

Addendum Applied Data Science

In mei 2019 is een werkgroep samengesteld met betrokkenen van verschillende hogescholen die een brede kennis en ervaring hebben met data-gerelateerde onderwerpen in het HBO-I onderwijs. Doel van de werkgroep was het toegenomen belang van data in het ICT-werkveld beter onder te brengen in de Domeinbeschrijving Bachelor of ICT.

Uit de analyse van het Edison Framework als ook uit de ervaringen van de deelnemers aan de werkgroep en hun achterban, bleek dat binnen het terrein van Applied Data Science een veelheid van steeds veranderende beroepsrollen ontstaan is. Dit maakt het verzamelen en generiek beschrijven van de beroepstaken ingewikkeld. Daarom is gekozen Applied Data Science te beschouwen als een continuüm van beroepen en beroepsrollen met aan de ene kant de **Data Engineer** die ruwe data verzamelt en bewerkt zodat het geschikt is voor analyse, dan de **Data Analyst** die met geschoonde data analyses en voorspellingen doet en als laatste de **Machine Learning Engineer** die op basis van data modellen traint.

Voor elk van deze **beroepsrollen** is onderzocht op welke wijze de matrices gewijzigd moeten worden om deze adequaat te beschrijven. Hierbij is in principe alleen het hoogste bachelor niveau, niveau 3, beschreven. Niveau 1 en 2 volgen hier logisch uit en zijn pas bij een volledige opname in de domeinbeschrijving van belang. Zo ook niveau 4.

Data Analyst

De data analyst is een ICT-er die data producten gebruikt en realiseert om inzichten te verkrijgen die waarde opleveren voor een organisatie; de data analyst doet dus de taken beschreven bij "Organisatieprocessen" én de taken beschreven bij "Software".
 Andere benamingen voor dit profiel zijn "(Applied) Data Scientist", "Business Analyst", "Data Analyst", "BI Consultant", "Informatie Analyst".

		Architectuurlagen				
		Gebruikersinteractie	Organisatieprocessen	Infrastructuur	Software	Hardware interfacing
Activiteiten	Analyseren	Benchmarken van functionaliteit, user interaction en UX design ten behoeve van een opdracht die uitgevoerd wordt binnen een bedrijf. Analyseren van kernwaarden opdrachtgever, product of dienst, user needs en hoe die tot uiting komen in product of dienst. Evalueren van de projectvoortgang vanuit het perspectief van de gebruiker. [Niveau 2]	Analyseren van de huidige acceptatie van nieuwe technologie en eventuele weerstanden. Analyseren van gestructureerde en ongestructureerde interne en externe data m.b.v. de data karakteristieken: Volume, Velocity, Variety, Variability, Varicity, Validity, Volatility en Visualisation (8 V's) .	<niet van toepassing>	Uitvoeren van een requirementsanalyse voor een data-oplossing met verschillende belanghebbenden in een context van bestaande systemen. Definiëren van acceptatiecriteria aan de hand van kwaliteitseigenschappen en een uitgevoerde risicoanalyse met onder andere aandacht voor security aspecten.	<niet van toepassing>
	Adviseren	Geven van een goed gemotiveerd concreet advies over de te gebruiken interactieve technieken en/of interactieconcepten. Voorstellen doen over de realisatiekeuzes, zoals de te gebruiken technologieën, daarbij rekening houdend met gebruikers- en bedrijfscontext. Adviseren over de doelstellingen van de huidige en volgende iteraties. [Niveau 2]	Adviseren over een veranderkundige aanpak bij de invoering van nieuwe ICT-mogelijkheden o.b.v. informatie direct uit de data . Adviseren in oplossingen voor gestructureerde en ongestructureerde data rekening houdend met ethische en juridische aspecten .	<niet van toepassing>	Adviseren met betrekking tot de keuze voor softwarearchitectuur of bestaande software frameworks, waarbij kostenaspecten en kwaliteitskenmerken zoals beschikbaarheid, performance, security en schaalbaarheid een rol spelen. Adviseren over de aanpak bij het verwerken en raadplegen van grote hoeveelheden data met aandacht voor privacy. "Adviseren over de inrichting van een softwareontwikkelproces, waaronder het testproces." Adviseren over het gebruik van prototypes bij het valideren of exploreren van eisen.	<niet van toepassing>
	Ontwerpen	Vertalen van adviezen in een ontwerp van gedetailleerde gebruikersinteractie met verschillende prototyping- technieken. Ontwerpen van een testdesign waarmee de doelstellingen van de iteratie geëvalueerd worden. [Niveau 2]	Ontwerpen van de architectuur van organisatieprocessen, besturingsmodellen en veranderproces , inclusief bijbehorende beheersing met behulp van data oplossing(en) .	<niet van toepassing>	Opstellen van een data-architectuur voor een data-oplossing , opgebouwd uit bestaande en nieuwe data-bronnen , rekening houdend met meerdere belanghebbenden en kwaliteitskenmerken, waaronder security en schaalbaarheid. Opstellen van teststrategie voor data .	<niet van toepassing>
	Realiseren	Realiseren en kwalitatief testen van eenvoudige interactieve producten of diensten op basis van een ontwerp waarbij gebruik wordt gemaakt van gangbare tools, design guidelines en/of huisstijl . [Niveau 2]	Realiseren van invoering en acceptatie van gewijzigde organisatieprocessen op basis van een implementatieplan. Inrichten van oplossingen voor gestructureerde en ongestructureerde data.	<niet van toepassing>	Opstellen van een data-architectuur voor een data-oplossing , opgebouwd uit bestaande en nieuwe data-bronnen , rekening houdend met meerdere belanghebbenden en kwaliteitskenmerken, waaronder security en schaalbaarheid. Opstellen van teststrategie voor data.	<niet van toepassing>
	Managen & controleren	Gebruiken en correct toepassen van standaarden (design guidelines m.b.t.interactie ontwerp, protocollen en methoden) passend binnen bedrijfscontext. [Niveau 2]	Richten en actualiseren van principes, business rules en modellen van procesarchitectuur. Proactief signaleren van behoefte aan verandering in alle organisatie processen en bijbehorende veranderprocessen in gang zetten.	<niet van toepassing>	Beheren en gebruiken van een ontwikkelstraat ter ondersteuning van softwareontwikkeling in teams, waardoor onder andere continuous integration tot de mogelijkheden behoort.Toepassen van methoden en technieken om een software-ontwikkelproces te managen en de kwaliteit ervan te borgen. Neemt ook data mee in de beheerprocessen en de ontwikkelomgeving. [Niveau 2]	<niet van toepassing>

