The Egmont Group
of Financial Intelligence Units

Egmont Strategic Analysis Course

Session 7

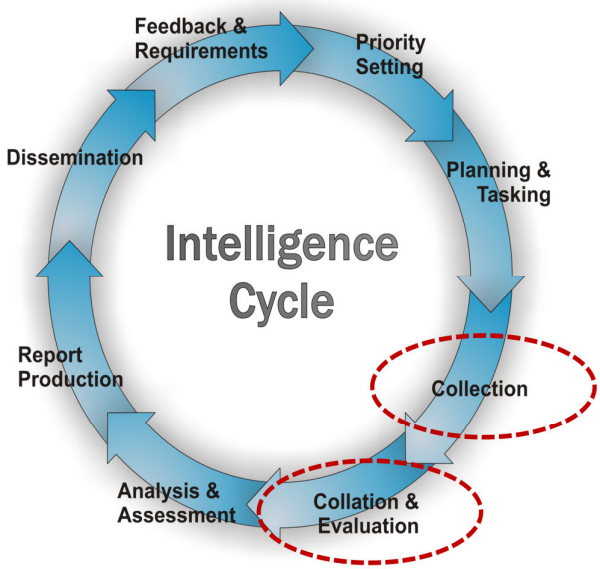
Collection, collation and evaluation

Participant Manual

Version 30 October 2012

1. Introduction

This session discusses sources of information the collection and collation of Information as well as a look at data diagnostics as a way of evaluating the information that has been collected to undertake the analysis.



2. Information

Information is interpreted data that is specific and organized for a purpose. It is presented within a context that gives it meaning and relevance, and that can lead to an increase in understanding.

Graded and processed information forms the basis of intelligence.

Simply stated, information is a message received and understood.

For example:

Currency transaction reports contain data, but this data can be used to produce **information** about the use of cash in certain regions.

a) Types of information

There are two different types of information that we may use in analysis.

- Quantitative information
- Qualitative information

Quantitative information

Quantitative information focuses on numbers and frequencies rather than on meaning and experience.

Quantitative methods provide information which is easy to analyse statistically and is fairly reliable. It can be displayed using tables, charts, histograms and graphs.

Quantitative methods are associated with a scientific or experimental approach and are often criticised for not being descriptive enough.

For example:
Significant rise or fall in the amount of STR's, money flows, certain types of reports, reports made by a specific reporting sector etc. If a rise or fall is significant, it tells us that this rise or fall is not caused accidentally.

Qualitative information

Qualitative information provides a description that can be observed but not measured. It provides meaning rather than a statistical inference. Qualitative information may not be as reliable but it is often more validity as it provides a more in-depth and rich description.

For example:
Information gained through interviewing experts or through questionnaires.

You will find that in many cases both types of data are collected, because often they prove to be complementary to each other.

b) Sources of information

Open and closed

Information can be obtain formation through either open or closed sources.

- **Open sources** of information are accessible to the general public such as newspaper, websites etc.
- **Closed sources** are only accessible to those with authority to access the information, such as police records, law enforcement data, FIU information

Primary and secondary

Sources of information can also be categorized into two types.

- **Primary** sources are original materials. They provide direct evidence and have not been filtered through interpretation or evaluation. Primary sources are original materials on which other research is based

  eg STR database, original documents, records of organizations, surveys.

- **Secondary** sources are less easily defined but are generally accounts written after the fact with the benefit of hindsight.

  They are interpretations and evaluations of primary sources. Secondary sources are not facts, but rather commentary and a discussion of facts.

e.g., Tactical and Operational Analysis on SAR's/CTR's, Bibliographical works, textbooks, websites, dictionaries, journal articles

Where possible it is always advisable to use primary sources of information. Often secondary sources will cite the primary sources from where they extracted their information. As an analyst you should go to the primary source so that you can make your own interpretations and not rely on the opinion of another person.

c) Lists of sources

There are a number of sources where an analyst will find both quantitative and qualitative information whether it be a primary or secondary source.

i. FIU database information

The FIU database is one of the main sources of information available to an analyst. Suspicious Transaction Reports and Currency Transaction Reports may give information about transaction trends and patterns by:

- Type of client
- Type of product or service
- Zones
- Reason for suspicion
- Type of business

This information can be used to answer the analytical questions. Moreover patterns and trends observed in reports could be the starting point for a strategic analysis product.

ii. Government agencies

We can get information from Governmental agencies, such as:

- Central Bureau of Statistics,
- Tax authorities,
- Supervisors,
- Customs
- Law enforcement agencies

iii. Foreign information

You may also need to seek information regarding another jurisdiction.

