

# A question of identity: detecting organised fraud

A Detica white paper

# A question of identity: detecting organised fraud

Organised fraud is on the increase in the UK and represents a significant proportion of the total cost of crime in the UK. Much organised fraud is perpetrated through some form of identity fraud. Detecting and eliminating even a modest percentage of this crime would have a significant effect on both public and private sector finances.

**A Norwich Union report in 2005 put the cost of fraud in the UK at £16bn**

**A Cabinet Office report in 2003 put the cost of identity fraud (alone) in the UK at £1.3bn**

**The two biggest classes of perpetrator were managers and organised crime, which together accounted for almost 90% of the cases**

**CIFAS reported in September 2005 that false identity fraud is once again the area that has shown the most significant increase, with a rise of almost 15%**

Whilst much of the recent press coverage around identity fraud has focussed upon the theft of identities from members of the public, there are a wide range of ways in which identity fraud occurs, for instance:

- the creation of false identities, perhaps using valid documents or the modification of data about a valid identity (e.g. alternative spellings when foreign names are Romanised);
- the theft of data pertaining to an identity and subsequent use of this identity without the owner's knowledge – commonly referred to as identity theft;
- the exaggeration of details tied to an identity or withholding of information (e.g. such as occurs in false benefit or insurance claims).

Once a false or stolen identity has been established, the fraudster will typically seek to maximise returns, targeting frauds against multiple public and private sector bodies. For example, Detica has evidence that suspected fraudulent retail bank accounts are also in receipt of government payments. Because of the nature of the crime, detecting this fraud in isolation is a difficult matter. By building a picture of behaviour across the data held in different systems, Detica has shown that it is possible to detect many instances of identity fraud and the crimes it underpins. This practical experience suggests that a strong and concerted effort between public and private sector bodies, sharing data and intelligence on fraudsters and their patterns of behaviour would yield significant benefits in the reduction of these crimes.

## Patterns of behaviour

Detecting identity fraud requires an understanding of the patterns and behaviours behind the activity. To do this we must break the hard link between individuals and the identities they present. This is essentially a 'many to many' relationship – it is possible for individuals to present more than one identity and also possible for an identity to be presented by more than one individual.

An individual will interact either directly or indirectly with a wide range of organisations through their normal life. Records of their identity information are therefore scattered throughout the systems those organisations use to record day to day transactions. Every interaction with every organisation involves the individual presenting a set of identity information. Different identity information is used at different times for different purposes.

Identity fraud exploits the fact that organisations only deal with limited identity information at each point in time. Criminals can create, confuse and control identities piece by piece and so long as nobody can see the big picture, they can keep their activities 'under the radar'. So a picture of the fraudster's 'true' identity can only be seen in full by joining up and collating all of the pieces of the identity jigsaw from the systems they have been recorded in.

In figure 1 we illustrate the practical steps for detecting identity fraud.

