



Ministerie van Veiligheid en Justitie

> Retouradres Postbus 20301 2500 EH Den Haag



Directoraat-Generaal
Rechtspleging en
Rechtshandhaving
Directie Juridische en
Operationele
Aangelegenheden
JBOZ

Turfmarkt 147
2511 DP Den Haag
Postbus 20301
2500 EH Den Haag
www.rijksoverheid.nl/venj

Contactpersoon



Datum 27 september 2016
Onderwerp beslissing op uw Wob-verzoek van 14 maart 2016

Geachte 

Ons kenmerk
784926

*Bij beantwoording de datum
en ons kenmerk vermelden.
Wilt u slechts één zaak in uw
brief behandelen.*

In uw brief van 14 maart 2016 aan FIU-Nederland, ontvangen door FIU-Nederland op 19 april 2016 heeft u met een beroep op de Wet openbaarheid van bestuur (hierna: Wob) informatie verzocht over werkzaamheden van FIU-Nederland met betrekking tot datamining. Bij brief van 9 mei 2016 van FIU-Nederland is u medegedeeld dat FIU-Nederland geen bestuursorgaan is in de zin van de Wob en dat de behandeling van uw Wob-verzoek wordt overgedragen aan het ministerie van Veiligheid en Justitie. Mijn ministerie heeft de ontvangst van uw Wob-verzoek bij brief van 25 mei 2016 bevestigd en de termijn waarbinnen op dit verzoek moet worden beslist verdaagd tot 22 juni 2016. Bij brief van 27 juni 2016 heb ik u bericht dat de behandelend ambtenaar van mijn ministerie verschillende vergeefse pogingen heeft gedaan om u telefonisch te bereiken. Zij had inhoudelijke vragen bij uw Wob-verzoek en wilde weten of eerdergenoemde brief van FIU-Nederland van 9 mei 2016 reeds antwoord gaf op uw vragen. Bij e-mail van 4 juli 2016 heeft u mij bericht nog steeds prijs te stellen op een inhoudelijke reactie naar aanleiding van uw Wob-verzoek.

Uw verzoek

U vraagt om informatie over werkzaamheden van FIU-Nederland met betrekking tot datamining.

Wettelijk kader

Uw verzoek valt voor zover u verzoekt om informatie vastgelegd in documenten, onder de reikwijdte van de Wob. Voor de relevante Wob-artikelen verwijs ik u naar bijlage I.

Inventarisatie documenten en beantwoording van de door u gestelde vragen

In uw brief heeft u 9 vragen gesteld over werkzaamheden van FIU-Nederland met betrekking tot datamining. Om u van dienst te zijn, zal ik uw verzoek zoveel mogelijk per door u gestelde vraag (eventueel geclusterd) beantwoorden. Veel van de door u gevraagde informatie is reeds openbaar. Waar dat het geval is zal ik hieronder bij de desbetreffende onderdelen verwijzen naar de relevante vindplaats.

1. Een lijst van alle databases van bedrijven en instellingen waartoe FIU-Nederland toegang heeft, met name de niet-meldplichtige bedrijven en instellingen.

2. Een deelop overzicht per database zoals bedoeld onder 1 van de categorieën informatie die in de databases toegankelijk zijn (ofwel de structuur van de database met indeling van de informatie in kolommen).

3. Informatie/documentatie waaruit blijkt of de gegevens in de Nederlandse database geanonimiseerd dan wel ongeanonimiseerd wordt opgeslagen in het kader van de binnenlandse raadpleging (dus los van de gegevensuitwisseling met buitenlandse FIU's).

Ik beschik niet over een actuele lijst van alle databases van bedrijven en instellingen waartoe FIU-Nederland toegang toe heeft. Ik verwijs u echter naar het Mutual Evaluation Report van de FATF over Nederland (2011) en het tweede follow up report over Nederland van de FATF (2014). Beide rapporten zijn openbaar en beschikbaar via de site van de FATF. [http://www.fatf-gafi.org/publications/7hf=10&b=0&q=mual+evaluation+report+netherlands&=desc\(fatf_releasedate\)](http://www.fatf-gafi.org/publications/7hf=10&b=0&q=mual+evaluation+report+netherlands&=desc(fatf_releasedate)). In het eerstgenoemde rapport is op pagina 90 een overzicht opgenomen van systemen en bronnen waartoe FIU-Nederland direct en indirect toegang heeft. Deze lijst is grotendeels nog actueel, zij het dat systemen zijn samengegaan om efficiënter te kunnen raadplegen.

4. Het aantal ongebruikelijke transacties dat jaarlijks is gemeld in de periode 2010 t/m 2015 en het jaarlijks aantal verdachte transacties in diezelfde periode.

Het aantal ongebruikelijke transacties dat jaarlijks gemeld wordt, is opgenomen in het jaaroverzicht van FIU. FIU-Nederland publiceert jaarlijks een jaaroverzicht. Deze jaaroverzichten –die teruggaan tot 2006- zijn gepubliceerd op de website van FIU-Nederland: <https://www.fiu-nederland.nl/nl/over-fiu/jaaroverzicht>. Uit het jaaroverzicht 2015 valt bijvoorbeeld op te maken dat in 2015, 312.160 ongebruikelijke transacties werden gemeld bij de FIU-Nederland door de verschillende instellingen met een meldingsplicht. Voor het tweede jaar op rij is dit een aanzienlijke stijging: het zijn er bijna 35.000 meer dan in 2014 en ongeveer 110.000 meer dan in 2013.

Ten aanzien van de als verdacht verklaarde transacties geldt eveneens dat hierover in de jaaroverzichten van FIU-Nederland informatie wordt gegeven. Uit het jaaroverzicht van 2015 blijkt bijvoorbeeld dat FIU-Nederland in 2015, 7.352 dossiers in onderzoek heeft genomen. Hiervan zijn 6.382 dossiers met 40.959 transacties verdacht verklaard en ter beschikking gesteld aan de opsporing. Voor 970 van de onderzochte en afgeronde dossiers met daarin 3.069 transacties werd (vooralsnog) geen of onvoldoende aanleiding gevonden om deze verdacht te verklaren.

5a. De methodes en technologieën waarmee FIU-Nederland verdachte transacties uit zijn informatiebestanden destilleert.

5b. Een overzicht van ingeprogrammeerde beslisregels waarmee de verdachte transacties gedestilleerd worden.

De FIU-Nederland analyseert alle transacties die bij haar worden gemeld. Dat gebeurt via verschillende werkprocessen, namelijk semi-automatische matches

Directoraat-Generaal
Rechtspleging en
Rechtshandhaving
Directie Juridische en
Operationele
Aangelegenheden
JBOZ

Datum
27 september 2016

Ons kenmerk
784926

met externe gegevensbestanden, onderzoeken naar aanleiding van verzoeken van opsporingsdiensten, eigen onderzoeken en verzoeken van buitenlandse FIU's. Via deze wegen kan aanleiding gevonden worden om ongebruikelijke transacties verdacht te verklaren. In het jaaroverzicht wordt in het hoofdstuk "Analyseren van ontvangen ongebruikelijke transacties" globaal toegelicht hoe de analyse van gegevens plaatsvindt; <https://www.fiu-nederland.nl/nl/over-fiu/jaaroverzicht>.

