




CATE

(Continuous Auditing based on Techn(olog)ical Evolution and Data Mining)

Eindverslag

Programma	
Aanbestedende diensten	
Externe begeleider	

Inhoud

1	Managementsamenvatting	4
2	Executive summary	5
3	Werkingskader en objectieven.....	6
3.1	De projectinitiatoren	6
3.1.1	Vlaamse Auditautoriteit van de Europese Structuurfondsen	6
3.1.2	Programma Innovatieve Overheidsopdrachten	6
3.2	Het voortraject: "CATE".....	6
3.2.1	Context	6
3.2.2	Doelstellingen van het voortraject.....	7
4	Inschatting innovatiepotentieel vanuit gebruikersstandpunt.....	9
4.1	Gevolgd proces.....	9
4.2	Resultaten.....	9
4.2.1	Use cases A rond de start van het project.....	9
4.2.2	Use cases B tijdens de duur van het project	11
4.2.3	Use cases C aan het einde van het project.....	14
4.3	Conclusie	14
5	State-of-the-art analyse	16
6	Inschatting van het innovatiepotentieel vanuit technologische invalshoek.....	17
6.1	Gevolgd proces.....	17
6.2	Resultaten.....	17
6.2.1	Use cases A rond de start van het project.....	18
6.2.2	Use cases B tijdens de duur van het project	22
6.2.3	Use cases C aan het einde van het project.....	29
6.3	Conclusie	30
7	Architectuur en implementatieplan.....	31
7.1	Value-Risk analyse	31
7.2	Architectuur finaliteit audit.....	32
7.3	Gefaseerd implementatieplan.....	34
7.3.1	Fase 1: Data verzameling en gestructureerde informatie-extractie	35
7.3.2	Fase 2: Extractie uit databronnen en standaardrapportering.....	36
7.3.3	Fase 3: Machine learning voor patroondetectie en risico's	36
7.3.4	Fase 4: Diepe patroondetectie en betere risicoinschatting	37
7.3.5	Fase 5: Volledige interpretatie van vrije tekst en accurate risicoinschatting	37

7.3.6	Parallele fase: conceptueel uitwerken risicomodel van de toekomst	38
8	Algemene conclusie.....	39
	Bijlage A: State-of-the-art analyse	41
	Bijlage B: Deelnemende partijen in de marktconsultatie	45

1 Managementsamenvatting

De Vlaamse Auditautoriteit voor Europese Structuurfondsen (VAA) wenst hun auditproces zowel op technologisch als op conceptueel vlak volledig te herdenken, vertrekkend van de mogelijkheden die machine learning en data mining bieden. Het CATE (Continuous Auditing based on Techn(olog)ical Evolution and Data Mining) project ingediend bij het Programma Innovatieve Overheidsopdrachten heeft als doel het auditproces te versnellen, het verlies van Europese middelen te verminderen en accuratere rapporteringen te leveren van hoge kwaliteit. Om tot een realistisch, kwaliteitsvol bestek te komen werd een voortraject georganiseerd dat liep vanaf 8 mei 2020 tot begin oktober 2020, waarvan dit eindverslag het proces en de resultaten beschrijft.

Het voortraject heeft als doel om zowel de behoeftes van de eindgebruikers, als de technologische haalbaarheid scherp te stellen. Hiervoor werden eerst use cases opgesteld en geprioriteerd door auditors en andere stakeholders die de gebruikersnoden capteren. In een tweede stap werd een state-of-the-art analyse uitgevoerd om de relevante technologieën in kaart te brengen en om de belangrijkste marktspelers te identificeren. In de marktconsultatie op 4 september 2020 werden vervolgens samen met de marktspelers voor elke use case de haalbaarheid en de risico's ingeschat, zowel vandaag als in de toekomst. Uit inzichten verkregen uit de gebruikersnoden en de marktconsultatie werd tot slot een plan van aanpak uitgewerkt voor het bestek. Uit deze analyse van het voortraject komen drie conclusies naar voor:

Een eerste conclusie uit het voortraject is dat het project innovatief is vanuit gebruikersstandpunt. Door maximaal in te zetten op de mogelijkheden van machine learning, krijgt het auditen van programma's een volledig nieuwe invulling. Dit vereist een nieuwe manier van werken en leidt tot kortere doorlooptijden, snellere feedback en minder verlies van (Europese) middelen. De eindgebruikers gaven bij de opgestelde use cases aan dat deze allemaal een zeer hoge toegevoegde waarde leveren.

Een tweede inzicht uit het voortraject is dat CATE heel innovatief is vanuit technologisch perspectief, en verder onderzoek in machine learning vereist om complexe use cases te implementeren. Tijdens de marktconsultatie werd aangegeven dat een deel van de use cases reeds vandaag mogelijk zijn, zoals het controleren op volledigheid van documenten en data-extractie uit gestructureerde documenten. Bij complexere use cases laat de huidige stand van machine learning en natural language processing echter nog geen volledige automatisatie toe. Zo is het vandaag niet mogelijk om in alle contexten projectrisico's in te schatten of om vrije tekst volledig te interpreteren. Door het intensieve onderzoek in deze domeinen schat de markt echter in dat deze use cases wel mogelijk zullen zijn over enkele jaren.

Een derde conclusie uit het voortraject is dat een gefaseerd implementatieplan is aangewezen voor CATE, vertrekkend vanuit een referentiearchitectuur. Om te vermijden dat het huidige auditproces gewoon geautomatiseerd wordt, werd een referentiearchitectuur opgesteld die de ideale eindoplossing capteert. Deze bestaat uit verschillende bouwblokken die elk gefaseerd kunnen uitgewerkt worden volgens de evoluties in de technologie. In het eindverslag wordt een implementatieplan voorgesteld in vijf fases waarin voor elk bouwblok wordt uitgestippeld hoe deze kan evolueren tegen 2026. Parallel met deze fases dient een track te lopen waarin het risicomodel voor de toekomst, dat centraal staat in de nieuwe aanpak, conceptueel wordt uitgedacht. Deze wordt vervolgens na elke ontwikkelingsfase getoetst en bijgeschaafd om de visie van het project te bewaken.

2 Executive summary

The Flemish Audit Authority for European Structural Funds (VAA) wants to completely rethink their audit process both technologically and conceptually, starting from the possibilities offered by machine learning and data mining. The CATE (Continuous Auditing based on Techn(ological) Evolution and Data Mining) project submitted to the Programme for Innovation Procurement aims to speed up the audit process, reduce the loss of European resources and provide more accurate and high quality reporting. In order to obtain a realistic, high-quality set of specifications, a preliminary process was organized that ran from 8 May 2020 to the beginning of October 2020, of which this final report describes the process and the results.

The aim of the preliminary process is to focus on both the needs of the end users and the technological feasibility. To capture the user needs, use cases were first drawn up and prioritized by auditors and other stakeholders. In a second step, a state-of-the-art analysis was carried out to identify the relevant technologies and to identify the most important market players. During the market consultation on 4 September 2020, the feasibility and risks for each use case were assessed together with market players, taking both today's and the future state of the art into account. Finally, a plan of attack was drawn up for the project, based on insights gained from the user needs and market consultation. Three conclusions emerge from the analysis in this preliminary process:

A first conclusion from the preliminary process is that the project is innovative from the user's point of view. By making maximum use of the possibilities of machine learning, the auditing of programs is given a completely new interpretation. This requires a new way of working and leads to shorter lead times, faster feedback and less loss of (European) resources. The end users indicated that all use cases that were drafted provide a very high added value.

A second insight from the preliminary process is that CATE is very innovative from a technological perspective, and requires further research in machine learning to implement complex use cases. During the market consultation it was indicated that part of the use cases are already possible today, such as checking for completeness of documents and data extraction from structured documents. For more complex use cases, however, the current state of machine learning and natural language processing does not yet allow full automation. For example, it is currently not possible to assess project risks in all contexts or to fully interpret free text. However, due to the intensive research in these areas, the market estimates that these use cases will be possible in a few years.

A third conclusion from the preliminary process is that a phased implementation plan is appropriate for CATE, starting from a reference architecture. In order to avoid that the current audit process is simply automated, a reference architecture has been drawn up that captures the ideal final solution. This consists of several building blocks, each of which can be developed in phases according to the evolutions in technology. In the final report, an implementation plan is presented in five phases, outlining for each building block how it can evolve by 2026. Parallel to these phases, a track should be run in which the risk model for the future system, which is central to the new approach, is conceptually conceived. This will then be fine-tuned after each development phase in order to safeguard the project's vision.

3 Werkingskader en objectieven

3.1 De projectinitiatoren

3.1.1 Vlaamse Auditautoriteit van de Europese Structuurfondsen

Binnen de Vlaamse overheid is de Vlaamse Auditautoriteit van de Europese Structuurfondsen (VAA) verantwoordelijk voor het auditen van de besteding van Europese en daaraan verbonden Vlaamse fondsen. De VAA doet dit in het kader van de structuurfondsen ESF (Europees Sociaal Fonds), EFRO (Europees Fonds Regionale Ontwikkeling) en Interreg Vlaanderen/Nederland.

De Vlaamse Auditautoriteit van de Europese Structuurfondsen (VAA) moet aan de Europese Commissie (EC) een redelijke mate van zekerheid kunnen bieden over:

- Het functioneren van het beheers- en controlesysteem van de instanties die het bijstandspakket beheren, alsook van de projecten die medegefinancierd zijn door de Europese Structuurfondsen. De instanties die het bijstandspakket beheren zijn de Managementautoriteiten (MA), de Certificeringsautoriteiten (CA) en Bemiddelende Instanties (BI).
- De juistheid van de uitgavenstaten ingediend bij de EC. Hoofdzakelijk wordt de wettig- en regelmatigheid van de onderliggende verrichtingen onderzocht.

3.1.2 Programma Innovatieve Overheidsopdrachten

Het Programma Innovatieve Overheidsopdrachten (PIO) van het Departement Economie, Wetenschap en Innovatie (EWI) heeft als doel de omvangrijke koopkracht van de Vlaamse overheid (en de bredere publieke sector in Vlaanderen) meer strategisch in te zetten voor innovatie. Hiertoe wil het PIO overheidsorganisaties in Vlaanderen stimuleren en helpen om een deel van hun aankoopmiddelen te besteden aan innovatieve overheidsopdrachten, d.w.z. het (laten) ontwikkelen en/of aankopen van innovatieve producten en diensten waarmee ze hun eigen werking en publieke dienstverlening kunnen optimaliseren en beter kunnen inspelen op de vele maatschappelijke uitdagingen waarvoor ze staan. Op die manier wil het PIO bijdragen tot een performantere overheid, competitievere ondernemingen en oplossingen voor uitdagingen van maatschappelijk belang (gezondheid, milieu en energie, veiligheid, etc.).

3.2 Het voortraject: “CATE”

3.2.1 Context

Vandaag bestaat het beheers- en controlesysteem van Europese Structuurfondsen uit meerdere stappen en deelnemende partijen. De twee belangrijkste instanties hierin zijn de Managementautoriteiten (MA), die instaan voor de eerstelijnscontroles, en de Certificeringsautoriteiten (CA) die de rekeningen opstellen.

De VAA controleert de MA en CA door gebruik te maken van 3 soorten audits: systeemaudits, audits op concrete acties en audits op de output. Bij systeemaudits gaat de VAA na of de MA en CA het operationeel proces volgens de Europese regels gevolgd hebben. Vervolgens voert de VAA zelf operationele audits uit op de begunstigde van een gesteund project door hetzelfde proces als de MA

te doorlopen en na te gaan of dezelfde conclusies als van de MA bereikt worden. Tot slot controleert de VAA ook de outputdocumenten (rekeningen en verslagen) van de MA en de CA.

Bij het uitvoeren van audits worden vandaag meerdere problemen en ongemakken vastgesteld:

- De audits worden pas twee jaar na de effectieve uitgaven uitgevoerd. Dit heeft als gevolg dat het gedeeltelijk recupereren van verworpen kosten moeilijk verloopt of zelfs niet mogelijk is.
- Operationele en financiële audits omvatten meerdere componenten die een terugkerende systematiek bevatten en bijgevolg weinig toegevoegde waarden hebben. Dit repetitief proces is weinig kosteneffectief en verhoogt het risico om inhoudelijke en/of procedurefouten te maken.
- Door de seizoensgebondenheid kunnen niet alle audits uitgevoerd worden met de beschikbare auditoren, met als gevolg dat een aanzienlijk deel van de audits wordt uitbesteed. De kost van deze uitbesteding is relatief hoog, en de opdracht wordt vaak door onervaren auditoren uitgevoerd. Dit heeft als gevolg dat de VAA nadien een zeer tijdrovende kwaliteitscontrole moet uitvoeren.

Om deze problemen op te lossen is het 'CATE'-project (Continuous Auditing based on Techn(olog)ical Evolution and Data Mining) in het leven geroepen. Het project heeft als doel het auditen van projecten volledig te herdenken – zowel op technologisch als conceptueel vlak – beginnende van de mogelijkheden die machine learning en data mining bieden. Het is bijgevolg niet de bedoeling om het huidige proces te automatiseren, maar om conceptueel na te denken over hoe het auditproces kan veranderen en vergemakkelijken. De kernvraag van het project is hoe de VAA tot auditresultaten kan komen op basis van gestructureerde en ongestructureerde data, waarbij maximaal automatisatie nagestreefd wordt. Hierbij wordt ook rekening gehouden met het tijdsaspect dat aangeeft welke oplossingen misschien niet vandaag mogelijk zijn, maar wel in de nabije toekomst. Deze mogelijkheden kunnen vervolgens afgestemd worden met eventuele auditrichtlijnen en -beperkingen.

De objectieven van het project kunnen samengevat worden als volgt:

- Het vroeger, sneller en continu kunnen monitoren/auditen zodat korter op de bal gespeeld wordt en eventuele verworpen uitgaven kunnen gerecupereerd worden.
- Kosteneffectief te zijn door het maximaliseren van nuttig werk en het elimineren van repetitief werk met lage toegevoegde waarde om zo het verlies van (Europese) middelen te beperken.
- Accurate rapporteringen te leveren, van een hoge kwaliteit met grotere zekerheid en meerwaarde, door het automatiseren van het auditproces.

3.2.2 Doelstellingen van het voortraject

Het PIO-voortraject 'CATE' heeft als doelstelling om, op basis van de functionele wensen van de VAA en technologische mogelijkheden, de aanpak uit te werken voor dit traject. Dit houdt in dat concreet bepaald wordt welke scope (i.e. welke use cases) moet nagestreefd worden en hoe hier kan naartoe gewerkt worden in de specifieke context van de VAA.

De doelstellingen van de voorbereidingsfase (behoefte- en marktanalyse) bestaan hoofdzakelijk uit:

- Bepalen van de noden van de eindgebruikers, zijnde de VAA en hun auditoren.

- Aangeven wat de mogelijke technologische oplossingen zijn met hun toegevoegde waarde en risico. Dit risico wordt bepaald door de huidige stand van de technologie en hoe lang het zou duren voor een volledige oplossing mogelijk zou zijn.
- Definiëren van het plan voor een innovatieve aankoop of samenwerking met een derde partij.