- Foreign liaison officers
- Other FIU's
- Egmont Group, FATF
- International Organisations such as IMF and World Bank

iv). FIU Operational and Tactical Areas

FIU operational and tactical colleagues can also provide you with information regarding operational or tactical cases. They can also provide feedback from strategic analysis products that have been previously delivered.

You can ask your FIU Operational and Tactical Areas for things like prosecution statistics, for intelligence reporting, sanitized cases, typologies, etc. This material gives you information that can be collated with other data/information and integrated into your strategic analysis report.

v. Commercial databases

There are also commercial databases that can help you to get information about persons or companies.

World-Check's research profiles individuals and entities deemed worthy of enhanced scrutiny, as well as occurrences of fraud, corruption, terrorism, money laundering, narcotics trafficking and many other types of financial crime.

LexisNexis provides access to documents and records from legal, news and business sources.

Dun & Bradstreet is a provider of business information.

vi. Private industry

Private industry can give you detailed information about particular technologies and provide market information, industry standards and business practices. It may also offer insights about potential money laundering or terrorism financing risks within the industry.

Examples of the private sector include –

- Trade associations
- Reporting institutions
- Private surveys

viii. Academia

Some academic articles and research methods gives you information about tools and methodologies to perform better strategic analysis products.

Some examples of useful journals are

- Journal of Financial Crime. Emerald Group
- Journal of Money Laundering Control. Emerald Group
- Journal of International Money and Finance. Elsevier

Academia is also known for establishing or refining various research methods and for engaging in extensive studies on topics such as crime. These can be extremely useful for analysts

d) Using the internet

There is a wealth of information available on the Internet, but as you can imagine, there are problems associated with gathering data from the web. Since anyone can post anything on the Internet there is no guarantee that

the information is accurate; it is up to the analyst to determine the validity and reliability of the source and data.

When conducting Internet searches, look for sites that have databases of information. A database will provide a more pointed and targeted search for information.

It is a good practice to save a copy of the web pages as you find them because the site may change when you return to the site at a later date. Similarly, always records your references and keep a list of sites that are useful. Set a time limit for searching on the Internet. It is easy to get sidetracked on the Internet if you do not follow a plan.

Keep in mind that people can see who accesses their websites. Most FIUs should have a policy about accessing websites to ensure that searches are not easily linked back to the organization.

3. Collecting information

When considering you information requirements it is important to ensure that you only collect data/information that will assist you to answer the key questions you've outlined for your project.

Within this parameter, information collection should be done as thoroughly and as broadly as possible, keeping in mind the timeframe/complexity of the task.

When collecting information it is important to-

- Create a list of sources
- Discuss ways to extract information from these sources
- Look into different types of information
- Create Information Collection Plan (ICP)

a) Information collection Plan (ICP)

An Information Collection Plan helps analysts to determine what kind of information is needed, and how that information can be collected and located.

In preparing an Information Collection Plan, you create a list of sources, discuss ways to extract information from these sources, and look into different types of information.

You will find a sample of an Information Collection Plan at the end of this session

b) Collection considerations

In the collection process you must also consider:

Storage

You will need different storage mechanisms depending on how you receive the information, eg is it electronic, in the form of a database, paper based. You may require different systems for different forms of information.

Retrieval

Information must be stored in a manner that allows for quick and easy access. A robust search facility will assist this.

Security

Security of information must be maintained at all times.

Individual logins for electronic (permissions and access based) key coded or locked filing cabinets

Confidentiality
You may also need to consider the need for the protection of the source of your information.

