare columns on other records.

A final column is required for cases where cards within groups are part of a series. This mainly applies to long names. Where one survey generates more than 60 columns of data within one card group, this can be split as though coming from separate surveys. It may therefore be possible to avoid the use of the extra column. In which case, the sort prior to input of the cards need only be made on the first two columns, since it will be necessary to scan the remaining records by program anyway to detect whether a particular one is present.

Date of information-received coding
The address and description cards, as well as the link cards, will each need one or two columns to indicate the date at which the last information was received. This cannot be restricted to one card, since whilst general inform- ation about an organization may be up to date, an address may have lapsed or membership etc. changed slightly

Source of information-received coding
Survey coding will be identified by the card code. It may however be necessary to consider an identification of soiree code for address and basic description coding.

Geographical codes

The system is designed to permit entities to be sorted on the basis of country or city geographical coding, whether this refers to the headquarters of an organiza- tion, the location of its members, or the location of a meeting.

Country coding

The country codes used are based on a List of Nations and Territories prepared by the International Peace Research Institute, Oslo on the basis of UN Statistical Office information. The list was intended as a 'description of the universe of nations, not as any kind of sample. In addition to all independent nations, non- independent territories are included if they fulfil **all** the following criteria :

1. separate administration, 2. indigenous and resident population, 3. geographical non-contiguity to a dominant nation.

The following are **not** necessary conditions for inclusion :

> 1. national independence, 2. the existence of an independence movement if the territory is a dependency, 3. a minimum size, 4. availability of data ..... Although certain small territories may have been left out due to incomplete information, we believe that this list comes close to the best definition that can be made of the universe of nations and territories at the present time'. ('The PRIO List of Nations and Territories',Oslo).

Three digit codes are used. The first digit representing a regional classification. 'Regions are purely geographical, and some that are habitually kept apart have been collapsed in order to get regions of roughly comparable size'.

100 Africa: 101-157 countries

200 America: 201-252 countries

300 Asia: 301-349 countries

400 Europe: 401-436 countries

500 Oceania and Australia: 501-528 countries.

This system does not provide for sub-regions which may be considered of importance, e.g. Caribbean, Latin America, South-East Asia, some of which are difficult to define. There are sufficient free codes within any region however to be able to identify such regions if necessary at a later date. In addition the standard search program should permit the user to define and select his own regions.

Similarly, since the codes 600-900 are free there should be no difficulty in defining bi-regional, or political regions if this is considered necessary.

This system does not permit areas such as 'Scotland' to be treated as countries or territories. Since some international bodies indicate members in Scotland as dis- tinct from those in England, this could create a problem.

City and within-country coding

In order to be able to extend the application of the system to national and local bodies and their activities, within-country regional coding and city codes are re- quired. Ideally this could be based on the telephone, telex or postal codes. Since these have not been developed in a consistent manner (Sterky, H. 'Is a common, world-wide numerical code for all countries Utopian or feasible' Union Postale, UPU, 1968,

5) and are not readily available for major cities and towns of the world, another system must be sought. Convenient system could have been usefully based on the latitude and longitude of the city in question, but this would involve 8 digits **for** reasonable accuracy, and does not take advantage of the grouping already achie- ved by the country coding. It would have been very useful for specifying regions within countries.

Name cards

1st card cols:

- 1/2 name card code
- 3/8 identification number
- 9/10 language of name
- 11 name translated, *if* necessary
- 12/79 name split into address line zones
- 80 second card exists; abbreviation in first end of name

If the name ends by col. 69, cols. 70/79 may be used for the name abbreviation, provided the appropriate code is given in col. 80, and provided the name abbreviation does not exceed 10 letters. 10 letters covers most abbreviations (international Initialese, UAl, 1963) but does not cover shortened names and alternative names cum abbreviations, which exist for some bodies. It may be necessary to provide for an alternative name option to pover this possibility. It is useful to include such names in order to complete indexes and as a short alternative for addressing purposes.

2nd card col.

- 1/79 as above
- 80 abbreviation in second card end of name

3rd card

In some cases a third card may be necessary. Since this is not frequent, it may be possible to omit the need for a third card

Problems

- how to identify which of several 'other' languages have been identified by which codes
- is there any advantage in attempting to include common parts of entity names as single letter standard abbreviations and regenerating the full word on print. Examples are: International (l), Commission (c), etc. What about other languages with common root: Internationale (l/e), etc. 10 such abbreviations would save much space
- would title abbreviations affect the KWIC index program

Processing

- list after sort on any combination of name, abbreviation, identification number, possibly grouped and/or selected according to codes on other cards
- KWIC index on names and/or abbreviations in selected language(s), possibly grouped and/or selected according to codes on other cards
- regeneration of full name if within-name abbreviations are used
- format options for printout

Address cards

1st card col.