De visie van de VAA is om uiteindelijk te werken via twee simultane trajecten: enerzijds een traject binnen het ESF-universum en anderzijds een traject binnen de omgeving Interreg Vlaanderen-Nederland. Dit eindverslag geeft een volledig overzicht van het gevolgde proces, de bekomen informatie en de conclusies van het voortraject.

4 Inschatting innovatiepotentieel vanuit gebruikersstandpunt

4.1 Gevolgd proces

Vanuit de expertise van de projectinitiators werden eerst de verscheidene gebruikersgroepen in kaart gebracht, hier hoofzakelijk bestaande uit de auditoren van de Vlaamse Auditautoriteit van de Europese Structuurfondsen (VAA), maar ook medewerkers van de Managementautoriteit (MA) en de Certificeringsautoriteit (CA). Vervolgens werden de belangrijkste gebruikersvereisten geïdentificeerd vanuit de expertise van deze gebruikersgroepen aan de hand van een niet-exhaustieve lijst van use cases.

De use cases die met de gebruikersgroepen waren opgesteld werden vervolgens geclusterd en verrijkt op een hoger abstractieniveau. Dit werd gedaan om de gewenste visie van de VAA, zijnde het conceptueel herdenken van het auditproces, te bewaken doorheen het voortraject. Bij use cases die zeer low-level blijven zou immers het risico bestaan dat het huidige proces gewoon geautomatiseerd wordt.

Om de use cases te prioriteren op basis van hun toegevoegde waarde werd *planning poker* gebruikt, waarbij de eindgebruikers bepalen hoe belangrijk elke use case is. De scores werden toegekend door dezelfde groep stakeholders als bij het opstellen van de use cases. De laagste waarde in de schaalverdeling is een 0 en betekent “houdt geen meerwaarde in”. Een waarde 20 en hoger dient dan weer geïnterpreteerd te worden als “houdt zeer veel meerwaarde in binnen dit traject”. Een waarde van 100 geeft aan dat de use case absoluut vervuld moet zijn binnen dit project, er is een negatieve impact indien deze use case niet voldaan is.

4.2 Resultaten

Onderstaande lijst presenteert alle use cases met hun bijhorende score van toegevoegde waarde voor de eindgebruiker. In een korte paragraaf worden de details van de use case toegelicht en wordt er aangegeven of de use case ook interessant is voor de MA en/of CA. De use cases zijn verdeeld in 3 categorieën, waarbij steeds een overkoepelende “Epic” use case gevolgd wordt door onderliggende use cases:

- Use cases A betreffen de start van het project.
- Use cases B betreffen de duur van het project.
- Use cases C betreffen het einde van het project.

4.2.1 Use cases A rond de start van het project

Use case Epic A: Als auditor kan ik de projectrisico's inschatten op basis van alle beschikbare informatie bij de start van het project en historische informatie, en aanbevelingen geven voor de zwakke punten in het project zodat ik meteen feedback kan geven om de slaagkansen van het project te verhogen.

Score van het innovatiepotentieel: 20-40

Voorbeelden van projectrisico's zijn dat het geld opraakt bij gedeeltelijke zelffinanciering of dat er niet wordt voldaan aan de deliverables. De informatie die bij de start van het project beschikbaar is bestaat uit het projectvoorstel met de doelstellingen, de partners, geplande deliverables, de financiën, enzovoort.

Use case A1: Als auditor kan ik binnenkomende projecten automatisch classificeren naar de juiste processtroom zodat ik de juiste stappen kan volgen en de juiste checks kan uitvoeren tijdens de audit.

Score van het innovatiepotentieel: 13-20

Het classificeren van binnenkomende projecten gebeurt vandaag op basis van de waarde van het project, de grootte van de organisatie, of het een private of publieke organisatie is, etc. De documenten van binnenkomende projecten zijn gestandaardiseerde vrije formulieren, Excel-bestanden en vrije bijlagen. Met juiste processtroom wordt verwezen naar de focus van de audit, die bijvoorbeeld eerder op overheidsopdrachten of op staatssteun kan gericht zijn.

Use case A2: Als auditor kan ik de huidige regelgeving automatisch inladen, actualiseren en interpreteren zodat ik checks kan uitvoeren op projecten om na te gaan of ze compliant zijn.

Score van het innovatiepotentieel: 40

De regelgeving duidt hier op verordeningen inzake overheidsopdrachten en staatssteun, Europese financieringsregels (CPR) & fonds specifieke regelgeving, de nationale regelgeving, etc.

Use case A3: Als auditor kan ik voor binnenkomende projecten automatisch controleren of alle noodzakelijke documenten in het juiste formaat aanwezig zijn zodat ik meteen ontbrekende elementen kan detecteren en signaleren.

Score van het innovatiepotentieel: 13-20

Het juiste formaat duidt hier op de juiste bestandstypen, of alle gevraagde bijlagen aanwezig zijn, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case A4: Als auditor kan ik voor binnenkomende projecten automatisch controleren of alle data in de juiste vorm aanwezig zijn zodat ik meteen ontbrekende of incorrecte elementen kan detecteren en signaleren.

Score van het innovatiepotentieel: 20

Alle data in de juiste vorm aanwezig houdt in dat alle secties ingevuld zijn, dat alle nodige cijfers zijn ingevuld, het BTW nummer de correcte vorm heeft, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case A5: Als auditor kan ik voor binnenkomende projecten automatisch bepalen hoe haalbaar de deliverables zijn zodat ik meteen onrealistische projecten kan identificeren.

Score van het innovatiepotentieel: 20-40

Deliverables die kunnen nagekeken worden zijn bijvoorbeeld het aantal deelnemers bij de trainingen en workshops, het aantal onderzoeksverslagen en wetenschappelijke publicaties, de realisatie van infrastructuurwerken, etc.

Use case A6: Als auditor kan ik automatisch, op basis van deliverables en andere doelstellingen, vergelijkbare historische projecten vinden zodat ik relevante historische informatie kan gebruiken voor mijn analyse.

Score van het innovatiepotentieel: 20

Relevante historische informatie in deze context kan op projecten slaan waarin gelijkaardige infrastructuurwerken zijn uitgevoerd, projecten die gelijkaardige doelstellingen hadden zoals het activeren van jongeren of de vaardigheden van arbeiders te verbeteren via trainingen.

Use case A7: Als auditor kan ik voor binnenkomende projecten de risico's van een project inschatten, waar ze zich bevinden en wanneer ze zich waarschijnlijk zullen manifesteren zodat ik gerichte controles kan uitvoeren en interventies kan plannen.

Score van het innovatiepotentieel: 40

Projectrisico's zijn bijvoorbeeld dat het geld opraakt bij gedeeltelijke zelffinanciering, dat er niet wordt voldaan aan de deliverables, etc.

Use case A8: Als auditor kan ik automatisch het beheersysteem en de procedures van de management- of certificeringsautoriteit downloaden en controleren of ze de juiste vorm hebben en volledig zijn zodat ik kan controleren waar aanpassingen nodig zijn.

Score van het innovatiepotentieel: 20-40

De juiste vorm in deze context houdt in dat elk document in het juiste formaat staat, dat alle secties ingevuld zijn, dat alle nodige bijlages toegevoegd zijn, etc.

Use case A9: Als auditor kan ik automatisch verifiëren of het beschreven beheersysteem en de procedures van de management- of certificeringsautoriteit in overeenstemming zijn met de regelgeving zodat ik kan controleren waar aanpassingen nodig zijn.

Score van het innovatiepotentieel: 40

Het beheersysteem en de procedures zijn compliant als ze voldoen aan de voorwaarden betreffende functiescheiding, overheidsaanbestedingen, staatssteun, etc. Het beheersysteem is beschikbaar als vast sjabloon met vrije tekst en bijlages, terwijl de procedures doorgaans in vrije tekst omschreven zijn.

4.2.2 Use cases B tijdens de duur van het project

Use case Epic B: Als auditor kan ik mijn inschatting van de projectrisico's updaten op basis van nieuwe inzichten en informatie tijdens projecten, en bijgewerkte aanbevelingen geven voor de zwakke punten in het project zodat ik directe feedback kan geven om de kans op succes van het project te vergroten.

Score van het innovatiepotentieel: 20-40

Nieuwe inzichten en informatie zijn bijvoorbeeld veranderingen in organisatie, vertragingen in deliverables, additionele kostdocumenten, etc.

Use case B1: Als auditor kan ik toegang krijgen tot de projectgegevens en kan ik automatisch mijn analyse starten zodra de bestanden door de projectpartner zijn geüpload in het geïntegreerde beheersysteem van het programma zodat ik snel data van het project kan analyseren, fouten detecteren en identificeren.

Score van het innovatiepotentieel: 20-40

Use case B2: Als auditor kan ik zodra een document is toegevoegd, automatisch de vorm ervan controleren (correct formaat, aanwezige bijlagen) zodat ik ten allen tijde een actueel beeld kan krijgen van de documenten binnen projecten.

Score van het innovatiepotentieel: 20

De documenten kunnen bestaan uit facturen, deelnemersregistraties, trainingcertificaten, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B3: Als auditor kan ik, zodra er data aan een project worden toegevoegd, automatisch de vorm ervan controleren en mijn projectoverzicht bijwerken zodat ik onmiddellijk ontbrekende of onjuiste elementen kan opsporen en signaleren.

Score van het innovatiepotentieel: 13-20

Alle data in de juiste vorm aanwezig houdt in dat alle secties ingevuld zijn, dat alle nodige cijfers zijn ingevuld, het BTW nummer de correcte vorm heeft, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B4: Als auditor kan ik zodra een nieuw document wordt toegevoegd, automatisch controleren of verschillende gegevensbronnen consistente gegevens bevatten zodat ik tegenstrijdigheden kan ontdekken.

Score van het innovatiepotentieel: 20

Gegevens die consistent moeten zijn, zijn bijvoorbeeld de werknemersnamen die moeten overeenstemmen tussen het arbeidscontract en de personeelslijst, de aangegeven kosten die moeten gelijk zijn aan het factureringsbedrag, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B5: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch controleren of de totalen van (numerieke) gegevens over verschillende documenten overeenkomen met de rapporten zodat ik onmiddellijk afwijkingen kan opsporen en een diepgaander onderzoek starten.

Score van het innovatiepotentieel: 40

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B6: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch gegevens vergelijken met historische gegevens (bv. van andere projecten) zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Score van het innovatiepotentieel: 20

Historische data van andere projecten zijn bijvoorbeeld facturen, inzichten rond projectpartners, enzovoort.

Use case B7: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch gegevens vergelijken met gegevens van de Vlaamse overheid (bv. KMO-portefeuille) zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Score van het innovatiepotentieel: 13

Use case B8: Als auditor kan ik zodra de gegevens beschikbaar zijn, gegevens automatisch vergelijken met openbare gegevensbronnen of gegevensbronnen van derden zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Score van het innovatiepotentieel: 13-20

Enkele voorbeelden van publieke gegevensbronnen: Graydon, financiële overzichten, online hulpmiddelen, ...

Use case B9: Als auditor kan ik de ingevulde inhoud van standaard gestructureerde sjablonen automatisch interpreteren van zodra deze beschikbaar zijn zodat ik deze inhoud kan gebruiken voor mijn verdere analyse.

Score van het innovatiepotentieel: 40

Een voorbeeld van standaard gestructureerde sjablonen zijn jaarrekeningen.

Use case B10: Als auditor kan ik de ingevulde inhoud van vrije formulieren automatisch interpreteren zodra deze beschikbaar zijn zodat ik deze inhoud kan gebruiken voor mijn verdere analyse.

Score van het innovatiepotentieel: 40

De ingevulde vrije formulieren zijn bijvoorbeeld projectvoorstellen, staatssteunanalyses, etc.

Use case B11: Als auditor kan ik vrije tekst automatisch interpreteren van zodra deze beschikbaar is zodat ik deze inhoud kan gebruiken voor mijn verdere analyse.

Score van het innovatiepotentieel: 40

De vrije tekst kan bestaan uit procedures van de management- en certificeringsautoriteit, krantenartikelen over het project, etc.

Use case B12: Als auditor kan ik een signaal krijgen in typische (bekende) gevallen van fraude zodra deze zich voordoen zodat ik onmiddellijk op de hoogte word gebracht en de informatie verder kan onderzoeken of doorgeven.

Score van het innovatiepotentieel: 40-100

Deze typische gevallen zijn bijvoorbeeld een groot aantal kleine facturen in een overheidsaanbesteding, mensen in de Raad van Bestuur waarvan verborgen wordt dat ze door de overheid zijn aangewezen om regels rond overheidsaanbestedingen te vermijden.

Use case B13: Als auditor kan ik automatisch controleren of projecten de juiste stappen in hun vooraf gedefinieerde proces volgen zodat ik onmiddellijk op de hoogte ben van eventuele afwijkingen en deze nauwlettend kan opvolgen.

Score van het innovatiepotentieel: 20

Controleren of projecten de juiste stappen volgen houdt in dat de vooropgestelde aanbestedingsplannen gevolgd worden, de mijlpalen in de juiste volgorde afgehandeld worden, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B14: Als auditor kan ik de deliverables in lopende projecten automatisch monitoren en te grote afwijkingen van de doelstellingen opsporen zodat ik onmiddellijk kan herkennen wanneer projecten fout dreigen te gaan en ik ze van dichtbij kan opvolgen.

Score van het innovatiepotentieel: 20-40

Er kunnen bijvoorbeeld grote afwijkingen zijn in het aantal gegeven trainingen en workshops, de vooruitgang van infrastructuurwerken, etc.

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B15: Als auditor kan ik het betrouwbaarheidsniveau van mijn beoordelingen berekenen en deze tijdens de projecten actualiseren met nieuwe informatie zodat ik gericht onderzoek kan doen naar zaken waar ik niet zeker van ben.

Score van het innovatiepotentieel: 20-40

Use case B16: Als auditor kan ik toegang krijgen tot documenten en beoordelingen van de management- en certificeringsautoriteit en automatisch mijn analyse starten zodra zij een definitieve versie op hun server uploaden zodat ik snel fouten kan opsporen en aangeven.

Score van het innovatiepotentieel: 20

De deelnemers van de workshop gaven aan dat deze use case ook voor de MA zeer zinvol is.

Use case B17: Als auditor kan ik automatisch controleren of de output van de management- of certificeringsautoriteit de juiste vorm heeft zodra deze beschikbaar is zodat ik onmiddellijk de ontbrekende items kan identificeren.

Score van het innovatiepotentieel: 13-20

De juiste vorm houdt in dat het document in het correct formaat moet staan, alle vereiste velden aanwezig zijn, etc.

Use case B18: Als auditor kan ik automatisch de output van de managementautoriteit controleren door middel van tests op specifieke projecten en verifiëren of ik tot hetzelfde oordeel kom zodat ik fouten van de managementautoriteit kan vinden en corrigeren.