Data engineer

Data engineer ontwikkelt, construeert, test en onderhoud data architecturen, waaronder master data management, data warehouses, -lakes en -swampsdatabases en data processing systemen. Zijn taken zijn gericht op het verzamelen, opslaan, optimaliseren, verspreiden, analyseren en monitoren van grote hoeveelheden data binnen de organisatie met aandacht voor kwaliteit, privacy, betrouwbaarheid en security. Door deze functie is de data engineer betrokken bij gebruikersinteractie voor vraag en levering data, organisatieprocessen gericht op data verspreiding, infrastructuur voor data registratie en distributie en software voor data modelleren en opvragingen. Andere benamingen voor dit profiel zijn "Information engineer" en "Data manager".

		Architectuurlagen				
		Gebuiersinteractie	Organisatieprocessen	Infrastructuur	Software	Hardware interfacing
Activeren	Analyseren	<p>Analyseren van de gebruiker, gebruikersinteractie en -ervaring, zowel individueel (fysiek, psychologisch, persoonlijke karakteristieken) als in grotere maatschappelijke context (sociaal / cultureel / ethisch / technologisch).</p> <p>Analyseren van actuele en state of the art interactieve technologieën.</p> <p>Analyseren per iteratie van het effect van de eigen interventie op/in de UX. Analyseren van de requirements van de gebruikers voor ETL-processen.</p>	<p>Analyseren van de consequenties van een (strategische) koerswijziging voor organisatie processen en hun informatievoorziening.</p> <p>Analyseren (kwantitatief en/of kwalitatief) van de huidige en toekomstige situatie op het gebied van bijvoorbeeld beleid, strategie, alignment en architectuur, met gebruikmaking van gangbare methoden.</p> <p>Analyseren van de huidige acceptatie van nieuwe technologie en eventuele weerstanden. Analyseren van gestructureerde en ongestructureerde interne, externe data en workflowdata.</p>	<p>Analyseren van de trends op het gebied van infrastructuur en deze vertalen naar gewenste of benodigde ontwikkelingen in bedrijfsinfrastructuur.</p> <p>Uitvoeren van een requirements-impact- of gap-analyse voor een enterprise infrastructuur om (niet-) functionele eisen, behoeftes en/of tekortkomingen, in kaart te brengen. Analyseren van potentiële incidenten, problemen en bedreigingen, waaronder security en data-integriteit volgens data governance.</p>	<p>Uitvoeren van een requirementsanalyse voor een softwaresysteem met verschillende belanghebbenden in een context van bestaande systemen.</p> <p>Definiëren van acceptatiecriteria aan de hand van kwaliteits eigenschappen en een uitgevoerde risicoanalyse met onder andere aandacht voor security aspecten.</p> <p>Analyseren over welke software geschikt is voor ETL-processen aan de hand van de beschikbare data en de beschikbare software.</p> <p>Pipelines en query's samenstellen ten behoeve van toekomstige data-analyses, aan de hand van de beschikbare data.</p>	<niet van toepassing>
	Adviseren	<p>Vertalen van analyse naar strategische aanbevelingen (op korte, middellange en langere termijn) voor het ontwerpen of verbeteren of onderzoeken van een UX met inzet van interactieve middelen. Hierbij wordt ook een onderbouwd advies gedaan voor het meest geschikte ontwerpproces (bijvoorbeeld UCD).</p> <p>Adviseren over de UX-interventie(s) in de huidige en volgende iteraties.</p> <p>Adviseren van gebruikers over potentiële data oplossingen voor ETL-processen.</p>	<p>Adviseren over de interne en externe afstemming tussen business en ICT (alignment en governance) rekening houdend met de doelstellingen van de organisatie (o.a. missie, visies, strategie en KPI's).</p> <p>Adviseren over een veranderkundige aanpak bij de invoering van nieuwe ICT-mogelijkheden.