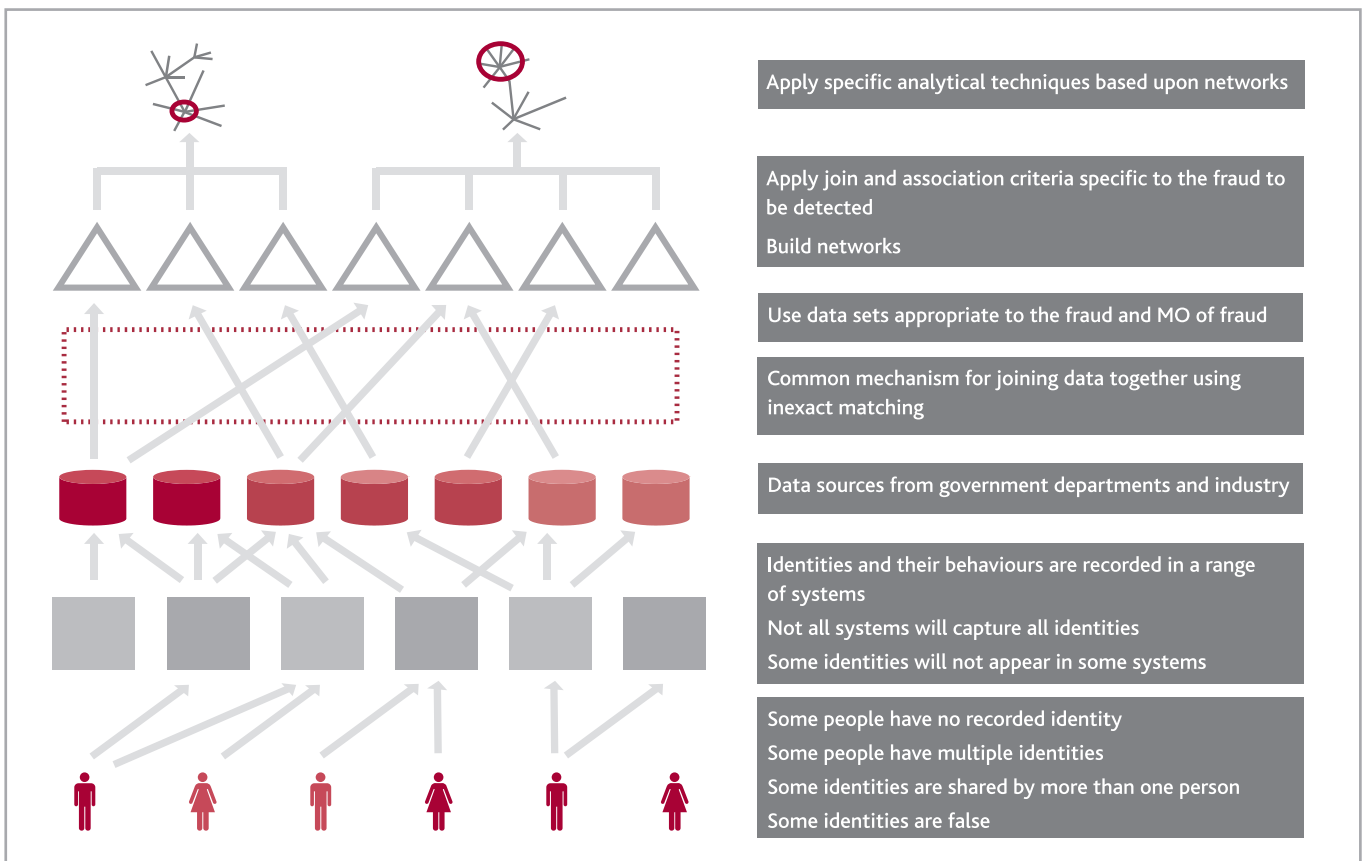


Figure 1 – Building a picture of identity fraud

KPMG's annual Fraud Barometer said 222 cases involving a total of £942m reached court in 2005, up from 172 cases worth £329m in 2004.

Taken alongside other estimates that put the cost of fraud at £20bn per annum, these figures suggest that less than 5% of frauds are successfully prosecuted.

KPMG's Fraud Barometer claims that organised crime "continues to dominate" large-scale fraud cases, with professional criminals involved in VAT scams, counterfeiting and tax evasion.

### Sources of data

Whilst there are a number of upcoming programmes that will deliver new or consolidated sources of data relating to identity, there are rich data sources already available within systems operated by the government upon which the analytics can be run today. Examples of government systems that contain data that can provide a rich picture of identity include:

- systems capturing information relating to an individual's identity, proof of identity or application for an identity token (eg immigration records, passport application records, driving licence application records etc);
- systems capturing information relating to the identity and make up of companies (eg Companies House data, VAT registration records etc);
- systems dealing with transactions between the individual or company and the government (eg benefits provision, healthcare provision, tax collection).