Score van het innovatiepotentieel: 40

De controle kan bijvoorbeeld bestaan uit het verifiëren of het geplande aantal workshops is bereikt, of infrastructuurwerken voltooid zijn, of kosten in aanmerking komen voor financiering, etc.

Use case B19: Als auditor kan ik automatisch de output van de certificeringsautoriteit (rekeningen met onderliggende Excel-bestanden) controleren door na te gaan of dezelfde resultaten worden verkregen op basis van dezelfde bronnen zodat ik fouten kan vinden en corrigeren.

Score van het innovatiepotentieel: 40

Use case B20: Als auditor kan ik gebruik maken van software die automatisch bijleert wanneer de evaluatie ervan afwijkt van de evaluatie van de managementautoriteit en als onjuist wordt beschouwd zodat audits steeds beter worden en er rekening wordt gehouden met historische inzichten.

Score van het innovatiepotentieel: 20-40

4.2.3 Use cases C aan het einde van het project

Use case Epic C: Als auditor kan ik automatisch de geleerde lessen extraheren en opslaan bij het afsluiten van een project zodat kennis en inzichten uit het verleden niet verloren gaan.

Score van het innovatiepotentieel: 20-40

De lessen worden getrokken uit bijvoorbeeld de projectresultaten, pijnpunten, projectpartners, etc.

Use case C1: Als auditor kan ik bij het afsluiten van een project automatisch controleren of alle correcties en aanbevelingen zijn uitgevoerd zodat er geen open punten overblijven wanneer het project wordt afgesloten.

Score van het innovatiepotentieel: 40

De correcties kunnen bestaan uit het verwijderen van afgekeurde kosten, het bijvoegen van ontbrekende bijlages en documenten, etc.

Use case C2: Als auditor kan ik automatisch een auditrapport en een opinie genereren zodat manueel werk tot een minimum wordt beperkt.

Score van het innovatiepotentieel: 40

Het auditrapport bevat inzichten op de subsidiabiliteit, staatssteun, overheidsaanbesteding, indicatoren, etc. De auditopinie doet dan weer een uitspraak over de systeemefficiëntie.

4.3 Conclusie

Over het algemeen werden zeer hoge scores gegeven vanuit gebruikersstandpunt, wat aangeeft dat de gebruikersgroepen alle use cases belangrijk tot onmisbaar vinden. Bij het herdenken van het

auditproces, zowel technologisch als conceptueel, dienen deze use cases bijgevolg maximaal afgedekt te worden.

Enkele use cases (A3, A4, B2, B3, B4, B5, B13, B14, B16) hebben zelfs nog meer toegevoegde waarde aangezien ze nuttig zijn voor meerdere gebruikersgroepen (VAA en MA). De software implementatie van deze use cases kan mogelijk gedeeld worden tussen de betreffende gebruikersgroepen.

5 State-of-the-art analyse

Vooraleer in dialoog gegaan wordt met marktspelers, werd een high-level overzicht gemaakt van de huidige stand van de technologie in de state-of-the-art analyse. Deze laat toe om de belangrijkste technologieën met hun beperkingen te identificeren.

Het state-of-the-art onderzoek gaat de huidige stand van de techniek na in volgende domeinen:

1. Risk management & assessment
 - Project dashboarding
 - Automatic risk management & assessment
2. Automatisatie & machine learning
 - Robotic process automation (RPA) & Optical character recognition (OCR)
 - Machine learning
 - Voor patroonherkenning
 - Voor Natural language processing (NLP)

Hierbij werd steeds nagegaan wat de implicaties van de huidige stand van de techniek zijn op het 'CATE'-project en wie de belangrijkste marktspelers zijn voor die technologie. Volgende bevindingen komen uit de state-of-the-art analyse naar voor:

- Er zijn platformen beschikbaar om audits/projecten te managen, deze kunnen in veel gevallen aangevuld worden met automatisatie (RPA & OCR) voor repetitieve taken met lage complexiteit.
- Patronen herkennen in (historische) data is mogelijk, maar vereist voldoende gestructureerde en betrouwbare data om het machine learning model efficiënt te trainen.
- Complexere taken (i.e. gerichte NLP of automatic risk assessment) vereisen specifiek onderzoek naar de mogelijkheden en grenzen van de technologie. Hiervoor is verregaande machine learning expertise noodzakelijk.

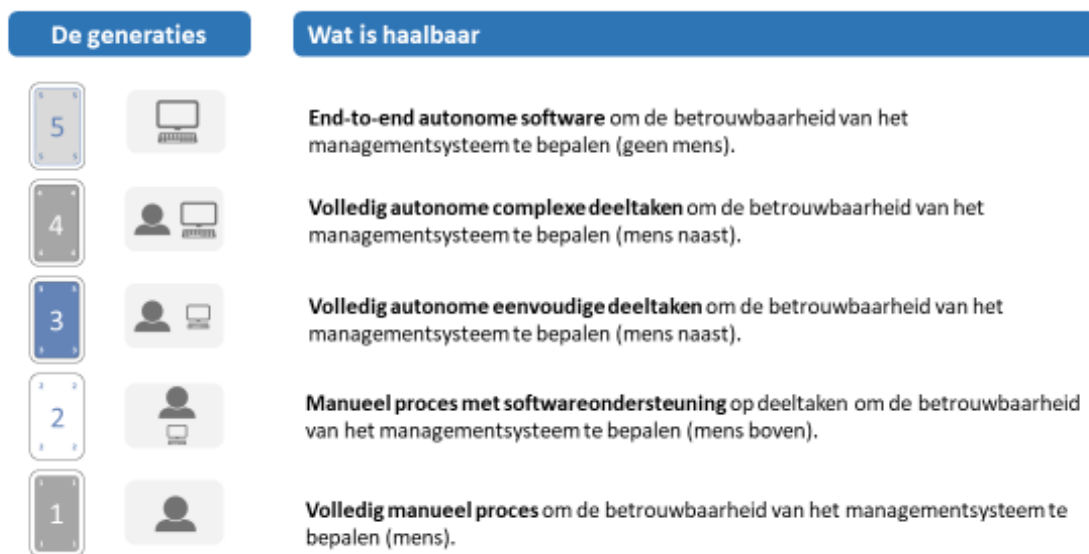
Deze analyse geeft bijgevolg aan dat er nog machine learning (research) noodzakelijk is om verder te gaan dan standaard 'copy-paste' automatisatie. Bovendien moet nagegaan worden of er voldoende data aanwezig zijn om zinvolle modellen te trainen.

Voor een gedetailleerd overzicht van de state-of-the-art analyse wordt verwezen naar Bijlage A. Op basis van de inzichten uit deze analyse werden marktspelers uit de relevante sectoren en marktsegmenten aangespoord om deel te nemen aan de open marktconsultatie om een voldoende ruime en relevante set deelnemende marktpartijen te bekomen.

6 Inschatting van het innovatiepotentieel vanuit technologische invalshoek

6.1 Gevolgd proces

Het innovatiepotentieel vanuit technologisch standpunt wordt ingeschat door middel van *planning poker*. Hierbij schatten experts van industriële partners en kennisinstellingen de technologische risico's in die vasthangen aan elke use case tijdens een marktconsultatie. Het bepalen van het technologische risico voor elke use case verloopt in twee stappen. In een eerste stap geven de experts onafhankelijk van elkaar een generatiescore van 1 t.e.m. 5, die aangeeft wat de huidige stand van de techniek is voor die specifieke use case:



In de tweede stap schatten de experts in hoeveel jaar het zal duren alvorens generatie 5 technisch haalbaar is. Hierbij kiest de expert tussen 1 jaar, 2 jaar, 3 jaar, 4 jaar, 5 jaar of een uitschieter van 10 jaar. De combinatie van de huidige generatie en het aantal jaar alvorens generatie 5 mogelijk is geeft een duidelijk beeld van de expert over het technologische risico dat een specifieke use case bevat.

De marktconsultatie vond plaats op 4 september 2020, onder de vorm van een MS Teams meeting. Ongeveer 70 deelnemers hadden zich vooraf aangemeld; een 45-tal mensen namen effectief deel waarbij sommige bedrijven of organisaties meerdere personen afvaardigden.

De volledige lijst van deelnemende partijen in de marktconsultatie kan gevonden worden in bijlage B.

6.2 Resultaten

Onderstaande lijst presenteert alle use cases met hun bijhorende generatiescore en het aantal jaar tot generatie 5 mogelijk is. In een korte paragraaf worden voor elke use case de belangrijke punten in de discussie van de experts toegelicht. Hierbij wordt voornamelijk gefocust op de argumenten die werden aangehaald voor een lagere of hogere score dan het gemiddelde.

6.2.1 Use cases A rond de start van het project

Use case Epic A: Als auditor kan ik de projectrisico's inschatten op basis van alle beschikbare informatie bij de start van het project en historische informatie, en aanbevelingen geven voor de zwakke punten in het project zodat ik meteen feedback kan geven om de slaagkansen van het project te verhogen.

Huidige generatie: 2 – 3

De consensus onder de deelnemers was dat het risico hier sterk afhangt van de mate van structuur in de inputdata (kwaliteit van data, structuur van de tekst, subjectieve nuances in de tekst, etc.). Doordat ongestructureerde data vaak variabel is, was de algemene opinie dat een menselijke expert idealiter in de loop blijft. Verschillen in toegekende generaties kwamen hoofdzakelijk door verschillende interpretaties van de randvoorwaarden. Deelnemers die een lagere score gaven haalden aan dat Natural Language Processing (NLP) technologie vandaag nog niet in staat is om nuances in tekst te capteren en dat daarom het proces vooral nog manueel is. Enkele deelnemers kenden een hogere generatie toe omdat deze use case perfect mogelijk is als de binnenkomende documenten voldoende gestructureerd en genormaliseerd zijn. Bovendien haalden enkele experts aan een soortgelijke applicatie geïmplementeerd te hebben bij diverse klanten.

Aantal jaren tot generatie 5: twee clusters: 3 – 4 jaar en 10 jaar

Naar toekomstperspectief toe werden twee clusters onderscheiden. Deelnemers die een kortere periode tot generatie 5 hebben geantwoord haalden aan dat de technologie klaar is maar dat de tijd vooral afhangt van het selecteren van de juiste technologie. Hierbij werd aangegeven dat mate van structuur in de input documenten het proces kan verkorten. Deelnemers die een langere periode aangehaald hebben gaven aan dat er vooral veel relevante historische data nodig zijn.

Use case A1: Als auditor kan ik binnenkomende projecten automatisch classificeren naar de juiste processtroom zodat ik de juiste stappen kan volgen en de juist checks kan uitvoeren tijdens de audit.

Huidige generatie: 4

Over het algemeen werd het risico laag ingeschat en lag het gemiddelde op generatie 4. Deelnemers die een lagere generatiescore gaven motiveerden dit op twee manieren. Enerzijds is er het risico dat machine learning modellen overfitten op de specifieke layout van ongestructureerde documenten, waardoor het model moeilijk veralgemeent naar andere documenten. Anderzijds kan het gebrek aan historische data efficiënte automatisatie verhinderen. Zo kan bijvoorbeeld een nieuwe process flow of een andere dynamiek met een nieuwe partner ervoor zorgen dat er niet genoeg historische data zijn voor het model om de nieuwe vorm van input correct te classificeren. Deelnemers die een hogere score gaven, haalden aan applicaties geïmplementeerd te hebben die projecten automatisch classificeren. De uitdaging bij deze use case ligt niet in de ongestructureerde data, maar in het toevoegen van nieuwe process flows zonder historische data.

Aantal jaren tot generatie 5: 3 jaar

Naar toekomstperspectief toe lag het gemiddelde op 3 jaar tot end-to-end automatisatie. Deze periode tot de ideale situatie is vooral afhankelijk van de verbeteringen van NLP. De experts die een kortere periode antwoordden gaven aan dat NLP ver genoeg stond voor deze use case en dat enkel nieuwe process flows voor problemen zouden kunnen zorgen. De experts die een langere periode antwoordden haalden aan dat het sterk afhangt van de structuur en taal van de binnenkomende documenten, NLP moet nog verder ontwikkeld worden.

Use case A2: Als auditor kan ik de huidige regelgeving automatisch inladen, actualiseren en interpreteren zodat ik checks kan uitvoeren op projecten om na te gaan of ze compliant zijn.

Huidige generatie: 2

Volgens het merendeel van de deelnemers moet de mens deze use case nog steeds leiden bij de huidige stand van de technologie. De use case is in de antwoorden opgedeeld in 2 delen omdat ze verschillen van moeilijkheidsgraad: enerzijds het inladen en actualiseren, anderzijds het interpreteren. Het automatisch inladen en actualiseren is vandaag zeker mogelijk. Het interpreteren van de regelgeving en daaruit conclusies trekken valt nog niet binnen de mogelijkheden van NLP. Eén van de partijen gaf een hogere generatiescore en geeft aan een actief programma te hebben dat deze zaken kan voor voldoende eenvoudige regelgevingen en voldoende hoge datakwaliteit.

Aantal jaren tot generatie 5: 5-10 jaar

Naar toekomstperspectief toe heeft het merendeel van de experts geantwoord met een periode van 5 tot 10 jaar vooraleer deze use case end-to-end geautomatiseerd kan worden. De technologie is vandaag in staat om veranderingen te detecteren, deze te markeren, maar is nog niet in staat om de verandering te begrijpen. Over de nodige tijd om de automatische interpretatie te verbeteren zijn de meningen van de deelnemers verdeeld. Enkele deelnemers die een kortere periode antwoordden gaven aan dat dit binnen enkele jaren mogelijk is. De deelnemers die een langere periode antwoordden vonden de limiterende factoren vooral NLP technologie en het volledig actueel houden van de regelgeving die vaak wijzigt.

Use case A3: Als auditor kan ik voor binnenkomende projecten automatisch controleren of alle noodzakelijke documenten in het juiste formaat aanwezig zijn zodat ik meteen ontbrekende elementen kan detecteren en signaleren.

Huidige generatie: 5

Volgens de meeste deelnemers is de automatisatie van deze use case perfect mogelijk zonder menselijke interventie. De deelnemers die een lagere score gaven argumenteerden dat sommige documenten nog steeds manuele checks zullen nodig hebben. Als je weet welk formaat te verwachten is het een simpele use case, maar er is soms variabiliteit in het formaat vanwege de inhoud van het document. Enkele experts twijfelden tussen generatie 4 en 5 omdat het toch een belangrijke taak is en dat er altijd een finale menselijke controle nodig is. Andere deelnemers gaven dan weer aan dat ze een volledig autonome applicatie actief bij een klant hebben lopen.

Aantal jaren tot generatie 5: /

Use case A4: Als auditor kan ik voor binnenkomende projecten automatisch controleren of alle data in de juiste vorm aanwezig zijn zodat ik meteen ontbrekende of incorrecte elementen kan detecteren en signaleren.