**Directoraat-Generaal
Rechtspleging en
Rechtshandhaving**
Directie Juridische en
Operationele
Aangelegenheden
JBOZ

Datum
27 september 2016

Ons kenmerk
784926

De overige informatie over methodes en technologieën en een overzicht van de ingeprogrammeerde beslisregels verstrek ik niet op grond van artikel 10, tweede lid, aanhef en onder c van de Wob (het belang van opsporing en vervolging). Op grond van artikel 10, tweede lid, aanhef en onder c, van de Wob blijft openbaarmaking van informatie achterwege voor zover het belang daarvan niet opweegt tegen het belang van opsporing en vervolging. FIU-Nederland heeft toegang tot diverse intelligence bronnen c.q. databestanden voortvloeiend uit verschillende samenwerkingsverbanden zoals het FEC¹ verband en het ICOV² samenwerkingsverband. Daarnaast beschikt FIU over, rechtspersoon- en bestuurdersgerelateerde informatie, fiscale informatie aangeleverd door de FIOD liaisons, maar ook intelligence afkomstig van buitenlandse FIU's. Ook beschikt FIU over informatie verkregen uit zogenaamde artikel 17 WWFT verzoeken (aanvullende vragen aan meldplichtige instellingen en de verplichte beantwoording daarvan). Deze informatie wordt betrokken bij het voorkomen en opsporen van misdaden.

Het openbaar maken van deze informatie zou inhouden dat de specifieke werkwijze van FIU-Nederland bekend zou worden bij het grote publiek en dat criminelen hierop zouden kunnen anticiperen. Het belang van opsporing en vervolging vind ik daarom zwaarder wegen dan het belang van openbaarheid.

5c. Informatie/documentatie waaruit blijkt of FIU-Nederland gebruik maakt van technologieën om geautomatiseerd patronen in hun informatiebestanden te ontdekken zoals machine learning, en zo ja: de onderliggende technologie en of het hierbij gaat om supervised dan wel unsupervised learning.

6.

a. Een rapportage van de effectiviteit van de gebruikte selectietechnologieën om verdachte transacties uit de informatiebestanden te destilleren.

b. Een uitwerking van de onder de a bedoelde effectiviteit, bij voorkeur per jaar in de periode 2010-2015, in de volgende categorieën:

1. true positives (terecht als verdacht aangemerkt),
2. true negatives (terecht niet als verdacht aangemerkt),
3. false positives (ten onrechte als verdacht aangemerkt), en
4. false negatives (ten onrechte niet als verdacht aangemerkt).

Er wordt door FIU Nederland geen gebruik gemaakt van dergelijke technologieën. (vraag 5c en 6).

7. Het aantal LOvJ-verzoeken (verzoeken van de Landelijk Officier van Justitie inzake witwassen) dat jaarlijks is gedaan in de periode 2010-2015.

Het aantal LOvJ-verzoeken dat jaarlijks is gedaan in de periode 2010-2015 is te vinden op de website van FIU-Nederland in de betreffende jaaroverzichten zoals

¹ FEC staat voor Financial Expertise Centrum

² ICOV staat voor Infobox Crimineel en Onverklaarbaar Vermogen

bijvoorbeeld het jaaroverzicht van 2015: https://www.fiu-nederland.nl/sites/www.fiu-nederland.nl/files/documenten/fiu-nederland_jaaroverzicht_2015_def.pdf

Directoraat-Generaal
Rechtspleging en
Rechtshandhaving
Directie Juridische en
Operationele
Aangelegenheden
J80Z

8. Welke onderliggende technologie wordt gebruikt om de hashes te maken (bijv. MD5, SHA-1 etc.)

9. Welke maatregelen heeft FIU-Nederland genomen om de hashes alsnog onherleidbaar te maken?

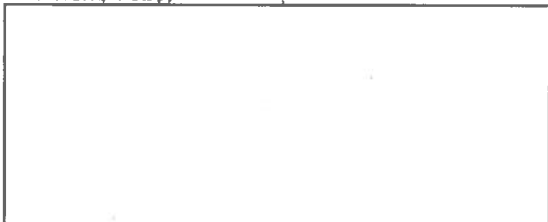
Datum
27 september 2016

Ons kenmerk
784926

Informatie hierover is te vinden in de paper "Ma3tch: Privacy AND Knowledge". Deze paper is te vinden op <http://ieeexplore.ieee.org/document/6691683/>. In overleg met de auteur doe ik u hierbij ten behoeve van uw afstudeerscriptie een kopie van deze paper toekomen.

Ik ga er vanuit u met het vorenstaande afdoende te hebben geïnformeerd en wens u veel succes met uw afstudeerscriptie.

Hoogachtend,
De Minister van Veiligheid en Justitie,
namens deze,



Tegen dit besluit kunt u binnen zes weken na de dag waarop dit bekend is gemaakt een bezwaarschrift indienen. Het bezwaarschrift moet door de indiener zijn ondertekend en bevat ten minste zijn naam en adres, de dagtekening, een omschrijving van het besluit waartegen het bezwaar is gericht en de gronden waarop het bezwaar rust. Dit bezwaarschrift moet worden gericht aan: de Minister van Veiligheid en Justitie, t.a.v. de directie Wetgeving en Juridische Zaken, Postbus 20301, 2500 EH Den Haag.

Ma³tch: Privacy AND Knowledge

'Dynamic Networked Collective Intelligence'

Udo Kroon, CIO
FIU.NET
Ministry of Security and Justice
The Hague, Netherlands
udo.kroon@fiu.net

Abstract—This paper introduces a new information technology: ma³tch (autonomous anonymous analysis). Ma³tch enables virtual information integration to build a 'dynamic networked collective intelligence' without infringing upon security, confidentiality, privacy and/or data protection regulations. It provides organizations with information and knowledge advantages.

The ma³tch technology is empowered by a decentralized information oriented architecture: a 'privacy by design' framework that uses distributed agents to facilitate decentralized but integrated information access, processing and analysis. It shapes a 'virtual information cloud' between autonomous organizations that enables secure, integral and intelligent real time information analysis. Relevant information and knowledge distributed between autonomous organizations is automatically detected and applied throughout the network as soon as it emerges.

The dynamic design principles allow practically any type of (cross domain) information to be virtually integrated: government, commercial, intelligence, law enforcement, financial, telecom, biomedical, compliance, etc., without infringing privacy, confidentiality, security or data protection rules and regulations. It advances both privacy AND knowledge beyond conventional limitations.

Keywords—decentralized information oriented architecture; privacy by design strategy; virtual information cloud; dynamic networked collective intelligence; autonomous anonymous analysis

I. INTRODUCTION

For executives and policy makers information technology is like a double edged sword. Collecting, combining and analyzing information offer huge benefits for business and governments. But at the same time when personal data are involved that information easily infringes upon the privacy of citizens. Balancing between privacy and knowledge is a continuous struggle.

Centralizing and combining data have many advantages: increased knowledge and insight, diagnoses, identification of trends and threats, strategic and competitive advantage, etc. But it also leads to severe risks of unlawful access, hacking,

identity theft, fraud, manipulation, exclusion from insurance, etc. As the amount of combined information grows, so do the risks and threats. 'In the 21st Century, bits and bytes can be as threatening as bullets and bombs,' said Deputy Defense Secretary William J. Lynn III during a speech after a cyber-attack at the Pentagon [1]. It emphasizes both the power and threat of information and the importance to protect that information [2], [3].