- 1/2 address card code
- 3/8 identification number
- 9/10 address type code
- 11 1st language preference
- 12 2nd language preference
- 13/15 country code
- 16/18 city/town code
- 19/79 address line zones
- 80 second card code
- end of address code

2nd card col.

- 1/10 as in 1st card
- 11/79 address line zones
- 80 end of address code

3rd card

a third card may have to be envisaged

Problems

- to retain mailing list flexibility and permit addresses to be used in many countries, the appropriate language must be used for the country name e.g. can Deutschland be used from the U.K. or France, and is an English sorted list of towns acceptable with Wien instead of Vienna
- should telephone and telex numbers be included in zones in the sane way as abbreviations as an aid in preparation of directories

Description cards

The first of the description cards carries the minimum coding necessary to distinguish between the various types of entity. The card will therefore probably be allocated the lowest card code. The minimum coding in fact only occupies the first part of the card. The remaining columns arc available for ore detailed coding.

col.

- 1/2 description card code 3/8 identification number 9/10 description card type 11/15 Yearbook of Int. Organizations number (for international)
- 16 permanent/temporary
- 17 independent/dependent
- 18 proposed/founded/active/passive/dead
- 19 type of entity
    - organization
    - meeting
    - treaty
    - periodical
    - bibliography
    - directory
    - etc.
- 20 governmental/non-governmental governmental
    - non-governmental, non-profit profit
    - government/NGO government/profit NGO/profit
    - government/NGO/profit inapplicable/uncertain
- 21 qualifiers on each type within col. 20
    - e.g. for non-governmental, non-profit: learned society trade association mass membership etc. membership
    - finance board members activities objectives

  A second description card planned will cover the serial number, year, city, and country (and possibly the number and origin of participants) of past meetings in a series.

To facilitate file maintenance, the first description card will carry a series of codes indicating action that needs to be taken to acquire further information, whenlast tried, etc.

Extracts from PRIO coding

Group I

- How long has Secretary-Gneral been Secretary-General 8
- For how long is Secretary-General aprointed/elected 10
- Onwhat initiative does Secretary-Gneral act 42
- Description of organization - activity initiated 57
- branch communication 58
- branch contact frequency 60
- membership basis 61
- origins 62
- source of membership 63 1st,2nd,3rd

Group II

- source of Sec-Gens 6 1st,2nd,3rd
- location of HQ 16 1st,2nd,3rd
- source of board members 26 1st,2nd,3rd
- source of staff 36 1st,2nd,3rd
- source of revenue 46 1st,2nd,3rd
- size of staff - paid 56
- vol 57
- size of budget 58-60
- source of non-member country finance 61 1st,2nd,3rd

Group III

- ranked sources of income 6-
- ranked sources of expend 14-21
- change of structure,

- future plans 22-25
- Contactswith other int. organizations
- shared secretariat facilit 321st,2nd,3rd
- exchange of board members 481st,2nd,3rd
- exchange of observers 6 lst,2nd,3rc
- joint meetings 22 1st,2nd,3rd
- Sec-Gen friendship 38 1st,2nd,3rd
- other 54 1st, 2nd, 3rd
- national governments 70
- Description of organization (contd}
- basis of membership 71-72
- membership as % of 73
- potential basic of activity 74

Group V

- ranked goals 6-17
- frequency of board mtg 63
- frequency of open cnf 64

Group VI

- membership of inclusive int. orgs 6 1st, 2nd
- status with UN Agencies 17
- activity importance 18
- problems 19-27

Group Supplementary

- Telephone, cable 6
- UN Representative 7
- Foundation date 8-9
- Foundation place 10-12
- No. of nations represented 13-15
- No. of individual members 16-18
- No. of group members 19-21
- General Ass. frequ*ency* 22
- Structural bodies 23
- Staff paid 24-26
- vol 27-29
- Languages used 31-37
- Budget 39-41
- Finance 42
- IGO Relations 43
- NGO Relations 44
- Number of activities 45-46
- No. of conferences 47-48 No. of historical publications 49
- No. of periodicals '50
- No. of non-period. '51 N*o*. of publication languages 52-54
- Date of information 55
- II: Nations represented (each) 1-50
- III ' ' 1-50
- IV ' ' 1-65
- UN Agency membership 69

