

Spraaktechnologie voor Verslaggeving Vlaams Parlement (ST-VP)

Eindverslag

Programma	
Aanbestedende diensten	
Externe begeleider	

Inhoud

1	Objectieven en werkingskader	4
1.1	Achtergrond en doelstellingen	4
1.2	Deelnemende partners.....	6
2	Samenvatting.....	7
3	Positioneren en definiëren.....	9
3.1	De projectinitiators.....	9
3.1.1	Vlaams Departement Economie, Wetenschap en Innovatie (EWI).....	9
3.1.2	Vlaams Parlement	9
3.2	Werkingskader van de marktconsultatie	10
3.2.1	De digitale uitdaging van het Vlaams Parlement	10
3.3	Het voortraject: “Spraktechnologie voor Verslaggeving Vlaams Parlement” (STVP)	10
3.4	Beoogde doelstelling van het voortraject	10
3.5	Deliverables van het voortraject	11
4	Inschatting van het innovatiepotentieel vanuit gebruikersstandpunt.....	11
4.1	De verschillende gebruikersgroepen en hun noden	11
4.1.1	Redacteuren van het woordelijk verslag	11
4.1.2	Verslaggevers van de dienst commissies	12
4.1.3	Beeldverslagen	12
4.1.4	Geïmpacteerde gebruikers	12
4.2	Opstellen van de gebruikersvereisten.....	13
4.3	Inschatting van het innovatiepotentieel	16
4.4	Aanvullende algemene beïnvloedingsfactoren.....	23
5	Inschatting van de stand van de techniek: innovatiepotentieel vanuit technologische invalshoek 24	
5.1	De referentieoplossing	24
5.2	Inschatting van de technologische risico’s	25
5.3	Algemene opmerkingen van de markt	33
6	Synthese: opstellen van een RFP.....	34
6.1	Fases en doelstellingen van het commercieel traject	34
6.2	Prioriteiten van de geïntegreerde oplossing	34
6.2.1	Fase 1: Innovatieve Aankoop van een proof-of-concept	35
6.2.2	Fase 2: Klassieke RFP	35
6.2.3	Fase 3: Innovatieve Aankoop van onderzoek.....	36

7	Voorwaarden tot succes.....	36
	Bijlage A: Technische omschrijving digitale opnames Vlaams Parlement	37
	Bijlage B: Redactieapplicatie Woordelijk Verslag.....	40
	Bijlage C: State of the Art analyse	49

1 Objectieven en werkingskader

1.1 Achtergrond en doelstellingen

Dit eindverslag beschrijft het gevolgde traject van de marktconsultatie ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’, presenteert de synthese van de resultaten en schetst de vervolgstappen. De doelstelling van dit eindverslag is het vormen van een basis voor het opmaken van een precommercieel en/of commercieel bestek. Hiertoe worden op de marktconsultatie de verschillende projectstakeholders samengebracht met als doel het verzamelen van kennis, inzichten en concrete voorstellen vanuit verschillende invalshoeken.

Dit eindverslag formuleert een nauwkeurig antwoord op de volgende vragen:

- Wat behelst het project ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’?
- Waar ligt het innovatiepotentieel van dit project, zowel vanuit gebruikerstandpunt als vanuit technologische invalshoek?
- Waar liggen de echte noden voor een prototype en een precommerciële aanbesteding in het kader van Innovatief Aanbesteden?

Addestino Innovation Management werd door de projecteigenaars, met name het Vlaams Departement Economie, Wetenschap en Innovatie (EWI) in samenwerking met het Vlaams Parlement, aangesteld om de marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’, in het kader van het Programma Innovatieve Overheidsopdrachten, te coördineren. Addestino heeft als missie om innovatie te leveren aan zijn klanten, *end-to-end*, vanaf het idee tot het eindresultaat. Deze missie wordt volbracht dankzij drie centrale pijlers:

- Addestino omvat een multidisciplinair team dat in staat is om innoverende doorbraken te realiseren doorheen business, strategie, technologie en gebruikerservaring.
- Addestino beheerst een iteratieve *end-to-end* methodologie waarmee risico gereduceerd wordt en product- en dienstontwikkeling succesvol versneld wordt.
- Addestino bezit een diepgaande technologische kennis en past die pragmatisch toe in verscheidende industriesectoren (telecom, gezondheidszorg, energie, transport, elektronica, enz.) en in verscheidene omgevingen (start-ups, kmo’s, multinationals, universiteiten en overheidsinstellingen).

Tijdens de marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ neemt Addestino de rol op van externe begeleider. Als externe begeleider faciliteert en coördineert Addestino de marktconsultatie, begeleidt en modereert de vergaderingen en stimuleert de nodige wisselwerking tussen de verschillende partijen. Als externe begeleider treedt Addestino steeds op in het algemeen belang, met als taak om de deelnemers aan de marktconsultatie resultaatgericht en op één lijn te krijgen. Daarnaast verschaft Addestino het nodige inzicht en de nodige ervaring in het innovatiegebeuren, mede door het aanwenden van een weldoordacht plan van aanpak voor de marktconsultatie en een methodologie voor de inhoudelijke discussies en denkprocessen tijdens de werksessies.

Dit eindverslag is aldus het resultaat van een gestructureerd proces om het innovatiepotentieel van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ te bepalen met als doel de vraagstelling naar aanbodsijde te verfijnen in het kader van een precommercieel of een commercieel bestek.

De marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ is effectief gestart op 22 september 2017 en eindigt op 27 november 2017. Het beslaat de volgende vijf iteraties doorheen 3 dimensies: gebruikers, technologie en ‘governance’:

1. **‘Scoping’ werksessie met medewerkers van de projectinitiator** waarbij getracht wordt het eindantwoord van de marktconsultatie neer te schrijven op basis van een combinatie van ervaring uit het verleden, wensen en hypothesen.
Deze sessie laat toe een eerste maal de maatschappelijke waarde en het innovatiepotentieel van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ te identificeren.
2. **Werksessie met een aantal gebruikersgroepen** om het standpunt van de (eind)gebruikers te belichten. Deze sessie laat toe het innovatiepotentieel vanuit gebruikersstandpunt verder uit te werken, in het bijzonder de gebruikersnoden en de naar toegevoegde waarde geprioriteerde use-cases.
3. **Marktanalyse uitgevoerd door Addestino** om de use-cases op te delen in modules naar gelijkaardige oplossingselementen en voor elke module een eerste referentiearchitectuur uit te tekenen door de stand van de technologie, belangrijke componenten en noodzakelijke koppelingen na te gaan.
4. **Werksessie met geïnteresseerde industriepartners en kennisinstellingen** om de stand van de techniek verder te specificeren, een inschatting te maken van de risico’s en het innovatiepotentieel vanuit technologische invalshoek te bepalen.
5. **Eindverslag van de marktconsultatie** opgemaakt door Addestino op basis van alle vergaarde informatie tijdens de werksessies.

1.2 Deelnemende partners

In het kader van de transparantie en het open karakter van de marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ worden hierbij de deelnemende bedrijven, overheidsdiensten en kennisinstellingen gelijst.

Vlaams Parlement
Departement Economie, Wetenschap en Innovatie
Zoom Media
Cedat 85
Universiteit Gent
Tweede Kamer der Staten Generaal
E-Deo
Telequip
Nuance Communications
Braincreators
Bumicom Telecommunicatie
Vrije Universiteit Brussel
MyForce
KU Leuven
Bertin IT
Tilburg University
Arbor Media
TeleCats
VRT
Limecraft
AVT Benelux

2 Samenvatting

Dit document is het eindverslag van de marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’, dat liep van 22 september 2017 tot 27 november 2017 en waarbij Addestino optrad als externe begeleider voor EWI. Het bevat de volledig uitgewerkte resultaten van het traject en schetst de mogelijke prioriteiten en vervolgstappen voor het opmaken van een RFP (Request for Proposal). Het Vlaams Departement Economie, Wetenschap en Innovatie (EWI), in samenwerking met het Vlaams Parlement, wenst door middel van een innovatieve aanpak van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ de huidige workflow van de redacteur aanzienlijk te verbeteren zodat meer werkbaar werk mogelijk wordt, zonder daarbij in te boeten aan kwaliteit van de gegenereerde verslagen. De conclusies van de marktconsultatie van ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ situeren zich op drie vlakken.

Als eerste conclusie: het project ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ is zeer innovatief vanuit gebruikersstandpunt. Het project adresseert heel wat verscheidene gebruikersgroepen (direct of indirect) en tracht voor elk van die gebruikersgroepen specifieke noden, behoeften, problemen of uitdagingen op te lossen. De vele use-cases met hoog ingeschatte toegevoegde waarde voor de eindgebruiker getuigen hiervan. In totaal werden een vijftigtal diverse use-cases opgesteld met relevante eindgebruikers, waarvan ruim de helft bestempeld werden door die eindgebruikers als “zeer belangrijk”. Het innovatieve karakter van deze use-cases werd ook bevestigd door de partners uit de industrie en kennisinstellingen.