</p> <p>Adviseren in oplossingen voor gestructureerde en ongestructureerde data en voor aanpassingen aan processen ten behoeve van het verkrijgen en/of verbeteren van workflowdata.</p>	<p>Adviseren omtrent het gebruik van de infrastructuur voor data, bijvoorbeeld dataopslag- en verwerkingsomgeving.</p>	<p>Adviseren met betrekking tot de keuze voor softwarearchitectuur of bestaande software frameworks, waarbij kostenaspecten en kwaliteitskenmerken zoals beschikbaarheid, performance, security en schaalbaarheid een rol spelen.</p> <p>Adviseren over de aanpak bij het verwerken en raadplegen van grote hoeveelheden data met aandacht voor privacy. Adviseren over de inrichting van een softwareontwikkelproces, waaronder het testproces.</p> <p>Adviseren over de te gebruiken software en/of de op te stellen pipeline(s) voor ETL-processen, aan de hand van de beschikbare data.</p>	<niet van toepassing>
	Ontwerpen	<niet van toepassing>	<p>Ontwerpen van de architectuur van organisatieprocessen en besturingsmodellen, inclusief bijbehorende beheersing, informatievoorziening en veranderproces.</p> <p>Ontwerpen van een veranderkundige aanpak met bijbehorende interventies. Ontwerpen van oplossingen voor gestructureerde, ongestructureerde data en voor procesveranderingen ten behoeve van het verkrijgen en/of verbeteren van workflowdata.</p>	<p>Ontwerpen van data opslag- en verwerkingsomgeving, waarmee big data opgeslagen, bewerkt en geanalyseerd kan worden.</p>	<p>Opstellen van een softwarearchitectuur voor een softwaresysteem, opgebouwd uit bestaande en nieuwe systemen, rekening houdend met meerdere belanghebbenden en kwaliteitskenmerken, waaronder security en schaalbaarheid.</p> <p>Opstellen van teststrategie voor systeemtesten.</p> <p>Ontwerpen van software en/of scripts aan de hand van de beschikbare data en pipelines voor ETL-processen.</p>	<niet van toepassing>
	Realiseren	<niet van toepassing>	<niet van toepassing>	<p>Implementeren van een oplossing aan de hand van een opgezet ontwerp van een data opslag- en verwerkingsomgeving. Deze kan zich on-site en/of in de cloud bevinden.</p>	<p>Bouwen en beschikbaar stellen van een schaalbaar softwaresysteem dat aansluit bij bestaande systemen, eventueel in de cloud, volgens de ontworpen architectuur met gebruik van bestaande frameworks.</p> <p>Toepassen van testautomatisering bij het uitvoeren van testen.</p> <p>Bouwen van een pipeline of applicatie met beschikbare data, workflows en pipelines voor ETL-processen.</p>	<niet van toepassing>
	Manage & control	<niet van toepassing>	<niet van toepassing>	<p>Onderhouden, documenteren, beheersen en updaten van de infrastructuur voor data opslag- en verwerkingsomgeving. Deze kan zich on-site en/of in de cloud bevinden.</p>	<p>Uitvoeren van configuratie-, change- en releasemanagement in afstemming met infrastructuur-management.</p> <p>Inrichten van een ontwikkelstraat met geautomatiseerde build en test infrastructuur.</p> <p>Onderhouden, documenteren, beheren en updaten van de software en pipelines voor de ETL-processen. Deze kan zich on-site en/of in de cloud bevinden.</p>	<niet van toepassing>

Machine Learning (ML) Engineer

De machine learning engineer is een (embedded) software engineer die ML componenten in de software of hardware verwerkt; de ML engineer doet dus of de taken beschreven bij "Software" of de taken beschreven bij "Hardware Interfacing"; Soms ook een combinatie hiervan. De ML engineer stelt requirements op voor de benodigde data, bewerkt de data om ML te kunnen toepassen, traint ML-modellen en verwerkt de getrainde modellen in (embedded) softwaresystemen. In samenspraak met een Data-analist kan de ML engineer beslissen welke data verder geanalyseerd moet worden m.b.v. ML. Andere benamingen voor dit profiel zijn "Applied Data Scientist", "Data Science Engineer", "Artificial Intelligence (AI) Engineer".