Huidige generatie: 4 - 5

Over het algemeen werd deze use case met een lage risicoscore ingeschat. De enige redenen waarom de use case niet volledig automatiseerbaar zou zijn, zijn hoofdzakelijk variabiliteit in de specificaties van de data en de hoeveelheid van historische data die ter beschikking is. Duidelijke specificaties en historische data zijn cruciaal. De deelnemers die lagere scores gegeven hebben argumenteerden dat er vaak velden zijn die niet ingevuld moeten worden en dat de notatie van niet-genormaliseerde data soms kan verschillen volgens de specificaties. Veel experts gaven score 5 met de opmerking dat ze actieve applicaties hebben bij klanten.

Aantal jaren tot generatie 5: 1 jaar

Gemiddeld werd de tijd tot generatie 5 geschat op een korte periode van 1 jaar. Toch zijn er enkele experts die langer hebben geantwoord, omdat er volgens hen nog lang een mens gaat moeten zijn die uitzonderingen behandelt die de machine niet kan verwerken.

Use case A5: Als auditor kan ik voor binnenkomende projecten automatisch bepalen hoe haalbaar de deliverables zijn zodat ik meteen onrealistische projecten kan identificeren.

Huidige generatie: 3

De consensus tussen de experts is dat de stand van de huidige technologie in staat is om de mens bij te staan door simpele deeltaken automatisch uit te voeren maar dat de mens nog steeds de complexere beslissingen moet nemen. Deelnemers die een lagere generatiescore gaven haalden aan dat de beslissing of deliverables haalbaar zijn vaak van meerdere factoren afhangt. Sommige projecten met dezelfde specificaties in de projectdocumenten hebben niet dezelfde haalbaarheid wegens de achterliggende context van het bedrijf (e.g., de grootte van het bedrijf). Deze factoren beïnvloeden de complexiteit van de beslissing dat het model moet nemen. Deelnemers die een hogere score gaven argumenteerden dat het met de technologie van vandaag al mogelijk is om die factoren in rekening te brengen met data-driven of rule-driven modellen.

Aantal jaren tot generatie 5: 4 jaar

Naar toekomstperspectief toe werd gemiddeld geschat dat het 4 jaar zou duren vooraleer deze use case end-to-end geautomatiseerd kan worden. Hierbij werd aangegeven dat generatie 4 snel bereikt kan worden maar dat volledige automatisatie moeilijk is door de 5% van de gevallen waar het een complexe beslissing is.

Use case A6: Als auditor kan ik automatisch, op basis van deliverables en andere doelstellingen, vergelijkbare historische projecten vinden zodat ik relevante historische informatie kan gebruiken voor mijn analyse.

Huidige generatie: 4-5

Volgens het merendeel van de experts is het technisch haalbaar om deze use case volledig te automatiseren of toch complexe deeltaken door een machine te laten uitvoeren. De enige moeilijkheid die volledige automatisatie kan verhinderen is dat de deliverables gespecificeerd kunnen zijn door vrije tekst. Dit aspect werd extra benadrukt door deelnemers die een lagere score gaven. Enkele deelnemers met een hogere score gaven aan dat ze een dergelijk model reeds intern gebruiken of al geïmplementeerd hebben bij klanten.

Aantal jaren tot generatie 5: 1-2 jaar

De algemene conclusie is dat vele deelnemers al op generatie 5 zitten en dat andere inschatten dat het niet lang zou duren voor generatie 5 bereikt wordt.

Use case A7: Als auditor kan ik voor binnenkomende projecten de risico's van een project inschatten, waar ze zich bevinden en wanneer ze zich waarschijnlijk zullen manifesteren zodat ik gerichte controles kan uitvoeren en interventies kan plannen.

Huidige generatie: 3

De technische haalbaarheid aangegeven door experts lag gemiddeld op generatie 3. Basismodellen zijn in staat om informatie te vinden die de auditor kunnen helpen om de risico's in te schatten. Meer geavanceerde modellen kunnen zelf risicoscores berekenen om vervolgens suggesties aan de

auditor te geven. De eindbeslissing rond de impact van de risico's moet vandaag dus nog bij de mens liggen.

Aantal jaren tot generatie 5: 5 jaar

De vereiste tijd tot volledige end-to-end automatisatie van deze use case wordt als relatief lang (5 jaar) ingeschat. Dit is hoofdzakelijk door de randgevallen die vaak moeilijk zijn uit te voeren door machines. Deelnemers die een kortere periode aangaven baseerden zich op hun actieve modellen die al in staat zijn om risico's in te schatten. Het bepalen wanneer de risico's zich zullen manifesteren wordt door hen als de grootste uitdaging gezien. Enkel historische data kunnen hierbij helpen. Enkele deelnemers die 10 jaar antwoordden gaven aan dat het geen ideale eindsituatie is om dit helemaal te automatiseren, je wilt steeds een mens die nagaat waarom een hoog risico is gedetecteerd. Bovendien kan de complexiteit van randgevallen zeer hoog zijn.

Use case A8: Als auditor kan ik automatisch het beheersysteem en de procedures van de management- of certificeringsautoriteit downloaden en controleren of ze de juiste vorm hebben en volledig zijn zodat ik kan controleren waar aanpassingen nodig zijn.

Huidige generatie: 4-5

De conclusie van de experts was dat deze use case in sterke mate automatiseerbaar is. Deelnemers die een lagere score gaven hadden voornamelijk zorgen over de mate van formaatstandaardisatie die afgedwongen kan worden van de management- en certificeringsautoriteit. Als er steeds een vast formaat wordt gebruikt kan er veel geautomatiseerd worden, zo niet wordt de use case complexer om uit te voeren. Er zijn enkele deelnemers die aangaven dat de use case vrij voor de hand liggend was en dat ze al soortgelijke applicaties hadden geïmplementeerd.

Aantal jaren tot generatie 5: 1 jaar

De meeste aanwezige bedrijven zijn al dicht bij generatie 5 en de andere geven aan dat in minder dan 1 jaar generatie 5 mogelijk zou moeten zijn, er van uitgaande dat een consistent formaat kan opgelegd worden aan de betreffende autoriteiten.

Use case A9: Als auditor kan ik automatisch verifiëren of het beschreven beheersysteem en de procedures van de beheers- of certificeringsautoriteit in overeenstemming zijn met de regelgeving zodat ik kan controleren waar aanpassingen nodig zijn.

Huidige generatie: 2

Over het algemeen werd de huidige stand van de technologie op generatie 2 ingeschat. Om na te kijken of procedures in overeenstemming zijn met de regelgeving is de grootste limiterende factor veruit de huidige staat van NLP. Er is veel menselijke ondersteuning nodig om te bepalen of het systeem compliant is. NLP is vandaag enkel in staat aan te duiden waar interessante informatie zich bevindt om deze beslissing te nemen. Er is ook altijd enige subjectiviteit in de regelgeving die moeilijk te interpreteren is voor machines.

Aantal jaren tot generatie 5: 5-10 jaar

De periode om volledige automatisatie te bereiken werd door het gros van de experts als lang ingeschat. De hoofdreden hiervoor is dat de huidige staat van NLP niet geavanceerd genoeg is om de volledige use case te automatiseren. Enkele deelnemers die 10 jaar hebben geantwoord geven aan dat er altijd een klein percentage van het proces gaat zijn waar menselijke interactie noodzakelijk is wegens de variabiliteit van de procedures en de regelgeving. Eén deelnemer haalde zelfs aan dat NLP misschien niet in de nabije toekomst tot het nodige niveau zal geraken omdat de

computationele en ecologische impact van de vereiste berekeningen te groot is om praktisch toepasbaar te zijn. Deze deelnemer geeft bijgevolg aan dat dit probleem eerder vanuit een perspectief van data acquisitie moet opgelost worden.

6.2.2 Use cases B tijdens de duur van het project

Use case Epic B: Als auditor kan ik mijn inschatting van de projectrisico's updaten op basis van nieuwe inzichten en informatie tijdens projecten, en bijgewerkte aanbevelingen geven voor de zwakke punten in het project zodat ik directe feedback kan geven om de kans op succes van het project te vergroten.

Huidige generatie: 3

Over het algemeen was de consensus tussen de experts dat deze use case een complex model vereist. Deelnemers die een lagere score gaven waren hoofdzakelijk van mening dat het niet de bedoeling mag zijn om de project manager te vervangen, maar enkel de persoon te ondersteunen in het proces. Hoe dan ook is de hoeveelheid data die nodig is om een accuraat model te verkrijgen enorm. Deelnemers die een hogere generatiescore gaven, gaven aan dat de use case mogelijk is met de nieuwste state-of-the-art modellen. Het model zou een extensie zijn van het model in use case A7 met meer features voor het updaten van de risicoscores. Technisch zijn er meerdere manieren aangehaald om het probleem op te lossen: reinforcement learning, deep learning, regression learning, belief revision gebruikmakende van een bayesian approach, etc.

Aantal jaren tot generatie 5: 5 jaar

De algemene consensus onder de deelnemers was dat het relatief lang ging duren om volledige automatisatie te bereiken. Zo is geavanceerd onderzoek nodig om een computermodel te ontwikkelen dat in staat is om eigenhandig projectrisico's te updaten aan de hand van nieuwe informatie.

Use case B1: Als auditor kan ik toegang krijgen tot de projectgegevens en kan ik automatisch mijn analyse starten zodra de bestanden door de projectpartner zijn geüpload in het geïntegreerde beheersysteem van het programma zodat ik snel data van het project kan analyseren, fouten detecteren en identificeren.

Huidige generatie: 5

In het algemeen waren de experts het eens dat deze use case end-to-end te automatiseren is met de huidige technologie. Een deelnemer die een lagere score gaf haalde aan dat de use case zeker technisch haalbaar is maar dat de moeilijkheid voor volledige automatisatie afhangt van hoe de projectgegevens zijn geformatteerd. Voor het gros van de deelnemers was deze use case vrij standaard en simpel te implementeren. Meerdere experts refereerden naar geïmplementeerde applicaties bij klanten.

Aantal jaren tot generatie 5: /

Use case B2: Als auditor kan ik zodra een document is toegevoegd, automatisch de vorm ervan controleren (correct formaat, aanwezige bijlagen) zodat ik ten allen tijde een actueel beeld kan krijgen van de documenten binnen projecten.

Huidige generatie: 5

Alle deelnemers gaven aan dat deze use case technisch haalbaar is en vandaag al volledig autonoom geïmplementeerd wordt. De use case is uiteindelijk een kleine extensie van use case A3.

Aantal jaren tot generatie 5: /

Use case B3: Als auditor kan ik, zodra er data aan een project worden toegevoegd, automatisch de vorm ervan controleren en mijn projectoverzicht bijwerken zodat ik onmiddellijk ontbrekende of onjuiste elementen kan opsporen en signaleren.

Huidige generatie: 4-5

De experts waren akkoord dat de huidige staat van de technologie toelaat om deze use case volledig te automatiseren. De enige deelnemers die lagere scores gaven argumenteerden dat in de specifieke situatie van het controleren van compliance, de use case iets complexer wordt.

Aantal jaren tot generatie 5: 1 jaar

Omdat de huidige situatie al dicht tegen generatie 5 aanleunt, wordt de tijd tot volledige automatisatie zeer kort ingeschat. Er zijn geen zware barrières te doorbreken.

Use case B4: Als auditor kan ik zodra een nieuw document wordt toegevoegd, automatisch controleren of verschillende gegevensbronnen consistente gegevens bevatten zodat ik tegenstrijdigheden kan ontdekken.

Huidige generatie: 4

De algemene consensus tussen de experts is dat complexe deeltaken van deze use case technisch goed te automatiseren zijn. Een mogelijke limiterende factor hierbij is het formaat van de gegevensbronnen. Deelnemers die een lagere score gaven haalden aan dat voor het nagaan van regelgeving er altijd een mens nodig is in de loop. Een applicatie kan vandaag immers informatie extraheren met een bepaald vertrouwensniveau maar kan geen 100% accuraatheid garanderen. Deelnemers met een hogere score gaven aan dat als de templates van de gegevensbronnen gebruiksvriendelijk zijn en er al veel bronnen zijn opgeslagen, deze use case dan geen moeilijkheden inhoudt.

Aantal jaren tot generatie 5: 2-3 jaar

Wat door alle deelnemers werd aangehaald is dat de tijd tot volledige automatisatie sterk afhangt van de vereiste accuraatheid. De aanvaarde foutmarge binnen audits is gespecificeerd als 2%. Na deze informatie werd de consensus van 2-3 jaar bereikt bij de deelnemers. Het is vrij evident om tot 95% accuraatheid te geraken bij automatisatie, het moeilijke en tijdrovende deel is die laatste 5%.

Use case B5: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch controleren of de totalen van (numerieke) gegevens over verschillende documenten overeenkomen met de rapporten zodat ik onmiddellijk afwijkingen kan opsporen en een diepgaander onderzoek starten.

Huidige generatie: 4

De huidige stand van de technologie werd door de deelnemers als voldoende geavanceerd beschouwd om complexe deeltaken van deze use case te automatiseren. Als de gegevens echter ongestructureerd zijn is het aanzienlijk complexer om end-to-end automatisatie te realiseren. Deelnemers die een lage score gaven haalden aan dat numerieke gegevens markeren eenvoudig is maar het identificeren en/of berekenen van totalen een stuk moeilijker. Als de gegevens gecategoriseerd zijn in een template dan is het al veel haalbaarder om de use case volledig te automatiseren. Deelnemers die een hoge score gaven namen de assumptie dat de numerieke gegevens duidelijk gestructureerd zijn in categorieën. Als deze gegevens in vrije tekst verwerkt staan dan wordt de interpretatie ervan een graad moeilijker.

Aantal jaren tot generatie 5: 3

Het gemiddelde van het aantal jaren tot volledige end-to-end automatisatie kwam neer op 3 jaar. Het zijn voornamelijk randgevallen die ervoor zorgen dat enkele experts een langere periode hebben aangehaald. Ongestructureerde en niet-gecategoriseerde numerieke gegevens zullen nog lange tijd voor modellen moeilijk blijven om te controleren.

Use case B6: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch gegevens vergelijken met historische gegevens (bv. van andere projecten) zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Huidige generatie: 4

Over het algemeen was de consensus van de experts dat de technologie zelf voldoende ver staat om complexe deeltaken van deze use case te automatiseren. Deelnemers die een lagere score gaven haalden aan dat er problemen kunnen ontstaan bij historische data die variëren in formaat van vrije tekst. Het is mogelijk om met de huidige technologie soortgelijke documenten terug te vinden en te markeren waar de gelijkenissen en verschillen zitten, maar er is nog steeds een mens nodig om de resultaten van het model na te kijken en om anomalieën te bevestigen. Deelnemers die een hogere score gaven veronderstelden dat historische data van goede kwaliteit aanwezig is. In dit geval is het een eenvoudige use case want de documenten zijn voordien geanalyseerd en gestructureerd. Enkele deelnemers gaven aan een werkende applicatie geïmplementeerd te hebben bij klanten.

Aantal jaren tot generatie 5: 3

Gemiddeld schatten de experts in dat het 3 jaar zou duren voor generatie 5 bereikt zou worden. Opnieuw is de aanwezigheid van vergelijkbare historische data de voornaamste factor die invloed heeft op de tijd. Goede historische data verlagen de vereisten van het NLP model.