Als tweede conclusie: het project ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ is zeer innovatief vanuit technologisch standpunt. Uit discussies met experts uit de industrie en kennisinstellingen blijkt dat er vandaag al heel wat technologische ondersteuning voor de automatische transcriptie van spraak naar tekst bij parlementaire vergaderingen haalbaar is en zelfs commercieel beschikbaar. Verder kwamen ruim 20 use-cases naar voor die elk een bepaald innovatiepotentieel (en dus risico) met zich mee dragen. Die facetten omvatten veelal dezelfde functionaliteiten als de vandaag reeds beschikbare, alleen zijn ze niet rechtstreeks toepasbaar op de omgeving van het Vlaams Parlement zonder aanpassingen op maat. De voornaamste technologische uitdagingen voor ‘Spraktechnologie voor Verslaggeving Vlaams Parlement’ zijn automatische spraakherkenning die sprekersonafhankelijk werkt en een juiste betrouwbaarheidsindicatie aanlevert; het ondersteunen van spraak naar tekst in de parlementaire omgeving, inclusief regionale accenten, vakjargon, onderbrekingen en discussies en vlotte workflow voor het automatisch herkennen van sprekers en het introduceren van nieuwe sprekers.

Kwaliteit van de transcriptie staat in deze altijd voorop. Een volautomatisch verslaggeving lijkt daarom uit den boze. De redacteur dient niet vervangen te worden door een automatisch proces, maar eerder ondersteund te worden in het verwerkingsproces.

Als derde conclusie: de belangrijke mate waarin transcriptie van spraak naar tekst een hoge toegevoegde waarde levert aan verschillende groepen van eindgebruikers in combinatie met een zeker risico dat op korte termijn gereduceerd kan worden, vereist een precommerciële ontwikkelingsfase. In dit traject worden een aantal kleinere precommerciële aanbestedingen gelanceerd met als doel de huidige performantie en kwaliteit van bestaande bouwblokken die inkoppelen op de centrale applicatie te evalueren en significant te verbeteren.

De voorgestelde vervolgstappen na dit traject omvatten het opstellen van een gedetailleerde tijdslijn en budget voor de precommerciële fase en het samenstellen van evenwichtige consortia van aanbieders. Deze zaken vallen echter buiten de context van dit document. Tot slot wenst Addestino het EWI te bedanken voor het gestelde vertrouwen en het Vlaams Parlement voor de geleverde expertise en inspanningen en alle partners voor hun deskundigheid en enthousiaste medewerking!

3 Positioneren en definiëren

3.1 De projectinitiators

3.1.1 Vlaams Departement Economie, Wetenschap en Innovatie (EWI)

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

Het PIO loopt van januari 2017 t.e.m. december 2019. In 2017 en 2018 worden in het kader van het PIO meerdere projectoproepen gelanceerd. 2019 dient om de lopende projecten af te ronden. Jaarlijks (2017 en 2018) worden een 15-tal projecten verwacht. Voor de uitvoering van het PIO wordt jaarlijks een budget van 5 miljoen euro voorzien.

Het PIO biedt aan overheidsorganisaties in Vlaanderen begeleiding en cofinanciering bij de ontwikkeling en validering van innovatieve oplossingen. Dit kunnen nieuwe of sterk verbeterde producten of diensten zijn maar ook nieuwe manieren van werken en organiseren.

3.1.2 Vlaams Parlement

Het Vlaams Parlement belichaamt de wetgevende pijler en de democratie in Vlaanderen. Het staat in voor de totstandkoming van de Vlaamse decreten en voor de goedkeuring van de Vlaamse begroting. Het Vlaams Parlement benoemt de Vlaamse Regering en controleert de Vlaamse Regering op de uitvoering van die decreten.

Het Vlaams Parlement wil het centrum van een democratisch Vlaanderen zijn en de burgers nauw bij de politieke besluitvorming betrekken, en streeft daartoe naar maximale transparantie en optimale communicatie. Het wil een modern parlement zijn dat slagvaardig, vooruitziend en dynamisch werkt, op basis van de principes van effectiviteit, efficiëntie en zuinigheid.

Door het uitdragen en aanmoedigen van burgerzin wil het Vlaams Parlement het democratisch draagvlak van de samenleving versterken en Vlaanderen profileren als een moderne, vooruitstrevende, verdraagzame en solidaire Europese regio.

De managementovereenkomst van de secretaris-generaal (2012-2018) geeft uitvoering aan deze missie en beoogt de verdere ontwikkeling van het Algemeen Secretariaat (de ondersteunende diensten van het Vlaams Parlement) tot een klantgerichte en professionele organisatie.

Strategische doelstellingen daarbij zijn onder meer een “kwaliteitsvolle dagelijkse ondersteuning van de 124 Vlaamse volksvertegenwoordigers op inhoudelijk, administratief, logistiek en technisch vlak” en een “actieve medewerking aan de uitvoering van de strategische keuzes van het Vast en Uitgebreid

Bureau voor wat de werking en ondersteuning van het politieke niveau betreft”. Een en ander houdt onder meer de verdere omvorming van het Vlaams Parlement tot een e-parlement en dus een digitalisering van de backoffice en de werkprocessen in. Bovendien moet het Vlaams Parlement en zijn werking (beter) bekendgemaakt worden en de instelling op de politieke kaart van Vlaanderen geplaatst.

Verder onderzoek naar de inzet van ‘speech to text’-spraaktechnologie, met andere woorden automatische spraakherkenning, bij de productie van verslagen van parlementaire debatten en bij de ondertiteling van de beeldverslagen van de plenaire en de commissievergaderingen op de parlementaire website kan een bijzonder waardevolle bijdrage leveren om deze doelstellingen te bereiken

3.2 Werkingskader van de marktconsultatie

3.2.1 De digitale uitdaging van het Vlaams Parlement

Met dit project beoogt het Vlaams Parlement in eerste instantie verder onderzoek naar een automatische transcriptie, ‘speech to text’. In het Nederlands taalgebied loopt op dit moment al een onderzoeksproject bij de openbare omroep omtrent automatische transcriptie van nieuwsuitzendingen en documentaires.

Het Vlaams Parlement wil verder bouwen op de beschikbare technologieën en spraakherkenning gebruiken om de audio-/video-opname van de diverse vergaderingen in het Vlaams Parlement automatisch naar tekst om te zetten en op die manier de werkprocessen van de redacteurs/verslaggevers te optimaliseren. Bijkomend beoogt het Vlaams Parlement (zo mogelijk live) ondertiteling van de beeldverslagen op de website van het Vlaams Parlement.

3.3 Het voortraject: “Spraktechnologie voor Verslaggeving Vlaams Parlement” (STVP)

3.4 Beoogde doelstelling van het voortraject

In dit project wordt prioritair gepoogd om de arbeidsomstandigheden van de redacteurs van de dienst Woordelijk Verslag en de dienst Commissies kwalitatief te verbeteren. Dit moet gebeuren door een accurate ‘speech-to-text’ oplossing die de redacteurs ondersteunt bij het verwerken van vergaderingen en toespraken en op deze manier tijdswinst oplevert. Bij de huidige stand van technologie is de volledige vervanging van de redacteur door technologie geenszins aan de orde. Om dit alles te verwezenlijken worden de technologische grenzen afgetast in de ontwikkeling van een oplossing die kwaliteitsvolle tekst oplevert en gebruiksvriendelijk is.

Een tweede doel is het live en volledig automatisch ondertitelen van Parlementaire vergaderingen en zittingen. Aangezien er op dit moment geen dergelijke ondertiteling wordt aangeboden is de tussenkomst door redacteurs niet mogelijk. Ten allen tijde heeft de accuraatheid van het systeem voorrang op de snelheid. Ondertiteling zal dan ook enkel worden opgenomen indien dit geen invloed heeft op de kwaliteit van de andere objectieven.

Idealiter werkt de marktconsultatie naar een ‘black box’ oplossing, met spraak als input en tekst als output, die zelflerend is en zich vlot aanpast aan de veranderende omstandigheden op het gebied van spraak binnenin het Vlaams Parlement.

3.5 Deliverables van het voortraject

Tijdens de marktconsultatie wordt toegewerkt naar een eindverslag:

Het voorliggende eindverslag dat alle inzichten van de marktconsultatie beschrijft waaronder:

1. de afstemming tussen gebruikersbehoeften en het innovatieve aanbod
2. een goede inschatting van het innovatiepotentieel
3. de haalbaarheid en de vereiste investeringen van de oplossing

4 Inschatting van het innovatiepotentieel vanuit gebruikersstandpunt

Vanuit de expertise van de projectinitiators worden de verscheidene gebruikersgroepen en hun noden in kaart gebracht. De gebruikersgroepen omvatten zowel de eindgebruikers die met de beoogde oplossing voor spraak- en taaltechnologisch ondertitelen aan de slag moeten gaan, als de doelgroepen die onrechtstreeks geïmpacteerd worden.

Vervolgens worden de belangrijkste gebruikersvereisten in kaart gebracht vanuit de expertise van de gebruikers zelf aan de hand van een niet-exhaustieve lijst van use-cases. Op deze manier wordt getracht een eerste algemeen beeld te schetsen van de geambieerde capaciteiten van het project.

4.1 De verschillende gebruikersgroepen en hun noden

Er zijn drie groepen gebruikers, elk met hun eigen behoeftes. De aangeboden oplossing voor de redacteurs van het Woordelijk Verslag (WV) en de verslaggevers van binnen de dienst Commissies is prioritair. Indien technologisch mogelijk mag de applicatie ook de vergaderingen live ondertitelen. Dit is echter niet prioritair.