4. Collation of information

Once you have collected all of the available information, it is important to collate it in a way that will facilitate your analysis.

The questions you need to answer will shape the way you group your information. For example, you can assemble your data and information against the analytical questions.

You can also use summaries, sorting mechanisms, or groupings to make data/information more manageable. For instance:

- Summarise data from the same source over a period of time.
- Summarise data from the same data elements but from different sources.

While in many cases information can be managed using simple methods, working with large quantities of data poses additional challenges that may require advanced solutions.

Software

There are a number of software programs that will assist the analyst to store and sort large quantities of information from a variety of sources, in a range of formats.

The right software should also have the ability to quickly retrieve data over a given range and provide an appropriate level of security.

Different types of software include -

- Document management systems
- Access
- Cognos *(business intelligence software)*
- Business Intelligence software *(many others available on the internet)*
- Auto report generating for a given set of rules
- Statistical Analysis software

  Statistical analysis software enables the user to create descriptive statistics.

  Descriptive statistics can provide a summary and analysis of both qualitative and quantitative data, using calculations such as frequency, percent change, mean, median, mode, and measures of variance (SD and SE).

  Inferential statistics are used for random sampling of populations (as in a survey) and allow the user to inferentially apply the results to the population from which the sample was drawn.

Spreadsheets

Simple spreadsheets such as Excel allow you to manipulate data to facilitate the analysis process through the use of –

- Formulas function
- Lookup function
  You can use the **VLOOKUP** function to search the first column of a range of cells
- Sort function for chronological listings
- hyperlinks to different types of electronic information (such as other spreadsheets, word documents, PDFs)

Pivot Tables

Among other functions, pivot-table tools can automatically sort, count, and total the data stored in one table or spreadsheet and create a second table (called a "pivot table") displaying the summarized data.

Pivot tables are also useful for quickly creating cross tabs. The user sets up and changes the summary's structure by dragging and dropping fields graphically.

Geographic Depictions

- Google earth
- Google maps
- Heat maps
- Mapping/GIS software can aid in mapping and in the analysis of data points or areas related to crime or other patterns (dates, times, hot spots of activity, buffer zones). This software can also be used to look at patterns of movement relating to crime to delineate "hunting zones" and reduce the number of leads in predatory crimes

Charting

Charting tools such as i2 can use the analytical capability to identify key information.

Data grouping and sorting

Often during Strategic Analysis we need to organize information that comes from different sources by subject/person/company/geographical.

Standardise information

To do this, first you have to standardize the information so that the structure is the same for each category.

For example, you can generate a dictionary with equivalent text lines or eliminate abbreviations, double spaces, periods.

Sort data into categories

The data must then be sorted into categories.

A series of tools for organizing data, may include some of the following:

Data profiling - initially assessing the data to understand its quality challenges

- Data standardization - a business rule engine that ensures that data conforms to quality rules
- Geocoding - for name and address data.

Matching or Linking

Marching or linking is a way to compare data from slightly different records. Matching may use reasonable logic to find duplicates in the data, for example, 'Bob' and 'Robert' may be the same individual.
It might also be able to find links, such as, between husband and wife at the same address,

5. Evaluation of data

During the collection and collation stages it is important to continually evaluate your information.

All information is not created equal and its quality will depend on various factors including:

- Source (reliability of the provider)
- Relevance (information must be relevant to the goal of the analysis)
- Accuracy (if your information is not accurate, your analysis will not be reliable)
- Timeliness (information must not be too old as in not relevant to the current situation)

Source

When evaluating the source of information consideration should be given to-

- What agency published the report and what is the mandate/political leaning of that organisation?
- Has this agency produced useful material in the past?
- Who wrote the report, and what role do they fill in the organisation?
- Who controls the source of the information?

Relevance

With regard to the relevance of the information, ask yourself -

- Does the data relate to the main question?
- Does this relate to associated issues to the main question?
- Does this make sense with what we know?

Evaluating information

When evaluating information, you should consider the types of data or information you may have.

There are three types of information we will consider.

Fact

- A fact is something that is known to have happened or existed. It is a truth that can be verified from experience or observation.