When information is distributed between (multinational) organizations and/or governments many implementation barriers emerge (business, organization, information and technology). These barriers significantly limit the information parties are allowed, willing and able to combine. Even at a national level it is often impossible to combine and analyze distributed government databases. What about international collaboration where national interests and (conflicting) national legislation come into play?

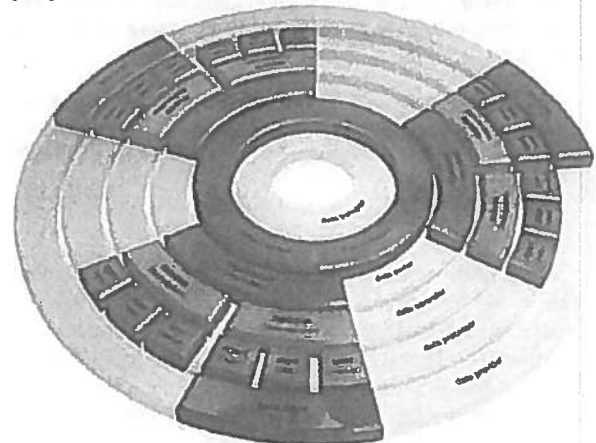
This paper introduces the ma³tch technology and the 'privacy by design' framework used by EU Member State Financial Intelligence Units, that mitigate these barriers. They enable virtual integration and real time analysis of distributed information, knowledge and technology capabilities, without the need for parties to move sensitive information beyond their own premises. Although the technology is designed for intelligence purposes, the dynamic architecture makes it suitable for many other industries and cross domain integration: government, commercial, financial, telecom, biomedical, telecom, law enforcement, compliance, etc.

II. PARADOX: PRIVACY AND KNOWLEDGE

Financial Intelligence Units (FIUs) are central, national intelligence agencies responsible for receiving, analysing and disseminating disclosures of financial information to the competent authorities (e.g., law enforcement or prosecutorial authorities) in order to combat money laundering and terrorism financing.

Decentralisation guarantees autonomous information control and it physically enforces '*privacy by design*' regarding storage, processing, analysis and exchange. Information is only physically stored and processed at the premises of the information owner, and is optionally (controlled by the owner) distributed where and when access

The *decentralized information oriented architecture* (Fig. 1) solves this by shaping a virtual enterprise architecture [6]. It directly *connects* data owners (*organisations*) without 3rd party involvement. Custom data from the *back office* are virtualized by *stores* into dynamic structured *virtual data elements*. All (distributed) data elements can be hierarchically linked and are governed by their information owner using dynamically configurable *information resources*. Resources specify custom business rules (dynamic data structures, data retention time, prior consent, information handling, etc) for each of the local information sources the FIU has access to and negotiate bilateral or multilateral capabilities. This enables both manual (*user driven*) and automated (*agent driven*) iterative case building and real time integral analysis of decentralized information and knowledge. Like Lego blocks, the virtual architecture can be *dynamically* adjusted to the legal, organisational information and technology needs of each party.



The 'privacy by design' decentralized information oriented architecture builds a virtual information cloud between all parties connected to the network and enables parties to quickly adapt to emerging changes and opportunities. The resulting

virtual platform bridges and harmonizes legal, organizational, informational and technological differences in a complex system [7] of autonomously governed parties. It establishes a dynamic information architecture and virtual enterprise architecture without infringing upon local governance, privacy, security and confidentiality.

In many ways it is the opposite of conventional information architectures and solutions. Instead of enforcing standards, standards are dynamic and virtualized. Instead of relational data storage, data are stored hierarchically. Unlike federated databases [8] there is no remote access or (costly) heterogeneous implementation. Compared to service oriented architectures [9] where information moves between services, information remains where it is. Instead, services (agents) are moved to the information. Compared to conventional analysis where information is first combined and then analyzed, information is first analyzed and only relevant information is combined by pushing or pulling information. Unlike cloud solutions, there is complete control where information is stored and processed. Instead of storing and processing information deep 'inside' boundaries of a service provider [10], the information architecture is reversed: in a sense it is a cloud inside out! The physical (vertical) decentralized processes flow of the information oriented architecture consists of 5 (horizontal) virtually integrated layers. Fig. 2 visualizes a generic information flow from an information provider (A) to one or more information consumers (B, C):

At both parties the information flows through all layers:

1. The data access layer connects any kind of data or service. Parties can connect local information 'as is' to the virtual information cloud (reduces time, requisites and costs).
2. The data virtualization layer virtualizes diverging data domains or services into uniform (standardized or generic) virtual data elements (harmonizes information).
3. The processing layer virtually integrates information with other parties, but also provides standardized local data processing and analysis capabilities (harmonizes technology).
4. The governance layer guarantees local autonomy and maximizes information that parties are able, allowed and willing to contribute (harmonizes governance and processes).
5. The exchange layer directly connects parties (direct bilateral connections guarantee that there is no intermediary data storage or third party that imposes or limits capabilities).

Since most information consumers are also information providers, similar information flows emerge in opposite direction. Sensitive information remains decentralized under full (physical and logical) control of each information owner. This guarantees that other parties cannot access the information but does allow information and capabilities to be virtually integrated and analyzed (when and where allowed) to identify relevant 'need to know' information and knowledge. Information is only physically stored and accessible where and when access to that information is needed and allowed.