4.1.1 Redacteurs van het woordelijk verslag

De dienst Woordelijk Verslag staat in voor het integraal geredigeerd verslag van de plenaire vergaderingen en van interpellaties en vragen om uitleg in commissievergaderingen. Dat verslag wordt zo snel mogelijk op de website van het Vlaams Parlement gepubliceerd. De huidige situatie laat zich als volgt samenvatten:

- Dat Woordelijk Verslag heeft een vaste vorm: het is een chronologisch verslag in de directe rede. In de huidige situatie spelen redacteurs geluidsopnames af van vergaderingen die ze hebben bijgewoond. Aan de hand van de audio en gemaakte notities typen ze het verslag.
- Dit is arbeidsintensief werk waarbij 5 minuten vergaderen ongeveer 1 uur de tijd vergt om omgezet te worden in een verslag. De spraaktechnologie moet dan ook dienen ter ondersteuning van de redacteurs en het lichter maken van de fysieke arbeid.
- De redacteurs gebruiken al Dragon NaturallySpeaking wat hen toelaat om het verslag in te spreken eerder dan in te typen. Het herinspreken zorgt niet voor tijdswinst en verlaagt de tijdsdruk op de redacteurs dus niet.

4.1.2 Verslaggevers van de dienst commissies

De dienst Commissies staat in voor het (synthese)verslag van besprekingen van parlementaire en regeringsinitiatieven en verzoekschriften en van hoorzittingen en gedachtewisselingen in commissie. Qua vorm benaderen verslagen van hoorzittingen en gedachtewisselingen het meest het Woordelijk Verslag. Verslagen van besprekingen daarentegen zijn meer samenvattend en (bij voorkeur) eerder logisch-thematisch dan chronologisch gestructureerd.

4.1.3 Beeldverslagen

Alle commissie- en plenaire vergaderingen worden live op de website van het Vlaams Parlement uitgezonden. Het ondertitelen van de vergaderingen is echter ondergeschikt aan de verslagen van plenaire zittingen en commissievergaderingen.

4.1.4 Geïmpacteerde gebruikers

Naast het verlichten of oplossen van problemen voor de actieve gebruikers bestaat er ook een doelgroep die geen invloed hebben op de verwerking maar wel gebruik maken van het eindresultaat. De groep kunnen we opdelen aan de hand van twee objectieven:

- Gebruikers van de verslagen;
 - Mogen geen verlaging van kwaliteit van de verslagen ondervinden.
 - Mogen geen verschillen merken in de robuustheid van het verwerkingsproces. I.e. de publicatie van een verslag moet minsten even snel gebeuren als in de huidige situatie en de technologie mag niet systematisch uitvallen.
- Gebruikers van de ondertiteling zoals doven, slechthorenden en mensen in publieke ruimtes zonder geactiveerde geluidsinstallatie;
 - Moeten een voldoende hoge kwaliteit van ondertiteling krijgen om de vaak technische discussies vlot te kunnen volgen.
 - Moeten dit live te zien krijgen via de website van het Vlaams Parlement.

4.2 Opstellen van de gebruikersvereisten

Onderstaande lijst van een dertigtal use-cases, opgesteld samen met vertegenwoordigers van het Vlaams Parlement, brengt de gebruikersvereisten in kaart.

Om de leesbaarheid te verhogen worden redacteurs van het Woordelijk Verslag hier benoemd als redacteur en verslaggevers van de dienst Commissies als verslaggever.

De betekenis van de Score kolom wordt uitgelegd in sectie 4.3.

ID	Als...	kan ik...	zodat...	Score
1.1	Redacteur	een Word-document openen met daarin de transcriptie van één van mijn beurtten	ik de audio niet moet herbeluisteren en snel een afgewerkt verslag kan maken	20-40
1.2	Verslaggever	een Word-document openen met daarin de transcriptie van één van mijn beurtten	ik de audio niet moet herbeluisteren en snel een afgewerkt verslag kan maken	3
2.1	Redacteur	het percentage van accuraatheid zien van woorden of zinsdelen	ik efficiënt kan controleren welke onderdelen manueel gecontroleerd moeten worden	13
2.2	Verslaggever	het percentage van accuraatheid zien van woorden of zinsdelen	ik efficiënt kan controleren welke onderdelen manueel gecontroleerd moeten worden	5
2.3	Redacteur	kiezen tussen verschillende alternatieven van woorden of zinnen	ik minder typewerk heb in geval van transcriptiefouten	20
2.4	Verslaggever	kiezen tussen verschillende alternatieven van woorden of zinnen	ik minder typewerk heb in geval van transcriptiefouten	13
3.1	Verslaggever	duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 100% accurate sprekersherkenning	ik dit snel en efficiënt in mijn verslag kan opnemen	40
3.2	Verslaggever	duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 80% accurate sprekersherkenning	ik dit snel en efficiënt in mijn verslag kan opnemen	3
3.3	Redacteur	een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf spreekt of als voorzitter	dit snel en efficiënt in het verslag wordt opgenomen	8
3.4	Verslaggever	een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf spreekt of als voorzitter	dit snel en efficiënt in het verslag wordt opgenomen	1
4.1	Beheerder van het systeem	automatisch éénmalige sprekers laten introduceren in het systeem	de sprekersherkenning up to date blijft voor externe sprekers	40
4.2	Beheerder van het systeem	automatisch éénmalige sprekers laten introduceren in het systeem op basis van de eerste minuten die ze spreken	de sprekersherkenning up to date blijft voor externe sprekers	20
4.3	Beheerder van het systeem	éénmalige sprekers laten introduceren in het systeem door hen een trainingstekst te laten inlezen	de sprekersherkenning up to date blijft voor externe sprekers	1

4.4	Beheerder van het systeem	gemakkelijk grote groepen nieuwe sprekers introduceren in het systeem	de sprekersherkenning up to date blijft bij de 5 jaarlijkse verandering van het parlement	20-40
5.1	Verslaggever	redactionele richtlijnen meegeven waar het systeem rekening mee houdt	ik minder verbeterwerk heb	20-40
5.2	Redacteur	redactionele richtlijnen meegeven waar het systeem rekening mee houdt	ik minder verbeterwerk heb	40
5.3	Verslaggever	kan ik gebruik maken van een lexicon met specifiek politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties	ik minder verbeterwerk heb	40
5.4	Redacteur	kan ik gebruik maken van een lexicon met specifiek met politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties	ik minder verbeterwerk heb	40
5.5	Redacteur of Verslaggever	een lexicon aanpassen naargelang het vakgebied	ik minder verbeterwerk heb	13
6.1	Redacteur	kan ik voor elk stukje tekst zien aan welk tijdstip van het audiofragment dit gelinkt is	ik dubieuze transcripties snel kan opzoeken	20
6.2	Verslaggever	kan ik voor elk stukje tekst zien aan welk tijdstip van het audiofragment dit gelinkt is	ik dubieuze transcripties snel kan opzoeken	20
7.1	Redacteur	live transcriptie krijgen en aanpassingen maken	ik na afloop van de vergadering de volledig juiste tekst bezit	13
7.2	Verslaggever	live transcriptie krijgen en aanpassingen maken	ik na afloop van de vergadering de volledig juiste tekst bezit	5
8.1	Redacteur	nauwelijks vertraging oplopen door het uitvallen van het systeem	nooit deadlines mis	20
8.2	Verslaggever	nauwelijks vertraging oplopen door het uitvallen van het systeem	nooit deadlines mis	13
9	Vlaams Parlement	de video's gepubliceerd op de website van het Vlaams Parlement live laten ondertitelen	ik een ruimer publiek kan aanspreken	13
10.1	Redacteur	ondervind ik geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm	de kwaliteit optimaal blijft	20
10.2	Verslaggever	ondervind ik geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm	de kwaliteit optimaal blijft	20
11.1	Redacteur	ondervindt mijn transcriptie applicatie geen last van regionale accenten	de kwaliteit optimaal blijft	20
11.2	Verslaggever	ondervindt mijn transcriptie applicatie geen last van regionale accenten	de kwaliteit optimaal blijft	20
11.3	Redacteur	ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers	de kwaliteit optimaal blijft	20

11.4	Verslaggever	ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers	de kwaliteit optimaal blijft	20
11.5	Redacteur	ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers	de kwaliteit optimaal blijft	20
11.6	Verslaggever	ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers	de kwaliteit optimaal blijft	20
12	Redacteur en Verslaggever	krijg ik grammaticaal correcte zinnen	ik sneller en efficiënter kan werken	20
13.1	Redacteur en Verslaggever	krijg ik semantisch correcte zinnen	ik sneller en efficiënter kan werken	40
13.2	Redacteur en Verslaggever	krijg ik semantisch correcte zinnen met gesuggereerde alternatieven	ik sneller en efficiënter kan werken	13
14	Verslaggever	krijg ik een logische en gestructureerde samenvatting van de commissievergaderingen	ik sneller en efficiënter kan werken	100
15.1	Redacteur	een vooraf gegenereerde begeleidende tekst van een toespraak meegeven	de kwaliteit van de speech-to-text omhoog gaat	13
15.2	Verslaggever	een vooraf gegenereerde begeleidende tekst van een toespraak meegeven	de kwaliteit van de speech-to-text omhoog gaat	13
16.1	Redacteur	rekenen op een systeem waar accuraatheid boven snelheid gaat	ik minder verbeterwerk heb	
16.2	Verslaggever	rekenen op een systeem waar accuraatheid boven snelheid gaat	ik minder verbeterwerk heb	
17.1	Redacteur	een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties	het gebruiksvriendelijk is	20
17.2	Verslaggever	een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties	het gebruiksvriendelijk is	1
18	Redacteur	op het systeem rekenen om op een correcte manier om te gaan met meerdere parlementsleden die elkaar onderbreken	Zodat ik minder verbeterwerk heb	8
19	Redacteur	op het systeem rekenen om op een correcte manier om te gaan met nodeloze herhalingen... in functie van mijn persoonlijke instelling	Zodat ik minder verbeterwerk heb	3
20	Redacteur	op het systeem rekenen om te gaan met haperingen, stottingen, stopwoorden...	Zodat ik minder verbeterwerk heb	20

4.3 Inschatting van het innovatiepotentieel

Het innovatiepotentieel van elke use-case wordt toegekend door middel van een *planning poker* techniek, waarbij de gebruikers een score geven aan elke use-case. De laagste waarde in de schaalverdeling is een 0 en betekent “totaal niet innovatief voor de eindgebruiker”. Een waarde 20 en hoger dient dan weer geïnterpreteerd te worden als “uitermate innovatief voor de eindgebruiker”.