Eg Data on reports in FIU database

Opinions

- An opinion is a judgement or belief not founded on a certainty, although it may be based on some factual evidence.

  Eg "24-year old men who buy expensive cars must be criminals"

Allegations

- An allegation is a story which is declared as being true but which has not yet been proven.

Consideration should also be given to the possibility of incorrect data entry.

Incorrectly spelled or punctuated information, information entered into the wrong field or duplicated data can be an obstacle for strategic analysis research.

Scales of Reliability and Validity

The Admiralty or NATO System is a method for evaluating information. The system comprises an alpha numerical code used to evaluate the reliability of the source of information and the accuracy of information.

The scales of reliability and validity ensure that you evaluate the source of information and the accuracy of the information independently.

The appraisal of one should not influence the appraisal of the other. It is extremely important to separate reliability of source and accuracy of information.

Reliability

Reliability of information sources is judged primarily from previous experience. It represents an estimate of how often reports from a given source are accurate.

When determining the accuracy of information you will want to consider the consistency. Is the source or date contradictory? Is the source consistent over time? Is the source's information consistent with previously known facts of the case?

For example, imagine stepping on your bathroom scale and weighing 140 pounds only to find that your weight on the same scale is 180 pounds an hour later and 100 pounds an hour after that. Based on the inconsistency of this scale, any research relying on it would certainly be unreliable.

Validity

Validity refers to the content of a report and is not a measure of whether the information was reported accurately or reliably.

It is an indication of how well an assessment actually measures what it is supposed to measure. Information is valid when it is free of 'systemic errors' meaning errors which influence a result consistently in the same direction eg, a watch which is always ahead one minute per 24 hours.

Scale of Reliability	Scale of Validity
A - Completely Reliable	**1** - Confirmed
B - Usually Reliable	**2** - Probably True
C - Fairly Reliable	**3** - Possibly True
D - Not Usually Reliable	**4** - Doubtfully True
E - Unreliable	**5** - Improbable Report
F - No Judgment	**6** - No Judgment

6. Decomposition and visualization

As Heuer and Pherson state the issue, "The limitations of human thought make it difficult, if not impossible, to do error-free analysis without the support of some external representation of the parts of the problem that is being addressed."

Two common approaches for coping with this limitation of our working memory are:

- **decomposition** – breaking down the problem, issue or data into their component parts, so that each part can be considered separately, and
- **visualization** – placing all the parts on paper or a computer screen in some organized, usually graphical, manner designed to facilitate understanding of how the various parts interrelate.

All structured analytic techniques, by definition, are directed at externalizing one's thinking and employ these approaches, in one way or another. Some of the more basic SATs are directed only at performing that function.

a) Data decomposition

So, decomposition entails breaking down the data into their component parts, so that each can be considered separately.

If one part of a data set is indistinguishable from another, then there is little to say about anything but the whole of the set – no specific signals can be identified, no trends can be analysed; in fact, it is essentially impossible to conduct any analysis, at all!

What if the amount of natural light, the air temperature and pressure, humidity and other atmospheric aspects were uniform – all the same around the world, all of the time? What we call "weather" would not exist and there would be no need or basis for weather forecasting. Of course, that is not the case and those who try to understand and forecast the weather have far more than enough possible information on which to conduct their analyses.

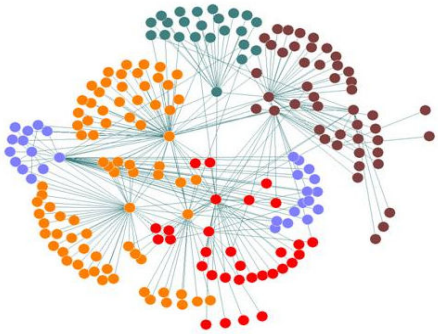
Examples of data decomposition include -

- Chronologies and timelines – organize data on events or actions in relation to their relative timing or sequence
- Sorting by categories or sub-categories – by subset groups with the same or similar characteristics
- Ranking, scoring and prioritizing – on such measures as relevance, importance, immediacy or impact
- Matrices – to identify the relationships between two or more sets of variables, or within a single set