Similarly, the private sector runs systems that represent other very good sources of identity related data both in the form of:

- applications for products and services (eg retail financial products, or telecommunication services);
- transactional data that provides rich pictures of behaviour (eg finance, retail and telecommunications providers all capture data around product usage).

In combination these data sets will contain a more complete behavioural picture that can assist in the identification of identity fraud.

### Analysing patterns of behaviour

#### Bringing data together

The first step in the analysis of patterns of behaviour is to establish how the disparate data sets can be joined together. When joining data sets we are normally dependent upon matching specific unique reference fields. But the systems involved will use different reference fields (e.g. Passport number, NI number, customer reference number etc). So the problem shifts to matching based upon fields held in common by systems such as name fields, date of birth, address fields etc. This poses a number of challenges as in many cases variable data quality and storage rules make such fields unreliable for exact matching.

Inexact matching is necessary to achieve a meaningful join and is an important tool in detecting identity fraud. Inexact matching finds the subtly changed identities often used by fraudsters. The techniques applied will depend upon the nature of the data and the cause of the differences (e.g. typing mistakes, phonetic spelling differences or deliberate obfuscation). Each match becomes an assertion of a relationship, rather than a definite fact and inexact matching rapidly increases the number of links between data items. These assertions are not something that traditional IT systems are well equipped to deal with.

### Network analysis

Network analysis comprises of a range of techniques designed to allow analysts to work with the characteristics of inexact matching and variable data quality. It allows analysts to visualise large numbers of asserted relationships in the data. Groups of such assertions will reinforce each other, allowing the analyst to detect strong 'macro' behavioural assertions.

To build a network, we decompose the individual datasets into entities and their relationships. For example we need to break the hard link between individuals and identities – as already noted this is a 'many to many' relationship. Entities have relationships to others, to addresses, businesses and assets etc. Relationships between entities can be built up through a range of techniques as described in figure 2. This leads to networks of matched entities, which can then be subjected to detailed analysis. Networks change with time and have a shape that is representative of the behaviour of the individuals in the network.

Creating credible identities or hijacking existing ones involves risk, and criminals act to maximise the returns against their risk. This tends to generate behaviour that distinguishes identities involved in criminal activity from those engaged in 'normal' behaviour. By collating disparate identity information to find the individuals behind the identities and then detecting abnormal patterns of behaviour amongst them, a criminal modus operandi (MO) can be established.

Once built, networks take the analysis to a new level. Instead of working with data such as accounts, transactions or applications in isolation, the analysis is now performed on networks. The result is significantly better insight and more accurate isolation of activities such as identity fraud.

### Organisational challenges

Bringing multiple data sources together, from across a range of public and private sector bodies will facilitate the analysis of behavioural patterns and reveal identity fraud. There are a number of factors often perceived to prevent this happening in the most efficient manner:

- the privacy laws eg Data Protection Act (DPA) prohibit such data sharing;
- the technical complexity required to integrate multiple data feeds from diverse sources presents what might be perceived as unjustified costs;
- there is a lack of common organisational purpose.

To deliver effective measures to detect and reduce identity fraud, the government and private sector must address these challenges.

### Privacy laws

The DPA is often cited as presenting obstacles to data sharing as it requires that data is only put to the use for which it was collected and consent has been given by the individual. However, in his report into the police handling of the Soham murders, Sir Michael Richard highlighted procedural failings not DPA restrictions for the failure to share information, suggesting that in many cases there may be greater grounds for information sharing than is commonly perceived.

In fact the Data Protection Act provides a framework for the sharing of information and, with the Crime and Disorder Act, facilitates data sharing for the pursuit of a reduction in crime. It is the Data Governance and standards that often prove a barrier to sharing.

In the Home Office green paper "New Powers Against Organised and Financial Crime", the authors suggest that privacy laws should

not represent a constraint for data sharing in the fight against fraud, but that departmental administrative law might do so. It is clear that the government needs to act to address the issue, to establish clear practical guidelines to departments as to when and how data sharing is an appropriate response to a threat of fraud.