Special coding for entity types

Meetings

- meeting series field of interest coding as organization
- meeting theme ' ' '
- open or closed '
- body responsible
- travel strategy
- periodicity
- average meeting budget ' ?
- average meeting attendance (participants/countries) ' ?
- report publication policy
- past meetings (place, year, country, number)

Universities

- number of teachers
- number of students
- library volumes
- budget

Libraries

- type breakdown (reference, public, specialized, etc.)
- number of volumes
- number of accessions per year
- number of periodicals received
- number of bibliographicals received
- publication of accessions list/bibliographies
- budget
- language range
- subject range (books and periodicals separately); general and special
- receptive to publishers catalogues
- branches

Periodicals

- type breakdown (newsletter, journal, yearbook)
- frequency
- circulation
- some measure of size
- Subscription price or cost
- acceptance of advertising
- circulation breakdown by class
- bibliography

Bibliographies

- frequency
- circulation
- number of entries
- subscription price or cost
- circulation breakdown by class
- abstracts
- index
- started; stopped
- subject coverage geographical coverage

Directories

- number of addresses
- geographical coverage
- cost
- frequency
- subject coverage
- circulation

Treaties/Agreements/Conventions

- proposed subject coverage
- drafted UN Treaty Series no.
- in process of ratification
- ratified
- date for renewal or termination
- administering organization
- depository
- signatories
- minimum number of signatories for effect

Programmes/Projects

- proposed
- voted
- period of activity
- budget
- subject coverage

- code number for project (cf. Unesco)
- associated organisations
- geographical area
- body responsible
- body reported to

Multinational business enterprises

- assets (with % in non-HQ country)
- capital (with % in non-HQ country)
- turnover (with % in non-HQ country)
- profit (with % in non-HQ country)
- subsidiaries (with % in non-HQ country)
- associated (with % in non-HQ country)
- branches (with % in non-HQ country)
- manufacturing countries(with % in non-HQ country)
- nationalities on board
- staff nationalities
- number employees

Foundations/Trusts/Funds

- assets
- allocations per year
- geographical area
- subject area
- annual deadline

Roles/Positions

- period of tenure
- time required
- voting power/policy formulation
- information received
- paid/unpaid
- honorary
- office entitlement

Individuals

- age
- background
- roles

Class link cards

These are required where information on an entity's links can only be given in general terms. This may be either because the detailed information is confidential or that it would be impracticable to list a large number of entity links, e.g. mai- ling lists, individual members, etc.

col.

- 1/2 class link card code
- 3/8 identification number
- 9/11 link type coding
- 12/80 Class coding.

In the case of field of interest/subject keywords the 12/80 field could be subdi- vided as follows :

- 12/36 five 5 digit zones : digit
    - 1 range value of importance of field
    - 2/4 3 digit general code (corres- ponding approximately to the first three digits of the Uni- versal Decimal Classification)
    - 5 detail on 2/4 if necessary.
- 37/78 three 14 digit zones to be used for detailed keywords not covered by the hierarchy coding in 12/36.

In the case of classes such as

- 1000 major libraries
- 50% of dental association in Europe
- 250 medical journals.

or of distribution lists such as mailing list of 1000 split

- 50% chemists in Asia

- 20 % hospitals in Asia
- 10 % government departments
- 10 % international bodies
- 10 % miscellaneous.

a satisfactory met hod of coding must be arrived at. It would be useful to make use of the coding conventions used in the description card when defining types of orga nization.

Uses

Aside from its use for subject or field of interest indexing, these links can be used :

- to describe mass membership or mass circulation of information
- the hierarchical structure of the UDC coding permits greater flexibility in selection.

Problems

- is the attempt to indicate relative strength of interest in a particular subject by range values satisfactory.
- it would be possible to have a historical record of past fields of interest by keeping this type of card in the system. Is this too close to documentation of the past.
- is the treatment of field of interest as a class forced and should it be treated separately.
- is the development of a class coding practicable. It must include subject field, type of entity, geographical location and the link must be described in terms of type (information, membership, policy, etc.), qualification ontype (e.g. for information :official report, journal, bibliography etc.), frequency of contact, a measure of the importance (e.g. number of items distributed, etc.)
- if certain classes of entity are themselves treated as entities, this may simplify the situation by permitting the use of the entity identification number.

Comment on keyword system

In choosing a simple numerical code, it was hoped to avoid the complications of comprehensive indexing met with in the U.D.C. Each general subject or aspect of a subject would receive its own code. No relationship between two or three such key- codes would be attempted at the time of indexing. Any indexing difficulties at the detail end would be dealt with by using keywords.

The 3 digits reserved for the main coding should give sufficient logical structure. The thesaurus required could advantageously be a modified form of the U.D.C. coding to avoid the problems of distortion to due to rapid development of particular fields of knowledge.