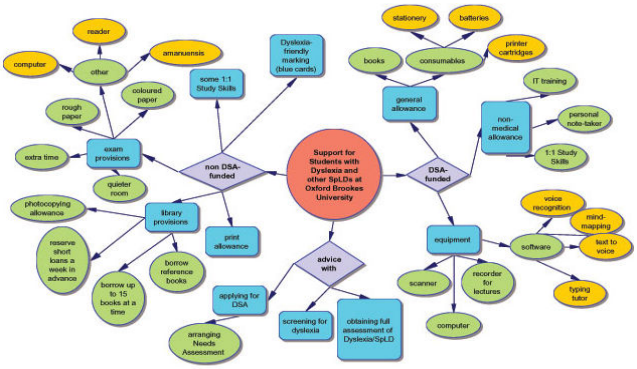
b) Data visualization

Network, association, link analysis

Network diagram, showing clustering and relationships.



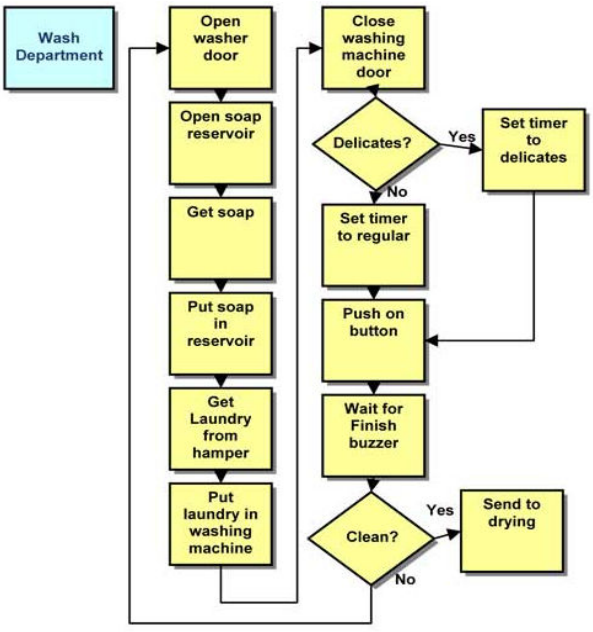
Mind maps and concept maps



Here is a simple, but powerful mind map, using shapes and colours to distinguish levels of ideas within each branch, but still allowing growth to capture new ideas.

Process maps and Gantt charts

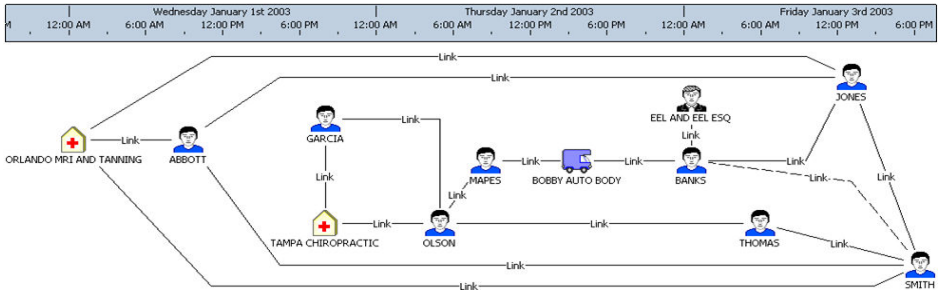
Here is a process map showing the actions and decisions in the first stage of a laundry process.