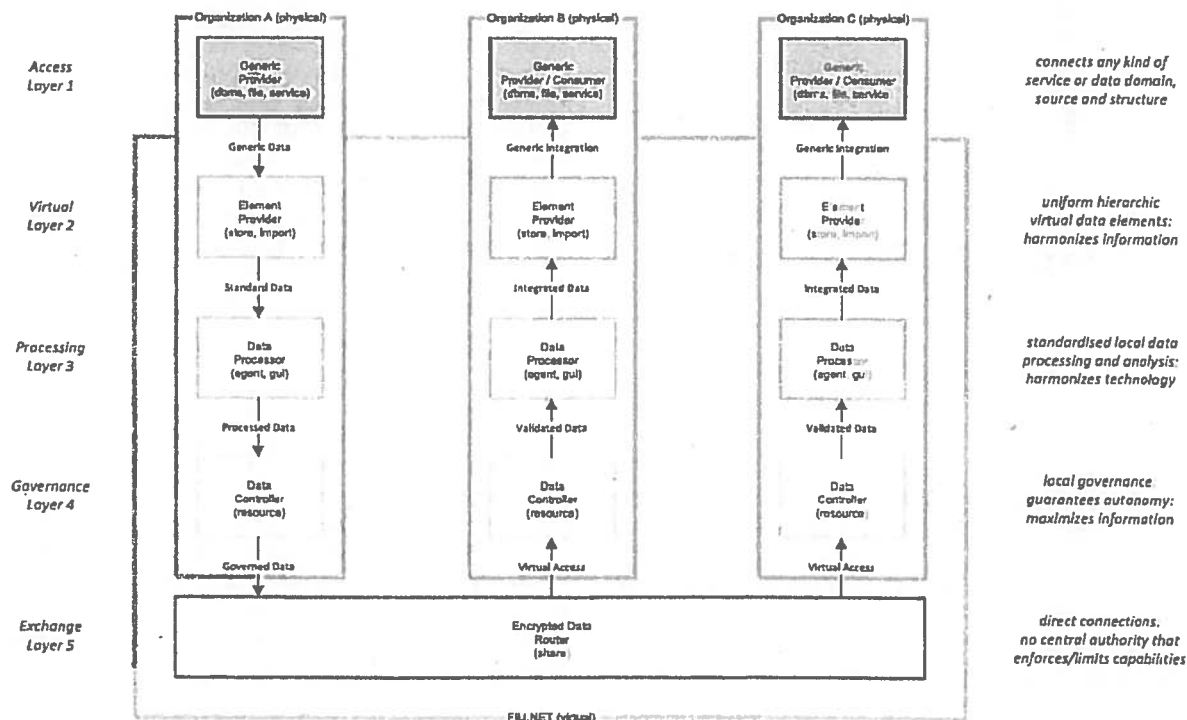


Fig. 2. Decentralized Information Oriented Architecture Processes

III. MA³TCH PROCESS AND PRINCIPLES

Ma³tch (autonomous anonymous analysis) is the generic analysis process to accomplish *virtual information integration*: the ability to combine information, knowledge and capabilities from multiple distributed organizations without bringing the data physically together. It is based upon the same core 'privacy by design' principles of the information oriented architecture: autonomy, decentralization and flexibility.

Ma³tch virtually integrates information and/or knowledge using distributed agents: programs with small and well-defined (analysis) tasks that run distributed in the processing layer. Local agents transform information and personal data into anonymized filters that can be (selectively) shared: the filters only contain characteristics of the original data, not the actual data itself. At their destination agents re-integrate the characteristics of the anonymous filters with local information resources. Ma³tch agents can be used for any kind of local, shared or distributed integral analysis (operational, strategic, social networks, geo-tagging, prediction, correlation, entity extraction, etc.).

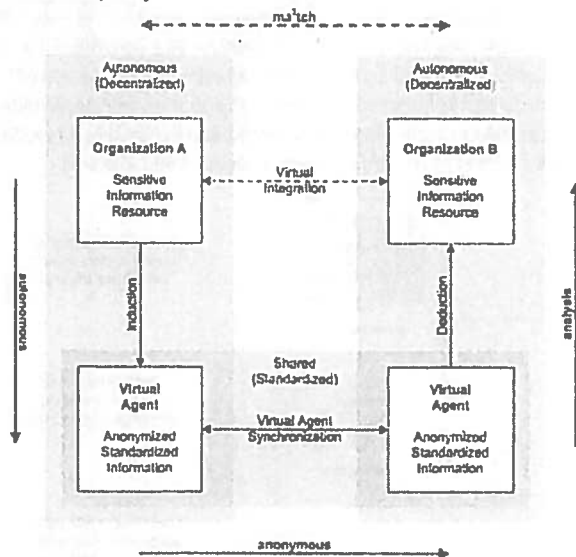


Fig. 3. Ma³tch process to accomplish virtual information integration.

The virtual integration process (Fig. 3) derives knowledge from the local resource (induction). The resulting knowledge is used locally or is distributed and integrated with remote resources (deduction). The ma³tch process identifies relevant information and knowledge links between all parties in real time. Even though no sensitive information is exchanged, all organizations can identify where and when other relevant information is available. Each organization increases its information position when it connects new information, resulting in a dynamic voluntary [11] and reciprocal [12] [13] knowledge network.

Ma³tch achieves cross-organization information and knowledge pull [13]: the ability to find, access and attract information, knowledge and resources that are relevant and valuable, even if currently all parties are unaware that the information exists. The process involves 3 steps: autonomous (access), anonymous (attract), analysis (achieve).

1. Autonomous (access)

Autonomy is guaranteed by the information oriented information architecture, and maximizes the information parties are allowed, willing and able to connect. Each information owner controls what data are included in the filter, how long the filter is valid, what the precision of the filter is, with which parties the filter is shared, and after a hit if, when and what personal data are exchanged. The fact that information is only stored when and where access to that information is needed maximizes proportionality and subsidiarity [14:151], but also empowers availability by enabling (virtual) information access.

2. Anonymous (attract)

Anonymization is non-trivial. Contrary to encryption, anonymization must be irreversible. Privacy sensitive data is considered anonymized when it cannot be directly or indirectly connected to an identified or identifiable individual [15]. Pseudonymization, for example using hashing values, does not suffice, as it still may be used to identify an individual. But even when all personal data is replaced with random numbers, or is completely removed, the remaining data may still be used to link it back to an individual [16]. To guarantee anonymity, individual dimensions and records are minimized and aggregated in such way that it is impossible to distinguish or link to individual personal records. Only characteristics, captured in the anonymous filters are shared with selected parties to 'magnetize' and *attract* relevant information.

3. Analysis (achieve)

Analysis integrates received anonymous filters with local information. This provides each organization with unique operational, tactical and strategic knowledge products integrated with their own sensitive information. Local information in each organization is 'tagged' in real time with the virtual emerging information and knowledge products: where other relevant information can be found, trends, threats, risks, etc. This analysis process requires flexibility (agents) to allow continuous dynamic (iterative) adjustment to emerging results and *achieve* integrated results.

The ma³tch process virtualizes analysis existing techniques to ma³tch: information, knowledge and technology.

IV. MA³TCH EXAMPLES

Full technical details of the analysis methods are out of scope, but basic principles and processes are briefly discussed.

1. Information Ma³tch

Examples of anonymization, aggregation and comparison algorithms that can be used to ma³tch distributed information are: fuzzy logic, hash tables, Bloom filters, transliteration, n-grams, and approximation techniques. The algorithm used in FIU.NET was initially designed for the Intersect project [17] for real time concept extraction and text analysis. It allows fast anonymous distributed cross matching: a million records easily fit in a 1 MB anonymous data structure and can be matched in less than 0.3 milliseconds on a single 3Mhz processor core, making it possible to check hundreds of distributed information resources in real time. As simplified example, an information resource contains:

- Philip Tataglia (12/28/16)
- Luka Brasi (3/13/26)
- Johnny Fontane (10/7/27)

The anonymization algorithm minimizes these 3 individual records into a single combined anonymous 4 character fuzzy logic data structure: 'tnUG'. This 4 character code captures the 'characteristics' of the combined original sensitive information, making it impossible to recover the individual records. The extreme data minimization enables (configurable) false positives (collisions) that further enhance anonymity. In addition, the information owner controls which data are included in filter, and if, when and where filters are shared (multiple filters can be created for a single dataset, for example with lower accuracy for sensitive data). Other parties that receive the filter can use it to ma³tch local sensitive data against the anonymized data structure 'tnUG' without knowing the underlying data. The example code 'tnUG' allows matching of transliterations, permutations and approximations with 99% accuracy: all following examples match positively with the data structure, even though they use different punctuation marks, date formats, and/or contain spelling errors. Positive hits are optionally or automatically followed up for (anonymous) validation, compliance check, and/or a fully detailed 'need to know' information exchange:

- Fillip Tataglia (12/28/16)
- Tattaglia Philip (28-12-1916)
- BRASI, 13-03-1926 (Luca)
- Luka Anthonius Braazzi (3-13-1926)
- Jonnie Fontarië (07-10-1927)
- Джони Фонтане (19270710), etc

Fig. 4 visualizes the process where organization A generates an autonomous match filter. Only the characters 'tnUG' are exchanged. Organization B locally checks its sensitive information against the anonymous filter 'tnUG' to identify what local information matches with organization A: Information links between the organizations are identified in real time as soon as they emerge.