An alternative as considered, which is more consistent with the treatment of a field of interest as a 'class' with which the entity has some form of relationship. This has the advantage of avoiding the necessity of building up a rigid logical struc- ture.

Each keyword would be given an identification number which would be selected on a sequential basis, identical to that of entities, but would avoid duplicates. In this way each field of interest is defined *as* an entity. The field of interest coding required would then amount to a description of the membership relationship of the entity coded to this new type of entity.

These new entities could then have their own descriptive coding by which they could be cross-linked into a complex hierarchy in a number of different ways.

This option is theoretically attractive but raises problems of space, (because each keyword is treated as an entity) and processing time for a selection of entities in any field of interest.

Keyword Processing

- sort on keycodes or keywords with printout of name of entity
- sort into hierarchy order
- merge with KWIC indexing of names, particularly with data card input giving key- code/keyword conversion in selected language.

Geographic link cards

Each card of the five card set refers to one geographical region (either within the world, or at some later stage within a country). Only one card need however be used if the countries are all concentrated within one area.

Col.

- 1/2 geographical link card code
- 3/8 identification number
- 9/10 link type code (e.g. membership, finance, policy, information, etc. and the direction of flow, from or to the geographical area in question).
- 11/12 link frequency and strength codes
- 13/15 geographical region code (e.g. for countries : 100 Africa 200 America 300 Asia 400 Europe 500 Oceania)
- 16/79 area coding. A single column is available per area. In this a qualitiative code or range value for the strength or frequency of the link may be punched.
- 80 reserved.

An option is being considered for those cases where the links are limited to a num- ber of widely dispersed areas. Using the above system, much space would be wasted. The option consists of giving the full 3 digit code of the countries (instead of *a* range value in a

single fixed position column) in cols. 16/79. This could be accom- panied by a single digit field for each country, to give range values making a 4 digit code per country.

An option is being considered to permit more detail on a single country. In this case cols. 13/15 would contain the code of the country. This could permit mixed national/international coding.

Uses

Could be used for :

- spread of membership by country, with some range value of numbers per country
- funds received from each country (possibly of a particular type) with some amount range value
- funds expended in each country with an amount range value
- type of program activity in each country
- amount of information distributed to each country
- nationality of board members
- etc.

Problems

- the single column of coding per country does not permit much detail. Budgets can only be indicated as ranges, not more accurately using the first two budget digits and a power of ten code in three columns.
- there are 36 European and 28 Oceanian countries. It would be an advantage to combine these onto one card to save space. This would however destroy the meaning of the country numbering.
- should the from or to sense of the flows be indicated in a separate column to facilitate flow-chart processing.
- update processing in the case of the 4 digits per country option. The field must be scanned.

Inter-entity link cards

col.

- 1/2 inter-entity link card
- 3/8 identification number
- 9/10 link type card
    - membership
    - policy
    - finance
    - information
    - consultation
    - etc
- 11/75 eight 8 digit zones:
    - 1 link frequency code
    - 2 volume/amount code
    - 3/8 linked entity identification number

Uses

Could be used for specifying any type of link or flow between entities

- 'members' of the entity in question
- entities of which the entity is a 'member' (e.g. an early use will be to specify with which United Nations Agencies the international organization has consultative status, which type of consultative status, and an evaluation of the importance of the link)
- entities from (to) which policy/recommendations/information/finance/etc. of specified types is received (sent)
- in the case of 'business enterprises, the two extra columns will be used to indicate the allocation of stock/voting power in 2 digit percentage terms

As previously, a new type of link can be specified at the time the set of data is input, provided that there is sufficient worthwhile information to warrant allocating a code to permit this

Problems

- as with the geographical link cards, it might be an advantage to reserve a single supplementary column to col. *9/10* in order to indicate the direction of flows to facilitate flow-chart preparation
- as before, a single column of code, or even two, does not permit adequate coding of the flows of funds between two entities. A third column to permit power of ten coding on the 2 digit field may not however be justifiable
- when updating a link card, the other end of the link must also be updated. This could create processing problems unless both changes were introduced separately

Processing

- suppression or addition of specified identification numbers
- scan to check the presence of a particular identification number

Processing requirements

The information is coded in such a manner as to facilitate the storage, retrieval and analysis of entities and their relation: hips in many ways. The main program will be designed to facilitate file maintenance, mailing list preparation and general research. This is described in the first section 'below. Special programs may be used for more detailed research and display of information. These poss- ibilities are described in later sections.