Use case B7: Als auditor kan ik zodra de gegevens beschikbaar zijn, automatisch gegevens vergelijken met gegevens van de Vlaamse overheid (bv. KMO-portefeuille) zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Huidige generatie: 4

De bevroegde experts concludeerden dat complexe deeltaken van deze use case technisch volledig te automatiseren zijn. Dit in de veronderstelling dat er een efficiënte API bestaat waarmee de auditor aan de meest actuele gegevens kan van de Vlaamse overheid. Deelnemers die een lagere score gaven merkten op dat het een uitdaging is om de data van de Vlaamse overheid actueel te houden en dat dit sterk afhangt van de interface naar die data. Deelnemers die een hogere score gaven zeiden dat het model van use case B6 hier ook gebruikt kan worden. Het enige verschil tussen beide use cases is waar de data vandaan komen (KMO-portefeuille t.o.v. historische gegevens). Daarnaast gaven ze wel aan dat het inderdaad belangrijk is om een efficiënte API ter beschikking te hebben om aan de gegevens van KMO te kunnen.

Aantal jaren tot generatie 5: 1-2 jaar

Over het algemeen zeiden de experts dat het niet lang zou duren om end-to-end automatisatie van deze use case te bereiken aangezien het vooral zou afhangen van de beschikbaarheid van een efficiënte API. Een deelnemer die een hogere score gaf argumenteerde dat het langer kan duren als de data van de overheid ongestructureerd zouden zijn of niet volledig beschikbaar zouden zijn.

Use case B8: Als auditor kan ik zodra de gegevens beschikbaar zijn, gegevens automatisch vergelijken met openbare gegevensbronnen of gegevensbronnen van derden zodat ik anomalieën kan opsporen en daaruit conclusies trekken.

Huidige generatie: 4

Vanuit technologisch perspectief is deze use case vergelijkbaar met de vorige. De use case is volledig te automatiseren, de enige vereiste is een goed ontwikkelde API die de openbare gegevensbronnen gestructureerd binnenbrengt. Deelnemers die een lagere score gaven haalden aan dat de automatisatie van deze use case mogelijk is maar dat er nog steeds assistentie van een persoon nodig is als de publieke gegevensbron ongestructureerd is. Deelnemers die een hogere score gaven vergeleken de use case met de vorige en bekeken het probleem meer als een data engineering probleem. Het gaat immers over het ontwikkelen van een goeie API zodat de openbare gegevensbronnen gestructureerd binnenkomen. Enkele deelnemers gaven nog als opmerking dat je het gebruik van derde partijen het best vermijdt omdat de data mogelijk niet correct zijn.

Aantal jaren tot generatie 5: 3 jaar

De deelnemers schatten dat er ongeveer 3 jaar nodig is om end-to-end automatisatie te bereiken in deze use case. Hierbij werd vooral gefocust op de tijd die nodig is om al de nodige APIs op punt te stellen voor alle gebruikte (en mogelijk ongestructureerde) gegevensbronnen van derden.

Use case B9: Als auditor kan ik de ingevulde inhoud van standaard gestructureerde sjablonen automatisch interpreteren van zodra deze beschikbaar zijn zodat ik deze inhoud kan gebruiken voor mijn verdere analyse.

Huidige generatie: 4-5

Over het algemeen is interpretatie van gestructureerde sjablonen volgens het gros van de experts technisch te automatiseren. Er bestaan enkel nuances rond de diepgang van de vereiste interpretatie, gezien deze kan variëren van een samenvatting van de context tot diepe conclusies.

Aantal jaren tot generatie 5: /

Use case B10: Als auditor kan ik de ingevulde inhoud van vrije formulieren automatisch interpreteren zodra deze beschikbaar zijn zodat ik deze inhoud kan gebruiken voor mijn verdere analyse.

Huidige generatie: 3

De deelnemers gaven een gemiddelde generatiescore van 3, wat betekent dat simpele deeltaken autonoom kunnen uitgevoerd worden. Deelnemers die een lagere generatiescore gaven haalden aan dat het interpreteren van vrije tekst en daaruit conclusies trekken nog steeds een zeer complexe taak is voor een machine en dat dit hoofdzakelijk door een mens wordt uitgevoerd. Deelnemers die een hogere score gaven focusten eerder op het beschrijven en samenvatten van vrije tekst en haalden aan dat ze al modellen hebben geïmplementeerd die dit kunnen.

Aantal jaren tot generatie 5: twee clusters: 2 jaar en 5-10 jaar

Bij deze use case werden twee clusters van antwoorden onderscheiden door verschillende interpretaties van de experts over hoe ver de interpretatie moet gaan. Deelnemers die een kortere periode gaven dachten aan basisinterpretatie van context. Deelnemers die een langere periode gaven dachten aan een diepere interpretatie om conclusies te trekken uit de vrije tekst.

Use case B11: Als auditor kan ik vrije tekst automatisch interpreteren van zodra deze beschikbaar is zodat ik de inhoud kan gebruiken voor mijn verdere analyse.

Huidige generatie: 2-3

Het gros van de deelnemers gaf een lage score omdat deze use case nog complexer is dan de vorige. De huidige staat van NLP is niet geavanceerd genoeg om voor volledige automatisatie te zorgen. Een deelnemer met een hogere score gaf aan dat ze actieve programma's hebben die in staat zijn om de context, sleutelwoorden en samenvatting te extraheren uit vrije tekst. Het interpreteren en begrijpen van procedures is echter veel complexer.

Aantal jaren tot generatie 5: 5-10 jaar

De limiterende factor volgens het merendeel van de experts is NLP technologie. Wegens het vele actief onderzoek in NLP met relatief weinig grote resultaten schatten de experts dat het nog lang zal duren alvorens NLP op punt staat voor deze use case.

Use case B12: Als auditor kan ik een signaal krijgen in typische (bekende) gevallen van fraude zodra deze zich voordoen zodat ik onmiddellijk op de hoogte word gebracht en de informatie verder kan onderzoeken of doorgeven.

Huidige generatie: 4

De experts waren het eens dat er volledige autonome modellen bestaan die in staat zijn om fraude te signaleren. De voornaamste reden dat enkele experts een lagere score gaven was dat de use case complexer wordt als er interpretatie bij komt kijken. Deelnemers die een lagere score gaven, gaven aan dat de technologie voor fraudedetectie beschikbaar is maar dat info van publieke instanties soms ingewikkelder is om te onderzoeken. Bovendien gaat technologie voor fraudedetectie vaak enkel de informatie vergaren om niet-geïnterpreteerd door te geven aan een persoon. Eén deelnemer met een hogere score gaf aan dat ze een gelijkaardig systeem hebben bij een bank dat volledig automatisch fraude gevallen signaleert. De accuraatheid van dit model is tot 10 maal beter als de mens. Het model kan bovendien bredere gevallen aan dan enkel transacties.

Aantal jaren tot generatie 5: 3

De schatting van het aantal jaar tot end-to-end automatisatie is gemiddeld 3 jaar. Deelnemers die een kortere periode gaven gingen uit van een model dat in staat is om info rond fraudedetectie te verzamelen en te signaleren. Deelnemers die een langere periode gaven gingen uit van een model dat daarenboven in staat is om een interpretatie door te geven.

Use case B13: Als auditor kan ik automatisch controleren of projecten de juiste stappen in hun vooraf gedefinieerde proces volgen zodat ik onmiddellijk op de hoogte ben van eventuele afwijkingen en deze nauwlettend kan opvolgen.

Huidige generatie: 3

Deze use case kreeg een gemiddelde generatiescore van 3. Dit is echter niet omdat de technologie niet in staat is om deze toepassing volledig te automatiseren. Het probleem zit in het frequent, continu ontvangen van informatie over het gevolgde proces. Deelnemers die een lagere score gaven haalden immers aan dat de uitdaging niet zit in de technologie maar eerder in de manier waarop processen gemonitord worden gezien de meeste rapporteringen van projecten periodiek zijn. Deelnemers die een hogere score gaven haalden aan dat het *timestampen* van processen om een bepaald pad te monitoren zeker mogelijk is. Het meest ingewikkelde deel is op frequente basis informatie van het project te vergaren.

Aantal jaren tot generatie 5: 5 jaar

Gemiddeld kwamen de deelnemers tot de conclusie dat het toch een periode van 5 jaar zou duren voor deze use case volledig te automatiseren is. De huidige situatie is nog ver verwijderd van het continu monitoren van projectprocedures en is voornamelijk nog periodiek.

Use case B14: Als auditor kan ik de deliverables in lopende projecten automatisch monitoren en te grote afwijkingen van de doelstellingen opsporen zodat ik onmiddellijk kan herkennen wanneer projecten fout dreigen te gaan en ik ze van dichtbij kan opvolgen.

Huidige generatie: 3-4

Automatiseren van simpele en complexe deeltaken met een model is technisch haalbaar. De mate van automatisatie die mogelijk is hangt vooral af van de structuur van de tekst die de deliverables beschrijft. Deelnemers die een lagere score gaven haalden aan dat de betreffende rapporten vaak ongestructureerd zijn. Een model kan belangrijke info extraheren maar zelfstandige interpretatie is vandaag nog moeilijk. Deelnemers die een hogere score gaven namen hierbij aan dat de datainput manueel wordt gedaan, waardoor de use case simpel wordt. Als de input van de informatie ook moet geautomatiseerd worden is het ontwerp van een model veel moeilijker. Het monitoren zelf is echter wel technisch haalbaar.

Aantal jaren tot generatie 5: 5 jaar

Het volledig automatiseren van deze use case hangt af van de mate van structuur in de documenten die de deliverables beschrijven. De deelnemers gaven aan dat 5 jaar evolutie in de technologie een veilige periode is om ook minder gestructureerde documenten te kunnen verwerken.

Use case B15: Als auditor kan ik het betrouwbaarheidsniveau van mijn beoordelingen berekenen en deze tijdens de projecten actualiseren met nieuwe informatie zodat ik gericht onderzoek kan doen naar zaken waar ik niet zeker van ben.

Huidige generatie: 3

De experts gaven aan dat met de technologie van vandaag het technisch mogelijk is om betrouwbaarheidsniveaus te berekenen en te actualiseren met nieuwe informatie. De mate van mogelijke automatisatie hangt vooral af van de gewenste complexiteit in de waarde. Deelnemers die een lagere score gaven hadden vragen over de definitie van betrouwbaarheidsniveau en in hoeveel detail er moet gegaan worden. Het niveau van detail bepaalt immers het aantal inzichten die je moet verwerken in het betrouwbaarheidsniveau, en daarmee de moeilijkheidsgraad van de use case. Deelnemers die een hogere score gaven argumenteerden dat het autonoom berekenen en actualiseren van betrouwbaarheidsniveaus mogelijk is met de technologie van vandaag. Een model zorgt voor de initiële beoordeling en wordt dan nagekeken door een mens.

Aantal jaren tot generatie 5: 3-4 jaar

Gemiddeld kwam men tot de conclusie dat het een periode van 3-4 jaar zou duren voor deze use case volledig te automatiseren is. Dit is hoofdzakelijk afhankelijk van de gevraagde diepgang van het betrouwbaarheidsniveau.

Use case B16: Als auditor kan ik toegang krijgen tot documenten en beoordelingen van de management- en certificeringsautoriteit en automatisch mijn analyse starten zodra zij een definitieve versie op hun server uploaden zodat ik snel fouten kan opsporen en aangeven.

Huidige generatie: 5

De technologie van vandaag is in staat om bovenstaande use case end-to-end te automatiseren.

Aantal jaren tot generatie 5: /

Use case B17: Als auditor kan ik automatisch controleren of de output van de management- of certificeringsautoriteit de juiste vorm heeft zodra deze beschikbaar is zodat ik onmiddellijk de ontbrekende items kan identificeren.

Huidige generatie: 5

De technologie van vandaag is in staat om bovenstaande use case end-to-end te automatiseren.

Aantal jaren tot generatie 5: /

Use case B18: Als auditor kan ik automatisch de output van de managementautoriteit controleren door middel van tests op specifieke projecten en door te verifiëren of ik tot hetzelfde oordeel kom zodat ik fouten van de managementautoriteit kan vinden en corrigeren.

Huidige generatie: 3-4

De technische haalbaarheid van deze use case hangt af van assumpties rond de structuur van de templates, de type checks die je wilt uitvoeren en in hoeverre je wil nakijken voor gemaakte fouten. Gemiddeld werd geconcludeerd dat de technologie van vandaag in staat is om de mens te assisteren door simpele tot complexe deeltaken uit te voeren. Deelnemers die een lagere score gaven haalden aan dat de moeilijkheid van de use case afhangt van de structuur van outputtemplates van de MA. Voornamelijk het detecteren van in aanmerking komende kosten voor financiering is een moeilijke opgave voor machines. Deelnemers die een hogere score gaven haalden aan dat resultaten vergelijken eenvoudig is als verondersteld wordt dat de templates een vaste structuur hebben. Het is echter moeilijker om ook na te kijken of er fouten zijn gemaakt.

Aantal jaren tot generatie 5: 4 jaar

Door de meerdere complexe voorwaarden voor volledige automatisatie mogelijk is, is de gemiddelde schatting van de deelnemers een periode van 4 jaar tot end-to-end autonomie.

Use case B19: Als auditor kan ik automatisch de output van de certificeringsautoriteit (rekeningen met onderliggende Excel-bestanden) controleren door na te gaan of dezelfde resultaten worden verkregen op basis van dezelfde bronnen zodat ik fouten kan vinden en corrigeren.

Huidige generatie: 5

De technologie van vandaag is in staat om bovenstaande use case end-to-end te automatiseren.

Aantal jaren tot generatie 5: /

Use case B20: Als auditor kan ik gebruik maken van software die automatisch bijleert wanneer de evaluatie ervan afwijkt van de evaluatie van de managementautoriteit en als onjuist wordt beschouwd zodat audits steeds beter worden en er rekening wordt gehouden met historische inzichten.

Huidige generatie: 4-5

De technologie van vandaag is in staat om bovenstaande use case end-to-end te automatiseren. De enige factor die volledige automatisatie kan verhinderen is slechte historische inzichten wat het bijleren bemoeilijkt.

Aantal jaren tot generatie 5: /

6.2.3 Use cases C aan het einde van het project

Use case Epic C: Als auditor kan ik automatisch de geleerde lessen extraheren en opslaan bij het afsluiten van een project zodat kennis en inzichten uit het verleden niet verloren gaan.

Huidige generatie: 3

De algemene conclusie is dat de huidige staat van de technologie toelaat om simpele deeltaken van deze use case door een machine te laten uitvoeren. Deelnemers die een lagere score gaven haalden aan dat hun technologie in staat was om terugkerende inzichten te extraheren maar dat het automatisch extraheren van lessen moeilijker is en vooral manueel gebeurt. Deelnemers die een hogere score gaven beweerden dat hun technologie in staat was pijnpunten te extraheren uit tekst die ze beschrijft. Als de pijnpunten niet expliciet beschreven staan is het een zeer moeilijke opgave voor een model om lessen te trekken.