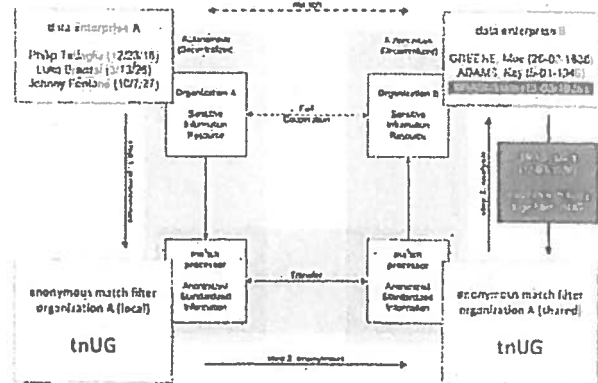


Fig. 4. Anonymous information ma³tch filter between 2 organizations.

2. Knowledge Ma³tch

Besides virtual information integration, ma³tch technology also allows virtual knowledge integration. Knowledge discovery (data mining) is the process of discovering patterns from large data sets using statistics and artificial intelligence. Intelligent data analysis can identify, describe and predict trends, threats, profiles, behaviors, etc. Existing data mining techniques can be extremely powerful in combination with the ma³tch: it enables global networked knowledge discovery and distribution. An example of a predictive data mining technology is a neural network. Neural networks capture and apply knowledge hidden within data. Like the human brain, they learn from experience. The more often a certain combination leads to a specific outcome, the stronger the associations and certainties get (simplified example in Fig. 5).

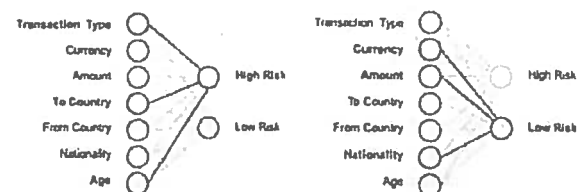


Fig. 5. neural networks (simplified) to capture knowledge.

Often neural networks are biased by internal processing logic, data quality and noise of each organization. As a result many predictions are unreliable, and if they are correct they 'explain' what the organization already implicitly knows (or thinks it knows).

Ma³tch allows parties to integrate external knowledge with

their own local information, and identify which data potentially have a higher risk based on that external knowledge. Existing local data are labeled with external knowledge tags. The ability to combine the results of multiple neural networks from multiple organizations in real time, leads to far more reliable and useful knowledge products.

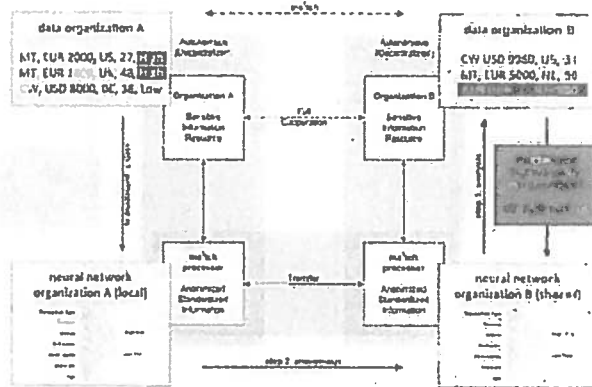


Fig. 6. Neural network ma³tch example between 2 organizations.

Ma³tch virtually integrates the emerging collective knowledge between organizations. Even better, organizations not only learn and apply what they already implicitly know, but they learn and apply the distributed collective knowledge: a high risk transaction is identified (Fig. 6) because transactions with similar patterns are identified in 5 other organizations as high risk. This transaction would not have been identified based on local knowledge only. Similarly other data mining technique can ma³tch (e.g., clustering, classification, decision trees, n-grams, etc). Prototyped examples include prediction of the gender or nationality of a person, based on first or last names of a person. All organizations locally train the ma³tch using their own local data, and utilize the combined *networked intelligence* to 'predict' missing information. There are countless other examples, but the essence is to jointly build, share and utilize common standardized knowledge without the need to disseminate sensitive information.

3. Strategic Ma³tch

Statistical ma³tch analysis enables strategic decision making. Statistics are automatically (locally) extracted from sensitive information, and virtually integrated to achieve integral knowledge products. Ma³tch analysis of money flows identifies hidden distributed financial trends in real time, for example to provide insight in local (known) versus global (unknown) money flows (Fig. 7). It reveals money flows that are locally not available and otherwise would have remained undetected (i.e., FI and PL are unaware of the money flows from NL, NL is unaware of the money flow to BG).

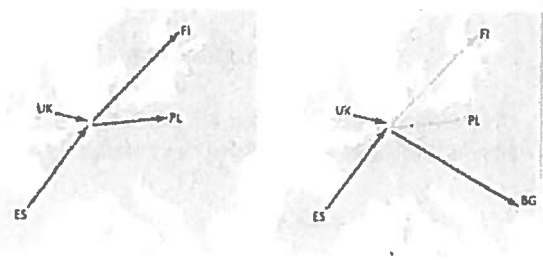


Fig. 7. Local money flows versus global money flows (perspective NL)

Trends, threats and irregularities can be detected as soon as they emerge, and automatically integrated with local data. For decision making this real time global view provides huge strategic advantages to all organizations participating in the ma³tch analysis. Combining the results of strategic ma³tches with directed neural network ma³tches or anonymous ma³tches, or even the underlying sensitive data, provides operational and knowledge insight into these strategic results.

4. Social Network Ma³tch

Social network ma³tch enhances social network analysis by creating knowledge about missing entities and links. Social network analysis maps and measures relationships and flows between people, groups, organizations and other connected information or knowledge entities. Common measures that provide insight into the various roles and groupings in a social network include: degree (number of direct connections), betweenness (number of indirect connections), closeness (distance to entities in network) and bridges (connecting different networks). But a single piece of missing information severely impacts these measures. Virtual integration (combining the information-, knowledge- and strategic ma³tch), identifies new and/or similar entities, links and relations between (otherwise disconnected) networks (Fig. 8).

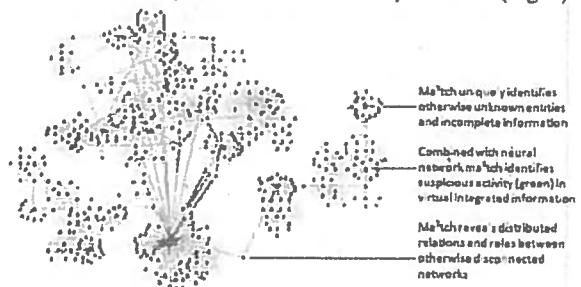


Fig. 8. Example where ma³tch completes 'loose ends'.

It allows parties to identify information links with other information (distributed across databases) to determine which role that information has in the 'global' network. Parties can even identify that two or more entities that locally seem to be disconnected, are actually related to each other based on information available at the other party (missing links).

