

Making the Auditor-General's Report More Accessible, Integrated and Usable



Mr. Aw Kong Koy

Mr. Aw Kong Koy, graduated with 1st Class Honours in mechanical engineering from Universiti Malaya in 1973. He was a former council member of IEM and the Founder Chairman of the IT Special Interest Group. He has been a full time programmer since 1988, with an interest in applying ICT to address societal problems.

This article describes the motivation, challenges and implementations in the development of the project “AG’s Report Information Graph” (Information Graph) to make the Malaysian Auditor-General’s report more accessible, integrated and usable. This project can be framed as a problem in managing strategic information, an important component of knowledge management systems.

The objective of the Auditor-General’s office is to strengthen the accountability, transparency and integrity of government and public sector entities. It also has to demonstrate ongoing relevance to citizens, Parliament and other stakeholders. Even though the audit reports contain valuable information that is not available elsewhere, the difficulty faced in using the information often results in the information being used on a piecemeal basis, under-utilised or even ignored. This is a shame as so much effort and resources went into auditing the public sector entities and producing the reports. At the same time, many issues and leakages are reported repeatedly.

The government machinery is huge and intricate and, as such, the audit reports have to cover a wide range of activities, performances and issues. It is a challenge just to understand the structure and functions of the government machinery and to develop a strategy to organise the disparate information.

It is difficult to expect people to read the annual AG’s reports from start to end as it is just too boring. People need to be able to

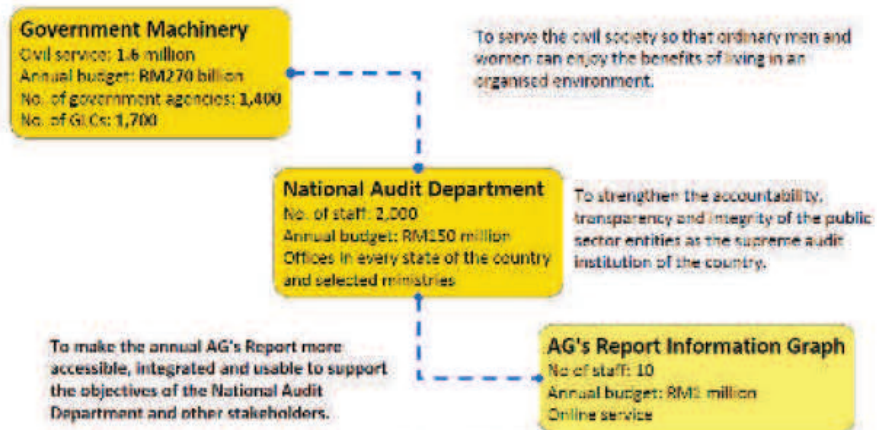
explore the reports to look at specific issues or subjects. They want the information to be arranged to suit their purposes. The static arrangement provided by documents as in the case of the AG’s reports will need to be supplemented by something more dynamic, a system that will allow the information to be rearranged in many ways, even interactively. This is to allow the users to discover trends and relationships. Our solution is to use the concept of the information graph to provide the required flexibility and adaptability.

Graph computing is a relatively new computing concept. Internet giants like Facebook, Google, LinkedIn and Amazon are all big users of graph databases. We are fortunate to have in-house graph computing engine which we started developing in 1995.

THE AG’S REPORT INFORMATION GRAPH IN CONTEXT

The Information Graph is not intended to duplicate the work of the National Audit

The Information Graph in Context

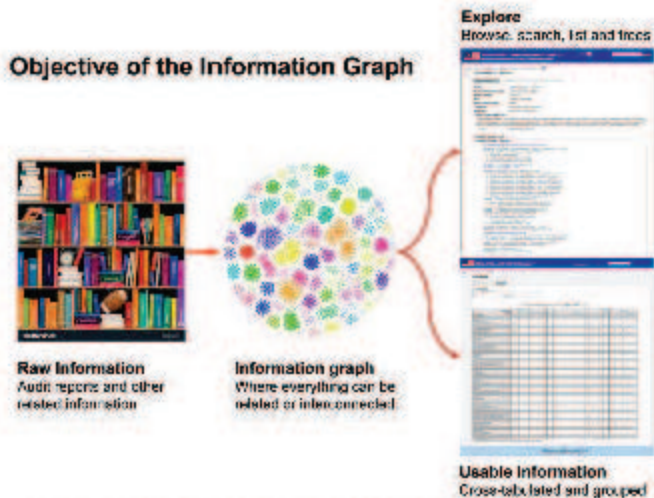


(All figures are approximate).

Department (NAD) but to complement it. Development of the Information Graph requires abstract thinking and appropriate technology as precedents are hard to find. The development will be interactive and opportunistic, focusing on the value added rather than the predefined system requirements.