For circumstances when full data sharing is not possible, Detica have developed a number of techniques which can provide significant benefit in the detection of fraud:

- Searching records without disclosure – allows data sets to be searched without sharing the results. This can reduce the effectiveness of inexact matching significantly, but does allow some level of search across unshared data sets.
- Partial disclosure – sharing limited information facilitating data matching, with further data only disclosed when further investigation is merited.
- "Honey dipping" – sharing lists of suspected fraudsters from one organisation and matching against full data held by another. Results are disclosed only when meaningful matches are achieved.

The Crime and Disorder Act 1998 provides a legal basis for data sharing whilst the Data Protection Act 1998 provides a legal framework for good practice in handling personal information. These acts should facilitate responsible information sharing between agencies in pursuit of a reduction in crime and disorder. They should be seen as regulating rather than prohibiting.

### Technical complexity of integration

Gaining access to the large scale systems that contain the records of identities and transactions is often considered to be prohibitively complex and expensive, but need not be so. The common reason for this perceived expense is the need to modify the source system to provide some form of data feed. Careful consideration needs to be given to the necessity of such a feed against the use to which the data is going to be put. Detica has successfully undertaken a number of complex fraud detection activities on data provided as a single database export onto tape. This low cost route allows the value of analysis to be shown through effective investigations and for the level of system change necessary to be assessed. Appropriate levels of integration can then be planned and implemented.

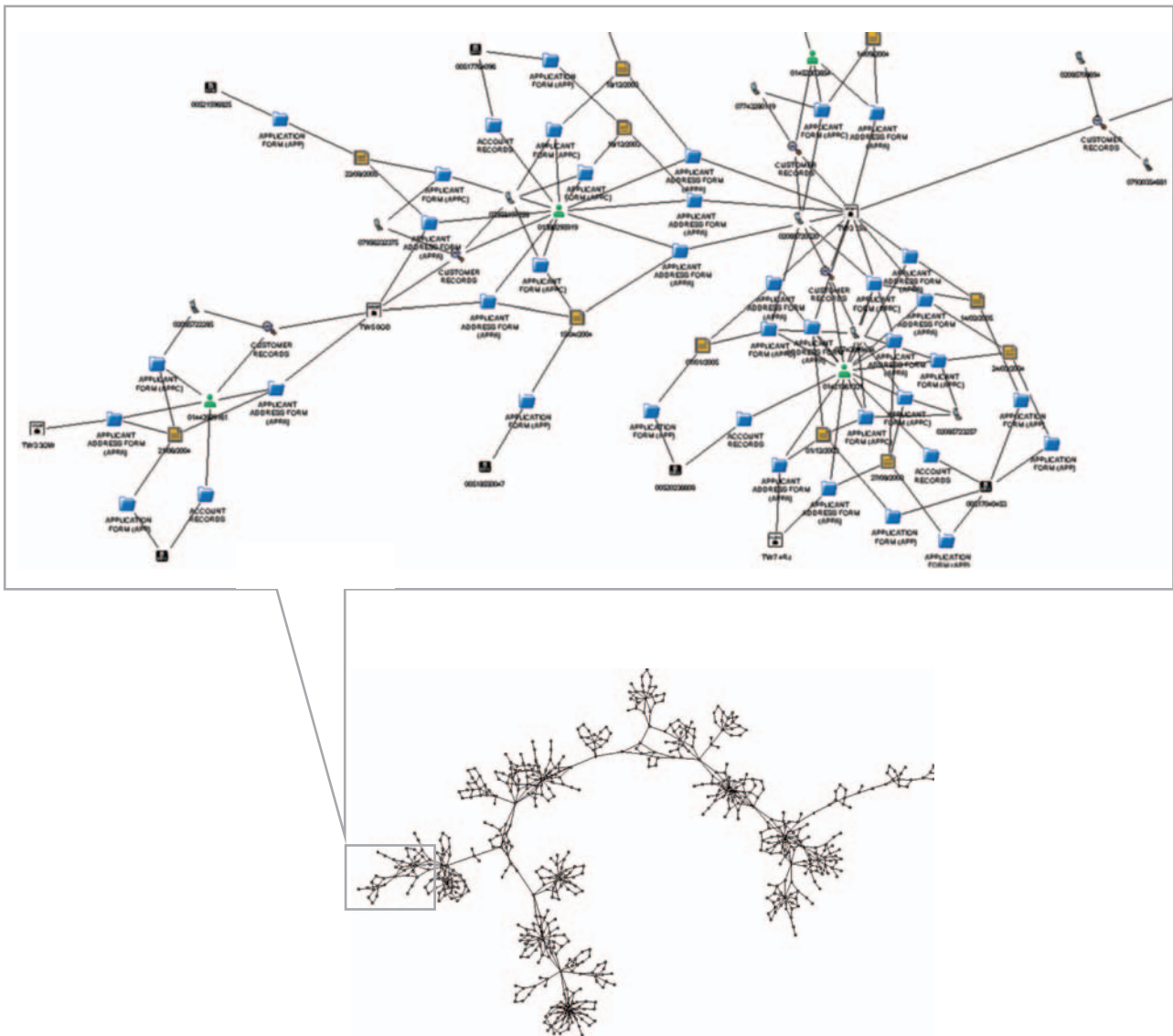
A second common technical issue raised is that of data quality. As described above, variable data quality and the necessity of using inexact matching can defeat conventional approaches. Detica's approach, applying network analysis, allows us to deal with this variable data quality and has been shown in many situations to deliver greater than expected value out of such data.