5. Technology Ma³tch

Ma³tch minimizes financial, legal, organizational, technical and informational implementation barriers, and maximizes the information and knowledge position of all parties connected to the virtual cloud. As connected parties share a common virtual enterprise architecture (and virtual information architecture), they can jointly develop or acquire new (commercial) information sources and services cost effectively, and seamlessly integrate these with their internal systems and databases supporting privacy, security or confidentiality. It reduces individual costs and further enhances and harmonizes their capabilities. It enables connected parties to act and operate as a single virtual enterprise.

This enhances the local technology capabilities of the parties involved, and enables access to information and technologies otherwise too difficult or too costly to achieve. Examples are (anonymous) access and integration with (commercial) databases, but also with generic services like entity-extraction, geo-tagging, content translation, rule engines, etc.

V. VIRTUAL INFORMATION CLOUD INFRASTRUCTURE

The physical infrastructure of the information oriented architecture and ma³tch is dynamic as well, and can run on low-end servers, and is vendor independent. In principle multiple nodes can run on a single (virtual) server. Each time privacy, security, or information access is a concern, a new physical distributed node can be added. Organizations can directly or indirectly connect to each other (Fig. 9). Organizations may share a physical server (Organization A and B), have a dedicated server (Organization C), or use delegated servers for each of their resources (Organization D).

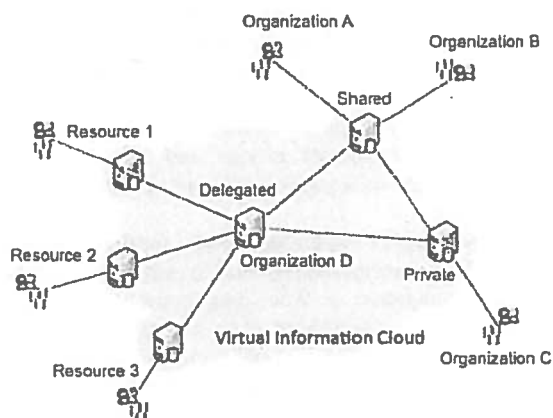


Fig. 9. Virtual Information Cloud Infrastructure

VI. MA³TCH OPPORTUNITIES

With ma³tch there is no need to exchange bulk information to achieve integral data fusion and analysis between parties, and there is no need to entrust sensitive data to a third party. Timeliness increases as relevant information is identified in real time, and is pushed or pulled when needed: no unnecessary information needs to be processed, cleaned or reviewed, which also prevents huge backlogs. Quality and reliability are enhanced as information can be processed, interpreted, amended and checked for compliance at the source.

Although the technology is designed for intelligence purposes, the dynamic architecture (due to the large variety of information FIUs process) makes it suitable for many other industries, and cross domain integration: government, commercial, financial, telecom, biomedical, telecom, law enforcement, compliance, etc.

At a national level many distributed government agencies deal with privacy sensitive information. They are expected to work together but are often (not without reason) limited by privacy regulations. This may for example prevent or limit linking weapon permits to mentally ill patients, detecting abuse of social security systems, linking sexual offences to social services. When incidents happen, in many cases the information was available, but not connected where and when it was needed. Ma³tch makes it possible to find where this (otherwise unknown) relevant information is available in real time. This can help social security or tax officials to identify fraud, hospitals to identify emerging health threats, banks to conduct joint compliance and risk assessments, etc. It can be used to analyze distributed telecom/internet traffic to combat cybercrime, and could provide an alternative to centralized DNA or electronic patient file databases; at a national level but also at an international level.

At international level benefits even further increase. Countries have lots of comparable agencies that perform similar but distributed activities in each jurisdiction. Ma³tch allows agencies of countries with a shared mission to identify relevant remote information, and use each other's knowledge to learn, cooperate and collaborate. For example to anonymously 'match' convictions between countries in order to identify criminal activities moving across borders. When privacy sensitive data are involved, for example in international collaboration between intelligence agencies on money laundering and terrorist financing, ma³tch identifies links, threats, trends and shifts between countries that otherwise remain undetected. Similarly hospitals could ma³tch biomedical data to detect trends and threats in healthcare in real time. Ma³tch shifts the focus from reactive to proactive.

International law enforcement agencies (like Interpol and Europol) may use ma³tch for real time identification of relevant (need to know) information on ongoing cases. It provides virtual access to information when and where it is needed (financial, telecom, intelligence, etc.) without the need for parties to expose sensitive information. The dynamic architecture allows matching virtually any kind of information from any source, at minimal costs. Also cooperation between Europe and the United States [18] (like PNR, TFTP) could benefit from the ma³tch, and build a collective intelligence and knowledge to identify relevant information, trends and threats as soon as they emerge, without infringing upon security, confidentiality and privacy.

Ma³tch and the information oriented architecture empower data minimization, proportionality and subsidiarity: the guiding principle within the European Union that matters ought to be handled by the smallest, lowest or least centralized competent authority [14]. It also enables the availability principle and the objective of the European Union for Member States to obtain information rapidly from private or public authorities of another Member State without the use of coercive measures. Despite of different legal systems and national regulations ma³tch enables collaboration opportunities far beyond existing limits and barriers.

The information oriented architecture has proven its value as the foundation for FIU.NET: it enables autonomous governed information access, exchange and analysis between the national FIUs (Financial Intelligence Units) of the EU member states. FIUs can virtually integrate their information, without the need to place sensitive information beyond their organization's premises. Between the EU FIUs there are over 550 information resources, each with their own information standards, policies, processes and data protection rules and regulations. It provides FIUs with secure (virtual) access to government and commercial databases through a customized unified interface, and allows them to innovate and acquire new services and information resources as one virtual enterprise.

With the information ma³tch FIUs can detect relevant information between each other, but also with commercial data sources, without leaving any traces. Within this diverse and complex environment, ma³tch technology enables joint information analysis and innovation capabilities otherwise impossible to achieve. But also at a national level FIUs can benefit from ma³tch to enhance local analysis capabilities, and by using it for (cross border) reporting by exchanging ma³tch filters with reporting institutions (reducing administrative burdens, and enhancing national investigations).

FIUs have already named the emerging information and knowledge 'scarefully powerful', and now face a new challenge: to deal with the emerging streams of real time information and knowledge.

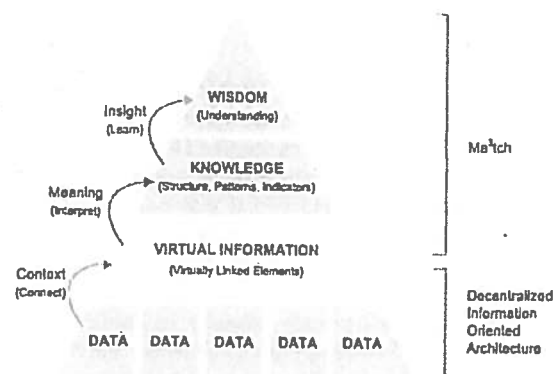


Fig. 10. Building a dynamic collective networked intelligence

VII. CONCLUSIONS

Ma³tch builds a *dynamic networked collective intelligence* (Fig. 10) that provides participating parties with tremendous information and knowledge advantages. Relevant information links and knowledge distributed between parties are detected in real time as soon as they emerge. It gives governmental and commercial organizations a unique information and knowledge position that otherwise would be impossible to achieve. It provides a global overview for strategic decision making and generates knowledge that otherwise would remain distributed, hidden and undetected. Integral operational and strategic analysis on distributed data without infringing upon privacy or security advances results beyond what is otherwise possible. Privacy and knowledge do not conflict but strengthen each other, maximizing both privacy AND knowledge.

