



# Process Mining

---

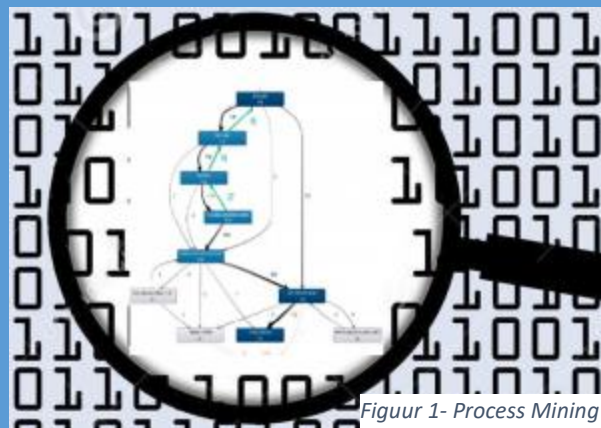
PROFESSIONALS STAAN VOOR DE UITDAGING OM DE COMPLEXITEIT VAN HET GEBRUIK VAN PROCESS MINING TE ONDERKENNEN, DAN WEL HET TOEPASSEN EN/OF TE CONTROLEREN HIERVAN.

**15 DECEMBER 2019**

---

Vrije Universiteit Amsterdam  
Premaster IT audit, compliance & advisory

Paul van Egmond



Figuur 1- Process Mining

*Turning event data into value, responsibly!*

## Titelblad

**Titel:** Turning event data into value, responsibly!  
**Subtitel:** Professionals staan voor de uitdaging om de complexiteit van het gebruik van Process Mining te onderkennen, dan wel het toepassen en/of controleren hiervan.

**Naam:** Paul van Egmond, BEc  
**Studentnummer:** 2685821

**Onderwijsinstelling:** Vrije Universiteit Amsterdam  
**Opleiding:** Premaster IT audit, compliance & advisory  
**Module:** Professional Judgement  
**Studie jaar:** 2019

**Vakdocenten:** De heer dr. P. Frijns  
De heer S. Rieske MSc

**Eerste lezer:** De heer dr. P. Frijns  
De heer S. Rieske MSc

**Tweede lezer:** De heer prof. dr. ing. A. Shahim, RE MSc  
De heer S. Schinagl, RE MSc CISA

**Publicatiedatum:** 15-12-2019  
**Versie:** Definitieve versie

## Voorwoord

Mijn interesse voor Process Mining is in 2014 gewekt door de presentatie van Pieter de Kok (RA) bij Hogeschool Inholland Alkmaar. Daarnaast ben ik enkele jaren actief geweest als consultant data-analist bij Coney (Discovery your data) in Amsterdam. Als consultant heb ik zelf de complexiteit van Process Mining mogen ervaren, met name de complexiteit vanuit de business dan wel praktisch. Ik wil met dit onderzoek de technische- (en business) beperkingen van Process Mining onderzoeken.

Dit onderzoek past in het domein van informatietechnologie-auditing (IT-auditing), met de volgende drie onderdelen:

- Techniek – De techniek achter Process Mining doorgronden (algoritmen);
- Analyse – Data-analyse van het ontsluiten van brondata en het vormen van een eventlog en daarnaast gebruikmaken van de data-analyse tools;
- Assurance en advisory – Process Mining kan toegepast worden in het financiële auditdomein, maar ook als procesoptimalisatie waar de IT-audit een belangrijke rol kan spelen in een project.

Mijn doel is professionals informeren over de technische- en businessrisico's die zich voor kunnen doen bij het toepassen van Process Mining. De informatie wordt weergegeven in de vorm van een risicomatrix. Ik breng hiermee het landschap van Process Mining in kaart en geef hiermee een helicopterview, zodat de professional zelf kan inschatten welke risico's er zijn. Vervolgens kan de professional gericht zijn werkzaamheden hierop aanpassen.

Graag zou ik alle begeleiders bedanken die betrokken waren bij de Pre Master module Professional Judgement. In het bijzonder Pieter Frijs voor de gespreken die we hebben gehad over de inhoud van het onderwerp en opbouw scriptie. Daarnaast bedank ik graag de medestudenten die in dezelfde periode de Pre Master hebben gevolgd. Bedankt voor jullie kritische blik, inbreng en gezelligheid!

Paul van Egmond

Julianadorp, 15 december 2019

## Abstract

Process Mining is een techniek, tool en een methode om geautomatiseerde bedrijfsprocessen te visualiseren. Professionals staan voor de uitdaging om de complexiteit van Process Mining te onderkennen en toe te passen. Technische - en businessrisico's zorgen ervoor dat een Process Mining Project faalt. Dit leidt tot de volgende onderzoeksvraag: *Welke risico's van Process Mining worden ondervangen door Process-Mining-Project-Methodology?* Met deze onderzoeksvraag wordt onderzocht hoe een Process Mining Project kan slagen wanneer rekening gehouden wordt met deze projectmethode.

De complexiteit van Process Mining is met name het verkrijgen van tijdige, relevante en accurate data. Daarnaast zijn er verschillende granulariteitsniveaus van de data mogelijk. Elk (detail)niveau van data geeft een andere weergave van - en inzicht in het bedrijfsproces.

De granulariteit van de ontsloten data moet aansluiten met de (beoogde) eindgebruiker. De interpretatie van het bedrijfsproces en data is van belang om de data te ontsluiten en te analyseren. Hierbij is de teamsamenstelling dan ook divers met verschillende professionals.

Voor het analyseren kan zowel een opensource- als commerciële Process-Miningtool worden ingezet. Bij de laatste ontbreekt de standaardisatie. De commerciële tools gebruiken bevatten eigen normen en naamgeving en hierbij verschillende algoritmes met verschillende uitkomsten.

De Process-Mining-Project-Methodology (PM2-methode) bestaat uit zes fases. Met behulp van deze fases kan een projectgroep geholpen worden om op gestructureerde wijze een project uit te voeren. Wanneer de planningsfase bijvoorbeeld wordt overgeslagen, is de kans groot dat later in het project vertraging ontstaat of zelfs stagneert. In de evaluatiefase wordt nagegaan of de analysebevindingen correct zijn geïnterpreteerd. Daarnaast worden deze geverifieerd en gevalideerd. Deze stap komt ten goede aan de kwaliteit van de uitkomsten en verbeteringen van Process Mining voor de organisatie.

Uit het onderzoek blijkt dat PM2 de meest voorkomende risico's afdekt. Echter, niet alle risico's worden afgedekt. Daarom moeten aanvullende beheersingsactiviteiten worden uitgevoerd om alle risico's die in de literatuur zijn gevonden af te dekken.

De aanvullende beheersingsactiviteiten bevinden zich ten eerste in de planningsfase, waar de keuze voor de teamsamenstelling wordt gemaakt. Ten tweede in de keuze van data-analysetools die tijdens het project worden toegepast. Aanvullend kan een training worden gevolgd om de benodigde kennis en inzicht te krijgen. Als laatste dienen in een aanvullende fase documenten en rapporten te worden toegevoegd om de keuzes en uitkomsten van het project goed te kunnen overdragen aan belanghebbenden.

Process Mining wordt steeds belangrijker, maar ook ingewikkelder om toe te passen. Een oorzaak hiervan is dat de techniek (algoritmes) niet transparant zijn bij commerciële tools. Het moet worden vastgesteld welke partij de beste tool biedt voor de analyse die uitgevoerd moet worden. Dit huidige onderzoek leidt tot aanbevelingen voor vervolgonderzoek naar welke commerciële tools er zijn en om de gebruikte algoritmes te toetsen. Daarnaast is vervolgonderzoek nodig om aan te geven met welke toepassingsmogelijkheid de tool zich onderscheidt, bijvoorbeeld in het analyseren op doorlooptijden. Daarnaast blijft de techniek, de methodes en tools van Process Mining in ontwikkeling. Ook hier geldt dat vervolgonderzoek verricht moet worden om de bewustwording en complexiteit over te dragen naar het bedrijfsleven en (eind)gebruikers die de techniek, methode en tool inzetten.

*Key woorden: Process Mining, Process-Mining-Project-Methodology, Technische- en Business risico's.*

## Figuren- en tabellenlijst

Figuur 1 - Process Mining .....	1
Figuur 2 - Indeling hoofdstukken .....	8
Figuur 3 - leeswijzer.....	8
Figuur 4 - Landkaart (Van der Aalst, 2012) .....	12
Figuur 5 - Process-Mining-Project-Methodology (Van der Aalst, 2015).....	14
Figuur 6 - Process Mining in de praktijk.....	19
Figuur 7 - House of IT auditing.....	20
Figuur 8 - Inrichting process van bedrijfsproces en process mining aanpak (Haasnoot, 2012).....	27
Figuur 9 - Onderzoek timeline Process Mining.....	28
Figuur 10 - Process Mining in de praktijk.....	28
Tabel 1 - Resultaten Process Mining risico's en Process-Mining-Project-Methodology. ....	16
Tabel 2 - Conclusie Process-Mining-Project-Methodology als beheersingsmaatregel.....	22
Tabel 3 - Volwassenheid activiteitenlog (Daniel, Barkaoui, & Dustdar, 2011) .....	27

## Verklarende woordenlijst

Activiteitenlog:	Een bestand waarin alle handelingen worden opgeslagen die plaatsvinden in de informatiesystemen.
Algoritme:	De techniek achter Process-Miningtools.
Bedrijfsproces:	Geautomatiseerde processen die voorkomen binnen een organisatie.
BPM:	Business Process Management.
Case ID:	Uniek nummer identificatienummer.
Controlebewijs:	Informatie die volledig en betrouwbaar is.
Controledossier:	De vastlegging van de werkzaamheden van de accountant tijdens de controlefasen.
Controle-informatie:	Informatie waar de accountant zijn controleverklaring op baseert.
Data-analist:	De persoon die zich bezighoudt met het creëren van het event log.
Data:	Informatie die wordt vastgelegd in de informatie systemen.
Event log:	Een bewerkte data set in CSV bestand, die geschikt is voor Process-Miningtool.
Granulariteit:	Geeft aan in welke mate er detail gegevens aanwezig zijn.
Informatiesysteem:	Een systeem waar informatie in wordt vastgelegd.
Meta-data:	De aanvullende data naast case ID, activiteit en tijdstempel.
Process Mining model:	De weergave van het Process Mining model op basis van de data (eventlog).
Process Mining:	Is een tool, methode en techniek om proces data visueel weer te geven.
ProM:	Een Process-Miningtool, geschikt voor de academische wereld (open-source).
Robuuste IT-omgeving:	Een informatie-omgeving die ver geautomatiseerd is.
Simulatie:	Het visueel maken van Process informatie uit fictieve data.
Tijdstempel:	Het moment dat de activiteit is vastgelegd in het informatiesysteem.
Verwachtingskloof:	De (eind)gebruiker die meer zekerheid uit Process Mining analyse haalt. Dan wel een andere verwachting heeft van de techniek van Process Mining.

## Inhoudsopgave

Titelblad .....	2
Voorwoord .....	3
Abstract.....	4
Figuren- en tabellenlijst .....	5
Verklarende woordenlijst .....	5
Inleiding.....	7
Aanleiding .....	7
Doelstelling .....	7
Onderzoeksvraag .....	7
Criteria en afbakening.....	7
Leeswijzer.....	8
Methodologie.....	9
Theoretisch kader .....	10
Resultaten .....	15
Beheersmaatregelen.....	17
Aanvullende beheersingsmaatregelen .....	17
Process Mining praktijk ervaring.....	19
Process Mining professionals.....	19
Conclusie .....	21
Reflectie .....	23
Discussie en aanbevelingen .....	24
Literatuurlijst.....	25
Bijlage 1 – Figuren en tabellen.....	27
Bijlage 2 - Reflectie.....	29



## Inleiding

### Aanleiding

Process Mining wordt steeds belangrijker, maar ook ingewikkelder, aldus hoogleraar professor Van der Aalst in het vakblad Agconnect (Agconnect, 2018). Het belang van Process Mining neemt toe, omdat steeds meer organisaties de toegevoegde waarde zien van het gebruik van data-analyse. De afgelopen jaren zijn de data toegankelijker geworden. Daarnaast zijn er inmiddels dertig softwareleveranciers die een Process-Miningtool aanbieden. Process Mining is een relatief nieuwe techniek in vergelijking met data-analyse. Hierdoor zijn niet alle mogelijkheden, uitdagingen en valkuilen bij iedereen bekend (Daniel, Barkaoui, & Dustdar, 2011).

De IT-professional (auditor) staat voor de uitdaging om de complexiteit van Process Mining te onderkennen, toe te passen en of te controleren hiervan. Zo wordt een bepaalde mate van zekerheid verschaft met betrekking tot de uitkomsten en hiermee indirect het gebruik van Process Mining. De onderzoeker wil de professional informeren over de technische - en businessrisico's die zich voor kunnen doen bij het toepassen van Process Mining. Hiervoor wordt gebruikgemaakt van Process-Mining-Project-Methodology (PM2-methode).