Use-case 1.1: Als redacteur van het Woordelijk Verslag kan ik een Word-document openen met daarin de transcriptie van één van mijn beurten

Score van het innovatiepotentieel: **20-40**

Automatische transcriptie levert zeker tijdswinst op voor de redacteurs van het Woordelijk Verslag. Desondanks zal het even duren alvorens het vertrouwen van de gebruikers kan gewonnen worden. Tijdens de periode waarin de gebruikers dit vertrouwen nog niet hebben, zullen ze de audiobestanden herbeluisteren.

Use-case 1.2: Als verslaggever van de dienst Commissies kan ik een Word-document openen met daarin de transcriptie van één van mijn beurten

Score van het innovatiepotentieel: **3**

Aangezien de verslaggevers als doel hebben een samenvattend en gestructureerd verslag te maken herbeluisteren ze altijd de audiobestanden. Het verslag wordt door hen gemaakt aan ongeveer de zelfde snelheid als de audio. Het verkrijgen van de exacte verwoording in tekstuele vorm levert nauwelijks tot geen voordelen op.

Use-case 2.1: Als redacteur van het Woordelijk Verslag kan ik het percentage van accuraatheid zien van woorden of zinsdelen

Score van het innovatiepotentieel: **13**

Indien de redacteurs aanwezig zijn in de zaal tijdens de zittingen is dit voldoende om fouten of merkwaardigheden snel op te merken in de aangeleverde transcriptie. Het is echter wel zo dat de aanwezigheid van een inschatting van de accuraatheid de nood om naar de zaal te gaan sterk verlaagt. In de veronderstelling dat de redacteur niet meer in de zaal aanwezig zal zijn wordt het innovatiepotentieel als relevant ingeschat.

Use-case 2.2: Als verslaggever van de dienst Commissies kan ik het percentage van accuraatheid zien van woorden of zinsdelen

Score van het innovatiepotentieel: **5**

Door de nood aan herbeluisteren van de audiobestanden door de verslaggevers (zie use-case 1.2) is het innovatiepotentieel laag.

Use-case 2.3: Als redacteur van het Woordelijk Verslag kan ik kiezen tussen verschillende alternatieven van woorden of zinnen

Score van het innovatiepotentieel: **20**

In de huidige applicatie voor het dicteren van tekst (Dragon Naturally Speaking) is een gelijkaardige functionaliteit opgenomen die door de gebruikers als een meerwaarde wordt ervaren.

Use-case 2.4: Als verslaggever van de dienst Commissies kan ik kiezen tussen verschillende alternatieven van woorden of zinnen

Score van het innovatiepotentieel: **13**

Bij het maken van een samenvattende tekst, zitten verslaggevers soms verlegen om synoniemen of andere woorden. Indien deze ook worden aangeboden als alternatieven kan dit zorgen voor een vlottere manier van werken.

Use-case 3.1: Als verslaggever van de dienst Commissies kan ik duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 100% accurate sprekersherkenning

Score van het innovatiepotentieel: **40**

Bij het herbeluisteren van de audio, wat verslaggevers altijd doen (zie use-case 1.2), kunnen sommige sprekers meteen herkend worden. Toch zijn er ook velen die niet meteen herkenbaar zijn door een onopvallende stem of vlakke manier van spreken. Indien de voorzitter van de Commissievergadering vergeet aan te geven wie aan het woord is, moet in deze gevallen de naam van de spreker opgezocht worden in een lijst. Het overslaan van deze opzoekingsstap biedt absolute meerwaarde.

Use-case 3.2: Als verslaggever van de dienst Commissies kan ik duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 80% accurate sprekersherkenning

Score van het innovatiepotentieel: **3**

Aangezien er niet vertrouwd kan worden op de kwaliteit van de sprekersherkenning, moet het opzoekingswerk omschreven in use-case 3.1 toch gebeuren. Bijgevolg heeft dit nauwelijks meerwaarde.

Use-case 3.3: Als redacteur van het Woordelijk Verslag kan ik een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf of als voorzitter spreekt

Score van het innovatiepotentieel: **8**

Dit komt regelmatig voor en kan soms moeilijkheden opleveren. Bij het bepalen of de voorzitter uit eigen naam spreekt kunnen zich drie situaties voordoen. Sommige voorzitters benoemen de situatie aan het begin van hun betoog, soms valt het uit de context op te maken en in sommige gevallen moet de redacteur een inschatting maken naar eigen goeddunken. Technologische hulp hierbij kan het werk verlichten.

Use-case 3.4: Als verslaggever van de dienst Commissies kan ik een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf of als voorzitter spreekt

Score van het innovatiepotentieel: **1**

Dit komt niet voor in het verslag en heeft dus geen meerwaarde.

Use-case 4.1: Als beheerder van het systeem kan ik automatisch éénmalige sprekers laten introduceren in het systeem

Score van het innovatiepotentieel: **40**

Het introduceren van nieuwe sprekers in het systeem kan niet manueel gebeuren net voor vergaderingen door de hectiek van de situatie. Dit vooral omdat het te tijdrovend is voor de sprekers zelf. Het automatiseren van deze handeling is dus een voorwaarde.

Use-case 4.2: Als beheerder van het systeem kan ik automatisch éénmalige sprekers laten introduceren in het systeem op basis van de eerste minuten die ze spreken

Score van het innovatiepotentieel: **20**

Het nadeel tegenover use-case 4.1 is dat deze methode extra werk oplevert tijdens de eerste minuten van de bezoeker aangezien die nog niet herkend wordt.

Use-case 4.3: Als beheerder van het systeem kan ik éénmalige sprekers laten introduceren in het systeem door hen een trainingstekst te laten inlezen

Score van het innovatiepotentieel: **1**

Zoals omschreven in use-case 4.1 is hiervoor geen tijd voor éénmalige bezoekers. Deze functionaliteit zou in de praktijk niet kunnen gebruikt worden en dus geen meerwaarde opleveren.

Use-case 4.4: Als beheerder van het systeem kan ik sprekers uit grote groepen laten introduceren in het systeem door hen een trainingstekst te laten inlezen

Score van het innovatiepotentieel: **20-40**

Deze situatie is van toepassing bij de vijfjaarlijkse verandering binnen het parlement. De parlementairen moeten bij het ingaan van een nieuwe ambtstermijn administratie in orde brengen (bv. pasfoto laten nemen). Tijdens deze procedure kan een trainingszin worden ingelezen.

Use-case 5.1: Als redacteur van het Woordelijk Verslag kan ik redactionele richtlijnen meegeven waar het systeem rekening mee houdt

Score van het innovatiepotentieel: **20-40**

Deze afspraken worden gemaakt samen met een taaladviseur. Enkele voorbeelden van redactionele richtlijnen zijn: procent of percent, hoofdlettergebruik, geen Mr. Of Mvr., lay-out, het juist gebruiken van uitdrukkingen etc. Dit alles wordt omschreven in een schrijfwijzer. Het volgen van deze richtlijnen dient door de redacteurs nauwlettend in de gaten gehouden te worden. De juiste opname van deze zaken in het verslag is dus van groot belang.

Use-case 5.2: Als verslaggever van de dienst Commissies kan ik redactionele richtlijnen meegeven waar het systeem rekening mee houdt

Score van het innovatiepotentieel: **40**

Idem als use-case 5.1.

Use-case 5.3: Als redacteur van het Woordelijk Verslag kan ik gebruik maken van een lexicon met specifiek politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties

Score van het innovatiepotentieel: **40**

Aangezien dit taalgebruik zeer vaak voorkomt in de parlementaire debatten is dit van groot belang om de vlotheid van het werk te garanderen. Deze use-case wordt beschouwd als een minimum.

Use-case 5.4: Als verslaggever van de dienst Commissies kan ik gebruik maken van een lexicon met specifiek politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties

Score van het innovatiepotentieel: **40**

Idem als use-case 5.3.

Use-case 5.5: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies kan ik een lexicon aanpassen naargelang het vakgebied

Score van het innovatiepotentieel: **13**

Ten eerste dient te worden opgemerkt dat dit enkel dient geïmplementeerd te worden indien deze aanpassing eenvoudig kan gebeuren. Indien dit kan, biedt het zeker een meerwaarde aangezien vele afkortingen voorkomen in meerdere vakgebieden maar daar een verschillende betekenis hebben. Deze overlappings moeten per onderwerp de juiste equivalent krijgen in het verslag.