The information oriented architecture provides a dynamic and flexible foundation for decentralized information access, exchange, and analysis that enables unique strategic advantages. It enhances data minimization, proportionality and subsidiarity and maximizes voluntary and reciprocal communication, coordination, cooperation and collaboration. It strengthens coherence and bridges legal, organizational, informational and technological differences between autonomous parties. It shapes a dynamic virtual enterprise that enables parties to operate as one, and quickly and cost effectively adapt to emerging changes and opportunities.

Understanding the ma³tch technology requires a paradigm shift, not just in understanding *what* is possible, but perhaps even more important in *how* organizations, governments, central authorities, legislators, data protection offices and information managers look at collaboration and information technology problems and solutions. Information from any national and/or international data source can be virtually integrated. It advances both privacy AND knowledge beyond conventional limitations: ma³tch opportunities start where other possibilities stop.

VIII. REFERENCES

- [1] W. J. Lynn, DOD Releases First Strategy for Operating in Cyberspace, National Defense University at Fort Lesley J. McNair, Washington, 2011.
- [2] Council of Europe, 81/679/EEC: "Commission Recommendation relating to the Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data", Strasbourg: *Official Journal of the European Communities*, 1981.
- [3] Council, European Parliament, "Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data", *Official Journal of the European Communities*, 1995.
- [4] World Bank, IMF, *Financial Intelligence Units - An Overview*, Washington, D.C.: International Monetary Fund Publication Services, 2004, pp. 9-22.
- [5] E. Schreuders, "The Legal Aspects of Cooperation between FIUs using FIU.NET", 2009. Unpublished.
- [6] J. Hoogervorst, *Enterprise Governance and Enterprise Engineering*, Heidelberg: Springer, 2009, pp. 297.
- [7] G. Bellinger, "Systems Thinking", Internet: <http://www.systems-thinking.org>, 2004 [2012].
- [8] M. A. Heimbigner, "A Federated Architecture for information management", *ACM Transactions on Information Systems*, pp. 253-278, 1985.
- [9] N. M. Josuttis, *SOA in Practice: The Art of Distributed System Design*, Sebastopol: O'Reilly Media, 2007, pp. 65-79, 261-283.
- [10] V. J. Winkler, *Securing the Cloud: Cloud Computer Security Techniques and Tactics*, Elsevier, 2011, pp. 55-85.
- [11] S. Gillinson, *Why Cooperate? A Multi-Disciplinary Study of Collective Action*, London: Overseas Development Institute, 2004, p. 7.
- [12] D. Gambetta, *Opportunism, Trust and Cooperation*, Oxford: Blackwell, 1998.
- [13] J. Hagel, J. S. Brown and L. Davison, *The Power of Pull*, New York: Basic Books, 2010, pp. 89-94.
- [14] S. Wolff, F. A. Goudappel and J. W. d. Zwaan, *Freedom, Security and Justice after Lisbon and Stockholm*, The Hague: Asser Press, 2011, pp. 137, 140, 146, 151.
- [15] P. Balboni, U. Kroon and M. Macenaite, "Data Protection and Data Security by Design Applied to Financial Intelligence", *Information Security Solutions Conference (ISSE)*, 2013.
- [16] A. Narayanan en V. Shmatikov, "Robust De-anonymization of Large Sparse Datasets", *IEEE Symposium on Security and Privacy*, pp. 111-125, 2008.
- [17] U. Kroon, "Intersect Project", Internet: <http://intersect.crowndesign.nl> 2008 [2012].
- [18] K. Archick, "U.S.-EU Cooperation Against Terrorism", Council on Foreign Relations, Internet: <http://www.cfr.org/counterterrorism/crs-us-eu-cooperation-against-terrorism/p30552>, 2013 [2013].

Bijlage 1 – Relevante artikelen uit de Wob

Artikel 1

In deze wet en de daarop berustende bepalingen wordt verstaan onder:

- a. document: een bij een bestuursorgaan berustend schriftelijk stuk of ander materiaal dat gegevens bevat;
- b. bestuurlijke aangelegenheid: een aangelegenheid die betrekking heeft op beleid van een bestuursorgaan, daaronder begrepen de voorbereiding en de uitvoering ervan;
- c. intern beraad: het beraad over een bestuurlijke aangelegenheid binnen een bestuursorgaan, dan wel binnen een kring van bestuursorganen in het kader van de gezamenlijke verantwoordelijkheid voor een bestuurlijke aangelegenheid;
- d. niet-ambtelijke adviescommissie: een van overheidswege ingestelde instantie, met als taak het adviseren van een of meer bestuursorganen en waarvan geen ambtenaren lid zijn, die het bestuursorgaan waaronder zij ressorteren adviseren over de onderwerpen die aan de instantie zijn voorgelegd. Ambtenaren, die secretaris of adviserend lid zijn van een adviesinstantie, worden voor de toepassing van deze bepaling niet als leden daarvan beschouwd;
- e. ambtelijke of gemengd samengestelde adviescommissie: een instantie, met als taak het adviseren van één of meer bestuursorganen, die geheel of gedeeltelijk is samengesteld uit ambtenaren, tot wier functie behoort het adviseren van het bestuursorgaan waaronder zij ressorteren over de onderwerpen die aan de instantie zijn voorgelegd;
- f. persoonlijke beleidsopvatting: een opvatting, voorstel, aanbeveling of conclusie van een of meer personen over een bestuurlijke aangelegenheid en de daartoe door hen aangevoerde argumenten;
- g. milieu-informatie: hetgeen daaronder wordt verstaan in artikel 19.1a van de Wet milieubeheer;
- h. hergebruik: het gebruik van informatie die openbaar is op grond van deze of een andere wet en die is neergelegd in documenten berustend bij een overheidsorgaan, voor andere doeleinden dan het oorspronkelijke doel binnen de publieke taak waarvoor de informatie is geproduceerd;
- i. overheidsorgaan:

1°. een orgaan van een rechtspersoon die krachtens publiekrecht is ingesteld, of

2°. een ander persoon of college, met enig openbaar gezag bekleed.

Artikel 3

1. Een ieder kan een verzoek om informatie neergelegd in documenten over een bestuurlijke aangelegenheid richten tot een bestuursorgaan of een onder verantwoordelijkheid van een

bestuursorgaan werkzame instelling, dienst of bedrijf.

2. De verzoeker vermeldt bij zijn verzoek de bestuurlijke aangelegenheid of het daarop betrekking hebbend document, waarover hij informatie wenst te ontvangen.

3. De verzoeker behoeft bij zijn verzoek geen belang te stellen.

4. Indien een verzoek te algemeen geformuleerd is, verzoekt het bestuursorgaan de verzoeker zo spoedig mogelijk om zijn verzoek te preciseren en is het hem daarbij behulpzaam.

5. Een verzoek om informatie wordt ingewilligd met inachtneming van het bepaalde in de artikelen 10 en 11.

Artikel 6

1. Het bestuursorgaan beslist op het verzoek om informatie zo spoedig mogelijk, doch uiterlijk binnen vier weken gerekend vanaf de dag na die waarop het verzoek is ontvangen.