### Common organisational purpose

The mandates of individual government departments frequently provide them with powers to execute a specific line of duty, but not to a wider aim. Similarly most private sector organisations are in pursuit of their individual goals, and many see their ability to detect fraud as a business differentiator and are hence protective of their approaches. Furthermore, many private sector organisations are reluctant to prosecute, put off by the attention and publicity that such activity generates.

This results in a lack of a common organisational purpose to detect and prosecute fraud across the public and private sector. In practice a wide number of organisations have a real and vested interest in detecting and disrupting or prosecuting those perpetrating fraud – the fraudsters will frequently be defrauding a wide range of organisations in both public and private sector.

Figure 2 – Network analysis



Networks are best visualised graphically. Each node in the diagram is an entity and the lines are relationships between entities. Entities can be almost anything contained within the data such as names, addresses, cars, documents, locations and transactions.

Relationships between entities can be built up through a range of techniques:

- references stored in a data set (e.g. name and address held in one system);
- close or exact match of one or more data fields (e.g. name match);
- proximity in space and time (postcode match, or time and location of a transaction);
- similarities between a set of factors or properties of an identity or transaction.

The topology of the network says much about how it was created (i.e. the lifestyle or behaviour of the individuals or companies it involves). Patterns of networks will emerge which reflect the modus operandi (MO) of the individuals concerned.

The tools used to analyse networks need to be sophisticated tools, allowing the analyst to unlock the secrets of the network. For example, time variance within the network may be important and if you have spotted an MO, you can then start to search for examples of the same MO at differing stages of maturity. Networks are organic, change with time and are Darwinian in some respects – when something happens to threaten its existence many will adapt and change, otherwise they will be eliminated.

There are examples of cross industry co-operation, in this area. Detica provide services to the Insurance Fraud Bureau, operated by the Association of British Insurers on behalf of the UK retail insurance industry. This initiative is successfully identifying patterns of organised fraudulent behaviour within the industry not visible to the individual companies. In other areas such as money laundering, the private sector provides the government with targeted datasets of suspicious activity.

Whilst convictions are rising, much of crime perpetrated through the fraud of identity remains undetected and unprosecuted:

- the discrepancy between KPMG’s figures of £1bn successful fraud convictions in 2005, against other estimates of £16bn of perpetrated fraud per annum points to a large undetected body of fraud;
- Norwich Union’s statement that they had only prosecuted 12 of the 4000 most serious cases of insurance fraud in 2004.