### Doelstelling

De uitkomsten van het onderzoek worden in de vorm van een risicomatrix weergegeven. Hiermee wordt het landschap van Process Mining in kaart gebracht. Middels een helicopterview kan de professional zelf inschatten welke risico's van toepassing zijn per projectfase. Vervolgens kan de professional zijn of haar werkzaamheden hier gericht op aanpassen.

### Onderzoeksvraag

Op basis van het bovenstaande is de volgende onderzoeksvraag geformuleerd: *Welke risico's van Process Mining worden ondervangen door Process-Mining-Project-Methodology (PM2)?*

Om de onderzoeksvraag te beantwoorden, zijn vier deelvragen geformuleerd, onderverdeeld in beschrijvende, analyserende en beschouwende deelvragen.

#### Beschrijvende deelvragen:

- Wat zijn de risico's van Process Mining?
- Wat is de Process-Mining-Project-Methodology als beheersmaatregel?

#### Analyserend deelvraag:

- Welke risico's van Process Mining worden door Process-Mining-Project-Methodology als beheersmaatregelen in oenschouw genomen?

#### Beschouwend deelvraag:

- Wat is de toegevoegde waarde van Process-Mining-Project-Methodology voor het beheersen van risico's bij Process Mining?

### Criteria en afbakening

#### De koksmuts (Frijns, 2019).

Om het onderzoek af te bakenen, is in het vooronderzoek gebruikgemaakt van de koksmuts. Bij het opstellen van de koksmuts zijn probleemvelden naar boven gekomen die niet tot de scope van het onderzoek behoren.

#### De scope

Het onderzoek is beperkt tot de belangrijkste kenmerken van Process Mining. De keuze welke algoritmes worden toegepast in een (commerciële) tool ligt buiten de scope van dit onderzoek. Ook Process Mining in relatie tot een specifiek beroep en tot nieuwe technologie behoren niet tot de scope van dit onderzoek. Gezien de beschikbare tijd en complexiteit van het onderwerp is dit onderzoek beperkt tot de hoofdlijnen van het gebruik van Process Mining en met name de technische - en businessrisico's die zich voor kunnen doen bij het toepassen van Process Mining. Deze risico's hebben uiteindelijk invloed op de (eind)conclusie die (in)formeel wordt gedaan door (eind)gebruikers van Process Mining.

## Aannames

In het uitgevoerde onderzoek worden de gevonden technieken en businessrisico's van Process Mining die beschreven zijn in de wetenschappelijke literatuurstudies als bewezen beschouwd. Hiervoor wordt geen vervolgonderzoek verricht.

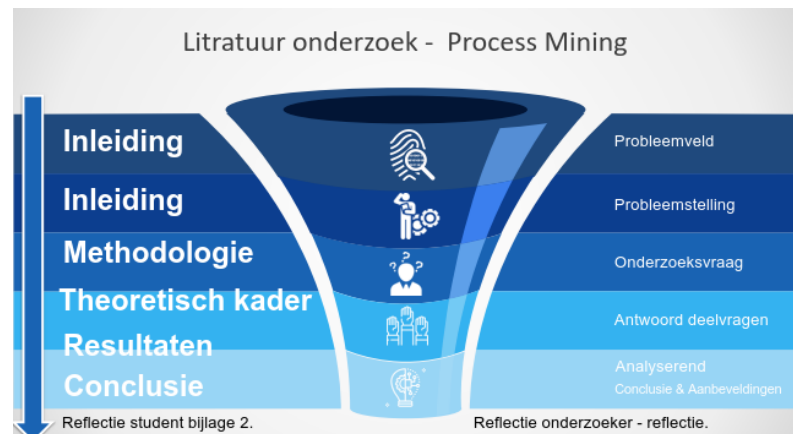
## Termen

Volgens het Process Mining Manifest (Daniel et al., 2011). Kan Process Mining als volgt gedefinieerd worden: Process Mining is een techniek, tool en een methode om geautomatiseerde bedrijfsprocessen te visualiseren. De technische risico's van Process Mining worden gedefinieerd indien de techniek (inclusief data) enkele beperkingen toont of dat er aandachtspunten zijn. Van businessrisico's van Process Mining is sprake indien de techniek, methode of gebruik van de tool op een verkeerde manier wordt toegepast of geïnterpreteerd.

## Leeswijzer

In de inleiding zijn de onderzoeks aanpak en de onderzoeksvragen geformuleerd. De uitkomsten van het literatuuronderzoek zijn in het theoretisch kader opgenomen. Vervolgens komen in de resultaten de risico's en beheersmaatregelen aan bod. Als laatste volgt de conclusie waar antwoord wordt gegeven op de onderzoeksvraag en de discussie en aanbevelingen.

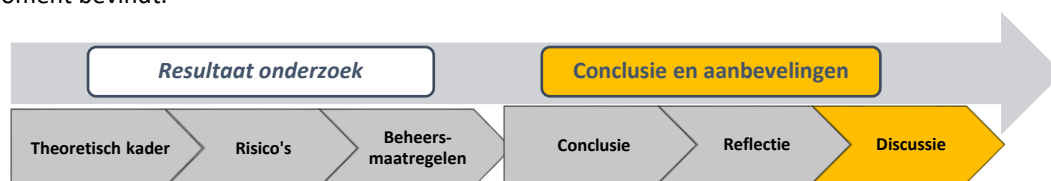
In de bijlage is de volgende tabel opgenomen: uitkomst analyserende- en beschouwende deelvraag. Daarnaast is het volwassenheidsniveau voor activiteitenlog opgenomen zoals deze herkend wordt in het Process Mining Manifest (Daniel et al., 2011). Als laatste bijlage is de studenten-evaluatie opgenomen.



Figuur 2- Indeling hoofdstukken

## Leesbaarheid

Om de leesbaarheid aan te geven, is gebruikgemaakt van het schema in figuur 3. Deze Afbeelding wordt in de voetnoot van elke pagina opgenomen. In het geel wordt aangegeven bij welke subparagraaf de lezer zich op dat moment bevindt.



Figuur 3- leeswijzer.



# Methodologie

## Literatuuronderzoek

### Soort onderzoek

Om de onderzoeksvraag te beantwoorden, is kwalitatief literatuuronderzoek verricht om antwoord te geven op de onderzoeksvraag en deelvragen. Hierbij is academische literatuur geraadpleegd. In het vooronderzoek is gebruikgemaakt van vakbladen en interviews.

### De bronnen

In de afgelopen twintig jaar is veel wetenschappelijke literatuur geschreven over Process Mining (Werner, 2014). De gevonden informatie komt overeen met de lifecycle van Process Mining. In de 20<sup>e</sup> eeuw veel onderzoek verricht, dat geleid heeft tot de eerste opensourcetool ProM 2014. Vanaf 2010 kwamen de eerste commerciële tools op de markt.

Vervolgens zijn naar aanleiding van het Process Mining Manifest (Daniel et al., 2011) meer casestudies verschenen over dit onderwerp. In de afgelopen vijf jaar is er meer gedaan aan de toepasbaarheid in combinatie tot een specifiek beroep, maar ook in relatie tot nieuwe technologie (Van der Aalst, 2014). In het onderzoek zijn voornamelijk relevante artikelen gevonden rond de introductietijd (2010 -2015) van Process Mining. In deze periode zijn voornamelijk de aandachtspunten van het gebruik van Process Mining besproken. In het vakblad Agconnect geeft hoogleraar en grondlegger professor Van der Aalst aan dat de problemen die hij vijftien tot twintig jaar geleden heeft benoemd nog steeds van toepassing zijn (Van Heur, 2019).

### Onderzoekers

Er is specifiek gezocht naar het werk van onderzoekers die veel geschreven hebben over het onderwerp Process Mining, namelijk:

1. Professor Van der Aalst: Van der Aalst wordt ook wel de godfather of Process Mining genoemd. Hij is de bedenker van de techniek en methode. Daarnaast heeft hij de afgelopen jaren veel onderzoek verricht en meerdere boeken geschreven. Veel van zijn onderzoeken zijn geciteerd in andere onderzoeken. Professor Van der Aalst is dan ook een van de meest geciteerde computerscientist in de wereld (Van der Aalst, 2019);
2. Doktor Rozinat: Rozinat heeft veel bijgedragen aan de ontwikkeling van Process Mining, beginnend met de ontwikkelingen in ProM. Vervolgens heeft ze haar eigen commerciële tool Disco gelanceerd (Rozinat, 2018);
3. Professor Jansen: Professor bij Hasselt Universiteit (BE). Jansen heeft in het verleden veel geschreven over Process Mining in combinatie met de auditpraktijk (Jansen, 2019).

### Zoektermen

De ingevoerde zoektermen in Web of Science en Google Scholar zijn: Process Mining; Lessons Learned; best practice; attention; interpretation en quality. Vervolgens is gekeken welke bronnen zijn gebruikt bij de literatuur die middels de zoektermen is gevonden. Ook is in het netwerk gevraagd naar artikelen die relevant zouden kunnen zijn voor het onderzoek.

De volgende onderzoeken zijn gebruikt:

1. Met Process Mining Manifest is de basis gelegd voor de beschrijvende deelvragen. De genoemde richtlijnen en uitdagingen zijn vertaald naar technische en business risico's (Daniel et al., 2011);
2. Het Process-Mining-Project-Methodology in (Van der Aalst, 2015). Hiermee is de beschrijvende deelvraag te beantwoorden;
3. Casestudies en overige onderzoeken zijn gebruikt om de technische en businessrisico's aan te vullen en nadere beeldvorming te creëren over het onderwerp Process Mining.

## Theoretisch kader

Dit theoretisch kader bevat allereerst een beschrijving van Process Mining. Vervolgens worden de risico's van Process Mining naar aanleiding van het onderzoek onderverdeeld in twee subcategorieën: de technische - en de businessrisico's. Hieronder worden beide subcategorieën behandeld.

### Process Mining

#### Definitie

Process Mining is een **techniek, tool** en een **methode** om geautomatiseerde bedrijfsprocessen te visualiseren. Door middel van data-analyse wordt informatie uit systemen ontsloten en geprepareerd tot een specifiek bestand (eventlog). In dit bestand dienen minimaal de activiteiten te zijn opgenomen die in het geautomatiseerde proces zijn vastgelegd. Daarnaast dient een uniek nummer opgenomen te worden die alle activiteiten aan elkaar bundelt (case-ID). Als laatste is een tijdstempel nodig om aan te geven wanneer deze activiteiten zijn vastgelegd. Ook kan aanvullende informatie worden opgenomen (meta-data). Een (commerciële) tool kan dit bestand vervolgens middels algoritmes vertalen naar een procesmodel. Hierdoor is het geautomatiseerde bedrijfsproces gevisualiseerd en worden automatisch lijnen getrokken tussen verschillende activiteiten. Met deze techniek komen data-analyse en procesmodelleren samen (Daniel et al., 2011).

#### Toepassingsmogelijkheden

Volgens het Process Mining Manifest (Daniel et al., 2011) zijn er drie toepassingsmogelijkheden voor Process Mining. **Ten eerste** kan Process Mining gebruikt worden om het geautomatiseerde bedrijfsproces te ontdekken. Vanuit de activiteitendata wordt het werkelijke gedrag ontdekt en als procesmodel gevisualiseerd. De **tweede** toepasbaarheid betreft de controle van het bedrijfsproces. Vanuit de vastlegging in informatiesystemen kan het werkelijke gedrag worden geanalyseerd, dat wil zeggen: in hoeverre het 'theoretische' model werkelijk wordt uitgevoerd.

**Als laatste** kan Process Mining gebruikt worden om het proces te verbeteren. Hierbij wordt het procesmodel verrijkt met aanvullende informatie uit de activiteitendata, teneinde het automatische bedrijfsproces te verbeteren (Daniel et al., 2011).

### De technische risico's van Process Mining

De technische risico's van Process Mining zijn onderverdeeld in: datavastlegging, informatiesystemen, extraheren van data uit informatiesystemen en als laatste Process-Miningtool en techniek.

#### Datavastlegging

In activiteitenlogs ontbreken vaak gegevens of de gegevens zijn onjuist, onnauwkeurig of irrelevant. Dit kan ontstaan door menselijke fouten en door technische fouten. Deze gegevens beïnvloeden de analyseresultaten (Daniel et al., 2011; Van der Aalst, 2012; 2015).

De data dient aantoonbaar betrouwbaar en valide te zijn. Dit betekent dat de activiteiten ook echt plaatsgevonden moeten hebben in het systeem. Daarnaast dient de activiteit uniek te zijn en niet te verwijzen naar meerdere afzonderlijke activiteiten. Ook dient de beveiliging en privacy gegarandeerd zijn (Daniel et al., 2011).

Het kan voorkomen dat de informatie uit verschillende informatiesystemen gehaald moet worden. De informatiebronnen kan een andere (unieke) identificeerbron gebruiken.