2. Het bestuursorgaan kan de beslissing voor ten hoogste vier weken verdagen. Van de verdaging wordt voor de afloop van de eerste termijn schriftelijk gemotiveerd mededeling gedaan aan de verzoeker.

3. Onverminderd artikel 4:15 van de Algemene wet bestuursrecht wordt de termijn voor het geven van een beschikking opgeschort gerekend vanaf de dag na die waarop het bestuursorgaan de verzoeker mededeelt dat toepassing is gegeven aan artikel 4:8 van de Algemene wet bestuursrecht, tot de dag waarop door de belanghebbende of belanghebbenden een zienswijze naar voren is gebracht of de daarvoor gestelde termijn ongebruikt is verstreken.

4. Indien de opschorting, bedoeld in het derde lid, eindigt, doet het bestuursorgaan daarvan zo spoedig mogelijk mededeling aan de verzoeker, onder vermelding van de termijn binnen welke de beschikking alsnog moet worden gegeven.

5. Indien het bestuursorgaan heeft besloten informatie te verstrekken, wordt de informatie verstrekt tegelijk met de bekendmaking van het besluit, tenzij naar verwachting een belanghebbende bezwaar daar tegen heeft, in welk geval de informatie niet eerder wordt verstrekt dan twee weken nadat de beslissing is bekendgemaakt.

6. Voor zover het verzoek betrekking heeft op het verstrekken van milieu-informatie:

- a. bedraagt de uiterste beslistermijn in afwijking van het eerste lid twee weken indien het bestuursorgaan voornemens is de milieu-informatie te verstrekken terwijl naar verwachting een belanghebbende daar bezwaar tegen heeft;
- b. kan de beslissing slechts worden verdaagd op grond van het tweede lid, indien de omvang of de gecompliceerdheid van de milieu-informatie een verlenging rechtvaardigt;
- c. zijn het derde en vierde lid niet van toepassing.

Artikel 7

1. Het bestuursorgaan verstrekt de informatie met betrekking tot de documenten die de verlangde informatie bevatten door:

- a. kopie ervan te geven of de letterlijke inhoud ervan in andere vorm te verstrekken,
- b. kennisneming van de inhoud toe te staan,
- c. een uittreksel of een samenvatting van de inhoud te geven, of
- d. inlichtingen daaruit te verschaffen.

2. Het bestuursorgaan verstrekt de informatie in de door de verzoeker verzochte vorm, tenzij:

- a. het verstrekken van de informatie in die vorm redelijkerwijs niet gevegd kan worden;
- b. de informatie reeds in een andere, voor de verzoeker gemakkelijk toegankelijke vorm voor het publiek beschikbaar is.

3. Indien het verzoek betrekking heeft op milieu-informatie als bedoeld in artikel 19.1a, eerste lid, onder b, van de Wet milieubeheer, verstrekt het bestuursorgaan, zo nodig, en indien deze informatie voorhanden is, tevens informatie over de methoden die zijn gebruikt bij het samenstellen van eerstbedoelde informatie.

Artikel 10

1. Het verstrekken van informatie ingevolge deze wet blijft achterwege voor zover dit:

- a. de eenheid van de Kroon in gevaar zou kunnen brengen;
- b. de veiligheid van de Staat zou kunnen schaden;
- c. bedrijfs- en fabricagegegevens betreft, die door natuurlijke personen of rechtspersonen vertrouwelijk aan de overheid zijn meegedeeld;
- d. persoonsgegevens betreft als bedoeld in paragraaf 2 van hoofdstuk 2 van de Wet bescherming persoonsgegevens, tenzij de verstrekking kennelijk geen inbreuk op de persoonlijke levenssfeer maakt.

2. Het verstrekken van informatie ingevolge deze wet blijft eveneens achterwege voor zover het belang daarvan niet opweegt tegen de volgende belangen:

- a. de betrekkingen van Nederland met andere staten en met internationale organisaties;
- b. de economische of financiële belangen van de Staat, de andere publiekrechtelijke lichamen of de in artikel 1a, onder c en d, bedoelde bestuursorganen;
- c. de opsporing en vervolging van strafbare feiten;

- d. inspectie, controle en toezicht door bestuursorganen;
- e. de eerbiediging van de persoonlijke levenssfeer;
- f. het belang, dat de geadresseerde erbij heeft als eerste kennis te kunnen nemen van de informatie;
- g. het voorkomen van onevenredige bevoordeling of benadeling van bij de aangelegenheid betrokken natuurlijke personen of rechtspersonen dan wel van derden.

3. Het tweede lid, aanhef en onder e, is niet van toepassing voorzover de betrokken persoon heeft ingestemd met openbaarmaking.

4. Het eerste lid, aanhef en onder c en d, het tweede lid, aanhef en onder e, en het zevende lid, aanhef en onder a, zijn niet van toepassing voorzover het milieu-informatie betreft die betrekking heeft op emissies in het milieu. Voorts blijft in afwijking van het eerste lid, aanhef en onder c, het verstrekken van milieu-informatie uitsluitend achterwege voorzover het belang van openbaarmaking niet opweegt tegen het daar genoemde belang.

5. Het tweede lid, aanhef en onder b, is van toepassing op het verstrekken van milieu-informatie voor zover deze handelingen betreft met een vertrouwelijk karakter.

6. Het tweede lid, aanhef en onder g, is niet van toepassing op het verstrekken van milieu-informatie.

7. Het verstrekken van milieu-informatie ingevolge deze wet blijft eveneens achterwege voorzover het belang daarvan niet opweegt tegen de volgende belangen:

- a. de bescherming van het milieu waarop deze informatie betrekking heeft;
- b. de beveiliging van bedrijven en het voorkomen van sabotage.

8. Voorzover het vierde lid, eerste volzin, niet van toepassing is, wordt bij het toepassen van het eerste, tweede en zevende lid op milieu-informatie in aanmerking genomen of deze informatie betrekking heeft op emissies in het milieu.

Artikel 11

1. In geval van een verzoek om informatie uit documenten, opgesteld ten behoeve van intern beraad, wordt geen informatie verstrekt over daarin opgenomen persoonlijke beleidsopvattingen.

2. Over persoonlijke beleidsopvattingen kan met het oog op een goede en democratische bestuursvoering informatie worden verstrekt in niet tot personen herleidbare vorm. Indien degene die deze opvattingen heeft geuit of zich erachter heeft gesteld, daarmee heeft ingestemd, kan de informatie in tot personen herleidbare vorm worden verstrekt.

3. Met betrekking tot adviezen van een ambtelijke of gemengd samengestelde adviescommissie kan het verstrekken van informatie over de daarin opgenomen persoonlijke beleidsopvattingen plaatsvinden, indien het voornemen daartoe door het bestuursorgaan dat het rechtstreeks

aangaat aan de leden van de adviescommissie voor de aanvang van hun werkzaamheden kenbaar is gemaakt.

4. In afwijking van het eerste lid wordt bij milieu-informatie het belang van de bescherming van de persoonlijke beleidsopvattingen afgewogen tegen het belang van openbaarmaking. Informatie over persoonlijke beleidsopvattingen kan worden verstrekt in niet tot personen herleidbare vorm. Het tweede lid, tweede volzin, is van overeenkomstige toepassing.