Aantal jaren tot generatie 5: twee clusters: 3 jaar en 5-10 jaar

Er waren twee clusters van periodes tot volledige automatisatie. Het onderscheid tussen beide komt door een verschillende interpretatie van de randvoorwaarden. In een eerste cluster van 3 jaar werd aangenomen dat een expliciete tekst beschrijft wat de pijnpunten waren in het project en eventueel lessen beschrijft om uit het project te trekken. Een tweede cluster van 5-10 jaar nam aan dat er geen expliciete tekst bestond die dit beschreef, wat de use case complexer maakt.

Use case C1: Als auditor kan ik bij het afsluiten van een project automatisch controleren of alle correcties en aanbevelingen zijn uitgevoerd zodat er geen open punten overblijven wanneer het project wordt afgesloten.

Huidige generatie: 4

De experts waren het er over eens dat de huidige technologie al complexe deeltaken van deze use case autonoom kan uitvoeren. Het volledig automatiseren kan verhinderd worden door slecht beschreven projecten. Agile gerunde projecten, die gebruik maken van iteratieve sprints, vergemakkelijken wel het probleem.

Aantal jaren tot generatie 5: 1 jaar

De consensus van de experts was dat het hoogstens een jaar ging duren voor volledige automatisatie van deze use case mogelijk is.

Use case C2: Als auditor kan ik automatisch een auditrapport en een auditoordeel genereren zodat manueel werk tot een minimum wordt beperkt.

Huidige generatie: 3-4

Over het algemeen argumenteerden de deelnemers dat de technologie vandaag in staat is om simpele tot complexe deeltaken van de use case uit te voeren. Deelnemers die een lagere score gaven haalden aan dat technologie elementen kan genereren die nodig zijn in het finaal rapport maar niet in staat is om een gestructureerd rapport te bezorgen. Het eigenlijk synthetiseren van het rapport dient nog manueel te gebeuren. Deelnemers die een hogere score gaven namen aan dat alle nodige informatie voor het rapport op een gestructureerde manier te vinden is. Eén deelnemer geeft aan een product ontwikkeld te hebben dat in staat is om finale rapporten te genereren met een vast formaat.

Aantal jaren tot generatie 5: 3 jaar

Gemiddeld gaven de experts de use case 3 jaar om volledig end-to-end geautomatiseerd te worden. De antwoorden varieerden onder de deelnemers wegens verschillende assumpties rond de gevraagde structuur van het finaal rapport en de structuur van de input informatie.

6.3 Conclusie

Een eerste belangrijk inzicht is dat de overkoepelende use cases van de 3 types (Epic A, Epic B en Epic C) maximaal een generatiescore van 3 hebben en de periode tot generatie 5 geschat wordt op 5+ jaar. Dit geeft weer dat de experts van mening zijn dat volledige end-to-end automatisatie van audits nog enige tijd zal duren door beperkingen op technologisch vlak.

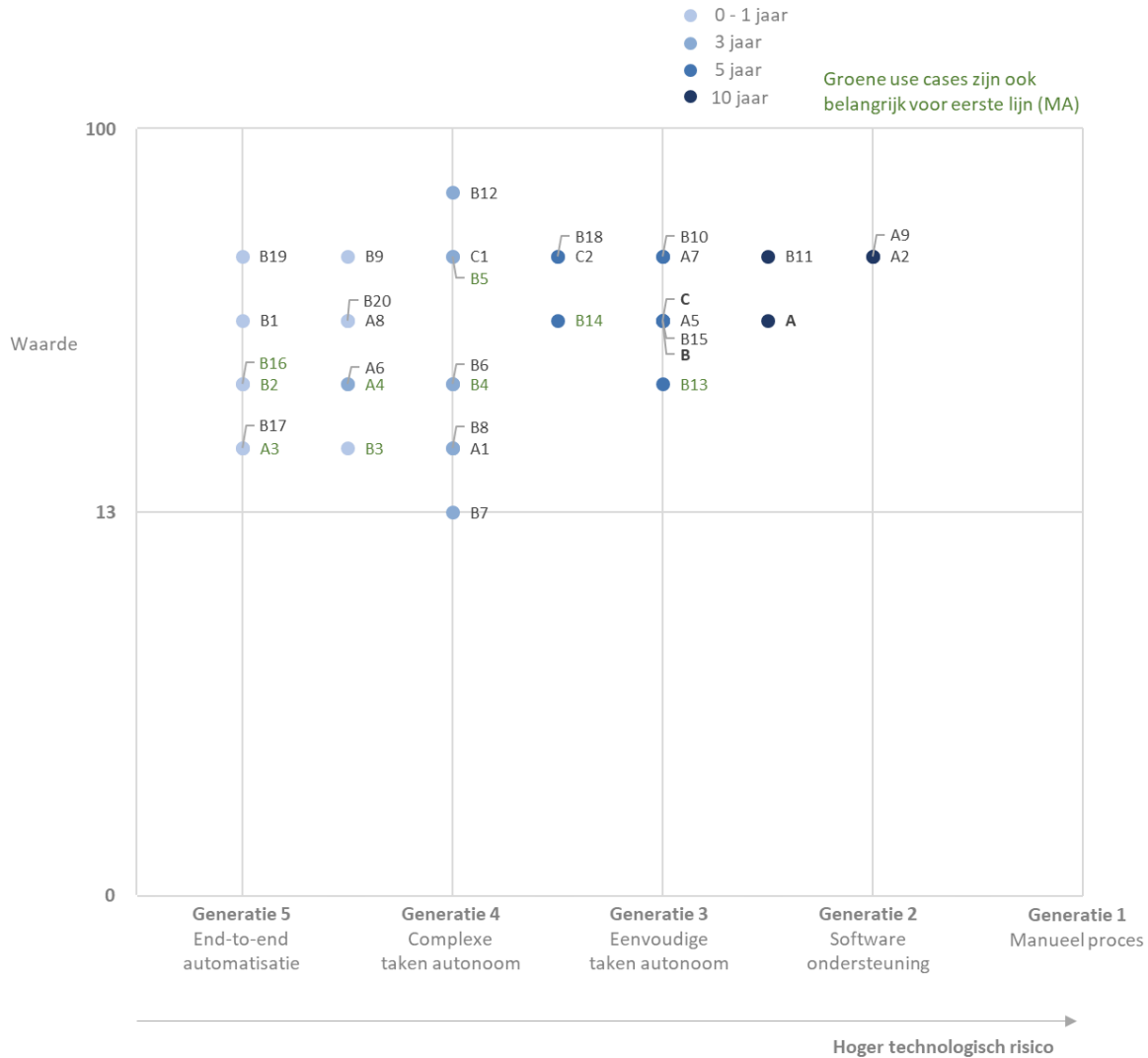
Zeer frequent werd door de experts aangehaald dat de technische moeilijkheid van de use case sterk afhangt van de randvoorwaarden: structuur van de input, hoeveelheid historische data, etc. Daarboven werd bij enkele use cases de opmerking gemaakt dat het volgens de marktspelers niet altijd voordelig is om nog de stap te zetten van generatie 4 naar 5. Dit kwam voornamelijk omdat ze vonden dat bij cruciale taken, zoals het inschatten van risico's, de kennis en ervaring van de mens onmisbaar is in het proces.

Er bestaan reeds meerdere automatische modellen op de markt voor de use cases over het nagaan van de volledigheid van documenten en het matchen van verschillende gegevensbronnen. Enkele marktspelers hebben zelfs (deel)oplossingen voor de moeilijkere opgaves rond diepere patroondetectie, opvolging en het inschatten van projectrisico's. Daarentegen was er een consensus tussen alle deelnemende partijen over het feit dat NLP technologie nog niet ver genoeg staat om volledig vrije tekst te interpreteren, al zeker niet in documenten rond regelgeving.

7 Architectuur en implementatieplan

7.1 Value-Risk analyse

Door de inzichten rond technologisch risico verkregen in de marktconsultatie te combineren met de meerwaarde vanuit gebruikersstandpunt, kan onderstaande value-risk matrix opgesteld worden. Deze geeft voor elke use case aan wat de toegevoegde waarde is voor de eindgebruiker (verticale as), wat de huidige stand van de techniek is (generaties op de horizontale as) en hoeveel jaar het nog zal kosten vooraleer de technologie voldoende matuur is om volledige automatisatie mogelijk te maken (kleur van de bol).



Volgende inzichten komen in deze value-risk matrix onmiddellijk naar voor:

1. Alle use cases worden door de eindgebruikers als belangrijk tot zeer belangrijk gezien. Er kunnen bijgevolg geen use cases weggelaten worden uit de scope van CATE zonder in te boeten op de toegevoegde waarde van de oplossing.
2. De huidige generatie correleert met het aantal jaar tot volledige automatisatie kan bereikt worden. Zo zal het langer duren voor een use case die vandaag in generatie 2 zit om generatie 5 te bereiken, dan voor een use case die vandaag reeds in generatie 4 zit.

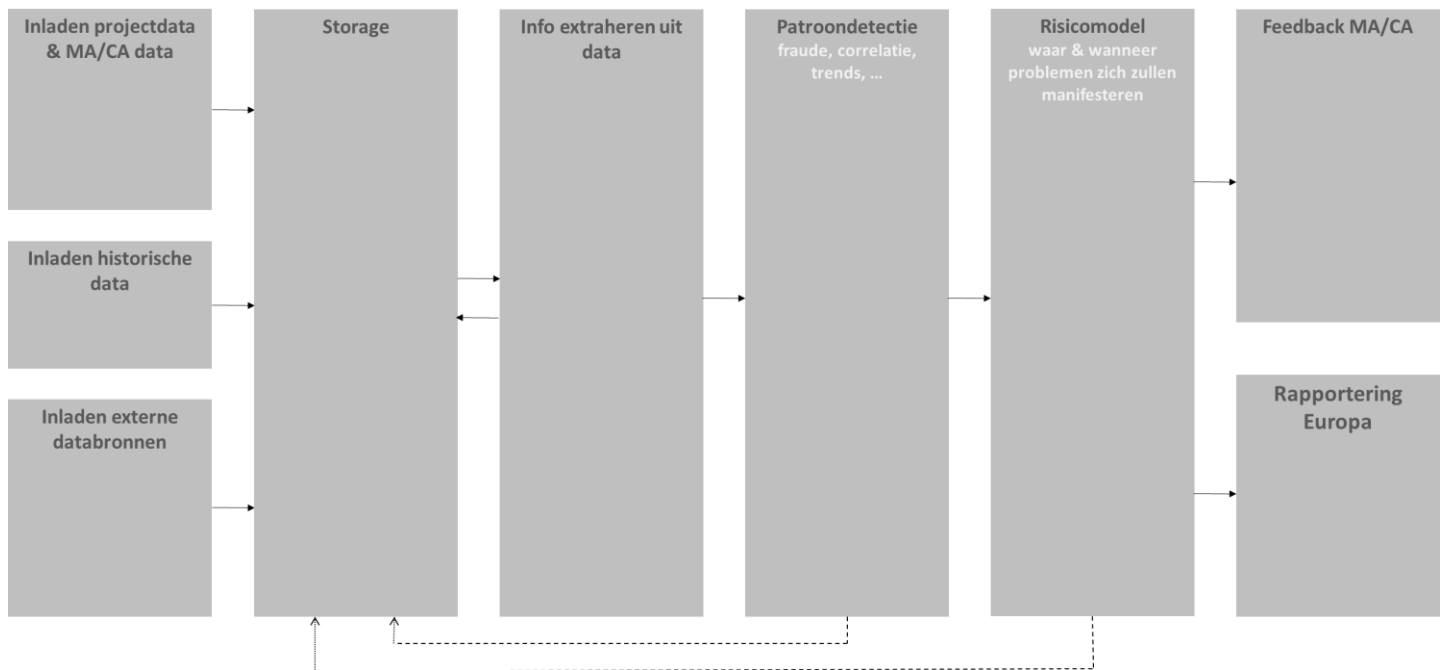
3. Voor 9 use cases die voor de eindgebruikers belangrijk zijn wordt expliciet aangegeven door de MA dat deze ook voor hen zinvol zijn. De software voor deze use cases kunnen, indien gewenst, ook ter beschikking gesteld worden van de MA.

De informatie uit de value-risk matrix kan gebruikt worden om een gefaseerd implementatieplan op te stellen, waarbij rekening wordt gehouden met de haalbaarheid van de use cases, zowel vandaag als in de toekomst. Om echter te verzekeren dat de implementatie geen automatisatie is van het huidige proces, wordt vertrokken vanuit een blauwdruk van de ideale eindsituatie.

7.2 Architectuur finaliteit audit

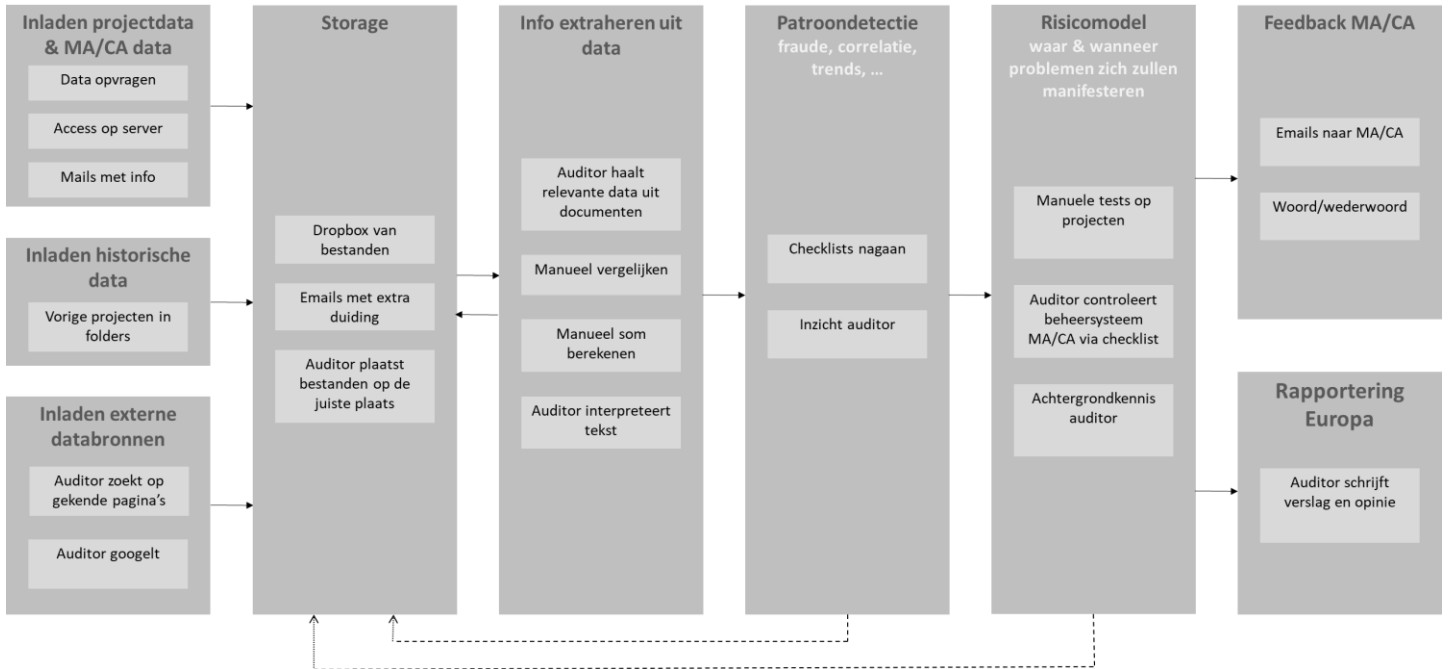
Het finale doel van het CATE project is om het auditproces conceptueel te herdenken vanuit de mogelijkheden van machine learning en data mining. Dit doel is vertaald in de overkoepelende use cases Epic A, Epic B en Epic C, zoals uitvoerig besproken in de vorige sectie.