Use-case 6.1: Als redacteur van het Woordelijk Verslag kan ik voor elk stukje tekst in het Word-document zien aan welk tijdstip van het audiofragment dit gelinkt is

Score van het innovatiepotentieel: **20**

In de huidige situatie wordt een manueel opgesteld overzicht geleverd van alle sprekers en de tijdstippen waarop ze spreken. De aanwezigheid van dit document is belangrijk. Bijgevolg is deze use-case in combinatie met sprekersherkenning dat ook.

Use-case 6.2: Als verslaggever van de dienst Commissies kan ik voor elk stukje tekst in het Word-document zien aan welk tijdstip van het audiofragment dit gelinkt is

Score van het innovatiepotentieel: **20**

Idem als use-case 6.1.

Use-case 7.1: Als redacteur van het Woordelijk Verslag kan ik live transcriptie krijgen en aanpassingen maken

Score van het innovatiepotentieel: **13**

Het live aanpassen en verbeteren van kleine zaken kan het werk na de vergaderingen verlichten.

Use-case 7.2: Als verslaggever van de dienst Commissies kan ik live transcriptie krijgen en aanpassingen maken

Score van het innovatiepotentieel: **5**

Omdat het werk voor verslaggevers voornamelijk bestaat uit samenvatten en structureren heeft de live weergave van een letterlijk verslag weinig nut.

Use-case 8.1: Als redacteur van het Woordelijk Verslag kan ik nauwelijks vertraging oplopen door het uitvallen van het systeem

Score van het innovatiepotentieel: **20**

Het niet uitvallen van het systeem onder normale omstandigheden is vooral van belang omdat niet meer kan teruggedrepen worden naar de oude manier van werken. Een afname van werk betekent immers ook een afname van mankracht. Bij falen van het systeem is dus niet voldoende mankracht meer aanwezig om al het werk manueel te verrichten.

Use-case 8.2: Als verslaggever van de dienst Commissies kan ik nauwelijks vertraging oplopen door het uitvallen van het systeem

Score van het innovatiepotentieel: **13**

Idem als use-case 8.1

Use-case 9: Als Vlaams Parlement kan ik de video's gepubliceerd op de website van het Vlaams Parlement live laten ondertitelen

Score van het innovatiepotentieel: **13**

Dit levert meerwaarde maar wordt zeker niet als eerste prioriteit van het project gezien.

Use-case 10.1: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm

Score van het innovatiepotentieel: **20**

Dit wordt geschat als waardevol. Wel moet er worden opgemerkt dat er zich soms situaties voordoen waarbij er gesproken wordt zonder dat er een persoonlijke micro aanstaat. Toch moet dit opgenomen worden in het verslag zoals dit in de huidige situatie ook gedaan wordt.

Use-case 10.2: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm

Score van het innovatiepotentieel: **20**

Idem als use-case 10.1

Use-case 11.1: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van regionale accenten

Score van het innovatiepotentieel: **20**

In het Vlaams Parlement komt een breed spectrum aan idiosyncratische en regionale spraakkenmerken en diverse spreekstijlen voor. Dit omzetten naar neutrale taal is van groot belang.

Use-case 11.2: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van regionale accenten

Score van het innovatiepotentieel: **20**

Idem als use-case 11.1.

Use-case 11.3: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers

Score van het innovatiepotentieel: **20**

Idem als use-case 11.1.

Use-case 11.4: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers

Score van het innovatiepotentieel: **20**

Idem als use-case 11.1

Use-case 11.5: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers

Score van het innovatiepotentieel: **20**

Idem als use-case 11.1 met als bijkomende problemen incoherente zinnen en woordenbreien.

Use-case 11.6: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers

Score van het innovatiepotentieel: **20**

Idem als use-case 11.5

Use-case 12: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik grammaticaal correcte zinnen

Score van het innovatiepotentieel: **20**

Grammaticale correctheid is van uitermate groot belang maar kan gemakkelijker door de redacteurs gecorrigeerd worden dan foutieve inhoud.

Use-case 13.1: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik semantisch correcte zinnen

Score van het innovatiepotentieel: **40**

Het omzetten van spreektaal die soms onsamenhangend is betekent veel werk. De automatisatie hiervan is dus een absolute meerwaarde.

Use-case 13.2: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik semantisch correcte zinnen met gesuggereerde alternatieven

Score van het innovatiepotentieel: **13**

Idem als use-case 13.1 met een verlaging van de productiviteit omdat er nog een keuze dient te worden gemaakt.

Use-case 14: Als verslaggever van de dienst Commissies krijg ik een logische en gestructureerde samenvatting van de commissievergaderingen

Score van het innovatiepotentieel: **100**

Indien dit mogelijk is wordt het werk van de verslaggever volledig geautomatiseerd en zijn alle mogelijke noden dus opgelost.

Use-case 15.1: Als redacteur van het Woordelijk Verslag kan ik een vooraf gegenereerde begeleidende tekst van een toespraak meegeven

Score van het innovatiepotentieel: **13**

In de praktijk gebeurt dit vaak al en biedt dit ook een meerwaarde. Vooral bij complexe onderwerpen gaat de kwaliteit van de verslaggeving omhoog.

Use-case 15.2: Als verslaggever van de dienst Commissies kan ik een vooraf gegenereerde begeleidende tekst van een toespraak meegeven

Score van het innovatiepotentieel: **13**

Idem als use-case 15.1

Use-case 16.1: Als redacteur van het Woordelijk Verslag kan ik rekenen op een systeem waar accuraatheid boven snelheid gaat

Dit onderdeel is niet gescoord maar dient ter illustratie naar de spelers op de markt toe. Er is geconcludeerd dat tijdswinst door minder werk opgeleverd wordt door de accuraatheid en niet door het moment waarop spraak verwerkt wordt. Het doel is dus altijd maximale accuraatheid, maar dit wel binnen een gangbare tijdspanne.

Use-case 16.2: Als verslaggever van de dienst Commissies kan ik rekenen op een systeem waar accuraatheid boven snelheid gaat

Idem als use-case 16.1

Use-case 17.1: Als redacteur van het Woordelijk Verslag kan ik een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties

Score van het innovatiepotentieel: **20**

In de huidige situatie bestaat er een webapplicatie met daarin alle vergaderingen. Een vergadering dient te worden opgezocht en geselecteerd, waarna een Word-document wordt gegenereerd met daarin de titels van de journaallijnen en de namen van de sprekers. In dit document redigeren de redacteurs de verslagen.

Er wordt geprefereerd om met Word te werken. Ondanks het feit dat kopiëren van tekst in Word een kleine handeling lijkt is dit niet gewenst door het groot aantal keren dat dit dient te gebeuren. De grote voordelen van Word zijn de spellingchecker, het aanmaken van Macro's en het automatisch aanvullen van veel voorkomende woorden en omzetten van afkortingen in volle woorden (bv MP tot Minister President).

Het nadeel van Word is dat dit individueel gebeurt en niet gemakkelijk met anderen kan worden gedeeld. Toch blijft het belangrijk dat ook iedereen zijn eigen voorkeuren in Word kan opnemen.

Use-case 17.2: Als verslaggever van de dienst Commissies kan ik een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties

Score van het innovatiepotentieel: **1**

Voor de verslaggevers bestaat in de huidige situatie geen applicatie. De verslaggeving wordt gemaakt direct van de audio (zonder aanwezigheid in de zaal) tot op het papier. Er bestaat geen relevante workflow en die is ook niet gewenst in de nieuwe toepassing.

Use-case 18: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met parlamentsleden die elkaar onderbreken

Score van het innovatiepotentieel: **8**

De plenaire etiquette zorgt ervoor dat een plenaire vergadering bestaat uit een mooie opvolging van spreekbeurten. Onderbrekingen uit de zaal gebeuren zonder dat de micro aanstaat en zijn niet te horen op de audio.

Use-case 19: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met nodeloze herhalingen... in functie van mijn persoonlijke instelling

Score van het innovatiepotentieel: **3**

Herhalingen kunnen zich voordoen onder twee omstandigheden. Het kunnen retorische stijlfiguren zijn om zaken te benadrukken en ze kunnen het gevolg zijn van een slechte spreekstijl of iemand die niet uit zijn woorden komt. Het is aan de redacteur om het onderscheid te maken.

Use-case 20: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met haperingen, stotteringen, stopwoorden...

Score van het innovatiepotentieel: **20**

Idem als use-case 11.1.

4.4 Aanvullende algemene beïnvloedingsfactoren

Naast de noden van de verschillende gebruikers, zijn er nog een aantal algemene factoren waarmee er rekening dient te worden gehouden. Deze zijn niet gelinkt aan een bepaalde groep gebruikers, maar zijn onderdeel van de werkomgeving en aanwezige faciliteiten.