Link chart example

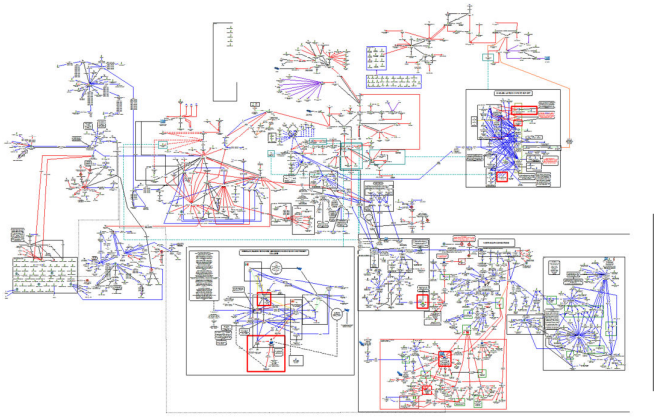
This is an example of a fairly simple link chart, showing individuals and businesses, with a number of undefined links between them. These links could show simple relationships among the individuals and between them and the businesses. These could range from such things as beneficial ownership to where they have gone for certain services. With a law office shown as one of the businesses, the links could relate to some legal process.

A time shows that the links must be some sort of events and in some way sequential over time, for example, Abbot's link with Orlando MRI and Tanning occurred before that of Garcia with Tampa Orthopractic, and so on.



Financial network linkchart

This slide demonstrates the use of financial network analysis of to identify and link a complex arrangement of key nodes in financial activity and to help us understand the key roles that some individuals and businesses may be playing, and their interrelationships in those financial activities…as well as identifying ML/TF patterns.



It is based on a large amount of data and almost any part of it, as identified by the various boxes, would provide for interesting analysis. However, it certainly focuses our attention and helps to set priorities for further work.

Activity 7.1 – Link chart

Your task is to construct a link chart using the Purpleana information.

Use as much of the relevant information as you can, but at a high (strategic) level. Show major groupings, key individuals and their interrelationships

Information Collection Plan

Project Name:
Prepaid Cards

Collection Objective (Questions)	Insight to be gained	Information Type	Information Elements	Information Source	Collection Technique	Outcome
What is the rationale for this project?	Current awareness of the link (potential and existing) between ML/TF and the use of prepaid cards	Reports/Papers	Existing Typologies, previous studies	Internet	Internet Search/Save documents by source	Summary of studies and their conclusions as to how ML/TF are/could be linked with the use of prepaid cards
	Identification of the gaps in the current understanding of the links between ML/TF and the use of prepaid cards	Reports/Papers	Unknown elements (references to what we don't know, what still needs to be done, etc)	Internet	Internet Search	Summary of the gaps in information as a rationale for this project
What do we know about the topic?	How prepaid cards work	Reports/Papers	Suppliers, types of prepaid cards, market size	Internet	Internet Search	List of key components
	Current risks related with the use of prepaid cards	Intelligence Reports	Summaries/Transaction Indicators	Operational/Tactical Department	Document search (confidential)	List of risk indicators identified from Intelligence Reports
		Commentary	Summaries/Overviews	Key Contacts (other FIU's, law enforcement agencies)	Informal Discussion	List of risk indicators identified from key contacts
		Statistics	Volume of STR's related with the use of prepaid cards and time series changes across zones and institutions	Internet	Internet Search	Patterns and Trends identified from Reports

Information Collection Plan

Project Name: _____

Collection Objectives (Questions)	Insight to be gained	Information type	Information Elements	Information Source	Collection Technique	Outcome
What is the rationale for this project?						
What do we know about the topic						

Information Collection Plan

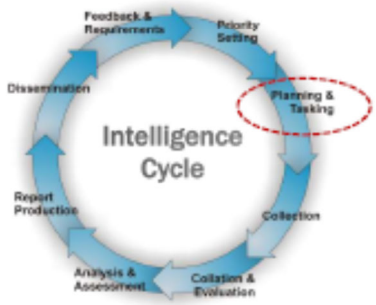
Project Name: _____

Collection Objectives (Questions)	Insight to be gained	Information type	Information Elements	Information Source	Collection Technique	Outcome
What is the rationale for this project?						
What do we know about the topic						

Strategic Analysis Course
Session 7
Collection and Evaluation

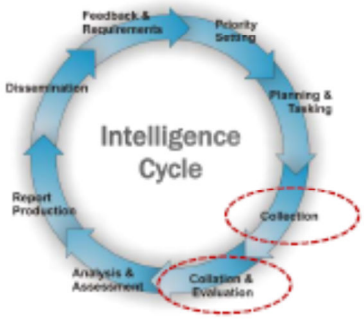
1

The Intelligence Cycle



2

The Intelligence Cycle



3

What is Information?