### Dealing with fraud

#### Taking the initiative

There is a real opportunity for the government to take the lead and coordinate cross industry activity to tackle fraud. Publications such as Lord Goldsmith’s Fraud Review, and the Home Office green paper “New Powers Against Organised and Financial Crime” show that the government is starting to accept this challenge.

For best effect, coordination needs to be maintained throughout the detection and intervention processes, and will require both data and information sharing and joint operational processes.

Such coordinated efforts would deliver several collective benefits:

- avoiding the “balloon effect” where closing down fraud in one area simply displaces it elsewhere;
- developing and testing shared intervention strategies to identify the most effective ways to reduce and disrupt fraud;
- allowing action to be prioritised against individuals in the most cost effective areas;
- allowing for the full weight and capability of public sector law enforcement agencies and the private sector to be focussed together to a mutual benefit.

This strategic vision will require strong leadership from senior industry and government leaders and a means of recognising the support and impact made by individuals, departments and companies against the common goal. The benefits to all organisations seem clear once individual competitive instincts are discounted.

Simple checks across multiple data sets can reveal important lines of investigation. For example:

- ANPR checking for lack of tax and insurance found large numbers of people wanted on other matters;
- border checks in eFrontiers validating VAT & VOSA information against van and the company owning the van is highlighting significant excise evasion.

#### Building a capability to detect organised fraud

Whilst the strategic vision requires co-ordination and co-operation between a large number of public and private sector bodies, the vision can be built through a series of smaller initiatives, such as a series of bilateral arrangements targeting specific areas of fraud and intent upon being broadened over time.

For a particular area of fraud, the approach will be to identify and analyse the most appropriate data sources from which to start. One or two data sources are often sufficient for some detailed analysis. An iterative process of hypothesis and investigation is then followed: analysts highlight anomalous or suspicious behavioural patterns and investigators work closely with the data analysts to explain the pattern.

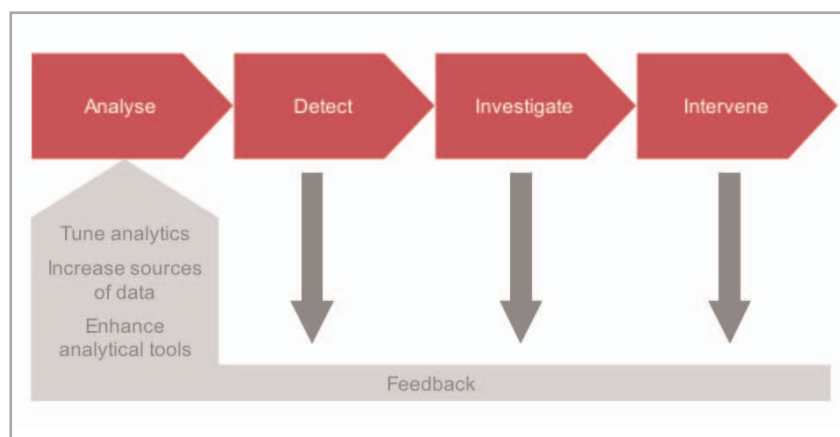
This incremental approach is essential. The analysts are building a thorough understanding of the data in question and the behaviours that it characterises. In practice some data sources will show themselves to be a rich source of behavioural content and others will not. This experimental approach is essential to ensure that, for instance, the expense of tighter integration of data sources is focussed upon those delivering most practical benefit. The approach also allows for investigators to build and hone approaches to intervention to maximise the impact on the reduction in fraud. This incremental approach is illustrated in Figure 3.

Through this incremental approach, the process of building a capability to address fraud need not be a complex or expensive business and allows funds and effort to be targeted at the most productive areas.