- In het Vlaams Parlement vormt de akoestiek van de commissiezalen geen echt probleem. Die van de Koepelzaal (plenaire vergaderzaal) is een grotere uitdaging. Nog moeilijker is de akoestiek van De Schelp, die echter maar in zeer uitzonderlijke omstandigheden als vergaderzaal voor parlementaire besprekingen wordt gebruikt.
- De geluidsinstallatie (microfoons, versterkers) is overal van hoge kwaliteit. De geluidsopnames worden digitaal bewaard in mp3-formaat, bitrate 48 kbit/seconde mono. De videobestanden zullen eerstdaags in minstens even hoge kwaliteit aangeboden worden. Een uitgebreide omschrijving van de digitale opnames is te vinden in Bijlage A.
- De redacteurs van het Woordelijk Verslag werken in een webapplicatie. In Bijlage B wordt deze met bijbehorende workflow omschreven.
- In het Vlaams Parlement zetelen 124 parlementsleden en maximum 11 ministers, zij behoren tot de regelmatig terugkerende sprekers. Er bestaat een cyclus van 5 jaar waarbij tot de helft van deze sprekers kunnen vervangen worden door nieuwe sprekers. Bovendien zijn er ook enkele tientallen, veelal eenmalige, externe sprekers.
- Het Vlaams Parlement heeft tientallen uren geannoteerd spraakmateriaal ter beschikking. Dat materiaal kan per spreker geselecteerd worden, zodat elke stem en accent beter kan worden herkend.
- Alle regels omtrent redactionele afspraken (i.e. use-cases 5.1, 5.2, 5.3, 5.4, 5.5, 12) dienen maximaal collectief in het systeem te worden geïntegreerd, en dus niet op basis van individuele voorkeuren van redacteurs en verslaggevers.

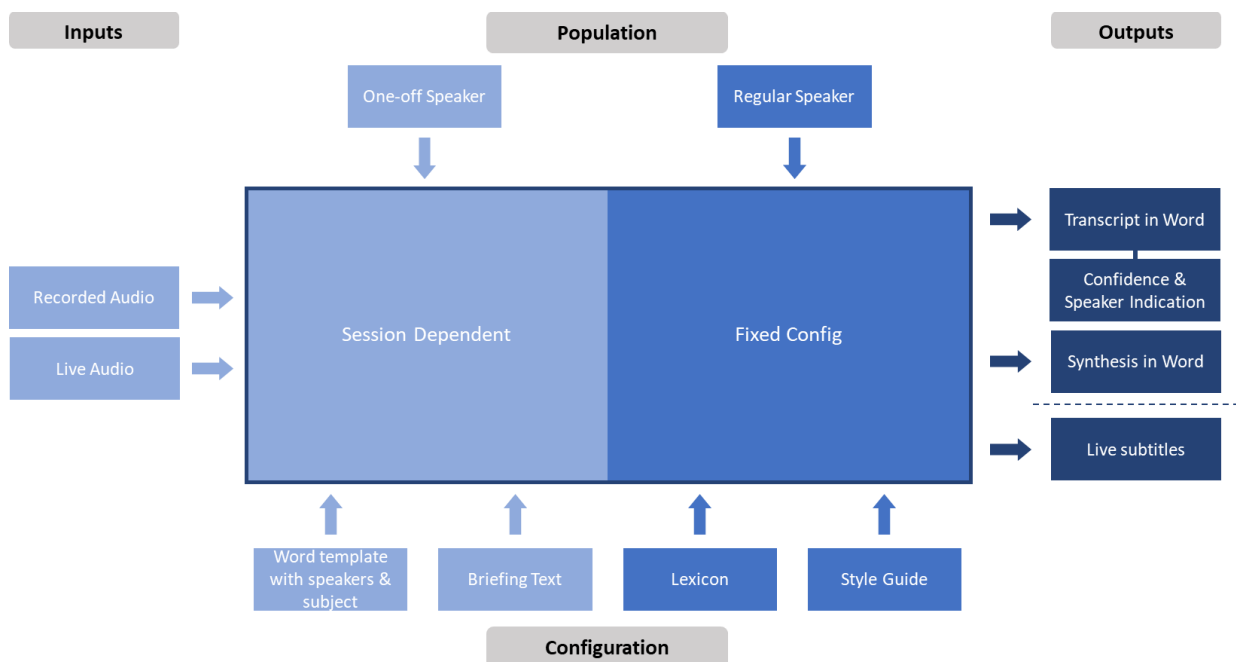
5 Inschatting van de stand van de techniek: innovatiepotentieel vanuit technologische invalshoek

5.1 De referentieoplossing

Figuur 1 illustreert vanuit functionele invalshoek de referentieoplossing. De referentieoplossing omvat een aantal inputs, configuratie en populatie, een aantal outputs en een functionele component. De inputs omvatten bestaande audio- en video-opnames van plenaire zittingen en commissievergaderingen, zowel live opnames als opnames uit het verleden. De functionele component zet audio om in tekst. De output bestaat in drie mogelijke vormen. Ten eerste is er de letterlijke tekstuele vorm van de audio, met een indicatie van accuraatheid van de transcriptie en weergave wie aan het woord is. Ten tweede is er een synthese van de opgenomen audio. Tot slot is er ook de live ondertiteling van plenaire zittingen en commissievergaderingen.

De werking van de functionele component kan worden beïnvloed door de configuratie ervan en de populatie van sprekers. De sessieafhankelijke beïnvloedingsfactoren zijn de eenmalige sprekers, een word-document met daarin de sprekers per vergadering, het onderwerp van die vergadering eventueel aangevuld door een bijhorend lexicon en een tekstbestand met daarin een voorbereiding van speeches of vergaderingen. De vaste beïnvloedingsfactoren zijn de 134 parlementsleden en ministers, een lexicon met afkortingen en vakjargon en een document met redactionele afspraken.

In de volgende secties worden de technologische risico's die vasthangen aan elk van de use-cases besproken.



Figuur 1 De referentieoplossing

De technologie omtrent spraak naar tekst en automatische ondertiteling wordt ontwikkeld door twee soorten spelers in de industrie. Ten eerste zijn er de SMEs en start-ups, die zich veelal op een specifieke markt richten en (vaak) custom made oplossingen aanbieden. Voorbeelden zijn Limecraft, Bertin IT, Braincreators, Bumicom... Ten tweede zijn er de grote technologiebedrijven die 'of the shelf'

oplossingen aanbieden, veelal oplossingen waarbij de spraak, door één persoon gedicteerd, wordt omgezet in tekst. Voorbeelden zijn Nuance, Google, IBM...

Het automatisch genereren van een synthese of samenvatting is dan weer het domein van kennisinstellingen en onderzoeksafdelingen van grote technologiebedrijven. Dit omdat op korte termijn nog geen toepasbare oplossingen in zicht zijn en deze instellingen geld kunnen vrij maken voor zuiver onderzoek.

Voor een volledige oplijsting van alle aanwezige industriepartners en kennisinstellingen tijdens de marktconsultatie wordt verwezen naar sectie 1.2. Voor een gedetailleerd overzicht van de markt wordt verwezen naar Bijlage C: State of the Art analyse.

5.2 Inschatting van de technologische risico's

Het innovatiepotentieel vanuit technologisch standpunt wordt ingeschat door middel van planning poker. Hierbij schalen experts van industriële partners en kennisinstellingen de technologische risico's die vasthangen aan elke use-case. De laagste waarde in de schaalverdeling is een 0 en betekent "totaal geen risico". De oplossing bestaat en kan zonder problemen worden toegepast. Een waarde van 8 dient de worden geïnterpreteerd als "een oplossing is mogelijk maar een inspanning is vereist om ze te bereiken". Waardes boven de 20 impliceren een kans dat een oplossing niet kan worden gevonden. De maximale waarde is 100 en houdt in dat "fysische grenzen dienen te worden verlegd" om een oplossing te ontwikkelen.

Use-case 1.1: Als redacteur van het Woordelijk Verslag kan ik een Word-document openen met daarin de transcriptie van één van mijn beurten

Vanuit technologisch standpunt valt deze use-case uiteen in twee situaties. 1.1a omvat de situatie waarin er geen fouten in de transcriptie staan en de gebruikers blindelings kunnen vertrouwen op de output. 1.1b omvat de situatie waarin geen garanties worden gegeven op kwaliteit.

Use-case 1.1a:

Risicoscore: 40

De huidige technologie is nog niet klaar om spraak om te zetten in tekst aan dezelfde kwaliteit die door mensen wordt behaald. Het risico om dit te bereiken wordt ingeschat als zeer hoog en is niet mogelijk in de nabije toekomst.

Use-case 1.1b:

Risicoscore: 0

Het gaat om een bestaande "piece of workflow" die dient te worden geïmplementeerd.

Use-case 1.2: Als verslaggever van de dienst Commissies kan ik een Word-document openen met daarin de transcriptie van één van mijn beurten

Aangezien use-case 1.2 vanuit technologisch standpunt identiek is aan use-case 1.1 wordt hiernaar verwezen. Ook in het vervolg van deze sectie verwijzen we altijd naar voorgaande use-cases indien enkel de gebruikersgroep verschillend is.

Use-case 2.1: Als redacteur van het Woordelijk Verslag kan ik het percentage van accuraatheid zien van woorden of zinsdelen

Risicoscore: **2**

De huidige technologieën gebruiken mechanismen waarin de accuraatheid van de woorden of zinsdelen zit ingebed. Er dient te worden opgemerkt dat deze subjectief te interpreteren zijn. Het percentage van accuraatheid dient best te worden weergegeven als indicatie (bv. met kleuren) en niet als een absolute waarde. Het weergeven van een absolute waarde die blindelings kan worden vertrouwd houdt een veel groter risico in en krijgt een risicoscore van 40.

Use-case 2.2: Als verslaggever van de dienst Commissies kan ik het percentage van accuraatheid zien van woorden of zinsdelen

Zie use-case 2.1.

Use-case 2.3: Als redacteur van het Woordelijk Verslag kan ik kiezen tussen verschillende alternatieven van woorden of zinnen

Risicoscore: **20**

Het weergeven van verschillende alternatieven houdt een wezenlijk risico in zowel op het niveau van woorden als van zinsdelen. De kwaliteit van de mogelijkheden gaat snel achteruit en indien de beste inschatting niet correct is, zijn de weergegeven alternatieven veelal fout.