Daarnaast kan bronformat anders zijn een voorbeeld hiervan is de vastlegging van de tijdstempel wanneer de activiteit heeft plaatsgevonden: dataformat 19-11-2019 of 19-11-2019 h12:m40s32.

Kortom, de data moeten gedestilleerd uit de systemen komen en hetzelfde format bevatten voordat deze worden samengevoegd (Andrews et al., 2019; Daniel et al., 2011).

Het bedrijfsproces hoort flexibel te zijn en kan daardoor aangepast worden. Dit kan gevolgen hebben voor de manier van vastlegging indien een tijdsperiode wordt gekozen waarin deze twee manieren van vastlegging zijn meegenomen. Hiermee dient rekening te worden gehouden tijdens de keuze van tijdsperiode (Van der Aalst, 2015). In het Process Mining Manifest worden enkele niveaus van volwassenheid van een eventlog weergegeven. (Daniel et al., 2011).

Het volwassenheidsniveau voor de activiteitenlog is opgenomen in bijlage 1.



### Extraheren van data

De informatie uit systemen ontsluiten uit datawarehouses kan als complex ervaren worden. Dit heeft te maken met de genoemde vastlegging in het datawarehouse, wat hierboven is benoemd. Daarnaast wordt deze handeling als tijdrovend beschouwd (Van der Aalst, 2012; 2015).

De volgende risico's hebben betrekking op het (juist) uitvragen van de data en het interpreteren van de data. Dit is van belang om goede aannamen te realiseren, wat vervolgens invloed heeft op de analyse en de uitkomsten van het project Process Mining.

### Samenwerking

Als eerste is de samenwerking van cruciaal belang. De verkeerde uitvraag kan ontstaan doordat de gradatieniveaus van data onduidelijk is, dan wel de interpretatie van data. Hierdoor is kennis nodig van een processpecialist en IT-specialist (system/applicatiebeheerder) om juist te interpreteren (Suriadi, 2013).

### Gradatieniveau

Ten tweede is het gradatieniveau van de data-ontsluiting van belang. Deze dient in lijn te zijn met de onderzoeksvragen en eindgebruikers (doelgroep). Het risico is dat er geen concrete onderzoeksvraag wordt geformuleerd in de beginfase, waardoor in eerste instantie onvolledige of onjuiste data worden ontsloten. Daarnaast wordt de onderzoeksvraag tijdens de analyse aangescherpt (Van der Aalst, 2012; 2015).

### Hoofdlijn

Als derde wil de eindgebruiker een bepaalde hoofdlijn of zelfs diepgang zien in een bedrijfsproces. Hierdoor moet een balans worden gevonden in welke informatie meegenomen wordt. Daarnaast moeten hierdoor soms activiteiten worden samengevoegd om deze uniek of minder complex te maken. Naar eigen inzicht van het team kunnen kleine handelingen in het systeem gemakshalve weggelaten worden. Deze keuzes moeten meegenomen worden in de onderzoekspresentatie (Van der Aalst, 2012). Het kan voorkomen dat in de gekozen tijdsperiode van de data niet alle case-ID's zijn voltooid. Dit wil niet zeggen dat deze activiteiten niet hebben plaatsgevonden. Mogelijk zijn deze niet in de dataset meegenomen of ze zijn buiten het systeem afgerond. (Daniel et al., 2011; Van der Aalst, 2012; 2015).

### Process-Miningtool en techniek

De aanleiding van het onderzoek van (Ailene, 2014) is dat door de komst van de opensourcetool ProM al snel ook commerciële tools op de markt kwamen. Dit was een positieve ontwikkeling voor de techniek Process Mining. Echter, het bracht ook meteen andere uitdagingen met zich mee. Het meest significante uit dit onderzoek was een gebrek aan standaardisatie. De commerciële tools gebruiken hun eigen normen en definities (naamgeving). Dit werkt verwarrend voor de (eind)gebruikers van de Process Mining tool (Ailene, 2014). Naar aanleiding van dit onderzoek kregen alle commerciële tools feedback. Er bestond op dat moment geen tool die alle toepassingsmogelijkheden had verwerkt en daarnaast met het gebruikte algoritme kon omgaan met alle complexiteit die van toepassing kan zijn bij Process Mining (Ailene, 2014).

De techniek van Process Mining techniek, algoritmes, zijn op een gebruiksvriendelijke interface verwerkt in (commerciële) tools. Hierdoor kunnen (eind)gebruikers door het plaatsen van enkele filters eenvoudig een procesanalyse verwerken. Door deze snelheid kan de techniek Process Mining op een verkeerde manier worden toegepast. Hierdoor worden onjuiste beslissingen genomen. (Silva, 2014).

Door deze gebruikersvriendelijke interface is het voor de eindgebruiker soms onduidelijk op welke data hij of zij heeft gefilterd waardoor andere informatie niet meer zichtbaar is in het model (Van der Aalst, 2015; Silva, 2014).

De (academische) opensourcetool ProM wordt voornamelijk door businessusers als ongebruiksvriendelijke tool beschouwd. Dit heeft te maken met de filtermogelijkheden en het zelf bepalen van het algoritme (plugins) (Van der Aalst, 2015). Mede om de complexiteit van het gebruik wordt deze tool niet gekozen voor commerciële doeleinden, omdat de kennis van deze tool vaak ontbreekt in de organisatie (Silva, 2014).



## De businessrisico's van Process Mining

Volgens het Process Mining Manifest (Daniel et al., 2011) wordt aangegeven dat beginnende gebruikers van Process Mining niet alle mogelijkheden dan wel beperkingen van de techniek en methode weten. In het manifest wordt gesproken over richtlijnen en uitdagingen bij het toepassen van de methode en de tool Process Mining (Daniel et al., 2011). De businessrisico's van Process Mining zijn onderverdeeld in Process Mining Team, Process Mining Gebruik en als laatste het Process Mining Model.

### Team

Het Process Mining team is een groep professionals die bepaalde kennis en vaardigheden heeft om het bedrijfsproces en data te doorgronden. Daarnaast dienen alle leden van het Process Mining team de complexiteit en de toepassingsmogelijkheden van het gebruik van Process Mining te onderkennen (Suriadi, 2013). Uit een casestudie (Van der Aalst, 2015) blijkt dat twee dingen fout gingen in de teamsamenstelling. **Als eerste** was de teamsamenstelling in de beginfase niet compleet. Hierdoor werd het bedrijfsproces in eerste instantie verkeerd geïnterpreteerd. Dit kwam omdat de procesanalist onvoldoende bedrijfskennis had van de organisatie en het gebrek aan interpretatie van de data. **Als tweede** bleek dat niet alle leden genoeg achtergrondkennis hadden van de methode, de techniek en mogelijkheden van Process Mining. Hierdoor was tevens de interpretatie van de bevindingen moeilijk en tijdrovend (Van der Aalst, 2015).

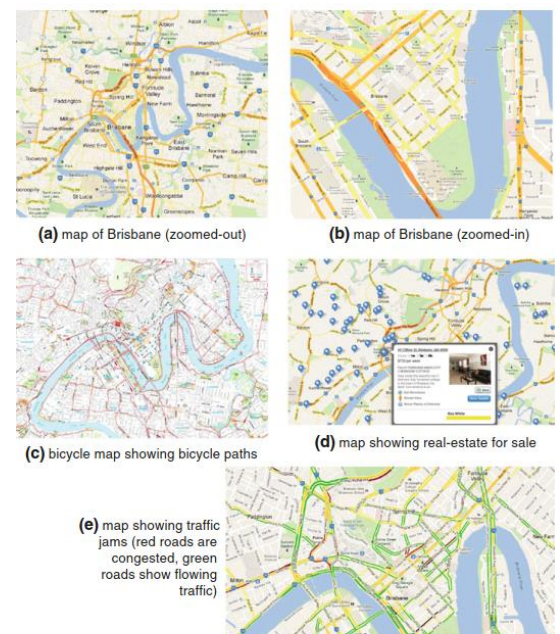
### Gebruik

Te beginnen met enkele misvattingen over Process Mining. Met name werd de techniek tijdens de introductiefase alleen het ontdekken van het bedrijfsproces gebruikt. Echter, dit betreft slechts één toepassingsmogelijkheid van Process Mining. Daarnaast moet vooraf duidelijk afgestemd worden welke data beschikbaar zijn, wat onderzocht gaat worden en welke concrete onderzoeksvragen beantwoord kunnen worden. Zonder hierover nagedacht te hebben, is het bijna onmogelijk om zinvolle eventdata te extraheren vanuit de informatiesystemen. Daarnaast kunnen de abstracte onderzoeksvragen tijdens de analyse worden aangepast om zo alsnog de vragen te beantwoorden (Daniel et al., 2011; Van der Aalst, 2015).

### Model

Het Process Mining Model wordt gevisualiseerd door middel van een procesmodel. Dit model dient gezien te worden als een landkaart. Er zijn verschillende landkaarten denkbaar voor verschillende eindgebruikers. Hierbij kan gedacht worden aan een stadskaat, fietskaart of autokaart. Elke kaart is een weergave van eenzelfde geografische gebied; er bestaat geen perfecte landkaart. Dit geldt ook voor Process Mining. Echter, het Process Mining Model moet aansluiten bij de gewenste eindgebruiker. Hiermee wordt bedoeld dat de data op (detail)niveau aansluit bij de procesanalyse en gerelateerd is aan de eisen van de eindgebruiker (Daniel et al., 2011; Aalst, 2012).

De Process Mining Tool geeft altijd een visuele weergave van de ingelezen eventlog (case-ID, activiteiten en tijdstempel). Dit wil niet zeggen dat het resultaat daadwerkelijk bruikbaar is. Daarnaast kan de eindgebruiker de output (het model) op een verkeerde manier opvatten, waardoor verkeerde conclusies worden getrokken (Daniel et al., 2011; Van der Aalst, 2012).



Figuur 4 - Landkaart (Van der Aalst, 2012)



## De Process-Mining-Project-Methodology

### Introductie

De Process-Mining-Project-Methodology is ontstaan naar aanleiding van het onderzoek van (Van der Aalst, 2015), waarin deze methode direct is toegepast in een casestudie. Dit onderzoek is uitgevoerd omdat de bestaande methoden, Process Diagnostics Methode (PDM) en The Life-cycle model, niet geschikt waren voor elk project. Daarnaast waren deze methodes te algemeen of benadrukten ze alleen een specifieke toepassingsbenadering, namelijk inzicht krijgen in bedrijfsprocessen (Van der Aalst, 2015).

Aan het einde van deze paragraaf is de Process-Mining-Project-Methodology opgenomen. De zes fases van Process-Mining-Project-Methodology zijn: planning, extractie, dataprocessing, mining & analyses, evaluatie en procesverbetering & ondersteuning. De fases worden in onderstaande alinea's nader toegelicht. Hierbij is gebruikgemaakt van het onderzoek van Van der Aalst (2015) en Suriadi (2013).

### Fase 1: Planning

Tijdens de planningfase wordt het project van Process Mining vormgegeven. Hier wordt de businessdrive om de techniek Process Mining in te zetten, bepaald. Hierbij komen dan ook de drie toepassingsmogelijkheden terug: het ontdekken, controleren en verbeteren van het bedrijfsproces. Vervolgens moet worden vastgesteld welk gedeelte van het bedrijfsproces geanalyseerd moet worden. Vervolgens wordt nagegaan of de data betrouwbaar en volledig zijn vastgelegd in de informatiesystemen. Indien dit het geval is, kunnen gerichte onderzoeksvragen worden geformuleerd (Van der Aalst, 2015).

Als laatste wordt het team samengesteld. Uit eerder onderzoek van (Suriadi, 2013) is gebleken dat de volgende professionals deel dienen te nemen aan Process Mining: de proceseigenaar, bedrijfseigenaar, systeembeheerder en procesanalist. Elke professional draagt bij binnen zijn of haar expertise.

### Fase 2: Extractie

In de extractiefase worden drie subfases herkend om de gebeurtenisgegevens uit (verschillende) systemen te verkrijgen. De eerste is de scopebepaling. Hier wordt dus de reikwijdte bepaald. Hierbij kan gedacht worden aan tijdsperiode, welk detailniveau (granulariteit) van de data dient ontsloten te worden en als laatste welke verbanden gelegd moeten worden om de data te verzamelen uit meerdere informatiesystemen. De tweede subfase betreft de extractie van de gebeurtenisgegevens tot één verzameling bestand. De laatste en derde fase heeft betrekking op het overbrengen van proceskennis om zo bijvoorbeeld meerdere activiteiten samen te voegen (Van der Aalst, 2015).