Due to the many types of structure which it is possible to treat as 'entities' and incorporate into the file, it is difficult to detail all the possible queries which may be answered. The file has however been designed to facilitate descrip- tion of the **state** or characteristics of entities and of their **dynamic** performance or modes of activity in order to permit flow analyses between entities in the most general sense.

The processing requirement and capability will depend clearly upon the amount of detail included on entities, which in turn will depend upon the priority of that Type of entity. Initially only sequential analysis will be possible. Once cross-reference or link cards are included more sophisticated processing will be possible for those entities for which such information has been collected.

Immediate Requirements

1. **Mailing and other lists**

- print mailing and other lists with choice of amount of information (abbrevia- tions, names, address, coding, etc.) and format, cased on combinations of codes in any of the cards grouped with the entity. This will mainly be used for selection by: country, city, language, field of interest, descriptive coding (size, type, membership, etc.) and also to select those entities for which more information is required
- index processing and listing based on names and fields of interest in chosen language as well as selective coding to group by geographical location or other criteria
- consideration will be given to the need to facilitate directory production by formatting in a manner suitable for a computer typesetting routine

2. **Calendar of meetings**

- print calendar of future meetings in all or chosen subjects and with the option of selecting on descriptive or past meeting codes
- mailing list to request information on new meetings based on past meeting frequency
- mailing list to request information on vast meeting reports
- index processing and listing as above

N.B. Although the routines required for this would permit preparation of a programme/project 'calendar', such calendars cannot be prepared until programme information is included at a later date)

3. **General research**

- total on any code or combination of codes
- most other general research requirements will be covered by the mailing list routines. Users will have to provide their own programs for more sophisticate, analyses

Medium-term Requirements - A

The following requirements should be covered by the same program used for the Immediate Requirements. It will not be possible to perform processing of the types mentioned below until membership and other link cards are coded, or in other cases until the demand is sufficient.

1. **Link research**

Selection and list of entitiesto which any given entity or group of entities is linked, based on the coding of the entity and on the coding of the entity linked.

This will mainly be used for: selection of national branches of international organizations; selection of international organizations with a given type of national branch; selection of organizations which have or have not held meetings in a given country or city

It could also be used for: selection of organizations with board members of a certain nationality; selection of organizations with programmes in a particular country

2. **Modification subscribers**

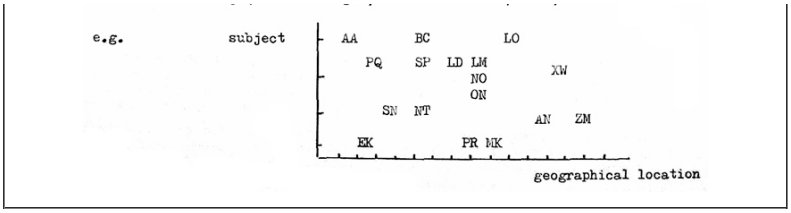
Selection and list of modifications, particularly change of address, based on *a* selection profile and a monthly or quarterly subscription

Medium-term Requirements - B

The following requirements should be covered by separate programs developed specifically for that purpose.

1. N**on-interactive graphic display**

a. **In the absence of cross-linking coding:** a program could be developed to print out the identification numbers of entities within a 2-dimensional frame. Both coordinates and scale to be specified by the user. Coordinates could be based on field of interest keycoding, geographical location coding, number of members, or any other descriptive coding (size of budget, foundation date, etc.)



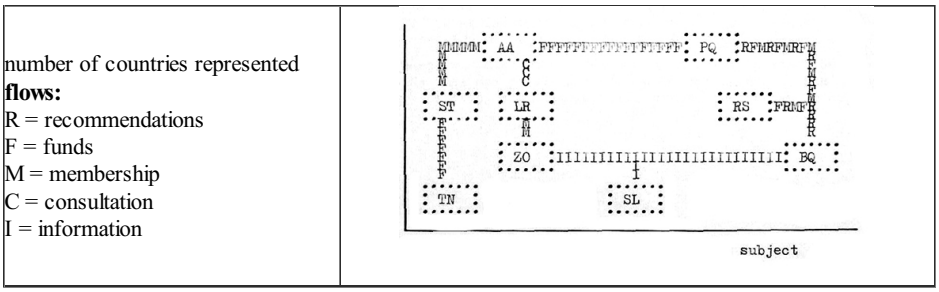
Identification numbers could possibly be printed out as 2 or 3 character codes with a conversion table or a direct list of names. Alternatively, abbreviations could possibly be used with codes only where necessary.

In a similar way, subjects (for example) not picked out by the coordinate framework, could be printed out as qualifiers on those subjects listed for organizations within the framework. Alternatively, such information could possibly be given with the conversion or look-up table.