Activiteiten	Architectuurlagen				
	Gebruikersinteractie	Organisatieprocessen	Infrastructuur	Software	Hardware interfacing
Analyse ren	<niet van toepassing>	<niet van toepassing>	<niet van toepassing>	<p>Uitvoeren van een requirementsanalyse voor een softwaresysteem met ML-componenten erin met verschillende belanghebbenden in een context van bestaande systemen.</p> <p>Analyseren van de data benodigd om het ML-systeem te trainen.</p> <p>Definiëren van acceptatiecriteria aan de hand van kwaliteitseigenschappen en een uitgevoerde risicoanalyse met onder andere aandacht voor security aspecten.</p>	<p>Specificeren van een gedistribueerd computersysteem met ML-componenten inclusief timing, resourcegebruik en performance. Beschrijven van security- aspecten van computersystemen die aan of via (openbare) netwerken gekoppeld zijn.</p> <p>Opstellen van een acceptatietestplan en een integratietestplan.</p>
Adviser en	<niet van toepassing>	<niet van toepassing>	<niet van toepassing>	<p>Adviseren met betrekking tot de keuze voor softwarearchitectuur of bestaande software frameworks, waarbij kostenaspecten en kwaliteitskenmerken zoals beschikbaarheid, performance, security en schaalbaarheid een rol spelen.</p> <p>Adviseren over de data-architectuur en de model-architectuur en bijbehorende frameworks.</p> <p>Adviseren over de aanpak bij het verwerken en raadplegen van grote hoeveelheden data met aandacht voor privacy.</p> <p>Adviseren van de opdrachtgever omtrent de consequenties van het inzetten van een ML-systeem.</p> <p>Adviseren over de inrichting van een software-ontwikkelproces, waaronder het testproces.</p>	<p>Uitbrengen van een technisch advies over een te realiseren (gedistribueerd) computersysteem, inclusief hardware- en software- componenten en koppelingen.</p> <p>Uitbrengen van een technisch advies over de toepassing van ML-componenten in een (gedistribueerd) computersysteem [dit is niveau 4 van Hardware Interfacing]</p>
Ontwerp en	<niet van toepassing>	<niet van toepassing>	<niet van toepassing>	<p>Opstellen van een softwarearchitectuur voor een softwaresysteem, opgebouwd uit bestaande en nieuwe systemen, rekening houdend met meerdere belanghebbenden en kwaliteitskenmerken, waaronder security en schaalbaarheid. Opstellen van teststrategie voor systeemtesten.</p> <p>Ontwerpen van de data-architectuur en de model-architectuur (i.e. trainen ML-algoritme); inclusief een teststrategie voor de data en de ML modellen.</p>	<p>Ontwerpen van een gedistribueerd computersysteem met gebruik van ML-componenten inclusief bepaling van actuatoren, sensoren, timing, resourcegebruik en performance. [Niveau 4]</p>
Realiser en	<niet van toepassing>	<niet van toepassing>	<niet van toepassing>	<p>Bouwen en beschikbaar stellen van een schaalbaar softwaresysteem met ML-componenten (i.e. deployen getraind ML-model) dat aansluit bij bestaande systemen, eventueel in de cloud, volgens de ontworpen architectuur met gebruik van bestaande frameworks.</p> <p>Toepassen van testautomatisering bij het uitvoeren van testen.</p> <p>Uitvoeren van een data quality check; valideren van ML-modellen [Niveau 4]</p>	<p>Realiseren van een compleet computersysteem met ML-elementen inclusief netwerk, hardware en systeemsoftware. Opstellen en uitvoeren van een acceptatieprocedure bijv. in een virtuele omgeving, inclusief aspecten als timing, resourcegebruik en performance. [Niveau 4]</p>
Manage & control	<niet van toepassing>	<niet van toepassing>	<niet van toepassing>	<p>Uitvoeren van configuratie-, change- en releasemanagement in afstemming met infrastructuur-management.</p> <p>Inrichten van een ontwikkelstraat met geautomatiseerde build en test infrastructuur.</p> <p>Neeemt naast software ook modellen en data mee in de beheerprocessen en de ontwikkelstraat</p>	<p>Opzetten en gebruikmaken van:</p> <ul style="list-style-type: none"> - versiebeheer, - releasemanagement, - teamworkondersteuning, - automated testing voor hard- en softwaresystemen. Neeemt ook modellen en data mee in de beheerprocessen en de ontwikkelomgeving