### Fase 3: Dataprocessing

Deze dataprocessingsfase is bedoeld om de data uit de extractiefase zo te verwerken dat optimaal gebruikgemaakt wordt van Process Mining (eventlog). In deze fase worden vier subfases gedefinieerd. **De eerste** betreft data zo te ontsluiten dat deze aansluit bij de specifieke analyse en eindgebruiker. **De tweede** subfase is het nagaan of activiteiten mogelijk kunnen worden samengevoegd. Er kan gekozen worden om in de dataset te kiezen voor een specifiek gedeelte van een subproces of toch een algemeen detailniveau. **De derde** subfase is het verreiken en aanvullen van de eventlog. Hierbij kan gedacht worden aan interne en externe informatie over de uitgevoerde activiteiten in de systemen. **Tot slot** is het handig om van tevoren na te gaan of alle gegevens in de eventlog verwerkt zijn om de analyse uit te voeren. Op welke (aanvullende) data informatie gefilterd kan worden en welke filtertechnieken gebruikt kunnen worden (Van der Aalst, 2015).

### Fase 4: Mining & analyse

In de fase mining & analyse wordt Process Mining toegepast (als tool) op basis van de samengestelde eventlog. Door de Process Mining wordt inzicht verkregen in de prestaties en naleving van het geautomatiseerde bedrijfsproces. Met deze inzichten worden (concrete) onderzoeksvragen geanalyseerd en tevens beantwoord. Daarnaast zou een andere dataminingstechniek of -visualisatie-analyse kunnen worden toegepast (Van der Aalst, 2015).





### Fase 5: Evaluatie

De evaluatiefase is bedoeld om het project te evalueren met het Process Mining team en hierbij te concluderen of de analysebevindingen correct zijn geïnterpreteerd. Daarnaast moet het team verifiëren en valideren dat er geen abnormale uitvoeringen zijn geconstateerd die niet overeenkomen met de data en of bedrijfsproces. Vervolgens kan worden besproken welke vervolgaanpassingen dan wel acties, verbeteringen en ideeën (nieuwe onderzoeksvragen) kunnen worden uitgewerkt (Van der Aalst, 2015).

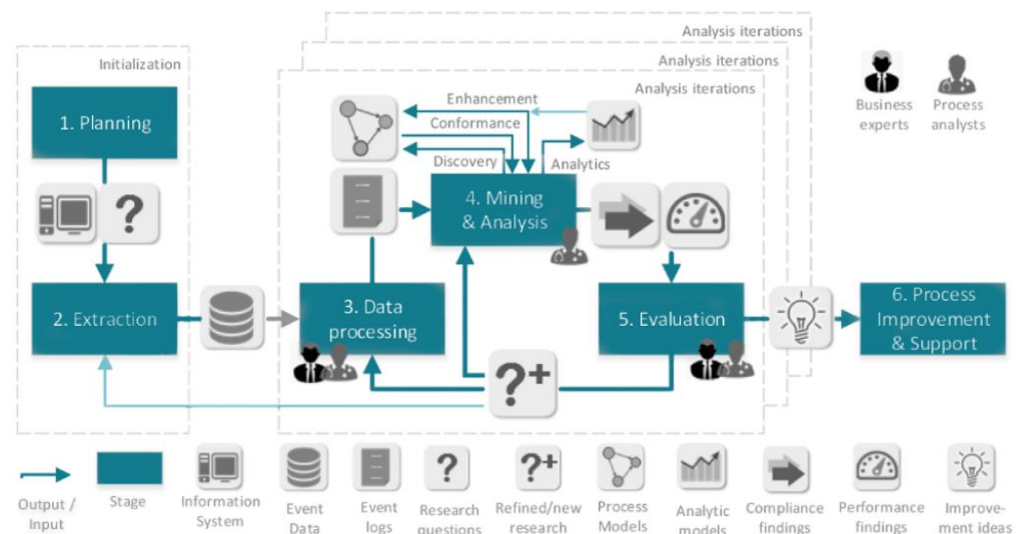
### Fase 6: Procesverbetering & ondersteuning

De laatste fase, procesverbetering & ondersteuning, is opgenomen om de inzichten van het Process Mining te gebruiken om het bedrijfsproces te verbeteren. Hierbij kan gedacht worden aan het lean-six-sigmatraject, maar ook de ondersteuning van het bedrijfsproces om de Process Mining vaker toe te passen (Van der Aalst, 2015).

### Het project

De Process-Mining-Project-Methodology kan het Proces Mining team helpen om op een gestructureerde wijze een project uit te voeren. Wanneer de planningsfase wordt overgeslagen, is de kans groot dat latere projectstadia vertragen of dat het project zelfs stagneert (Van der Aalst, 2015).

Met name in de evaluatiefase wordt door middel van de analysebevindingen nagegaan of deze correct zijn geïnterpreteerd. Daarnaast worden deze geverifieerd en gevalideerd. Deze stap komt ten goede aan de kwaliteit van de uitkomsten en verbeteringen voor Process Mining binnen de organisatie en voor de (eind)gebruikers. (Van der Aalst, 2015).



Figuur 5- Process-Mining-Project-Methodology (Van der Aalst, 2015).

### Deelconclusie

Uit bovenstaand literatuuronderzoek is naar voren gekomen dat zonder tijdige, relevante en accurate data geen betrouwbare (data-)analyse kan worden uitgevoerd.

Daarnaast is het van belang dat er een duidelijke en concrete analysevraag geformuleerd wordt. Hierbij moet het granulariteit ook nog eens aan sluiten met de eindgebruiker.

De interpretatie van het bedrijfsproces of data is van belang om de data te ontsluiten en te analyseren. Hierbij is de teamsamenstelling dan ook divers met verschillende specialiteiten.

Voor het analyseren kan zowel een open source als commerciële Process-Miningtool worden ingezet. Bij de laatste ontbreekt de standaardisatie. De commerciële tools gebruiken hun eigen normen en naamgeving. Hierbij wordt gebruikgemaakt van verschillende algoritmes, die verschillende uitkomsten hebben.

Er zijn enkele misvattingen over Process Mining. Daarnaast kunnen ook aanvullende data-analysetechnieken worden ingezet.





## Resultaten

### De risico's

In het theoretisch kader is beschreven wat Process Mining is en welke risico's zich kunnen voordoen. Ook is beschreven wat de Process-Mining-Project-Methodology inhoudt en hoe deze kan worden ingezet als beheersmaatregel. In deze paragraaf wordt antwoord gezocht op de analyserende deelvraag: *Welke risico's van Process Mining worden door Process-Mining-Project-Methodology als beheersmaatregelen in oogschouwen genomen?*

#### Introductie

De risico's van Process Mining kunnen in een latere fase van het Process-Mining-Project-Methodology ontdekt worden. Per fase is verschillende input nodig om vervolgens output te leveren aan de volgende procesfase. Hiermee kan een (verkeerde) keuze invloed hebben op een volgende fase en hiermee dus op het gehele project (Van der Aalst, 2015).

Daarnaast kan een (verkeerde) businesskeuze de technische risico's beïnvloeden en vice versa. In onderstaande alinea's wordt *per fase aangegeven welke risico's worden afgedekt met de Process-Mining-Project-Methodology*. In de conclusie is een matrix opgenomen waarin zichtbaar is welke risico's niet worden afgedekt met deze methode.

#### Fase 1: Planning

Tijdens de planningsfase worden twee risico's gedefinieerd: **Het eerste risico** de keuze van een (sub)bedrijfsproces geen concrete vragen worden geformuleerd en dat deze niet zijn gevalideerd. **Als tweede** risico is de teamsamenstelling onvolledig is, waardoor de data of het businessproces niet (betrouwbaar) kunnen worden geïnterpreteerd dan wel gevalideerd.

#### Fase 2: Extractie

Tijdens de extractiefase worden drie risico's gedefinieerd. **Het eerste** risico betreft dat de data niet of onvolledig is vastgelegd. **Het tweede** risico betreft een onjuiste business-interpretatie van het bedrijfsproces, waardoor bepaalde data niet worden meegenomen in de analyse. **Als derde** risico betreft dat het vorige risico zich voordoet bij het verkeerde detailniveau.

#### Fase 3: Dataprocessing

Tijdens de fase dataprocessing worden twee risico's gedefinieerd. **Het eerste** risico in deze fase is dat de activiteiten niet gecontroleerd zijn indien deze uniek zijn. **Als tweede** risico is dat de eventlog niet aansluit bij de gewenste datagranulariteit voor het detailniveau dat de eindgebruikers graag zouden willen zien.

#### Fase 4: Mining & analyse

Tijdens de fase mining & analyse worden drie risico's gedefinieerd. **Het eerste** risico dat de analysevragen niet beantwoorde kunnen worden, omdat ze niet concreet geformuleerd waren. **Het tweede** risico is dat de data-analyse en filtertechnieken binnen Process Mining onjuist worden toegepast. **Als derde** is het risico om direct nieuwe vragen te willen beantwoorden.

#### Fase 5: Evaluatie

Het overslaan van de evaluatiefase is ten nadele van de kwaliteit en de betrouwbaarheid van de uitkomsten van het project.



## Deelconclusie

Om de relatie tussen Process Mining risico's en de Process-Mining-Project-Methodology aan te geven, is er gebruikgemaakt van tabel 1. Deze tabel geeft aan welke risico's van Process Mining worden afgedekt in welke fases van Process-Mining-Project-Methodology.

Analyserende deelvraag			
	Risico's	PM2 methode	Bron
T e c h n i s c h e	<b>Data vastlegging</b> 1. Het niet op tijd inventariseren van de beschikbare data 2. Het ontbreken van (een) tijdige, relevante en betrouwbare data(bron) 3. Brondata samenvoegen welke geen gezamenlijke format en/of unieke identificeerbron heeft 4. Tijdsperiode, waarbij twee manieren van vastlegging zijn opgenomen	<b>Beheersingsfase</b> Planning en extractie Planning en extractie Planning en extractie Planning en extractie	(Daniel et al., 2011; Van der Aalst, 2015). (Daniel et al., 2011; Van der Aalst, 2012; 2015). (Hofstede, 2019; Daniel et al., 2011). (Aalst, 2015).
	<b>Extraheren van data</b> 1. Geen concrete analyse vraag, waardoor er onvolledig en/of onjuiste data wordt ontsloten 2. Granulariteit sluit niet aan bij eindgebruiker(s) dan wel de analyse vraag 3. Verkeerde data interpretatie van het extraheren van de brondata (teamsamenstelling) 4. Data is niet opgeschoond (samenvoegen en of verwijderen van activiteiten)	<b>Beheersingsfase</b> Planning en extractie Planning, extractie en data preparatie Planning, extractie en data preparatie Extractie en data preparatie	(Suriadi, 2013; Daniel et al., 2011). (Van der Aalst, 2012; 2015). (Suriadi, 2013; Van der Aalst, 2012; 2015). (Aalst, 2012).
	<b>Process Mining techniek en tool</b> 1. Keuze Process Mining tool en techniek (teamsamenstelling) 2. Gebrek aan standaardisatie commerciële Process Mining tool (normen en algoritme) 3. Data filtertechnieken onjuist gebruiken dan wel interpreteren	<b>Beheersingsfase</b>	(Ailene, 2014; Suriadi, 2013). (Ailene, 2014). (Suriadi, 2013; Van der Aalst, 2015).
B u s i n e s	<b>Process Mining team</b> 1. Er is onvoldoende kennis aanwezig om project Process Mining uit te voeren 2. Verkeerde bedrijfsproces interpretatie (teamsamenstelling)	<b>Beheersingsfase</b> Planning Planning en extractie	(Suriadi, 2013; Van der Aalst, 2015). (Suriadi, 2013; Van der Aalst, 2015).
	<b>Process Mining gebruik</b> 1. Geen concrete analyse vraag, waardoor er onvolledig en/of onjuiste data wordt ontsloten 2. Ideaal beeld van de beschikbare data 3. Ideaal beeld van de toepassingsmogelijkheden van Process Mining 4. Het ontbreken van documentatie over de gemaakte keuzes	<b>Beheersingsfase</b> Planning en extractie	(Suriadi, 2013; Daniel et al., 2011). (Ailene, 2014). (Ailene, 2014; Daniel et al., 2011).
	<b>Process Mining model</b> 1. Process Mining model wordt op een verkeerde manier geïnterpreteerd 2. Gradatie niveau sluit niet aan bij analyse vraag/ eindgebruiker 3. Er worden te (snel) conclusies getrokken over de Process Mining analyses/uitkomsten	<b>Beheersingsfase</b> Evaluatie Evaluatie Evaluatie	(Daniel et al., 2011; Van der Aalst, 2012). (Van der Aalst, 2012; 2015). (Silva, 2014; Van der Aalst, 2015).

Process-Mining-Project-Methodology (Van der Aalst, 2015).

Tabel 1 Resultaten Process Mining risico's en Process-Mining-Project-Methodology.

In de volgende paragraaf zijn aanvullende beheersingsactiviteiten opgenomen om die risico's op te nemen die niet worden afgedekt door Process-Mining-Project-Methodology.



## Beheersmaatregelen

De beheersmaatregelen bevatten een beschrijving welke aanvullende beheersingsactiviteiten nodig zijn wanneer Process-Mining-Project-Methodology wordt toegepast. Vervolgens wordt praktijkervaring gedeeld en als laatste wordt de rol van de IT- auditor behandeld.