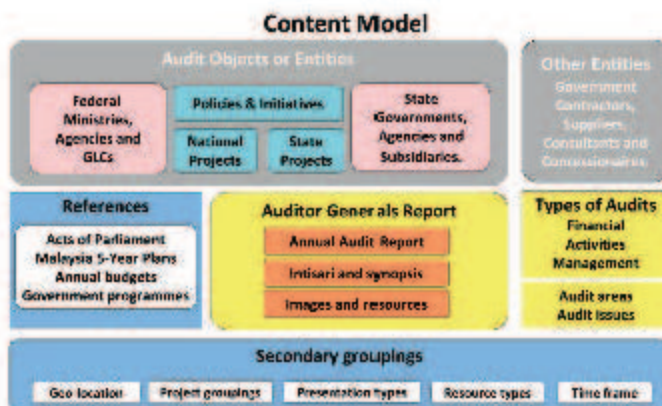
DIAGRAMMATIC REPRESENTATION OF OBJECTIVE

A simplified diagrammatic representation of the objective of the Information Graph is shown in the figure below.



Diagrammatic representation of the objective of the Information Graph

CONTENT MODEL



Content Model of the Information Graph

The content model for the Information Graph provides an overview of the information to be included in the Information Graph. This information will need to be acquired from various sources.

The annual Auditor-General's reports will need to be updated when new reports are released. They have to be decomposed and preprocessed before incorporation. The audit objects or entities are the public entities that the Auditor-General has the mandate to audit under the Federal Constitution and the Audit Act. They form the critical background information of the government machinery. Government contractors, suppliers and

concessionaires do not come under this mandate, but they are included for completeness.

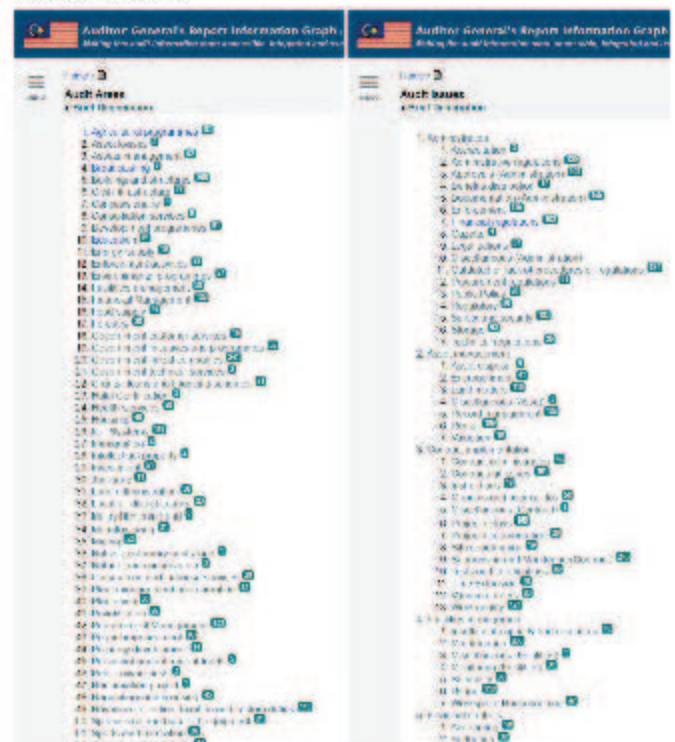
The reference information such as the Acts of Parliament, the 5-year plans, government programmes and annual budgets are important as the government machinery are required to comply with these acts and policies and to implement or enforce them as appropriate.

To provide another level of integration, a common thread that runs across all the audit information is required. We have identified two such threads: Audit areas and audit issues.

AUDIT AREAS AND AUDIT ISSUES

The audit area is a list of related business sectors or activities. Each audit may be related to one or more audit areas. By associating the audits to the audit areas, we can group these audits for comparison and analysis.

A wide range of issues is raised by the auditors in their reports. This provides another opportunity to bind the audit information together. You can see how these two lists are used to good effect in the usage scenarios that are described later.



Audit Areas and Audit Issues

Ideally, these two lists should be provided by the NAD, but as there are no published lists available, we have to construct them from information in the annual reports. The lists should be considered work-in-progress, and they will be revised and fine-tuned as more reports are processed. Ambiguity in the terminology used must also be removed, while conforming to the industry usage.

Manually matching the text of audit information against these two lists is very laborious and error prone. Researchers may have the motivation to do this, but we cannot expect this from content management staff. It is difficult

for humans to remember the two long lists and the definitions of the items to perform the match consistently. This is where a computer-aided facility can be helpful.

Technically, the items in the two lists are "named entities." For our purposes, we wrote a simplified named entity recognition utility using "match phrases with wildcards." It is a semi-automated process where the final decision has to be made by a person after viewing the match phrases in context. This utility needs to be trained for it to perform effectively.