If many organizations are represented by a particular combination *of* coordinates, a single code could be used, the complete list being given in the look-up table. This would be the only possible procedure *if* large scale coordinates were chosen.

Main uses for this form of print-out would be: supplementary indexes to directories (particularly the Yearbook of International Organizations); in meetings as a quick guide to other organizations operating in the same field (as a means of drawing attention to the need for liaison); as part of a questionnaire to organizations (a) to inform them of the existence of bodies operating in the same field, (b) to obtain information for inclusion in the file ('Please draw lines to those organizations with whom *you*are in contact', etc.)

b. W**hen cross-linking coding has been included: a** program could be developed as an extension of that above in order to print in the lines linking organizations represented in the 2-dimensional coordinate framework. This would be particularly useful after entities have been grouped into classes and averaged values for the links between classes determined by calculation. Instead of lines being-indicated simply by 'XX..' or '***' they could be given some mnemonic value oy using 'FF..' for fund flows, 'MM..' for membership link, etc., combined into 'FMFM..' for a double flow.e.g.



This represents one type of flow-chart analysis. Whilst preparing a print-out of this type, the program could also produce an analysis of those entities or classes of entity which were not in contact, according to some criteria of acceptability.

There are many possible extensions to a flow-chart analysis which could possibly be incorporated into the program and printed out in such a way as to increase the information content of the above diagram, e.g. contact frequency, flow volumes, flow quality rankings, etc.

2. **Interactive graphic display**

For research purposes it would be useful to develop a program which would permit the researcher to modify his query after each printout. In this way if he obtains some information from the first printout which indicates that he should explore a particular sub-system along a new dimension, he can immediately request such a printout. This cuts down all the usual delays associated with queuing requests for data processing and allows the researcher (or the decision-maker) to maintain thinking momentum!

Processing of this type is carried on by having the user interact with the computer via a typewriter keyboard. This may be in an office remote from the computer itself. Such processing would require that the file be transferred to a direct access device.

The user would then be in a position to 'work through' the file from some known entity (with a known identification number) exploring different flows between entities. He could group entities into classes to get a broad flowchart picture displayed or work at the inter-entity level, There are many useful possibilities to be examined.

A further development, instead of having diagrams or analyses printed out onto the typewriter (or a printer), would be to display the results on a cathode ray tube with which the user can interact via a light-pen, plus keyboard. Sortie of the possibilities in this area have been discussed in a separate note. Major advantages are the ease with which information can be added to the display, the higher information content possible with pseudo 3-D displays and the possibility of shading, blinking, rotating, zooming, etc.

In the long-term development of this type of system it is possible that the interactive visual display unit may prove to be the most convenient method of checking and updating the file (e.g. changes of address, membership, etc.)

System implications

Initially data on entities can probably best be stored on tape *in* identification number sequence. All card groups relating to a particular entity would then be together. This immediately raises, the problem of the variable length of the data strings tacked onto each entity. *A*

possible solution is to use two tapes, one for the constant minimum number of records relevant to each entity, and a second for the remaining variable numbers of records.

Whilst this solution may speed up the printout of mailing lists, it nevertheless increases the number of tapes without resolving the problem of the inconvenience of having variable length records on the second tape. In addition, specifying a maximum length for the variable length portion may prove to be an inconvenient limitation both by reducing available memory space and necessitating the maintenance of a count on the number of fixed length records within the variable length string.

A more satisfactory solution would be to work in card format, making use of any convenient blocking factor. The program could use a storage area to take care of overflow of the fixed length records pertaining to one entity from one block to the next. In this way no restriction is imposed on the length of the data strings tacked onto the entity. A second tape might however prove useful for part of the string at some stage.

File maintenance, lists and simple queries

Provided no cross-link processing is required this part of the required processing will be standard. Changes and selection masks will be input as cards, the tape will be updated and mailing lists printed.

A slight complication arises because each selection mask will result in a separate list. Clearly the first list can be printed immediately, unless it must first be sorted, whilst other lists will have to be prepared by sorting entities selected onto a scratch tape prior to printing them.

Cross-link processing

Inter-entity link cards can be stored at any stage. Simple selection using them poses the same problems discussed in the last section, processing becomes more complex however when it becomes necessary to list the entity to which the entity is linked, particularly when there are several, and their coding a factor.

This problem could have been avoided by tacking on the full information on linked entities after the main entity, e.g. all national branches of international organizations. This however creates another problem because a national 'branch' 'ay not be simply a dependent branch but a fully independent organization which is itself the member of other international organizations and may have other national organizations as its members.