### Aanvullende beheersingsmaatregelen

In deze paragraaf wordt antwoord gezocht op de analyserende deelvraag: *Wat is de toegevoegde waarde van Process-Mining-Project-Methodology voor het beheersen van risico's bij Process Mining?*

In de vorige paragraaf is gebleken dat de Process-Mining-Project-Methodology de meeste risico's afdekt als de projectmethodiek wordt gehanteerd. Echter, het is gebleken dat niet alle risico-onderdelen zijn besproken dan wel volledig zijn behandeld.

Hieronder wordt per fase aangegeven welke risico's niet worden afgedekt met Process-Mining-Project-Methodology. Vervolgens wordt aangegeven welke aanvullende beheersingsactiviteiten toegepast kunnen worden. Daarnaast is er één procesfase toegevoegd, namelijk 'presentatie en documentatie.' Het team dient andere belanghebbenden dan wel eindgebruikers goed te informeren welke keuzes (bedrijfsproces en dataset) ze hebben gemaakt tijdens het project. Hierbij is presentatie en documentatie onmisbaar en dus toegevoegd als extra fase aan de huidige beschreven Process-Mining-Project-Methodology.

#### Fase 1: Planning

Twee onderdelen worden toegevoegd aan de planningsfase: de teamsamenstelling en de projectvoorbereiding.

#### Teamsamenstelling

De teamsamenstelling komt in de Process-Mining-Project-Methodology als laatste. In dit onderzoek wordt deze activiteit als eerste opgenomen. Het is van belang dat teamleden, maar ook sponsors gevonden worden, zodat het project niet alleen kan worden gestart maar ook kan slagen (Rozinat & Geffen, 2015).

De teamsamenstelling is van belang vanaf de eerste fases van Process-Mining-Project-Methodology tot en met de laatste fase. Zo wordt misinterpretatie van het bedrijfsproces en data voorkomen. Tijdens de planningsfase dient ook in kaart te worden gebracht over welke kennis het team beschikt en welke kennis opgedaan moet worden of moet worden uitbesteed. In academisch onderzoek wordt gesproken over vier teamleden: de proceseigenaar, bedrijfseigenaar, systeembeheerder en procesanalist (Suriadi, 2013). In praktijk wordt toch vaak gesproken over twee aanvullende teamleden, namelijk (IT-)auditor of proces-verbeteraar /consultant om het proces te ondersteunen en hierbij een aanvullende expertise aan te bieden (KNVI, 2018).

#### Projectvoorbereiding

Bij de projectvoorbereiding is het van belang dat bij een eerste Process Mining Project gekozen wordt voor een relatief eenvoudig bedrijfsproces, die (eind)gebruikers goed (her)kennen en relatief eenvoudig aan data ontsloten kan worden. Door eenvoudig te beginnen, is het mogelijk om de techniek, methode en tool van Process Mining eenvoudiger te beheersen (Rozinat & Geffen, 2015).

Daarnaast is het van belang in deze fase na te denken over welke Process-Miningtool wordt toegepast. Als eerst dient te worden gekozen voor een opensourcetool of commerciële tool. In het artikel Agconnect geeft Van der Aalst aan dat er voor de toepassingsmogelijkheid nu twee wetenschappelijke denkwijzen zijn: één model dat heel specifiek is maar visueel niet leesbaar of een minder specifieke weergave maar wel visueel analyseerbaar. Van der Aalst geeft aan dat nog geleerd moet worden wanneer welk procesmodel gebruikt kan worden voor welke eindgebruiker(s) (Agconnect, 2018).

De uiteindelijke keuze voor een Process-Miningtool dient gebaseerd te zijn op de kennis en ervaring van het team. Daarnaast is het aan te raden om alle groepsleden een (gezamenlijke) Process Mining (basic)training te laten volgen. Hierdoor kan geen misvatting ontstaan van Process Mining als hulpmiddel en de toepassingsmogelijkheden. Daarnaast kunnen (toolspecifieke) filteranalyses worden uitgelegd.

De techniek van Process Mining kan door eindgebruikers verkeerde verwachtingen creëren. De Process Mining is bijvoorbeeld geen procesmodelleringstool, simulatietool, business-intelligencedashboard, implementatieproject of data mining. De laatstgenoemde kan soms zelfs sneller inzicht verschaffen dan Process Mining. Een combinatie van beide technieken is dan ook zeker waardevol tijdens een Process Mining Project (Rozinat & Geffen, 2015).



### Fase 2: Extraction

Er kan een irreëel beeld ontstaan van de vastgelegde en beschikbare data in de datawarehouse (Rozinat & Geffen, 2015). Daarom is het verstandig dat een proces eigenaar en IT- of applicatiebeheerder samen vaststellen welke data ontsloten dient te worden. Hierbij moet tevens rekening gehouden worden met de geldende wet- en regelgeving met name de Algemene verordening gegevensbescherming (AVG). Hier kan gekozen worden om de data te anonimiseren, zodat deze niet meer herleidbaar is naar een individu (Agconnect, 2018).

Indien de data op een hoger aggrergratieniveau wordt ontsloten uit systemen, moet de data hiervoor éénmalig worden ontsloten. In de volgende fases kunnen meerdere analysebestanden gemaakt worden met verschillende aggrergratieniveaus (Van der Aalst, 2015).

### Fase 3: Dataprocessing

De uiteindelijke ontsloten data van de vorige fase dienen verfijnd te worden voor het project. Data moeten geïnterpreteerd worden. Hierdoor is het van belang dat de juiste personen deelnemen aan deze fase. De procesanalist heeft in het algemeen geen specifieke proceskennis. Deze functie moet ondersteund worden door de proceseigenaar om zo de data juist te interpreteren en samen te komen tot een juiste en volledige dataset.

### Fase 4: Mining & analyse

De fase mining & analyse is relatief kort; het verkrijgen van de data betreft 80% en de data-analyse en presentatie is 20%. Het plaatsen van filter op filter verloopt eenvoudig en snel, wat de analyse bevordert. Echter, de eindgebruiker kan door dit filter soms moeilijker vaststellen welke data hiermee achterwege blijft. Dit komt mede doordat dit verwerkt is in de tool. De data zijn dus niet direct zichtbaar. Daarom is het verstandig om data-analyse naast de techniek van Process Mining toe te passen. Daarnaast worden hierdoor conclusies getrokken zonder dat de uitkomst is geëvalueerd met het team. Process Mining geeft nieuwe inzichten in het bedrijfsproces. Daardoor ontstaan nieuwe (relevante) vragen. Het (direct) willen beantwoorden van deze nieuwe vragen kost veel tijd, waardoor het project mogelijk niet wordt afgerond in het daarvoor bestemde tijdsbestek.

Vanaf de extractiefase tot en met de fase mining & analyse is het risico dat de in- en output van data vanaf de datawarehouse naar data-analysetool en vervolgens naar Process-Miningtool niet op elkaar aansluiten. Hierdoor is niet gecontroleerd of alle data volledig of juist zijn overgezet naar de volgende fase.

### Fase 5: Evaluatie

In de planningsfase is al aangegeven dat de teamsamenstelling van belang is vanaf de eerste fases van Process-Mining-Project-Methodology tot en met de laatste fase. Echter, wanneer per fase wordt geëvalueerd, wordt voorkomen dat een project opnieuw gedaan moet worden, omdat in deze fase pas wordt geconstateerd dat er verkeerde interpretaties zijn gedaan waardoor de bevindingen niet meer relevant en bruikbaar zijn voor de organisatie.

### Fase 6: Procesverbetering en support

Na dit project stoppen met Process Mining leidt tot gemiste kansen. Ten eerste komen de nieuwe inzichten niet tot hun recht. Ten tweede kan de eerdere data-ontsluiting (in fase 2) juist gebruikt worden voor andere analyses op hetzelfde bedrijfsproces.

Bij proces-optimalisatie kan ook in de beginfase gekozen worden voor een lean-aanpak. Hiermee wordt een open en vertrouwelijke omgeving gecreëerd, wat nodig is om het bedrijfsproces goed te interpreteren, observeren en optimaliseren. Tevens kan na het doorvoeren van de optimalisatie gemeten worden hoe het proces verlopen is. Hier kunnen vervolgens korte dataperiodes ontsloten worden.

### Fase 7: Presentatie en vastlegging

Een aanvullende fase is de vastlegging van de gemaakte keuzes per projectfase Process-Mining-Project-Methodology. Er zijn namelijk keuzes en interpretaties gedaan van de data en het bedrijfsproces. Deze moeten meegenomen worden in de bevindingen die voortkomen uit het Process Mining Project. Zo wordt voorkomen dat andere eindgebruikers buiten de projectgroep de uitkomsten onjuist interpreteren.



## Process Mining praktijkervaring

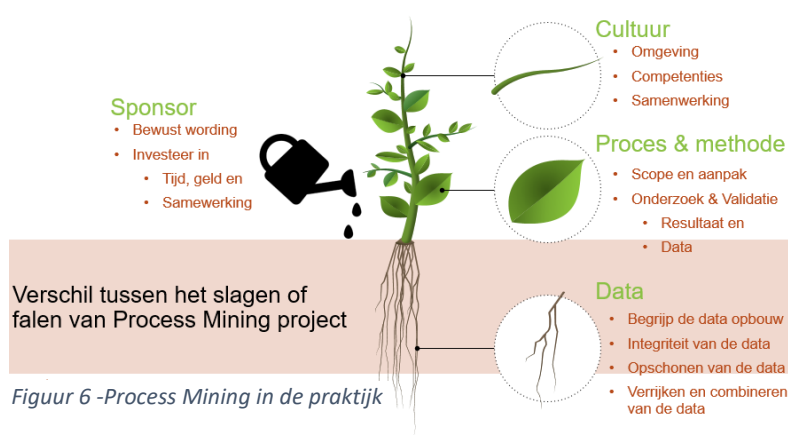
Process Mining is continue in beweging, door nieuwe inzichten uit onderzoeken en technische mogelijkheden. De wet- en regelgeving zoals de Algemene Verordening Gegevensbescherming. Kan het verwerven van data lastiger maken.

Process-Mining-Project-Methodology is bij de organisatie IBM in de praktijk toegepast (Aalst W. v., 2015). Hieruit is gebleken dat er een verschil is tussen de praktijk en theorie. De methodology helpt beide denkwijze, hoe een project kan worden uitgevoerd. De uitvoering kan anders zijn door praktijk omstandigheden. In figuur 6 is aangegeven welke praktijk omstandigheden er kunnen bijdragen bij het slagen dan wel falen van een project.

Process Mining begint meestal met een initiatief, vervolgens is het van belang dat er **Sponsor(s)** zijn die tijd en energie investeren om een eerste stap te kunnen zetten voor het gebruik van Process Mining in de praktijk. Hierbij is een manager en/of vennoot nodig voor een duurzaam project (in tijd).

Er moet bewustwording gecreëerd worden van de toepassingsmogelijkheden maar ook van de uitdagingen van Process Mining.

Vervolgens wordt Process Mining benaderd vanuit de **data** (IST-model) en niet vanuit de klassieke Business Process Management (BPM) aanpak (Haasnoot, 2012), (zie bijlage 1, figuur 8). Het is belangrijk om te begrijpen hoe de opbouw van data is en hoe deze tot stand is gekomen. Tijdens het inlezen van de data wordt er aansluiting gemaakt met vorige brondata, zodat er gecontroleerd kan worden of alles velden en regels ingelezen zijn.



Gedurende het hele **proces** moet er een balans zijn tussen de onderzoeksvraag en het detailniveau van de data. Met name de balans vinden tussen het gekozen bedrijfsproces en het aantal activiteiten (samenvoegen) is een uitdaging.

Een te eenvoudig bedrijfsproces, brengt beperkte nieuwe inzichten, een te complex proces maakt het lastig om de analyse uit te voeren. Er wordt voorzichtig omgegaan met de eerste resultaten van Process Mining, er wordt dan ook gesproken over observaties, wanneer er nog geen nader onderzoek dan wel validatie heeft plaats gevonden op de data en het bedrijfsproces.

Als laatste is de **cultuur** en omgeving waar het project wordt uitgevoerd erg belangrijk. De teamleden moeten allen in staat zijn om de gevraagde werkzaamheden uit te voeren, maar ook om samen te werken. Wanneer zij niet bereid zijn om samen te werken of dit verloopt moeizaam heeft een project minder kans van slagen. Daarnaast draagt een open en vertrouwelijke omgeving bij aan het leerproces. Het is belangrijk dat er in de organisatie en breed draagvlak is voor het toepassen van deze nieuwe manier van data-analyse. Wanneer er weinig vertrouwen is, is het bijna onmogelijk om tijdig relevante en accurate data te krijgen om een betrouwbare analyse te kunnen uitvoeren.