Op het niveau van zinsdelen helpt de structuur die aanwezig is in taal om alternatieven op te bouwen, ook al blijft punctuatie (of de afwezigheid ervan) een groot probleem. Bovendien bestaat de achterliggende technologie uit grafen en deze zijn heel moeilijk in een gebruiksvriendelijke weergave te gieten. De markt identificeert deze weergave dan ook als het grootste issue om winst in efficiëntie op te leveren en raad aan om andere oplossingen te gebruiken (vb. de accuraatheidsweergave met kleuren omschreven in use-case 2.1).

Use-case 2.4: Als verslaggever van de dienst Commissies kan ik kiezen tussen verschillende alternatieven van woorden of zinnen

Zie use-case 2.3.

Use-case 3.1: Als verslaggever van de dienst Commissies kan ik duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 100% accurate sprekersherkenning

Risicoscore: **5**

Ten eerste dient te worden opgemerkt dat een accuraatheid van 100% niet mogelijk is. Zo zijn tweelingen en familieleden heel moeilijk uit elkaar te halen. Ook indien er te veel sprekers door elkaar praten of indien sprekers regelmatig onderbroken worden is accurate sprekersherkenning moeilijk. Maar gezien de kwaliteit van de microfoons en het lage aantal individuen in de populatie zal het percentage van accuraatheid heel hoog zijn.

Bovendien kan deze nog sterk verbeterd worden door de aanwezigheid van metadata. Die data kan bestaan uit informatie over de tijdstippen waarop welke microfoons aan staan en videomateriaal waarin de bewegingen van de populatie kunnen worden gevolgd (een camera van voldoende kwaliteit die de zaal filmt volstaat).

Tot slot dient te worden opgemerkt dat de kwaliteit omhoog gaat met de lengte van de beurt van de spreker. Tien seconden of meer volstaan om een spreker te herkennen, vanaf drie seconden is het

mogelijk. Bovendien zijn er marktspelers die ervan overtuigd zijn dat het “een kwestie van milliseconden” is indien er genoeg rekenkracht aanwezig is.

Use-case 3.2: Als verslaggever van de dienst Commissies kan ik duidelijk zien wie aan het woord is bij elk stuk geleverde tekst door een 80% accurate sprekersherkenning

Risicoscore: **2**

Zie use-case 3.1 met als extra opmerking dat de lagere vereisten op kwaliteit het risico verder verlagen.

Use-case 3.3: Als redacteur van het Woordelijk Verslag kan ik een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf of als voorzitter spreekt

Risicoscore: **100**

Aangezien er niet kan worden gebouwd op vaste aankondigingen van een rolwijziging is het onmogelijk voor technologie om dit onderscheid te maken. Hiervoor dient men uitermate veel onderzoek te doen in het terrein van betekenis van tekst en omvorming tot contextuele informatie.

Use-case 3.4: Als verslaggever van de dienst Commissies kan ik een onderscheid maken tussen de voorzitter van een vergadering die voor zichzelf of als voorzitter spreekt

Zie use-case 3.3.

Use-case 4.1: Als beheerder van het systeem kan ik automatisch éénmalige bezoekers laten introduceren in het systeem

Risicoscore: **40**

Hier wordt uitgegaan van de situatie waarin er geen menselijke ingrepen zijn om aan te geven dat er een nieuwe spreker is. Bovendien gaat het hier over het herkennen van die spreker. De oplossing bevindt zich hier in het domein van open verzamelingen. Dit houdt in dat indien een verzameling aan sprekers in het systeem staan, het wel mogelijk is om in te schatten bij welke van deze sprekers een stuk audio het best past. Het is echter niet mogelijk om in te schatten of deze ‘best match’ ook werkelijk correct is.

Bovendien wordt er ook gerekend op een standaard tijd van 30 seconden tot 2 minuten om een stemprofiel te kunnen opbouwen van nieuwe sprekers. Onmiddellijke live herkenning van een nieuwe spreker is dus niet mogelijk.

Use-case 4.2: Als beheerder van het systeem kan ik automatisch éénmalige bezoekers laten introduceren in het systeem op basis van de eerste minuten die ze spreken

Voor deze use-case wordt uitgegaan van de volgende omstandigheden. De start van een spreekbeurt van nieuwe sprekers wordt aangekondigd of meegegeven aan het systeem (bv. door het gebruik van een knop).

Deze use-case valt uiteen in twee use-cases. Use-case 4.2a behandelt het omzetten van spraak naar tekst van een nieuwe spreker. Situatie 4.2b behandelt het herkennen van die nieuwe spreker.

Use-case 4.2a:

Risicoscore: **3**

Vele huidige technologieën kunnen gebruik maken van sprekersonafhankelijke transcriptie met een gemiddelde accuraatheid.

Use-case 4.2b:

Risicoscore: **5**

De grootste uitdaging ligt bij het ontwikkelen van een proces dat gebruiksvriendelijk is en weinig menselijke arbeid vergt. Verder zijn die risico's vrij laag door het lage aantal sprekers die in het systeem zitten. Er kunnen zich enkele uitzonderingen voordoen in de vorm van 'moeilijke' sprekers waarvan meer data nodig is.

Use-case 4.3: Als beheerder van het systeem kan ik éénmalige bezoekers laten introduceren in het systeem door hen een trainingstekst te laten inlezen

Risicoscore: **3**

Met de huidige technologie is dit perfect mogelijk. Het is veeleer een organisatorisch dan een technologisch probleem.

Een mogelijkheid bestaat erin om tijdens de eerste meetings met een iets lagere kwaliteit te werken en de sprekers door manuele aanduiding te herkennen. Hieruit kan het systeem dan leren en automatisch overnemen.

Use-case 4.4: Als beheerder van het systeem kan ik sprekers uit grote groepen laten introduceren in het systeem door hen een trainingstekst te laten inlezen

Risicoscore: **3**

Zie use-case 4.3.

Use-case 5.1: Als redacteur van het Woordelijk Verslag kan ik redactionele richtlijnen meegeven waar het systeem rekening mee houdt

Deze use-case valt uiteen in twee use-cases. In 5.1a wordt de situatie omschreven waar er woord voor woord vervangingen nodig zijn. In use-case 5.1b gaat het over verandering in stijl, woordvolgorde, punctuatie etc...

Use-case 5.1a:

Risicoscore: **0**

Indien er set van regels aanwezig zijn met een 1 op 1 relatie tussen de verkeerde en juiste vorm is dit een gemakkelijk op te lossen probleem.

Use-case 5.1b:

Risicoscore: **40**

Dit probleem bevindt zich in het domein van de betekenis van spraak. Eerst moet de betekenis worden begrepen door de technologie alvorens aanpassingen kunnen gemaakt worden. Dit is nog niet mogelijk.

De markt raadt aan om alle situaties gaande van 5.1a tot 5.2b (en alles ertussen) stap voor stap te behandelen. Los op wat er al kan worden opgelost en beweeg mee met de bestaande technologie om het proces te blijven verbeteren.

Use-case 5.2: Als verslaggever van de dienst Commissies kan ik redactionele richtlijnen meegeven waar het systeem rekening mee houdt

Ook deze use-case valt uiteen in 5.2a en 5.2b waarbij wordt gerefereerd naar 5.1a en 5.1b respectievelijk.

Use-case 5.3: Als redacteur van het Woordelijk Verslag kan ik gebruik maken van een lexicon met specifiek politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties

Risicoscore: **3**

De technologie kan dit faciliteren met de volgende opmerkingen. De woorden dienen geïntroduceerd te worden in het systeem door het aanbieden van audio aan het systeem met daarin de juiste uitspraak. Men dient er rekening mee te houden dat de uitspraak van de sprekers soms anders of fout kan zijn. Al deze vormen moeten ook in het systeem zitten om een goede accuraatheid te bekomen.

Meegeven aan het systeem wat het onderwerp van een vergadering is kan de kwaliteit verhogen. Ook het 'flaggen' van deze nieuwe woorden indien een redacteur ze tegenkomt zal de kwaliteit verhogen aangezien het systeem daaruit kan leren.

Use-case 5.4: Als verslaggever van de dienst Commissies kan ik gebruik maken van een lexicon met specifiek politiek-juridisch taalgebruik, inclusief verwijzingen naar instellingen en organisaties

Idem als use-case 5.3.

Use-case 5.5: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies kan ik een lexicon aanpassen naargelang het vakgebied

Risicoscore: **1**

Het kiezen van een lexicon dat gebruikt moet worden voor een specifieke vergadering is perfect mogelijk.

Use-case 6.1: Als redacteur van het Woordelijk Verslag kan ik voor elk stukje tekst in het Word-document zien aan welk tijdstip van het audiofragment dit gelinkt is

Risicoscore: **0**

Deze technologie wordt al in vele applicaties gebruikt en brengt geen problemen met zich mee.

Use-case 6.2: Als verslaggever van de dienst Commissies kan ik voor elk stukje tekst in het Word-document zien aan welk tijdstip van het audiofragment dit gelinkt is

Idem als use-case 6.1.

Use-case 7.1: Als redacteur van het Woordelijk Verslag kan ik live transcriptie krijgen en aanpassingen maken

Risicoscore: **2**

De risico's bevinden zich in het ontwikkelen van een applicatie die gebruiksvriendelijk is. Bovendien merkt de markt dat dit vaak voor problemen zorgt ivm de snelheid van de menselijke handelingen. Vaak is de mens niet snel genoeg om alle aanpassingen live te verwerken en wordt deze feature na verloop van tijd niet meer gebruikt.