Example of straightforward checks that could be performed through analysis of data joined together from a small number of sources:

- **Lifestyle inconsistent with means**  
For instance checking consistency of some or all of: car value (DVLA records), house value (Land Registry), tax records (HMRC Personal and Sole Trader), affluence index of residence (commercially available data).
- **Business inconsistent with peers**  
For instance for a retail or entertainment business validate the consistency in turnover (VAT and Corporation Tax records), business location, business type, business size (eg floor area).

Figure 3 – an incremental approach



## Summary and recommendations

Fraud costs the UK billions of pounds every year. To detect and tackle organised fraud requires an understanding of the behaviours underpinning the activities. This is best achieved by bringing together data from a variety of sources and applying sophisticated analytical techniques to highlight the behavioural patterns that underpin the fraud and to identify the inconsistencies or modus operandi of the individuals.

Fraud cuts across both public and private sectors and co-ordinated efforts across the sectors will yield the most effective results. Failure to co-ordinate is likely to simply shift the fraud around rather than combat it at source. Whilst this remains the case, organised fraudsters will continue to enjoy significant rewards for their activities.

The approach taken must address the technical, legal and organisational challenges highlighted in this paper. We make three recommendations below which we believe will represent a significant step in the detection and prosecution of fraud.

### Recommendation 1

Departments within the government should embark upon a series of low cost pilots focussed upon specific areas of fraud to establish and prove the approach and best practice for data sharing and behavioural analysis. For each pilot this will mean:

- targeting a specific area of fraud and establishing the business areas and data sources from which to begin analysis;
- building a combined analysis and investigations team formed of specialised data analytics experts with secondees with business area and domain investigations experience;
- building an analytical capability in an incremental way, supported by a process of analysis and investigations;
- developing recommendations for the ongoing detection and investigation of the fraud;
- identifying and piloting appropriate intervention strategies to disrupt or eliminate the fraud.

### Recommendation 2

The government should seek to engage cross industry representation to establish a task force tasked with detecting and dealing with fraud. This task force should:

- engage enforcement and detection units in both public and private sector organisations;
- engage business owners of data likely to help in the detection of identity fraud;
- develop and implement guidance as to the appropriate interpretation of the legislation that affects data sharing (eg DPA) and establish mechanisms under which data sharing can take place;
- establish a process of information sharing about suspect identities;
- establish processes for co-ordinated action against identified fraudsters.

### Recommendation 3

Recommendations 1 and 2 deliver tactical capabilities that will demonstrate the benefit of a co-ordinated approach to addressing the problem and impacts of fraud. This final recommendation is for the transition of this tactical goal into a long term strategic capability. Such a capability should encompass:

- legal expertise concerning the ongoing governance and sharing of data amongst organisations;
- an IT platform upon which data can be brought together and analysis performed;
- an efficient approach for integrating new sources of data;
- ownership of key commercially available data sources;
- a dedicated team of analysts and investigators focussed upon the specific challenge of detecting and prosecuting fraud;
- a facility for the secondment of domain experts and investigators in the support of specific campaigns;
- a facility for tracking and co-ordinating suspected identity fraud and its investigation;
- links with law enforcement agencies and regulators to manage action against offenders.

## Working with Detica

Detica is the only major business and IT consultancy to specialise in 'Information Intelligence'. The company helps large commercial organisations and government departments to convert complex data and information into relevant and useful intelligence.

Focusing in areas such as security and fraud, risk and compliance management, business intelligence, customer management and national security, Detica has a proven track record of delivering cost reduction, improved revenue growth, customer service, operational efficiency and regulatory compliance for its clients. Its services range from business and IT consulting through to software development, systems integration and system support and maintenance.



### Find out more:

If you require further information please contact:

Mark Harrison  
Detica Limited  
Surrey Research Park  
Guildford  
Surrey GU2 7YP  
UK

[mark.harrison@detica.com](mailto:mark.harrison@detica.com)

[www.detica.com](http://www.detica.com)

T +44 (0)1483 816210