STATUS OF THE INFORMATION GRAPH

A beta version of the Information Graph has been online since 1 November 2015 at <http://myagr.mcthosting.net>. Work started on the project in October 2013, and it took us some time to understand the requirements and how to organise the information meaningfully. At the time of writing this article, the Information Graph includes the annual reports of 2011, 2012, 2013 and 2014, and we are now processing the first series of the report of 2015.

USAGE SCENARIOS

To demonstrate the flexibility and usefulness of the Information Graph, we have written several articles in our blog at <https://myagrblog.wordpress.com>. The articles that addressed the following usage scenarios should be of particular interest:

- a) Who is audited?
- b) Government procurement
- c) Cross tabulation of audit information
- d) Government linked companies.

WHO IS AUDITED?

This is a simple legitimate question that is difficult to answer without the Information Graph. If you are looking for a single entity, you can try searching the AG's office website. However, for management and analytic purposes, we need to look at the audited entities more broadly and from different perspectives such as:

- a) National level
- b) Federal level
- c) State level
- d) Departmental level and
- e) Organisational level.

We may also want to know how many audits have been undertaken with any of the public sector entities and to be able to access directly the audit reports and summaries. The full article shows how this task becomes almost trivial with the Information Graph.

GOVERNMENT PROCUREMENT

This scenario uses procurement as one of the many cross-cutting issues in the government machinery to demonstrate how users can investigate it in the Information Graph.

We have identified 14 sub-issues in procurement, namely:

This article focused on one of the more prevalent procurement issue,

- | | |
|-------------------------------------|---------------------------|
| a) Approvals | h) Documentation |
| b) Bonds, Guarantees and Insurance. | i) Purchasing |
| c) Consultant selection | j) Scheduling |
| d) Contractor selection | k) Short or non-supply |
| e) Cost overrun | l) Specifications |
| f) Late delivery | m) Value for money |
| g) Miscellaneous (other issues) | n) Wastage in procurement |

i.e. specifications, and demonstrated how the user can drill down the Information Graph and use the interactive list processing facility to reorganise the information as part of the investigation. The same technique can be used on the other sub-issues or any other issue in the list of Audit issues.

CROSS TABULATION

For management purposes, we need to be able to have an overview of this intricate network of information. One way to do this is to cross tabulate the audit issues with the other facets available. This novel approach may help us unravel the issues encountered by the government machinery.

The article demonstrates how the various main audit issues can be cross tabulated against the audit areas and other audit entities' lists.

GOVERNMENT-LINKED COMPANIES

Government linked companies (GLCs) are grouped under "Federal Government-Linked Companies" and "State Government Subsidiaries." In this usage scenario, we are demonstrating how we can explore and compare the attributes of the lists of GLCs as an example of any other list of public sector entities.

In the Information Graph, the users can view the two lists of GLCs – the list grouped to the Federal Ministries or States and the Related Audits. The related audits can be cross tabulated against the audit issues headings

and the GLCs for the users to have an overview of the state of the GLCs.

DISCUSSIONS

We have demonstrated through this project that the annual AG's report, which is often used on a piece meal basis or ignored and underutilized, can be turned into a dynamic and valuable resource that can serve many purposes. This includes the needs of the audited public entities, the auditors, citizens, parliament, various stakeholders, management and others involved in public administration. Making the audit information more accessible, integrated and usable, will open up many other possibilities that are left to the imagination of the users.

The literature on the civil service is often based on conventional wisdom and conjecture rather than on evidence. This is not surprising as it is difficult to obtain data about the civil service. With the Information Graph, we are now able to pre-process the audit information into data. The information from the audit reports is not infallible but it is important information collected by many different people. We need the auditors to record their observations and findings judiciously. With more people scrutinising the audit reports, hopefully there will be motivation to improve the quality of the reports.

The civil service and other public sectors' entities are working within an environment that is getting more complex and challenging. With the current economic crisis,

the civil service will have to do more with less. There are issues in education, health, urban and rural infrastructure, human-resource management, foreign labour, poverty eradication, transportation, communications, water, energy, food and so on, that need to be addressed. All these issues are, in some way, interconnected and we need an integrated approach, starting with the way we manage information.

The usage scenario provided is representative of what is possible with the audit information available to identify areas of concern but they are unlikely to be sufficient to address the problems. For example, the specifications issue in procurement can have many dimensions. The specifications issues in different types of projects across the diverse spectrum, have very different characteristics. This can take several forms and it will require professionals in the various fields to analyse the sub-issues in detail and to formulate corrective actions. Some of the problems may be systemic and may only be effectively tackled at departmental or even national level. The list of audits related to the specifications issue shown in one of the usage scenarios will serve as a good starting point for further action.

The Information Graph is still in beta stage. We need to backtrack a few years and to incorporate the latest annual AG's reports as these are released, to make it more comprehensive and relevant. There is also a need to provide more processing and viewing facilities to support the diverse needs of the users. ■