The same problem arises where an address is common to two entities, e.g. a meeting, which may be the responsibility of a particular organization although overall responsibility may be shared between several independent organizations. Ifthe address *is* included twice (or more, where there are many such meetings), updating problems are created when the address changes. If it is included once, using the identification number as the cross-link from the other, then some sorting will be required. This alternative is considered preferable.

Indexes

Indexes would be prepared either directly from the tape of the updated file, usually with some sort of selection profile in the form of card input, or else from scratch tapes arising from a previous selection to tape.

Indexes would be based on a single language name or on the keyword coding in the class link cards. If English, these two could be merged. Indexes could be extended by inputing word equivalents of the numerical keycoding in the class link cards. There is no reason why this should not be possible in languages other than English, if necessary.

There is a possibility that it might be useful to format the indexes to facilitate computer typesetting.

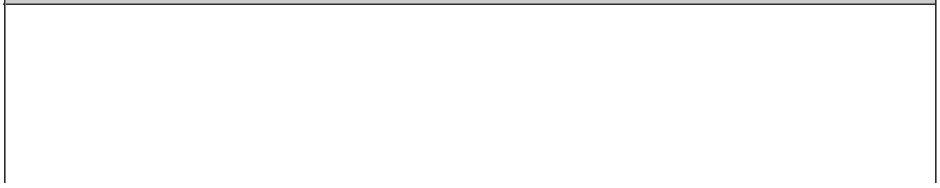
Graphic displays

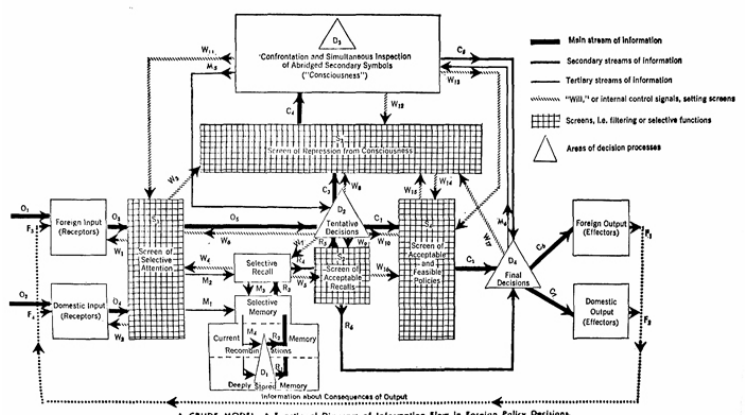
Eon-interactive displays are mainly aquestion of sorting and formating and should pose no problems.

Interactive displays would require that the entities of interest first be selected onto *a* direct access device. From this they could then be selected into memory as required, thus avoiding a sequential pass through the whole file. This implies that the identification numbers should be suitable, or at least adaptable by program, for direct access purposes. The cross-linking then becomes a means of avoiding sorts.

One unfortunate feature of depending upon the inter-entity link coding arises because fields of interest were not treated as entities. This lack of generality means that all linked entities can be traced without a sort and displayed, but entities with the same field of interest but out of contact can only be detected after a pass through the file selected to the direct access device. If the file is small this does not pose a problem, however, (It would be interesting to know for a time-shared interactive device whether it is possible to undertake two-types of processing simultaneously, one requiring a longer time than the other. In which case, a relatively lengthy sort could be requested from the same terminal from which interactive processing was being undertaken. The first program would then have to interrupt the second on completion of the sort or file scan.)

Crude model: Functional Diagram of Information Flow In Foreign Policy Decisions.
(Extract from Deutsch, K. Nerves of Government, 1963, Appendix)





MAIN INFORMATION FLOW (heavy arrows)

I. Current Information from outside the decision system

01 Current general information about foreign countries (part of external in- telligence)

O2 Current general information about domestic politics (part of internal in- telligence)

O3 Current foreign information, as selected by receptors O4 Current domestic information, as selected by receptors O5 Current foreign and domestic information, screened and combined

II. Past Information, recalled from storage within system

R1 Information recalled and recombined from deep memory

R2: Information recalled and recombined from recent or current memory

R3 Combined information from memory

R4 Combined information from memory, as selectively recalled

R5 Recalled information from memory, screened for acceptability in terms of culture, values, personalities, cognitive dissonance, etc., and transmitted to area of preliminary decision R6 Acceptable memories, transmitted to area of final decision

III. Combined Information, of memories and outside data

C1 Combined selected data and acceptable memories, moving toward final de- cision (e.g., 'action papers')