Information is interpreted data.

When presented within a context is gives it greater meaning.

→ *Graded and processed information forms the basis of* ***Intelligence***

4

Types of information

Quantitative	Qualitative
Numbers and frequencies	Meaning and experience
Easy to analyse statistically	Difficult to analyse statistically
Reliabile	Valid
Scientific	'Soft' but in-depth
Quantity	*Quality*

5

Types of sources

Open/closed

Open sources are accessible to the general public

Closed sources are only accessible to those with authority

Primary/secondary

Primary sources are firsthand, 'direct from the source'

Secondary sources are analyses from primary sources

6

List of sources

1. SAR's/STR's/CTR's
2. Governmental agencies
3. Foreign information
4. FIU Operational and Tactical Areas
5. Commercial databases
6. Private industry
7. Academia

7

FIU database information

Suspicious Transaction Reports (STRs) /Cash Transaction Reports (CTRs)

- ➡ Transaction trends/tatterns
- ➡ Use of cash (zones, business)

8

Governmental agencies

- Central Bureau of Statistics
- Tax authorities
- Regulatory /Supervisory bodies
- Customs
- Law enforcement agencies

9

Foreign information

- Foreign liaison officers
- Other FIU's
- Global bodies ➡ Egmont Group
  - ➡ FATF
  - ➡ IMF
  - ➡ World Bank
  - ➡ UN
  - ➡ Interpol

10

FIU Operational and Tactical Areas

- Information of operational/tactical cases ➡ Prosecution statisitcs
  - ➡ Intelligence reporting
- Feedback from delivered Strategic Analysis products ➡ Sanitised cases
  - ➡ Typologies

11

Commercial databases

- Worldcheck ➡ Research profiles
- Lexis Nexis ➡ Documents and record
- Dun & Bradstreet ➡ Business information
- Local PEP's list ➡ Potential individuals of interest
- World compliance

12

Private industry / sector

- Trade associations
- Reporting institutions
  - ➡ Information on new payment methods from commercial banks
  - ➡ Information on market trends and fluctuations from money remitters
- Private surveys

13

Academia

- Journals/Articles
  Journal of Financial Crime. Emerald Group
  Journal of Money Laundering Control, Emerald Group
  Journal of International Money and Finance. Elsevier
- Research Methods
- Crime studies

14

Using the Internet

Search engines

- Plan and track your searches
- Write down key words and ideas before you start
- Save relevant pages in Adobe Acrobat, notepad or some other software program
- Maintain list of internet sites and databases
- Set a time limit for searching and use site maps on web sites

Risks of using the internet?

- Stay organized
- Security

15

Collection of information

- Create a list of sources
- Discuss ways to extract information from these sources
- Look into different types of information
- Create Information Collection Plan (ICP)

16

Information Collection Plan (ICP)

- Brief description of the project/analysis
- Specific data needed
- What insight the data may provide
- Purpose of data once collected
- Matrix with specifics on the data, such as
  - ▶ Source
  - ▶ Type
  - ▶ Date sent/received
  - ▶ Collection technique
  - ▶ Collation number

17

Example of ICP

Collection of information

Take into consideration:

- Storage
- Retrieval
- Security
- Confidentiality

19

Collation of Information

- Assemble the information against the analytical questions
- Design a method to manage the collected data/information
- Group and sort your data

20

Working with Large Quantities of Data

- Binary 0's and 1's
- Telephone conversation
- Spreadsheet
- Texting
- Maps
- Good old paperwork

21

Working with Large Quantities of Data

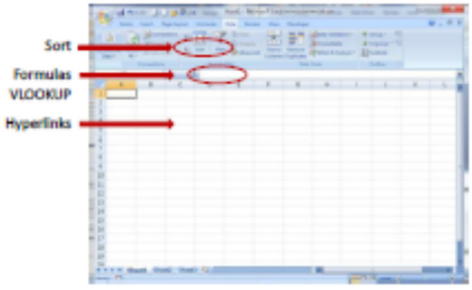
- Software
  - Document management systems
  - Access
  - Cognos (business intelligence)
  - Business Intelligence (BI) software
  - Auto report generating for a given set of rules
  - Statistical Analysis software