Om volledig end-to-end automatisch (i.e. generatie 5) een uitspraak te doen over de betrouwbaarheid van het managementsysteem, zijn de bouwblokken in onderstaande referentiearchitectuur vereist. Hierin dienen de bouwblokken allemaal volledige geautomatiseerd te worden.



Deze architectuur kan als volgt geïnterpreteerd worden. Eerst moeten alle relevante data worden ingeladen, zijnde projectdata en MA/CA data, historische data en externe databronnen. Deze ruwe (*raw*) datasets worden opgeslagen voor verder gebruik. Uit deze ruwe data kan vervolgens concrete informatie worden geëxtraheerd die relevant is voor de audit. Deze informatie wordt eveneens opgeslagen voor later gebruik en dient ook als input voor algoritmes en modellen gericht op patroondetectie. Het patroondetectiebouwblok is in staat om patronen te herkennen voor het detecteren van fraude, vinden van correlaties of trends, etc. De output hiervan wordt verder gebruikt door het risicomodel, dat in staat moet zijn om in te schatten waar en wanneer problemen zich zullen manifesteren binnen projecten. Op basis van deze inzichten, die continu verlopen tijdens het project, kan vervolgens automatisch feedback naar de MA/CA verstuurd worden of automatisch rapportering voor de Europese Commissie gegenereerd worden.

In de huidige aanpak van de VAA zijn alle bouwblokken van bovenstaande architectuur onderliggend aanwezig, maar worden vele stappen nog manueel uitgevoerd en is het onderscheid tussen de verschillende stappen minder duidelijk. Een voorbeeld is het extraheren van informatie uit de data: de VAA doet dit momenteel nog door manueel de juiste velden in documenten te zoeken, te interpreteren en er eventueel bewerkingen mee uit te voeren. Onderstaande figuur toont hoe de huidige, hoofdzakelijk manuele, acties van de VAA gemapt kunnen worden op de finale referentiearchitectuur:



Om tot volledige end-to-end automatisatie van het auditproces te komen, dient elk van de bouwblokken volledig geautomatiseerd te worden. Uit de marktconsultatie kwam naar voor dat de volledige automatisatie van meerdere blokken vandaag nog niet mogelijk is. Onderstaande figuur geeft per bouwblok weer wat de betekenis is van elke generatie. Hierin werd de bouwblok om informatie te extraheren verder opgesplitst volgens de aanwezige structuur in de data, gezien de marktspelers tijdens de marktconsultatie aangaven dat hier significante verschillen tussen zijn. In de figuur wordt tevens aangeduid met een blauw kader wat de huidige stand van de techniek is voor elk van deze blokken, gebaseerd op informatie uit de marktconsultatie en de state-of-the-art analyse.

Topics		Generatie 1 Manueel proces	Generatie 2 Software ondersteuning	Generatie 3 Eenvoudige taken autonoom	Generatie 4 Complexe taken autonoom	Generatie 5 End-to-end automatisatie
Inladen projectdata & MA/CA data		Auditor vraagt manueel bestanden op	Auditor krijgt melding wanneer nieuwe bestanden beschikbaar zijn	Partners kunnen (na aangeven auditor) bestanden pushen naar platform	Automatische periodieke data dump op platform van VAA	
Inladen historische data		Audit haalt manueel eigen oude bestanden terug of vraagt die van collega's	Auditor krijgt melding wanneer nieuwe bestanden (van collega's) beschikbaar zijn	Collega auditors kunnen data pushen naar gedeeld platform	Automatische periodieke data dump op platform van VAA	
Inladen externe databronnen		Auditor zoekt manueel op websites of via zoekmachine	Auditor krijgt suggestie voor relevante zoektermen	Crawlen van tekstuele informatie online	Crawlen van tekstuele informatie of afbeeldingen van tekst	
Storage		Auditor houdt alle bestanden manueel bij op pc	-	Gedeelde mappen voor auditdiensten met versiecontrole	-	
Info extraheren uit data	Gestructureerde data (Data extractie – RPA)	Auditor extraheert info uit bronnen	Kernwoorden highlighten (e.g. naam, BTW, ...)	Volledige extractie uit documenten die vaste template volgen	Volledige data-extractie uit gekende documentvormen (e.g. facturen, certificaten, ...)	End-to-end automatische uitspraken over betrouwbaarheid van het systeem
	Ongestructureerde data (Tekstinterpretatie – NLP)	Auditor leest en interpreteert alle bronnen	Kernwoorden en -getallen highlighten	Onderwerpen detecteren in factuele tekst	Interpreteren factuele tekst (geen beelddruk / wettelijk taalgebruik)	
Patroondetectie (E.g. vergelijkbare projecten, fraude, ...)		Auditor vergelijkt om patronen te herkennen	Correlaties berekenen / overeenkomsten aanduiden	Detectie vooraf gedefinieerde patronen	Patroondetectie rond gekende topics op basis van historische dataset	
Risicomodel		Auditor leest bronnen en schat in waar risico's zitten	Aanduiden van afwijkende waarden t.o.v. andere projecten	Risico's inschatten van specifieke aspecten	Risico's genereren binnen vooraf gedefinieerde topics (e.g. financieel, deliverables)	
Feedback MA/CA		Manueel versturen van mails	Melding wanneer er feedback is om te versturen	Feedback automatisch sturen bij standaardproblemen	Feedback en opvolging afhandelen bij standaardproblemen	
Rapportering Europa		Manueel schrijven van documenten	Standaardvelden automatisch ingevuld (e.g. datum, naam programma, auditor, ...)	Standaardsecties automatisch geschreven	Volledige verslagen worden geschreven voor gevallen zonder atypische opmerkingen	

Generatie die vandaag mogelijk is

Op basis van deze figuur kan bepaald worden wat vandaag reeds kan en welke automatische functionaliteiten nog moeten ontwikkeld worden om tot een volledig end-to-end automatisch systeem te komen dat in staat is om een uitspraak te doen over de betrouwbaarheid van het systeem. Hierbij moet zowel rekening gehouden worden met de huidige stand van de techniek als met de toekomstige mogelijkheden. De blauwe kaders in de figuur duiden per bouwblok de generatie aan die vandaag mogelijk is volgens de markspelers. Om volledige automatisatie te bereiken, zal bijgevolg stapsgewijs functionaliteit moeten toegevoegd worden, zoals aangegeven in het gefaseerd implementatieplan in de volgende sectie.

7.3 Gefaseerd implementatieplan

Zoals aangegeven in de vorige sectie, dient de implementatie van het CATE project gefaseerd te verlopen. In onderstaande figuur wordt een overzicht gegeven van hoe een dergelijk implementatieplan er kan uitzien. Voor elk bouwblok van de referentiearchitectuur, wordt in de figuur aangegeven welke graad van automatisatie (i.e. generatie) geïmplementeerd wordt in de betreffende fase. De betekenis van elke generatie voor een bepaald bouwblok kan steeds teruggevonden worden in de vorige figuur.

Topics		Fase 1	Fase 2	Fase 3	Fase 4	Fase 5
		2021	2022	2023	2024	2025
Inladen projectdata & MA/CA data		Generatie 4	Generatie 5	Generatie 5	Generatie 5	Generatie 5
Inladen historische data		Generatie 3	Generatie 5	Generatie 5	Generatie 5	Generatie 5
Inladen externe databronnen		Generatie 3	Generatie 4	Generatie 4	Generatie 4	Generatie 5
Storage		Generatie 5	Generatie 5	Generatie 5	Generatie 5	Generatie 5
Info extraheren uit data	Gestructureerde data (Data extractie – RPA)	Generatie 3	Generatie 4	Generatie 5	Generatie 5	Generatie 5
	Ongestructureerde data (Tekstinterpretatie – NLP)	Generatie 1	Generatie 2	Generatie 3	Generatie 4	Generatie 5
Patroondetectie (vergelijkbare projecten, fraude, ...)		Generatie 3	Generatie 3	Generatie 4	Generatie 5	Generatie 5
Risicomodel		Generatie 2	Generatie 2	Generatie 3	Generatie 4	Generatie 5
Feedback MA/CA		Generatie 1	Generatie 3	Generatie 4	Generatie 5	Generatie 5
Rapportering Europa		Generatie 1	Generatie 3	Generatie 4	Generatie 5	Generatie 5

Parallele fase: Conceptueel uitdenken risicomodel

7.3.1 Fase 1: Data verzameling en gestructureerde informatie-extractie

Fase 1 heeft een tweeledig doel. Enerzijds wordt sterk ingezet op het inladen van data rond projecten en opslaan van data in een data lake, wat de basis vormt voor alle volgende fases. Anderzijds wordt meteen waarde voor de auditor gecreëerd door eenvoudige deeltaken van gestructureerde informatie-extractie en patroondetectie te automatiseren. Deze blokken laten immers toe om automatisch data te extraheren uit documenten met een vaste template, en automatisch vooraf gedefinieerde patronen te herkennen in de data. Bovendien krijgt de auditor een signaal als deze geëxtraheerde waarden sterk afwijken ten opzichte van andere projecten.

In deze fase wordt voornamelijk gewerkt rond data lakes voor het opslaan van data, alsook rond RPA en OCR om de volledigheid van documenten na te gaan en informatie te extraheren. Patroondetectie bestaat in deze fase uit standaard algoritmes.

Op het einde van fase 1 kunnen volgende use cases deels geautomatiseerd worden:

A8	Controleren of het beheersysteem en procedures van de MA/CA juiste vorm hebben en volledig zijn
B1	Analyse starten zodra bestanden zijn geüpload door projectpartner
B6	Gegevens vergelijken met historische gegevens (bijv. van andere projecten)
B7	Gegevens vergelijken met gegevens van de Vlaamse overheid (bv. KMO-portefeuille)
B8	Gegevens automatisch vergelijken met openbare databronnen of databronnen van derden
B16	Analyse starten zodra MA/CA documenten en beoordelingen uploaden

Op het einde van fase 1 kunnen volgende use cases volledig geautomatiseerd worden:

A3	Controleren of alle documenten in het juiste formaat aanwezig zijn
A4	Controleren of alle data in de juiste vorm aanwezig is (BTW nummer, alles ingevuld)
B2	Vorm van een document (formaat, bijlagen aanwezig) controleren zodra deze is toegevoegd

B3	Controleren of vorm van toegevoegde data (BTW nummer, alles ingevuld) juist is
B17	Controleren of de output van de MA/CA de juiste vorm heeft (juiste bijlage, formaat)

7.3.2 Fase 2: Extractie uit databronnen en standaardrapportering

In fase 2 wordt ingezet op diepgaandere extractie van informatie uit de beschikbare databronnen. Bij gestructureerde bronnen kan data uit gekende documentvormen volledig geëxtraheerd worden, terwijl de extractie bij ongestructureerde bronnen in deze fase nog beperkt blijft tot het highlighten van kernwoorden en -getallen. Bovendien wordt standaardfeedback naar de MA/CA geautomatiseerd, alsook het schrijven van standaardsecties in de rapportering naar de Europese Commissie zodat de auditor zich meer kan focussen op randgevallen en complexere taken.

Ter ondersteuning van deze diepgaandere extractie wordt ook het inladen van data verder geautomatiseerd zodat nieuwe bestanden snel beschikbaar zijn. Enkel het inladen van externe data blijft in deze fase nog beperkt tot het crawlen van tekstuele informatie.

Net zoals in fase 1 spelen technologieën als RPA (en OCR) een belangrijke rol in de verdere extractie van informatie en bij de communicatie met MA/CA en de Europese Commissie. Bovendien zal basis NLP moeten gebruikt worden om kernwoorden in ongestructureerde bronnen te highlighten.

Op het einde van fase 2 kunnen volgende use cases reeds deels geautomatiseerd worden:

B4	Controleren of verschillende gegevensbronnen consistente gegevens bevatten
B5	Controleren of de totalen van numerieke gegevens over verschillende documenten overeenkomen met de rapporten
C1	Controleren of alle correcties en aanbevelingen zijn uitgevoerd bij afsluiten van projecten
C2	Automatisch een auditrapport en een oordeel genereren

Op het einde van fase 2 kunnen alle deels geautomatiseerde use cases van de voorgaande fase volledig geautomatiseerd worden. Daarenboven kunnen volgende use cases volledig geautomatiseerd worden:

A1	Binnenkomende projecten automatisch classificeren in processtroom
B9	Inhoud van standaard gestructureerde sjablonen interpreteren
B19	Output van CA controleren door na te rekenen op basis van zelfde bronnen

7.3.3 Fase 3: Machine learning voor patroondetectie en risico's

In fase 3 ligt de focus op de analysebouwblokken: extractie van informatie uit documenten, patroondetectie en risico inschatting. Er wordt meer intelligentie ontwikkeld voor deze bouwblokken om deeltaken van de auditor over te nemen. Voor ongestructureerde data wordt bijvoorbeeld volledig automatische informatie-extractie mogelijk, terwijl voor ongestructureerde data de onderwerpen van factuele tekst kunnen gedetecteerd worden. Via machine-learningmodellen worden complexe deeltaken van patroondetectie mogelijk, alsook deeltaken van het risicomodel zoals het geven van risicoscores op specifieke aspecten. In deze fase dient de auditor enkel in te springen bij taken waar veel interpretatie bij komt kijken (interpreteren vrije tekst met nuances, risico's in complexe situaties, communicatie in atypische gevallen).

Naar gebruikte technologieën toe worden state-of-the-art RPA & OCR in deze fase gebruikt voor het extraheren van informatie uit gestructureerde documenten. Voor ongestructureerde data en communicatie met derde partijen wordt gesteund op NLP modellen. Voor de evolutie in patroondetectie en het risicomodel wordt, zoals vermeld, gesteund op machine learning.