C2 Combined selected data and memories, as screened further for feasibility and acceptability as policies

C3 Abridged combined data, transmitted toward area of confrontation and simultaneous inspection

C4 Abridged combined data, screened for acceptability to consciousness

C5 Abridged data and memories, selected and combined at the level of con- scious confrontation, and transmitted to area of final decision

C6 Final policies selected and transmitted to effectors in foreign policy area

C7 Final policies selected and transmitted to effectors in domestic policy area Note: Policies need not always be consistent between C4 and C5, nor within C4 or *C5*, respectively. Thus the United States Congress mayvote a foreign policy resolution demanding greater anti-Communist efforts in the Western Hemisphere, and at the same time cut economic aid funds for Latin American countries; or the West German government could call upon Britain to aid in the defense of West Berlin while at the same time threatening British Trade with exclusion from the European Common Market.

Such inconsistencies might show up in advance in the recombinations and symbolic projection of information at the level of abridged simultane- ous inspection and 'consciousness'; or else they might be reported back only later in the feedback of information about the results of the first in- consistent actions taken under these policies in the outside world, but still early enough to permit correction of these policies at later stages.

IV. Feedback Information about the consequences of the actions of the system on its relations to the world outside it

F1 Feedback information about the results of foreign policy actions

F2 Feedback information about the results of domestic policy actions

F3 Feedback information gathered by foreign area receptors

F4 Feedback information gathered by domestic area receptors

V. The 'Will' System

MAIN SCREENS

S1 Screen of selective attention to current information

S2 Screen of acceptable recalls from memory

S3 Screen of acceptable summary information for confrontation and simul- taneous inspection ('consciousness')

S4 Screen of acceptable and feasible policies

MAIN INFORMATION FLOWS, ADJUSTING SCREENS

W1 Information which sets attention focus or 'tracking' pattern for foreign area receptors

W2 Information which sets attention focus or 'tracking' pattern for domes- tic area receptors

W3 Outside information, changing the screen of acceptability to conscious- ness

W4 Recalled information, changing screen of attention

W5 Selectively recalled information, changing screen of subsequent accepta- ble recalls

W6 Information about tentative decision, changing screen of attention (e.g. 'self-confirming policy')

W7 Information about tentative decision, changing search pattern for selec- tion of interesting recalls from memory (e.g., 'search for precedents ')

W8 Information about tentative decision, changing screen of acceptability to consciousness

W9 Information about tentative decision, changing screen of acceptable re- calls .

W10 Information about tentative decision, changing screen of acceptable and feasible policies

W11 information about results of simultaneous confrontation and inspection ('consciousness'), changing the screen of attention to outside informa- tion

W12 Information about results of simultaneous confrontation and inspection ('consciousness'), changing screen of acceptability to consciousness

\V13 Information about results of simultaneous confrontation and inspection ('consciousness'), changing screen of acceptable and feasible policies

W14 Information about results of simultaneous confrontation and inspection ('consciousness'), via screen of repression from consciousness, to screen of acceptable and feasible policies ('unthinkable')

W15 Information about feasibility and acceptability of policies, changing screen of acceptability to consciousness

W16 Acceptable recalled information, changing screen of acceptable and feasible policies

w17 Information about final decision, changing screen of repression from consciousness

VI. Minor or Secondary Information Flows

M1 Selected outside information, transmitted to memory for storage and pos- sible recall. This is a minor flow only as regards the making of immediate decisions. Its actual volume of information may be large

M2 Selected outside information, changing probabilities of recall ('that re- minds me . . .')

M3 Orders for recall, to memory

M4 Orders, or associative trails, or chain reactions, within memory

M5 Information about results of simultaneous confrontation and inspection ('consciousness'), transmitted to area of tentative decision

M6 Abridged information about final decision, which is being fed back to the area of simultaneous confrontation and inspection

VII. Consciousness

Feedback cycle C5-M6, on repeated run-throughs, would make the final deciion 'conscious'

VIII. Areas of Decision

D1 The area of dissociative and combinatorial memory is an implicit area of decision, since the forming of certain combinations, and the omission of others, functions indirectly as a series of partial decisions. Such combina- tions include not only data but also their patterns of configuration; they also include images and values

D2 Area of preliminary decision, where combinations between memory data and current intake function as explicit preliminary decisions

D3 The area of simultaneous confrontation and inspection, which functions indirectly as a decision area, since certain combinations between the simul- taneously presented data are formed, while other possible combinations are not, and the successful combinations have the effect of partial decisions

D4 The area of explicit final decision-which may, however, already have been prejudiced in its outcomes by the events at the earlier decision areas, D1-D3