Na een (succesvol) Process Mining project waarbij de data en cultuur (voedingsbodem) goed is kan er gekeken worden, welke aanvullende analyses er kunnen worden gedaan met dezelfde (of eerdere) data ontsluiting. Als laatste kan met dezelfde data-ontsluiting (query) de data opnieuw ontsloten worden om in het daaropvolgende jaar Process Mining nogmaals toe te passen.



## Process Mining professionals

Het viel Professor Van der Aalst op dat voornamelijk de oude generatie IT- auditors minder analytisch zijn opgeleid, waar de jongere generatie wel analytisch wordt opgeleid (Agconnect, 2018). De nieuwe generatie groeit op met data en heeft hierdoor meer analytisch vermogen.

Als professional bewaak je de interpretatie van de data dan wel de observaties en uitkomsten van Process Mining. De nuance van data-analyse dient ook terug te komen in de presentatie en documentatie.

### Techniek

Process Mining maakt gebruik van meerdere combinaties van algoritmes. Er zijn verschillende algoritmes in de academische wereld. Afhankelijk van het gekozen algoritme kan er een andere model als uitkomst worden gepresenteerd, met dezelfde data.

Dit heeft te maken met verschillende algoritmes die beter kunnen omgaan met verschillende omstandigheden (Ailene, 2014). De werkgroep Algorithm Assurance van Nederlandse Order van Register EDP- auditors buigt zich over de techniek, die steeds meer invloed heeft op onze dagelijkse activiteiten.

### Analyse

Een data-analist kan helpen met het ontsluiten, verarmen en/of verrijken van de brondata. Daarnaast beschikt de professional over competenties om verschillende data-analyse tools toe te passen.

### Assurance en advisory

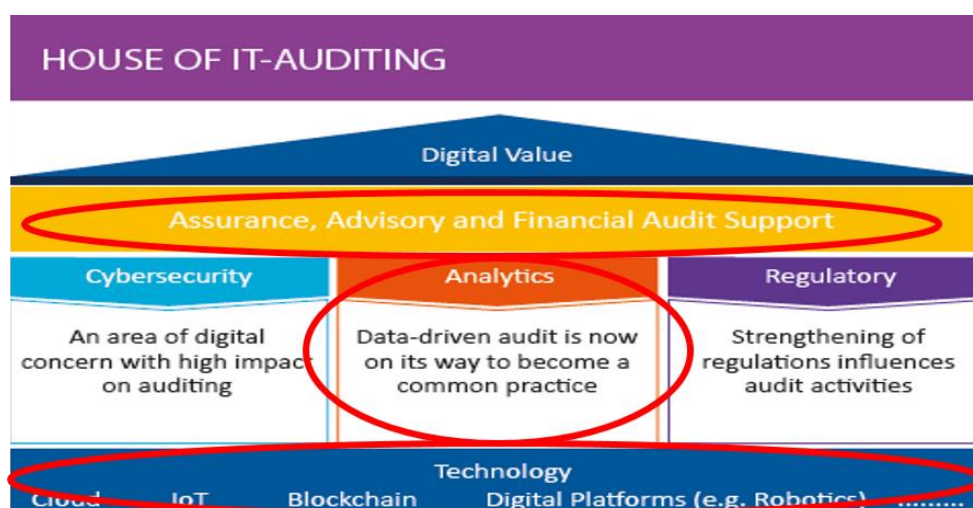
Process Mining kan toegepast worden in het financiële auditdomein, maar ook als procesoptimalisatie.

Waar in de Assurance een bepaalde mate van zekerheid wordt gegeven over de uitkomsten van Process Mining is het noodzakelijk dat de gegevens betrouwbaar en volledig zijn.

Een betrouwbaarheid van 100% is onhaalbaar, er is altijd een rest risico aanwezig. Er kan alleen beoordeeld worden wat je hebt gezien en gecontroleerd. Dit geldt niet alleen voor data-analyse-aanpak, maar ook voor de traditionele aanpak, zoals papier en lijncontrole.

Een verschil tussen Assurance en advisory is het normenkader, waar bij Assurance de NV COS 4000-4030 bij het gebruik van Process Mining van toepassing is, wordt ervoor advisory geen regelgeving opgelegd.

Er zijn binnen het Assurance domein twee groepen te onderscheiden. De eerste groep maakt gebruik van data-analyse technieken. De andere groep is afwachtend bij het toepassen van data-analyse. Zij vinden het met name moeilijk om de eerste stap te zetten, wanneer deze stap wel al is gezet kan het zijn dat er geen nieuwe inzichten zijn en/of dat het geen nieuwe audit bewijzen oplevert.



Figuur 7 - House of IT auditing.





## Conclusie

In dit onderzoek is gezocht naar een antwoord op de onderzoeksvraag: *Welke risico's van Process Mining spelen een rol bij Process-Mining-Project-Methodology?* Op deze onderzoeksvraag kan een antwoord worden geformuleerd middels vier deelvragen:

### Beschrijvende deelvragen:

- Wat zijn de risico's van Process Mining?
- Wat is de Process-Mining-Project-Methodology als beheersmaatregel?

### Analyserend deelvraag:

- Welke risico's van Process Mining worden door Process-Mining-Project-Methodology als beheersmaatregelen in ogenschouw genomen?

### Beschouwend deelvraag:

- Wat is de toegevoegde waarde van Process-Mining-Project-Methodology voor het beheersen van risico's bij Process Mining?

### Risico's van Process Mining

Uit literatuuronderzoek is naar voren gekomen dat de risico's kunnen worden onderverdeeld in technische - en bussinessrisico's. In onderstaande alinea's worden beide subcategorieën behandeld.

De **technische risico's** betreffen met name de datavastlegging, het extraheren van de data en de process-miningtechniek en -tools. Zonder tijdige, relevante en accurate data kan geen betrouwbare analyse worden uitgevoerd. Daarnaast is het van belang dat er een duidelijke en concrete analysevraag geformuleerd wordt om te waarborgen dat de juiste data worden ontsloten. Hierbij moet het granulariteitsniveau aansluiten bij de eindgebruiker. Als laatste zijn de risico's dat de data verkeerd worden geïnterpreteerd of niet worden opgeschoond. Waardoor de uitkomsten niet betrouwbaar of irrelevant zijn voor de organisatie.

Bij de **business risico's** zijn met name het process-miningteam, het gebruik en model van belang. Door een onvolledige teamsamenstelling kunnen data, bedrijfsprocessen en of data-analysetools verkeerd worden geïnterpreteerd. Het ontbreken van een concrete analysevraag maakt de analyse in data-analysetools onmogelijk. De Process Mining Tool geeft altijd een visuele weergave van de ingelezen eventlog. Dit betekent niet dat het resultaat daadwerkelijk betrouwbaar en bruikbaar is. Daarnaast kan de eindgebruiker de output (het model) op een verkeerde manier opvatten, waardoor verkeerde conclusies worden getrokken.

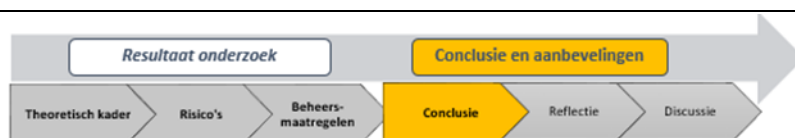
### Process-Mining-projectmethodology

Uit literatuuronderzoek is naar voren gekomen dat PM2 uit zes fases bestaat: planning, extractie, dataprocessing, mining & analyses, evaluatie en procesverbetering & ondersteuning. De PM2 kan het Process Mining team helpen om op een gestructureerde wijze een project uit te voeren. Wanneer de planningsfase wordt overgeslagen, is de kans groot dat latere projectstadia vertragen of dat het project zelfs stagneert. Met name in de evaluatiefase wordt door middel van de analysebevindingen nagegaan of deze correct zijn geïnterpreteerd. Daarnaast worden deze geverifieerd en gevalideerd. Deze stap komt ten goede aan de kwaliteit van de uitkomsten en verbeteringen voor Process Mining binnen de organisatie en voor de (eind)gebruikers.

### Matrix

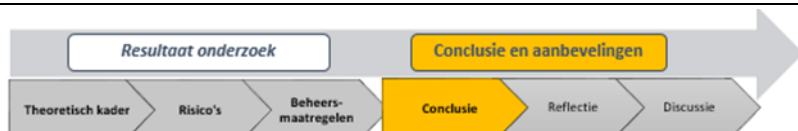
Om antwoord te geven op de onderzoeksvraag, is gebruikgemaakt van een matrix in tabel 2. Hierin wordt visueel weergegeven welke risico's van Process Mining worden afgedekt indien PM2 wordt uitgevoerd. Uit onderzoek is gebleken dat niet alle risico's worden afgedekt door deze methode. Hiervoor zijn aanvullende beheersingsactiviteiten geformuleerd. Deze activiteiten zijn benoemd middels de beschouwendende deelvraag.

Tabel 2 is opgenomen in de volgende pagina. Om de leesbaarheid van tabel 2 te rechtvaardigen, is de pagina liggend gepresenteerd.



Conclusie			
Risico's		PM2 methode	Aanvulling
T e c h n i s c h e	<b>Data vastlegging</b> 1. Het niet op tijd inventariseren van de beschikbare data 2. Het ontbreken van (een) tijdige, relevante en betrouwbare data(bron) 3. Brondata samenvoegen welke geen gezamenlijke format en/of unieke identificeerbron heeft 4. Tijdsperiode, waarbij twee manieren van vastlegging zijn opgenomen	<b>Beheersingsfase</b> Planning en extractie Planning en extractie Planning en extractie Planning en extractie	<b>Beheersingsmaatregelen</b>  Uniek aan PM2 project methode is dat de tussenstap extractie van data, om op een hoger niveau data te ontsloten voor de data preparatie waardoor er meerdere projecten uit kunnen voortvloeien.  Keuze tooling Project (DA - PM).  Eerlijkheid en transparantie over de techniek en beperkingen van commerciële Process Mining tools.  Keuze team (aanvulling extern).
	<b>Extraheren van data</b> 1. Geen concrete analyse vraag, waardoor er onvolledig en/of onjuiste data wordt ontsloten 2. Granulariteit sluit niet aan bij eindgebruiker(s) dan wel de analyse vraag 3. Verkeerde data interpretatie van het extraheren van de brondata (teamsamenstelling) 4. Data is niet opgeschoond (samenvoegen en of verwijderen van activiteiten)	<b>Beheersingsfase</b> Planning en extractie Planning, extractie en data preparatie Planning, extractie en data preparatie Extractie en data preparatie	
	<b>Process Mining techniek en tool</b> 1. Keuze Process Mining tool en techniek (teamsamenstelling)  2. Gebrek aan standaardisatie commerciële Process Mining tool (normen en algoritme)  3. Data filtertechnieken onjuist gebruiken dan wel interpreteren	<b>Beheersingsfase</b>	
B u s i n e s	<b>Process Mining team</b> 1. Er is onvoldoende kennis aanwezig om project Process Mining uit te voeren 2. Verkeerde bedrijfsproces interpretatie (teamsamenstelling)	<b>Beheersingsfase</b> Planning Planning en extractie	Keuze team (aanvulling extern). Keuze team (aanvulling extern). Keuze tooling Project (DA - PM). Aanvulling beheersingsfase.
	<b>Process Mining gebruik</b> 1. Geen concrete analyse vraag, waardoor er onvolledig en/of onjuiste data wordt ontsloten 2. Ideaal beeld van de beschikbare data 3. Ideaal beeld van de toepassingsmogelijkheden van Process Mining 4. Het ontbreken van documentatie over de gemaakte keuzes	<b>Beheersingsfase</b> Planning en extractie	
	<b>Process Mining model</b> 1. Process Mining model wordt op een verkeerde manier geïnterpreteerd 2. Gradatie niveau sluit niet aan bij analyse vraag/ eindgebruiker 3. Er worden te (snel) conclusies getrokken over de Process Mining analyses/uitkomsten	<b>Beheersingsfase</b> Evaluatie Evaluatie Evaluatie	

Tabel 2- Conclusie Process-Mining-Project-Methodology als beheersingsmaatregel.



## Reflectie

In het uitgevoerde onderzoek werden de gevonden technieken en businessrisico's van Process Mining uit de literatuurstudies als bewezen beschouwd. Hier is geen nader onderzoek naar verricht. Vervolgonderzoek dient dan ook nog gedaan te worden. Daarnaast zijn enkele onderwerpen buiten beschouwing gelaten tijdens het onderzoek. Dit heeft de uitkomst mogelijk beïnvloed.

In het onderzoek zijn verschillende bronnen gebruikt uit verschillende geografische gebieden. Process Mining is ontstaan uit onderzoek van de TU Eindhoven (TU/e). Professor Van der Aalst is de bedenker van de techniek en methode. Daarnaast heeft hij de afgelopen jaren veel onderzoek verricht en meerdere boeken geschreven. Veel van zijn onderzoeken zijn geciteerd in andere onderzoeken (Van der Aalst, 2019). Ondanks de diversiteit van bronnen in verschillende tijdperiodes, zijn deze indirect gelinkt met de professor, dit kan de uitkomst mogelijk beïnvloed hebben.