Cognos TM1

22

Working with Large Quantities of Data

- Spreadsheets



23

Working with Large Quantities of Data

- Pivot Tables
- Geographic Depictions
  - Google Earth/Maps
  - GIS software - ArcGIS
- Charting
  - i2

24

Data grouping and sorting

1. Standardize information:
   - Generate a dictionary with equivalent text lines: E.g. Sánchez/Sanchez, Jon/John
   - Remove abbreviations, double spaces, periods
2. Sort data into categories
3. Matching and Linking

25

Evaluation

Evaluate quality of information by:

- Source (reliability)
- Relevance (to goal of analysis)
- Accuracy (if not, analysis will not be reliable)
- Timeliness (relevance to current situation)

26

Evaluation

Source

- Nature of the agency/author/entity
- Motivation of the source

Relevance

- Does the data relate to the main question?
- Does this relate to associated issues to the main question?
- Does this make sense with what we know?

27

Evaluation

- **Fact**
  Data on reports in the FIU database
- **Opinion**
  ...."24-year old men who buy expensive cars must be criminals"
- **Allegation**
  Police allege that Mr X is a drug trafficker.
- **Corroboration**

28

Evaluating data

Reliability of information

Reliability of information sources is judged primarily from previous experience. It represents an estimate of how often reports from a given source are accurate.

A Completely Reliable
B Usually Reliable
C Fairly Reliable
D Not Usually Reliable
E Unreliable
D No Judgment

29

Evaluating data

Validity of information

An indication of how well an asessment actually measures what it is supposed to measure. Information is valid when it is free of 'systemic errors'.

1. Confirmed
2. Probably True
3. Possibly True
4. Doubtfully True
5. Improbable Report
6. No Judgment

30

Decomposition and visualization

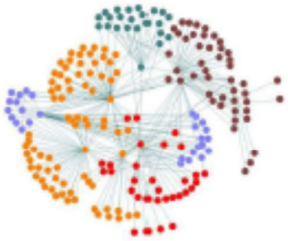
- Limitations of human thought affects analysis
- External representation of the problem required
- Two common approaches:
  - Decomposition – breaking down the problem, issue or data
  - Visualization – organizing the parts visually
- All structured analytic techniques employ such approaches, in some way

Data decomposition

- Chronologies and timelines – organize data on events or actions in relation to their relative timing or sequence
- Sorting by categories or sub-categories – by subset groups with the same or similar characteristics
- Ranking, scoring and prioritizing – on such measures as relevance, importance, immediacy or impact
- Matrices – to identify the relationships between two or more sets of variables, or within a single set

Data visualization

- Network, association, link analysis



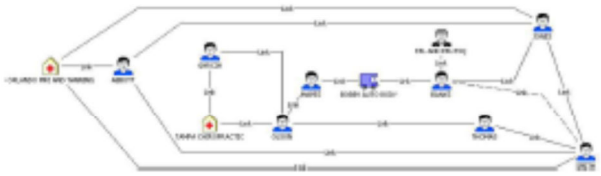
Data visualization

- Mind maps and concept maps

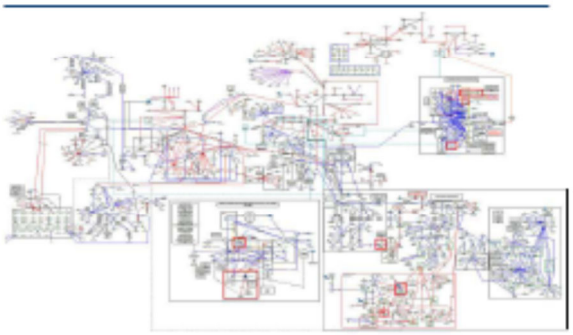
Data visualization

- Process maps and Gantt charts

Link chart example



Financial network linkchart



Activity

7.1 Link Chart

Questions

?