Op het einde van fase 3 kunnen volgende use cases deels geautomatiseerd worden:

B13	Controleren of projecten de juiste stappen in hun vooraf gedefinieerde proces volgen
B14	Deliverables van lopende projecten monitoren en afwijkingen detecteren
B18	Output van MA controleren door middel van tests op specifieke projecten
B20	Software gebruiken die bijleert bij aangegeven fouten
C	Lessons learned extraheren en opslaan bij het afsluiten van een project

Op het einde van fase 3 kunnen alle deels geautomatiseerde use cases van de voorgaande fase volledig geautomatiseerd worden. Daarenboven kunnen volgende use cases volledig geautomatiseerd worden:

A6	Vergelijkbare historische projecten vinden
B12	Signaal krijgen in typische (bekende) gevallen van fraude

7.3.4 Fase 4: Diepe patroondetectie en betere risicoinschatting

De focus in fase 4 ligt op het verder bouwen op de progressie van de vorige fase. In deze fase wordt volledige automatisatie bereikt van patroondetectie en communicatie met zowel de MA/CA als de Europese Commissie. Door machine-learning modellen te ontwikkelen die nabij de state-of-the-art liggen kunnen risico's in deze fase ook gegenereerd worden binnen vooraf gedefinieerde topics.

Ook voor de informatie-extractie uit ongestructureerde data wordt de state-of-the-art nagestreefd zodat niet-wettelijke teksten zonder beeldspraak automatisch kunnen geïnterpreteerd worden. Het bouwblok voor patroondetectie is nu in staat om diepe patronen tussen historische en recente documenten te detecteren. Het risicomodel is nu capabel om risicoscores te genereren binnen vooraf gedefinieerde topics (financieel, deliverables).

In deze fase wordt gebruik gemaakt van state-of-the-art gespecialiseerde machine learning modellen en NLP technologie. De limiterende factor is hier niet de implementatie in het VAA project maar de stand van de technologie in de tijdsperiode van deze fase. Er zal verder onderzoek nodig zijn in NLP en machine learning om end-to-end te kunnen automatiseren.

Op het einde van fase 4 kunnen volgende use cases deels geautomatiseerd worden:

A7	Risico's van een project inschatten (waar ze zitten en wanneer ze zich zullen manifesteren)
B	Projectrisico's en aanbevelingen updaten op basis van nieuwe inzichten en informatie

Op het einde van fase 4 kunnen alle deels geautomatiseerde use cases van de voorgaande fase volledig geautomatiseerd worden. Daarenboven kunnen volgende use cases volledig geautomatiseerd worden:

A	Projectrisico's inschatten bij start van het project
A5	Bepalen hoe haalbaar de deliverables zijn
B15	Het betrouwbaarheidsniveau van beoordelingen berekenen en actualiseren

7.3.5 Fase 5: Volledige interpretatie van vrije tekst en accurate risicoinschatting

De laatste fase focust op volledige automatisatie van de bouwblokken rond interpretatie van ongestructureerde documenten (en intake van externe databronnen) en het inschatten van risico's. Dit laat toe om ook wettelijke teksten en teksten met nuances te interpreteren en hier vervolgens conclusies uit te trekken. Betrouwbare risicoscores zullen in alle situaties berekend en periodiek

geüpdatet kunnen worden om in te kunnen schatten waar en wanneer er zich problemen zullen manifesteren.

In deze fase wordt ingezet op geavanceerde NLP en de nieuwste machine learning technieken om betrouwbare risicoscores in te schatten. In de marktconsultatie werd aangegeven dat dergelijke technieken pas vanaf 2025 voldoende matuur zullen zijn om volledige automatisatie mogelijk te maken, waardoor deze pas in fase 5 aan bod kunnen komen en onderzoek zullen vereisen.

Op het einde van fase 5 kunnen alle deels geautomatiseerde use cases van de voorgaande fase volledig geautomatiseerd worden. Daarenboven kunnen volgende use cases volledig geautomatiseerd worden:

A2	Huidige regelgeving automatisch inladen, actualiseren en interpreteren
A9	Controleren of het beheersysteem en procedures van de MA/CA overeenstemmen met de regelgeving
B10	Inhoud van vrije formulieren interpreteren
B11	Vrije tekst interpreteren

7.3.6 Parallele fase: conceptueel uitwerken risicomodel van de toekomst

In parallel met de vijf ontwikkelingsfases hierboven gedefinieerd, dient in een aparte track het risicomodel van de toekomst conceptueel uitgedacht te worden. Het risicomodel is een centraal punt van de nieuwe architectuur, waardoor de juiste visie reeds in het begin nodig is zodat niet wordt afgeweken van het uiteindelijke doel.

In elke ontwikkelingsfase dient de visie op het risicomodel afgetoetst te worden met de huidige stand van de technologie. Zo zal de visie mee bepalen hoe de implementatie in elke fase zal gebeuren, en zal omgekeerd de huidige stand van de technologie aangeven wat al mogelijk is en inspiratie geven om de visie verder bij te schaven. Op deze manier wordt de visie op het risicomodel continu bijgewerkt via interacties met de technologische evolutie.

8 Algemene conclusie

De Vlaamse Auditautoriteit voor Europese Structuurfondsen (VAA) wenst hun auditproces zowel op technologisch als op conceptueel vlak volledig te herdenken, vertrekkend vanuit de mogelijkheden die machine learning en data mining bieden. Het CATE-project heeft als doel het auditproces te versnellen, het verlies van Europese middelen te verminderen en accuratere rapporteringen te leveren van hoge kwaliteit. Om tot een realistisch, kwaliteitsvol bestek voor een overheidsopdracht te komen werd een voortraject georganiseerd met als doel de vraagstelling en scope te verscherpen, een analyse van de vraagzijde uit te voeren en deze finaal te combineren met een analyse van de aanbodzijde (marktspelers in audit & AI).

Door use cases op te stellen en te prioriteren in samenwerking met de auditoren van de VAA werden de noden van de vraagzijde in kaart gebracht en gerangschikt. Over het algemeen werden zeer hoge innovatiescores gegeven. Dit geeft weer dat de VAA alle opgestelde use cases belangrijk tot onmisbaar vond. Enkele use cases hebben zelfs nog meer toegevoegde waarde aangezien ze nuttig zijn voor meerdere gebruikersgroepen (e.g. MA), waardoor de software in principe kan gedeeld worden.

Na een state-of-the-art analyse, waarbij de belangrijkste technologieën en relevante spelers in kaart werden gebracht, werd een marktconsultatie georganiseerd om in dialoog te gaan met experts en marktspelers in audit, AI en machine learning. Hierbij werd voor alle use cases een generatiescore bepaald, duidende op de mate van automatisatie die mogelijk is met de huidige stand van de technologie, en hoe lang het zou duren voor end-to-end automatisatie haalbaar wordt. Use cases rond het nagaan van de volledigheid van documenten en het matchen van gegevensbronnen werden door de experts ervaren als volledig automatiseerbaar met de huidige technologie. De meest geavanceerde spelers beschikken zelfs over deels geautomatiseerde oplossingen voor de moeilijkere opgaves rond patroondetectie en het inschatten van projectrisico's. De marktpartijen gaven echter aan dat NLP vandaag nog niet ver genoeg staat om vrije tekst volledig te interpreteren. Frequent werd door de experts aangehaald dat de technische moeilijkheid van automatisatie sterk afhangt van de randvoorwaarden zoals de structuur van de inputdata, de hoeveelheid historische data, etc.

Tot slot werden de inzichten uit de innovatiescores van de stakeholders en de generatiescores van de marktspelers gecombineerd om tot een plan van aanpak te komen. Om te vermijden dat het huidige auditproces gewoon geautomatiseerd wordt, werd een referentiearchitectuur opgesteld die de ideale eindoplossing capteert. Deze bestaat uit verschillende bouwblokken die elk gefaseerd kunnen uitgewerkt worden volgens de evoluties in de technologie. Er werd een implementatieplan voorgesteld in vijf fases waarin voor elk bouwblok wordt uitgestippeld hoe deze kan evolueren tegen 2026. Specifiek worden volgende fases voorgesteld:

- Fase 1: Data verzameling en gestructureerde informatie-extractie
- Fase 2: Extractie uit databronnen en standaardrapportering
- Fase 3: Machine learning voor patroondetectie en risico's
- Fase 4: Diepe patroondetectie en betere risicoinschatting
- Fase 5: Volledige interpretatie van vrije tekst en accurate risicoinschatting

Parallel met deze fases dient een track te lopen waarin het risicomodel voor de toekomst, dat centraal staat in de nieuwe aanpak, conceptueel wordt uitgedacht. Deze wordt vervolgens na elke

ontwikkelingsfase getoetst en bijgeschaafd om de visie van het project te bewaken. Op basis van deze inzichten kan een bestek uitgewerkt worden voor de verdere uitwerking van het CATE-project.

Bijlage A: State-of-the-art analyse



State of the art onderzoek op twee relevantie domeinen

Management & risk assessment	<ul style="list-style-type: none"> • Project dashboarding • Automatic risk management & assessment
Automatisatie & machine learning	<ul style="list-style-type: none"> • Robotic process automation • Machine learning <ul style="list-style-type: none"> • Voor patroonherkenning • Natural language processing



Project Dashboarding

Technologie	<p>Oplossingen variëren van standaard projectmanagement tot auditspecifieke suites</p> <ul style="list-style-type: none"> • Globaal overzicht en historiek per project, programma, budgetten, enz. • Processtappen kunnen duidelijk gedocumenteerd worden in verschillende taken & deeltaken • Mogelijkheid tot snel genereren van rapporten rond alle opgeslagen data
Implicaties	<ul style="list-style-type: none"> • Dashboard kan vertrekpunt zijn en ondersteunende rol bieden • Het updaten van de status van taken is doorgaans niet geautomatiseerd • Externe bestanden (bv. facturen) worden doorgaans niet bijgehouden binnen de tool, dus geen volledige centralisatie
Marktspelers	<ul style="list-style-type: none"> • Veel mogelijke leveranciers, maar aantal geïndustrialiseerde suites is beperkt <div style="display: flex; align-items: center; margin-top: 10px;"> </div> <p style="text-align: right; margin-top: 10px;">Graad van industrialisatie</p>

• Merk op dat spelers met een hogere graad van industrialisatie meer mogelijkheden "out of the box" aanbieden voor o.a. data-integratie en (gelimiteerde) automatisatie



Automatic risk management & assessment

Technologie

Een overzicht genereren van aanwezige risico's en manieren om deze te beheersen

- Risico's identificeren gebeurt typisch manueel, breed overzicht maken wordt geautomatiseerd
- Beheersen van risico's gebeurt door standaard automatisatie (e.g. som van uitgaven checken met facturen) gebaseerd op harde regels
- Automatisch identificeren van risico's wordt amper gedaan (op enkele experimentele use cases na, e.g. automatisch bepalen of dataset privacygevoelig is na taggen van data)

Implicaties

- Mogelijkheden van automatische risk assessment zijn vandaag beperkt
- Specifiek research vereist om hiermee verder te gaan
- Management van risico wel mogelijk via automatische checks op kerngetallen

Marktspelers

- Spelers rond automatisatie en RPA (zie verder)
- Spelers met ervaring of platformen rond "risk management (in audit)"



Robotic Process Automation & OCR

Technologie

Robotic Process Automation (RPA), eventueel aangevuld met Intelligent Automation (IA)

- RPA laat toe om routinetaken in files met consistente structuren volledig te automatiseren
- RPA steunt vaak op OCR (Optical Character Recognition) om tekst uit afbeeldingen of foto's te extraheren
- IA-aanvulling voor audits staat nog in de kinderschoenen en is eerder tailored per use case

Implicaties

- Mature technologie voor standaard RPA en OCR, veel integratoren beschikbaar
- Repetitieve taken met lage complexiteit kunnen doorgaan geautomatiseerd worden
- Complexere taken worden uitdaging gezien beperkte maturiteit van intelligent automation

Marktspelers

Consultancy & integratoren



Standaard RPA-platform



Custom platform





Machine learning voor patroonherkenning

Technologie

Custom machine learning trajecten om patronen te detecteren of informatie te extraheren

- Patronen en trends uit historische gegevens leren en hiermee predicties maken
- Outliers in datasets detecteren (e.g. voor fraudedetectie)
- Kan "confidence level" meegeven bij predicties of classificaties

Implicaties

- Nood aan voldoende (gelabelde) data om zinvolle patronen te detecteren
- Researchproces: data verzamelen, data cleaning, preprocessing & feature selection, modelselectie & -validatie
- Specifiek traject vereist voor elke use case

Marktspelers

- Machine learning spelers met algemene expertise



- Machine learning spelers met algemene expertise en specifieke ervaring met fraudedetectie

BOLTZMANN



DEPARTEMENT
ECONOMIE, WETENSCHAP & INNOVATIE

Vlaamse
overheid



Natural Language Processing

Technologie

Natural language processing verwerkt en analyseert "natuurlijke" tekst (i.e. volle, ongestructureerde tekst)

- Vandaag nog in onderzoeksfase met focus op Engels (maar ook Nederlandse initiatieven, e.g. NLP4Gov)
- Herkennen van entiteiten in tekst (e.g. bedrijven of personen) is reeds behoorlijk accuraat
- Andere subdomeinen binnen NLP (automatisch samenvatten, kerninformatie extraheren) staan nog in kinderschoenen

Implicaties

- Bruikbaar voor eenduidige taken zoals entiteiten in tekst herkennen
- Zwaar researchproces voor complexere analysetaken
- Te combineren met andere ML / RPA technieken om geëxtraeerde informatie verder te gebruiken

Marktspelers

- Machine learning spelers met ervaring in NLP



RACCOONS



- Internationale softwarespelers (Google, Amazon, Facebook, ...) bieden open source tools en algoritmes aan

DEPARTEMENT
ECONOMIE, WETENSCHAP & INNOVATIE

Vlaamse
overheid



Machine learning (research) is noodzakelijk om verder te gaan dan standaard “copy-paste” automatisatie

Management & risk assessment

- Project dashboarding
- Automatic risk management & assessment

Automatisatie & machine learning

- Robotic process automation
- Machine learning
 - Voor patroonherkenning
 - Natural language processing

- Platformen beschikbaar om audit / projecten te managen, aangevuld met automatisatie (RPA & OCR) voor repetitieve taken met lage complexiteit
- Patronen herkennen in (historische) data vereist voldoende data machine learning traject
- Complexere taken (e.g. gerichte NLP of automatic risk assessment) vereisen specifiek onderzoek naar mogelijkheden en grenzen, alsook machine learning expertise

Bijlage B: Deelnemende partijen in de marktconsultatie

- Addestino Innovation Management
 - De Vlaamse Auditautoriteit van de Europese Structuurfondsen (VAA)
 - Europees Fonds voor Regionale Ontwikkeling (EFRO)
 - Europees Sociaal Fonds (ESF)
 - Het Programma Innovatieve Overheidsopdrachten (PIO) van het Departement Economie, Wetenschap en Innovatie (EWI)
 - Interreg Vlaanderen/Nederland
-
- AE – Architects for Business & IT
 - Alten
 - Arinti
 - Astoria consultancy
 - Datology
 - Deloitte
 - element61
 - EY
 - HLB
 - IBM
 - Info Support
 - InfoFarm
 - Keyrus
 - KPMG
 - ML2Grow
 - ML6
 - Paperbox
 - PwC
 - Radix
 - Realdolmen
 - Securitas
 - Sepia Solutions
 - Textgain
 - TME – Toyota Motor Europe
 - Weare4C
 - Yazzoom