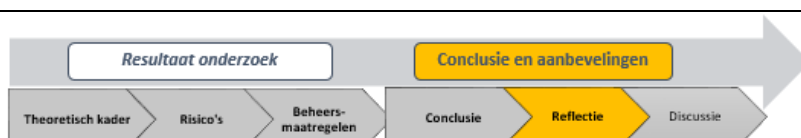
De gevonden technische - en businessrisico's kwamen in meerdere onderzoeken terug, met name in meerdere businesscases. Er is geen nader onderzoek geweest naar de interactie van meerdere risico's wat mogelijk kan leiden tot een nieuwe risico. Daarnaast is er bewust voor gekozen om geen zijstap te maken naar data-analyse, andere Process Mining methodologie en of businessprocesmanagement. Omdat dit niet paste binnen het huidige en gewenste tijdbestek.

Tijdens het onderzoek zijn de onderzoeksvragen eenmaal aangescherpt. Op basis van het toen uitgevoerde literatuuronderzoek is geen relatie gevonden tussen de risico's van Process Mining en de toepassingsmogelijkheden van Process Mining. Derhalve is gezocht naar de relatie met de risico's van Process Mining. Hieruit bleek dat er wel een relatie is met de Process-Mining-Project-Methodology. Daarom is deze vervolgens uitgewerkt in dit onderzoek.

### Suggesties voor vervolgonderzoek

Het advies is om nader onderzoek te doen naar de relatie van Process Mining in combinatie met nieuwe technologie zoals Robotic Process Automation (RPA), maar ook de wet- en regelgeving, zoals de Algemene Verordening Gegevensbescherming (AVG). Met de komst van AVG is data-uitwisseling niet alleen lastiger, maar de activiteitdata kunnen ook worden verwijderd (Van Heur, 2019).

Daarnaast is het advies om aanvullende onderzoek te doen naar welke specifieke rol de IT-auditor op zich kan nemen bij een Process Mining Project. Hierbij kan worden onderzocht welke bestaande werkzaamheden gebruikt kunnen worden voor Process Mining en welke aanvullende, nieuwe werkzaamheden dan van belang zijn om een bepaalde mate van zekerheid te geven voor de uitkomsten van Process Mining. Een betrouwbaarheid van 100% is onhaalbaar. Dit wordt veroorzaakt door restrisico. Dit geldt niet alleen voor data-analyse-aanpak, maar ook voor de traditionele aanpak, zoals papier en lijncontrole.



## Discussie en aanbevelingen

Professor Van der Aalst vergelijkt Process Mining met spraakherkenningssoftware. Deze kunnen worden gebruikt. Echter, er zijn verbeteringsmogelijkheden (Van Heur, 2019). De aanbevelingen zijn onderverdeeld in onderwerpen voor nader onderzoek. Daarnaast is er nagedacht over wat kansen zijn voor Process Mining in de toekomst.

### **Discussie 1:** Eerlijkheid en transparantie

De eerste aanbeveling betreft eerlijkheid en transparantie over de techniek achter Process Mining bij commerciële tools. Er zijn momenteel ruim dertig commerciële tools met alle hun eigen normen en naamgeving voor verschillende definities. Dit werkt verwarrend voor de (eind)gebruikers van Process Mining tool (Ailene, 2014). Commerciële maken gebruik van algortime, er zijn verschillende algoritmes in de academische wereld. Afhankelijk van het gekozen algoritme kan er een andere uitkomst zijn. Dit heeft te maken met verschillende algoritmes die beter kunnen omgaan met verschillende omstandigheden (Ailene, 2014).

**Aanbevelingen:** Door een eerlijke en transparante onderbouwing van de techniek en definities van Process Mining kan deze breder (in de Assurance) worden ingezet. Hiervoor dient een gezamenlijk normenkader te worden gecreëerd en hierop moet toezicht worden gehouden.

Het initiatief kan gestart worden door de werkgroep Process Mining binnen de Koninklijke Nederlandse Verenigingen van Informatieprofessionals (KNVI). Om drie redenen namelijk het creëren van een bewustwording, draagvlak en het vormgeven van het normenkader en hierop vervolgens toezien.

Commerciële tools moeten hun tool laten controleren op de gehanteerde techniek. Zoals is beschreven in het onderzoek Workflow mining: A survey of issues and approaches (Van der Aalst et al., 2003). Hierin werden meerde opensourcetools beoordeeld op negen punten: structure, time, basic parallelism, non free choice, basic loops, arbitrary loops, hidden en duplicate tasks en noise. In het onderzoek van (Ailene, 2014) zijn vier commerciële tools vergeleken om de sterke en zwakke punten van de tools nader te beoordelen op basis van een gehanteerd normenkader.

### **Discussie 2:** Standaard export

Het ontsluiten en creëren van data kost in de praktijk veel tijd (80% van het project). De analyse zelf vraagt het minste tijd. Bij een standaard export zal direct tijdwinst wordt behaald, beperking is daarop dat er wel data ontsloten wordt op een bepaald detail niveau, hiermee kunnen alleen bepaalde vragen worden beantwoord. Een standaard export kan ook gezien worden als basis ingrediënt voor een cocktail om zelf hier aanvullende data en/of activiteiten aan toe te voegen.

### **Aanbevelingen**

Datawarehouses ontwikkelen een standaard export voor een activiteitenlog met bepaalde standaard geïmplementeerde bedrijfsprocessen. Hierdoor wordt gedeeltelijk de drempel weggehaald bij het verkrijgen van de juiste en volledige dataset (activiteitenlog).

In de praktijk is dit lastig, omdat er ook commerciële belangen zijn en standaard scripts niet (gratis) worden gedeeld.

### **Discussie 3:** Meerdere ketens

De laatste discussie betreft de toepassing van Process Mining over meerdere ketens. Vaak zijn bedrijven afhankelijk van verschillende afdelingen of externe bedrijven. Om in een gehele keten het proces te optimaliseren, kan het gezamenlijk streven naar de optimale situatie voor alle partijen waar de klant, medewerkers en klimaat centraal staan.

### **Aanbevelingen**

De uitdaging zit hem in het delen van data Algemene Verordening Gegevensbescherming (AVG) maar ook in het vinden van een (unieke) identificeerbron om data aan elkaar te koppelen. Hier dient nog aanvullend onderzoek naar gedaan te worden.



## Literatuurlijst

- Agconnect. (2018, 4 januari). Waarde halen uit de groeiende berg aan eventdata. Geraadpleegd van <https://www.agconnect.nl/artikel/waarde-halen-uit-de-groeiende-berg-aan-eventdata>
- Ailenei, I. (2014). *Process Mining Tools: A Comparative* (Master Thesis). Geraadpleegd van <https://research.tue.nl/files/47023004/717803-1.pdf>
- Andrews, R., Wynn, M. T., Vallmuur, K., Ter Hofstede, A. H., Bosley, E., Elcock, M., & Rashford, S. (2019). Leveraging data quality to better prepare for Process Mining: an approach illustrated through analysing road trauma pre-hospital retrieval and transport processes in Queensland. *International journal of environmental research and public health*, 16(7), 1-25 .
- Daniel, F., Barkaoui, K., & Dustdar, S. (Eds.). (2011). *Business Process Management Workshops*. Berlijn, Duitsland: Springer.
- Frijns, P. H. A. M. (2019). *Onderzoek vanuit verschillende perspectieven*.
- Haasnoot, M. (2012). Lessons learned bij toepassing van Process Mining. Geraadpleegd van <https://www.deitauditor.nl/wp-content/uploads/2014/07/ITA-12-03-Lessonslearned.pdf>
- Jansen, M. (2019). *Curriculum Vitae Mieke Jans*. Geraadpleegd van [https://www.uhasselt.be/cv\\_pers/1555.pdf](https://www.uhasselt.be/cv_pers/1555.pdf)
- KNVI. (2018, 18 september). Seminar: Process Mining | hoe implementeer je dat nou? Geraadpleegd van <https://www.knvi.nl/nieuws/1909/Seminar-Process-mining-%7C-hoe-implementeer-je-dat-nou-%3FbackUrl=%2Finteressegroep%2F52b336ec6cb6fde6570001ce%2FProcess-Minin>
- Rozinat, A. (2015). Anne Rozinat en Frank van Geffen. Geraadpleegd van <https://www.agconnect.nl/artikel/succescriteria-voor-process-mining>
- Rozinat, A. (2018). Anne Rozinat. Geraadpleegd van <https://research.tue.nl/en/persons/anne-rozinat>
- Silva, L. F. (2014). *Process Mining: Application to a case study*. Geraadpleegd van <https://repositorio-aberto.up.pt/bitstream/10216/78160/2/34075.pdf>
- Suriadi, S. W. (2013). *Understanding Process Behaviours in a Large Insurance Company in Australia*. Berlijn, Duitsland: Springer.
- Van der Aalst, W. (2014). *Process Mining: A historical perspective*. Geraadpleegd van [http://www.processmining.org/\\_media/presentations/2013/8-process-mining-camp2013-history-of-process-mining.pdf](http://www.processmining.org/_media/presentations/2013/8-process-mining-camp2013-history-of-process-mining.pdf)

Van der Aalst, W. (2015). *PM2: A Process Mining Project Methodology*. Berlijn, Nederland: Springer.

Van der Aalst, W. (2019). Wil van der Aalst. Geraadpleegd van <http://www.padsweb.rwth-aachen.de/wvdaalst/>

Van der Aalst, W. M. (2012). What makes a good process model? *Software & Systems Modeling*, 11(4), 557-56.

Van der Aalst, W. M., Van Dongen, B. F., Herbst, J., Maruster, L., Schimm, G., & Weijters, A. J. (2003). Workflow mining: A survey of issues and approaches. *Data & knowledge engineering*, 47(2), 237-267.

Van Heur, R. (2019, april 30). Steeds meer bedrijven doen aan Process Mining. Geraadpleegd van <https://www.agconnect.nl/artikel/steeds-meer-bedrijven-doen-aan-process-mining>

Werner, M. (2014). *Business Process Analysis automation for financial audits*. Geraadpleegd van <http://ediss.sub.uni-hamburg.de/volltexte/2016/7663/pdf/Dissertation.pdf>



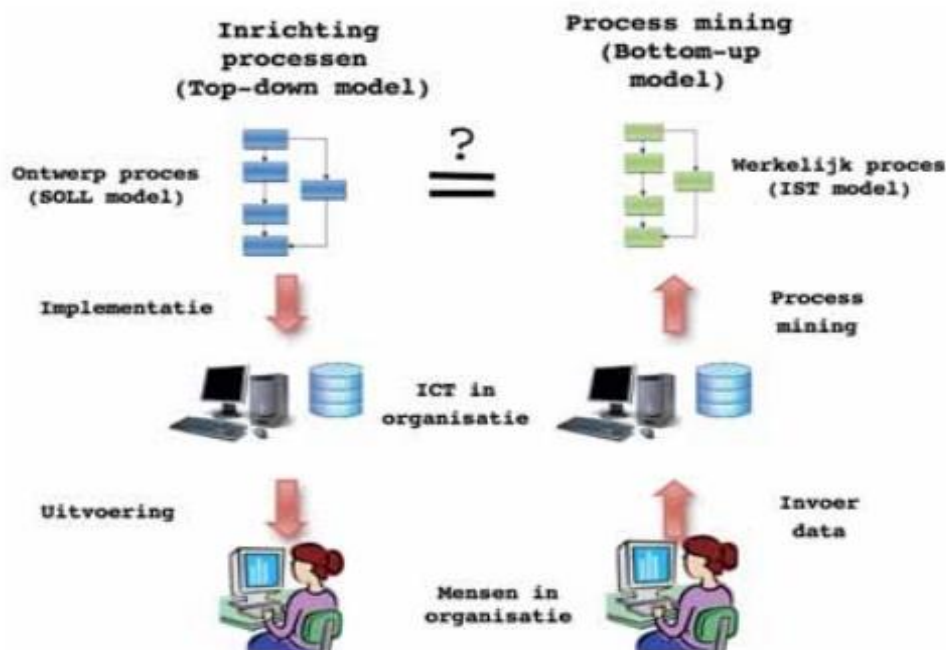
## Bijlage 1 – Figuren en tabellen

Process Mining: volwassenheid activiteitenlog:

Niveau	Karakterisering	Voorbeeld
★★★★★	Hoogste niveau: de event log is van excellente kwaliteit (d.w.z. betrouwbaar en compleet) en events zijn duidelijk gedefinieerd. Events worden op een automatische, systematische, betrouwbare en veilige manier geregistreerd. Privacy- en beveiligingsaspecten zijn adequaat in beschouwing genomen. Bovendien hebben de opgeslagen events (en alle bijbehorende attributen) een heldere semantiek. Dit impliceert dat er een of meer ontologieën bestaan waarnaar events en hun attributen naartoe verwijzen.	Semantisch geannoteerde logs van BPM systemen.
★★★★	Events worden automatisch en in een systematische en betrouwbare manier geregistreerd, d.w.z. logs zijn betrouwbaar en compleet. In tegenstelling tot de systemen die tot het ★★★ niveau behoren zijn noties zoals process instance (case) en activiteit expliciet ondersteund.	De events logs van traditionele BPM/workflow systemen.
★★★	Events worden automatisch geregistreerd, maar er wordt geen systematische aanpak gevolgd om deze event logs op te slaan. Echter, in tegenstelling tot de logs die zich op niveau ** bevinden is er een bepaalde garantie dat de geregistreerde events overeenkomen met de realiteit (d.w.z. de event log is betrouwbaar maar niet noodzakelijk compleet). Een voorbeeld hiervan zijn de geregistreerde events van een ERP systeem. Desondanks dat events vanuit verschillende tabellen genomen dienen te worden, kan aangenomen worden dat de informatie correct is (het is betrouwbaar om aan te nemen dat een opgeslagen betaling in het ERP systeem ook daadwerkelijk bestaat en vice versa).	Tabellen in ERP systemen, event logs van CRM systemen, transactie logs van messaging systemen, event logs van high-tech systemen, enz.
★★	Event logs worden automatisch opgeslagen, d.w.z. als bijproduct van een informatiesysteem. Het bereik kan variëren, d.w.z. er is geen systematische aanpak gevolgd om te besluiten welke events geregistreerd worden. Bovendien is het mogelijk om het informatiesysteem te omzeilen. Zodoende kunnen events ontbreken of deze zijn niet op een correcte manier geregistreerd.	Event logs van document management systemen en product management systemen, registraties van foutmeldingen van embedded systemen, worksheets van servicemonteurs, enz.
★	Laagste niveau: event logs zijn van slechte kwaliteit. Het kan voorkomen dat geregistreerde event logs niet overeenkomen met de werkelijkheid en events kunnen ontbreken. Event logs die handmatig geregistreerd zijn, hebben normaliter deze karakteristieken.	Sporen die nagelaten zijn in papieren documenten die door een organisatie heengaan ("yellow notes"), papieren medische dossiers, enz.

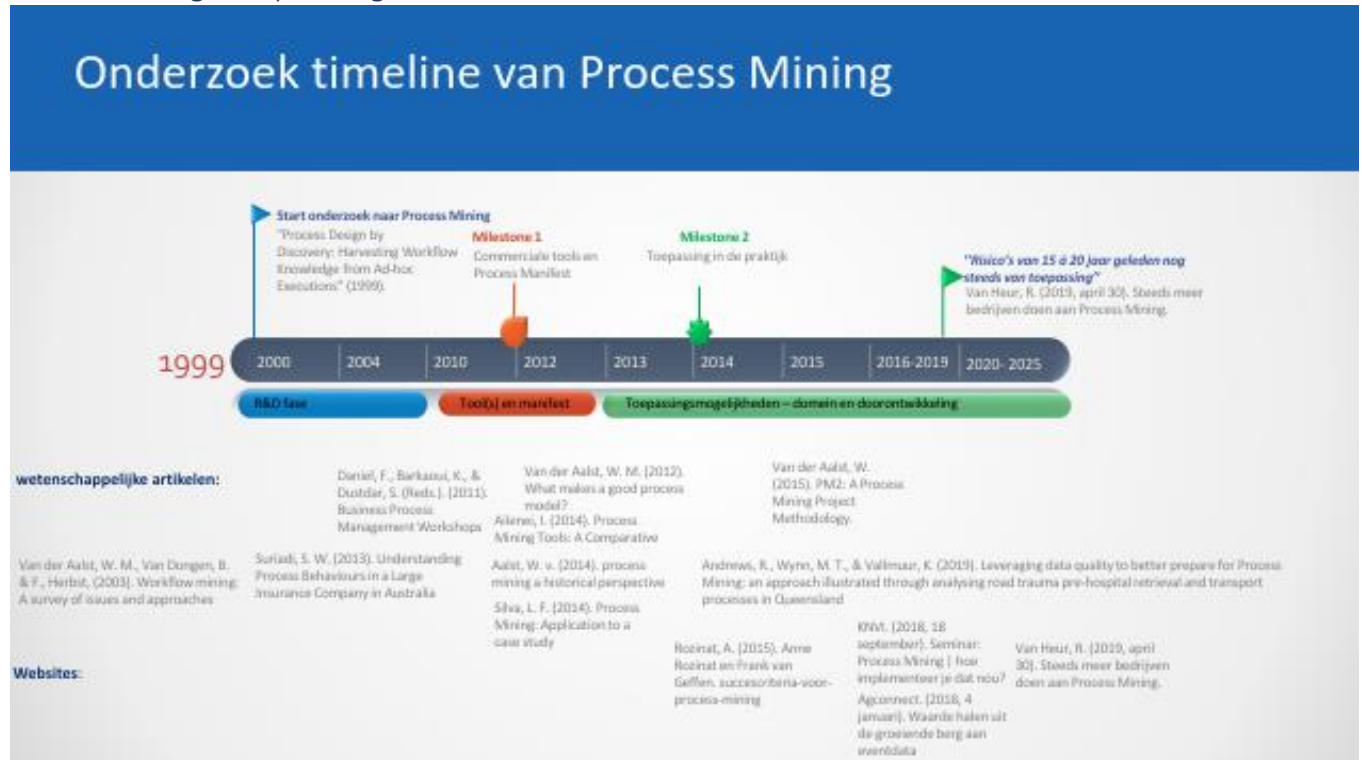
Tabel 3- Volwassenheid activiteitenlog (Daniel et al., 2011).

Process Mining: Inrichting processen van bedrijfsproces en Process Mining aanpak:



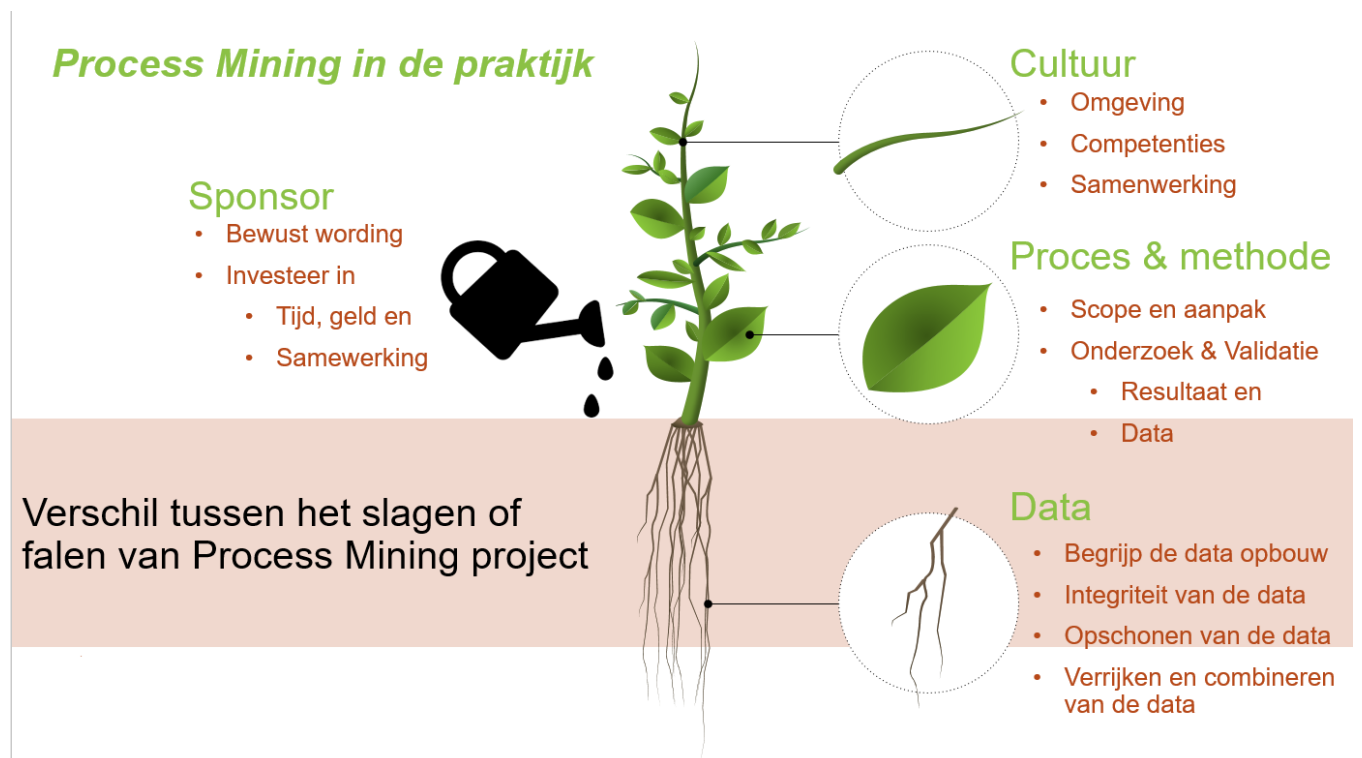
Figuur 8 - Inrichting proces van bedrijfsproces en Process Mining aanpak (Haasnoot, 2012).

Process Mining life cycle en gehanteerd bronnenonderzoek:



Figuur 9- Onderzoek timeline Process Mining

Process Mining in de praktijk:



Figuur 10 - Process Mining in de praktijk.

## Bijlage 2 - Reflectie

Voor de pre-masteropleiding IT-audit, Compliance & Advisory aan de Vrije Universiteit Amsterdam (VU) is een reflectie opgenomen. Hierin wordt de ervaring beschreven: wat goed is gegaan, maar ook de valkuilen en uitdagingen die de onderzoeker tegenkwam tijdens het schrijven van dit onderzoek.

De reflectie is ingedeeld in vier vragen. Deze zijn vervolgens beantwoord door de huidige situatie te benoemen en vervolgens hoe hiermee is omgegaan. De VU-logfile die periodiek verstuurd diende te worden naar de vakdocenten heeft geholpen om deze vragen achteraf zo goed mogelijk te beantwoorden.

### **Wat zou ik in een soortgelijke situatie weer doen of blijven doen?**

- In de module Professional Judgement moesten we elke les een PowerPointpresentatie geven. Hierdoor hebben we veel gesproken en gediscussieerd over het onderwerp.
- Het is van belang om nogmaals in het onderwerp te verdiepen. Tevens had ik tijdens mijn studie een SRA-training – Process Mining in de controlepraktijk. Door deze training heb ik de methode, techniek en tool weer scherp op mijn netvlies, wat hielp met de beeldvorming.
- In eerste instantie wilde ik niet kiezen voor Process Mining, omdat ik al enige voorkennis had en graag nieuwe dingen wilde leren. Ik heb veel nieuwe inzichten gekregen in Process Mining en daarnaast heel veel geleerd.
- Ondanks de overlap was het goed om meer literatuurstudies te lezen, omdat elke onderzoeker de uitdagingen, techniek, methode of tool net op een andere manier benadert. Hierdoor werd ik geprikkeld om op een andere manier te kijken naar hetzelfde risico.
- Aanscherpen van het onderzoek. Dit heeft bijgedragen aan de kwaliteit van het onderzoek.

### **Wat zou ik in een soortgelijke situatie niet meer doen, maar wel beter willen doen?**

- Tijdens (oriënterend) literatuuronderzoek heb ik artikelen gezocht zonder deze concreet te labelen. Daardoor werd een artikel soms pas na de tweede keer scannen juist was gelabeld. Structuur is erg van belang.
- Bij een studiedag(deel) is het van belang duidelijk aan te geven wat het plan is en hier dan ook naar te handelen. Dit was voor mij een uitdaging.

### **Wat zou ik in een soortgelijke situatie niet meer doen?**

- Het zoeken om het zoeken naar academische literatuur. Ik verdwaalde een beetje in alle onderzoeken naar Process Mining. Het is van belang om vast te houden aan een gemaakt plan.
- Wachten met schrijven en indien het schrijfproces toch is gestart, eerst de informatie opschrijven die al gevonden is. Dit proces van zoeken en uitschrijven liep in eerste instantie door elkaar. Dit heb ik opgelost door eerst de analyserende vraag op hoofdlijnen te beantwoorden hierdoor kreeg ik weer overzicht.
- Werken in een vast kader; hierdoor lette ik te veel op lay-out in plaats van op de inhoud.

### **Wat zou ik in een soortgelijke situatie anders of aanvullend doen?**

- Ik heb in een vroeg stadium al ideeën op papier gezet. Echter, ik had hierbij de bronnen nog niet bijgezet. Achteraf was het zoeken van meerdere bronnen een tijdrovende besteding. In het vervolg zal ik de kapstokmethode toepassen met hierbij de bronvermelding.
- Ik heb enkele studiedagen ingepland. Na de enkele weken na de verdedig van P1. Het plannen had ik eerder kunnen doen, waardoor minder druk zou liggen in de laatste maand. Daarnaast moest ik zoeken naar de juiste tijden en momenten om te studeren.