Use-case 7.2: Als verslaggever van de dienst Commissies kan ik live transcriptie krijgen en aanpassingen maken

Idem als use-case 7.1.

Use-case 8.1: Als redacteur van het Woordelijk Verslag kan ik nauwelijks vertraging oplopen door het uitvallen van het systeem

Risicoscore: **2**

Het risico wordt zeer laag ingeschat.

Use-case 8.2: Als verslaggever van de dienst Commissies kan ik nauwelijks vertraging oplopen door het uitvallen van het systeem

Idem als use-case 8.1

Use-case 9: Als Vlaams Parlement kan ik de video's gepubliceerd op de website van het Vlaams Parlement live laten ondertitelen

Risicoscore: **2**

Dit kan gemakkelijk geïmplementeerd worden en brengt weinig risico's met zich mee. De grootste moeilijkheid is het aligneren van tekst en audio omdat transcriptie 7-8 seconden nodig heeft. Indien dit niet gedaan wordt gaat de kwaliteit vanuit gebruikersstandpunt snel achteruit.

Use-case 10.1: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm

Risicoscore: **1**

Dankzij de goede kwaliteit van de microfoons kunnen storende geluiden gemakkelijk worden gefilterd. Wel wordt verondersteld dat sprekers zich niet plots omdraaien of weglopen. Indien ze dit wel doen blijft de kwaliteit van de audio niet behouden.

Use-case 10.2: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen invloed van omgevingsgeluiden, achtergrondlawaai en nagalm

Idem als use-case 10.1

Use-case 11.1: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van regionale accenten

Risicoscore: **5**

Ten eerste dient te worden opgemerkt dat er altijd invloed zal zijn van accenten, maar deze klein genoeg is om een hoge kwaliteit te garanderen. Bovendien gebeurt hier veel onderzoek naar en is er recentelijk een sterke vooruitgang geboekt in dit gebied, specifiek voor de Nederlandse taal.

Uitzonderingen en moeilijkheden zijn weinig voorkomende accenten waarvan niet voldoende data aanwezig is.

Use-case 11.2: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van regionale accenten

Idem als use-case 11.1.

Use-case 11.3: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers

Risicoscore: **5**

Hier dient men meerdere dimensies in rekening te brengen omdat kwaliteit van spraak niet enkel afhangt van snelheid, maar samenhangt met intonatie, volume, spreekstijl...

De resultaten zijn van voldoende hoge kwaliteit (ondanks dat er altijd invloed te merken is) indien de snelheid binnen bepaalde boven en ondergrenzen blijft. In het algemeen geven tragere sprekers betere resultaten dan snelle sprekers.

Use-case 11.4: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van snelle of trage sprekers

Idem als use-case 11.3

Use-case 11.5: Als redacteur van het Woordelijk Verslag ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers

Risicoscore: **5**

Idem als use-case 11.3.

Use-case 11.6: Als verslaggever van de dienst Commissies ondervindt mijn transcriptie applicatie geen last van spreekstijl tussen 134 terugkerende parlementsleden en ministers en de tientallen externe sprekers

Idem als use-case 11.3.

Use-case 12: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik grammaticaal correcte zinnen

Risicoscore: **20-40**

Hier zit men op de grens van betekenis van gesproken taal. Om zaken zoals tijden van werkwoorden (waarin verschillen in uitspraak soms heel klein zijn of niet bestaan) en punctuatie juist te hebben, moet men vaak de betekenis van een zin kennen. Op dat vlak dient nog veel onderzoek te gebeuren.

Use-case 13.1: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik semantisch correcte zinnen

Risicoscore: **40**

Opnieuw bevindt dit probleem zich in het vakgebied van de betekenis van gesproken taal. Zoals al eerder aangegeven staat de technologie hier nog niet ver genoeg. Bijkomende problemen bevinden zich in kleine woorden die niet worden opgevangen (zoals negatie) en de gelijkenissen tussen verschillende werkwoordsvormen.

Use-case 13.2: Als redacteur van het Woordelijk Verslag en verslaggever van de dienst Commissies krijg ik semantisch correcte zinnen met gesuggereerde alternatieven

Risicoscore: **40**

Zie use-case 13.1.

Use-case 14: Als verslaggever van de dienst Commissies krijg ik een logische en gestructureerde samenvatting van de commissievergaderingen

Risicoscore: **40-100**

Hier dient men niet enkel rekening te houden met de betekenis van een zin, maar van een hele tekst of betoog. Dit maakt het probleem nog moeilijker. Vooral het herstructureren van tekst is zo goed als onmogelijk.

Zaken die wel binnen de grenzen van het mogelijke liggen zijn: het weergeven van de belangrijkste punten, namen, het onderwerp en andere 'entiteiten'. Bovendien kunnen belangrijke stukken tekst en woorden met de meeste betekenis worden aangeduid.

De noden en het gebruik hiervan vormen echter een nieuwe verzameling van problemen en oplossingen en worden best buiten de scope van dit project gelaten. Deze use-case behoorde dan ook niet tot de eigenlijke doelstelling van het project.

Use-case 15.1: Als redacteur van het Woordelijk Verslag kan ik een vooraf gegenereerde begeleidende tekst van een toespraak meegeven

Risicoscore: **1**

Het meegeven van een tekst komt de kwaliteit van de transcriptie ten goede. De huidige technologieën brengen de woorden van deze tekst in rekening en weten dat er een verhoogde kans bestaat dat ze voorkomen in een meeting.

Use-case 15.2: Als verslaggever van de dienst Commissies kan ik een vooraf gegenereerde begeleidende tekst van een toespraak meegeven

Idem als use-case 15.1

Use-case 16.1: Als redacteur van het Woordelijk Verslag kan ik rekenen op een systeem waar accuraatheid boven snelheid gaat

Dit onderdeel is niet gescoord. Alle marktspelers kunnen dit garanderen.

Use-case 16.2: Als verslaggever van de dienst Commissies kan ik rekenen op een systeem waar accuraatheid boven snelheid gaat

Idem als use-case 16.1

Use-case 17.1: Als redacteur van het Woordelijk Verslag kan ik een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties

Risicoscore: **1**

Aangezien wordt gewerkt in Word zijn de integratievoorwaarden zo laag dat deze geen enkel probleem vormen.

Use-case 17.2: Als verslaggever van de dienst Commissies kan ik een transcriptie krijgen geïntegreerd in mijn huidige workflow en bestaande applicaties

Zie use-case 17.1.

Use-case 18: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met parlementsleden die elkaar onderbreken

Risicoscore: **40**

De kwaliteit van stukken audio waar overlap is tussen meerdere sprekers gaat snel naar beneden. De analogie met de mens wordt gemaakt, die dit ook niet kan. Technologie kan de mens vooralsnog niet verslaan op dit vlak.

In de huidige 'setup' van microfoons waar er maar één audiolijn aanwezig is, is het risico hoog. Een aanpassing van de faciliteiten naar meerdere audiolijnen kan dit verhelpen. Één marktspeler gaf aan dat een oplossing met een kleine aanpassing aan de faciliteiten dit kan verhelpen, maar kan deze oplossing niet delen. Deze speler schatte het risico in als 0,5.

Use-case 19: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met nodeloze herhalingen... in functie van mijn persoonlijke instelling

Risicoscore: **20-40**

Om het verschil te maken tussen herhalingen en retorische stijlfiguren dient men de betekenis en interpretatie te kunnen vatten. Bovendien zijn er soms situaties waar ook de mens niet kan inschatten of een herhaling bewust is of niet. Het risico hiervan is heel hoog.

Use-case 20: Als redacteur van het Woordelijk Verslag kan ik op het systeem rekenen om op een correcte manier om te gaan met haperingen, stottingen, stopwoorden...

Risicoscore: **1**

Haperingen en stottingen worden automatisch herkend als noise en dus ook verwijderd. De enige uitzonderingen bevinden zich in partiële woorden die herkend worden als andere woorden en niet worden verwijderd.

5.3 Algemene opmerkingen van de markt

Naast de risicoscores van de use-cases wou de markt ook nog enkele algemene opmerkingen meegeven waar het Vlaams Parlement rekening mee kan houden:

- YouTube ondersteunt al ondertiteling in het Nederlands. Indien deze van voldoende hoge kwaliteit is naar inschatting van het Vlaams Parlement kan dit gratis gebruikt worden en dient de markt niet te worden aangesproken. Een snelle steekproef leert echter dat de kwaliteit erg daalt naarmate de snelheid van de spreker toeneemt.
- Het Vlaams Parlement houdt best rekening met de groeiende vraag naar het doorzoeken van audio/videofragmenten. Dit geeft geïnteresseerde kijkers de optie om snel vergaderingen over bepaalde onderwerpen te vinden.
- De historische audio en bijbehorende verslagen kunnen gebruikt worden om semantische modellen te helpen ontwikkelen.
- Het uitschrijven van een RFP dient ook in het Engels te gebeuren indien men internationale spelers wil bereiken.
- Best wordt zo een systeem gradueel ingevoerd, om onder andere kwaliteit en snelheid van transcriptie te doen stijgen. Maar dit zal zich slechts langzaam in productiviteitswinsten vertalen, en zeker niet toelaten om op korte termijn bijvoorbeeld het aantal redacteurs te verminderen.

6 Synthese: opstellen van een RFP

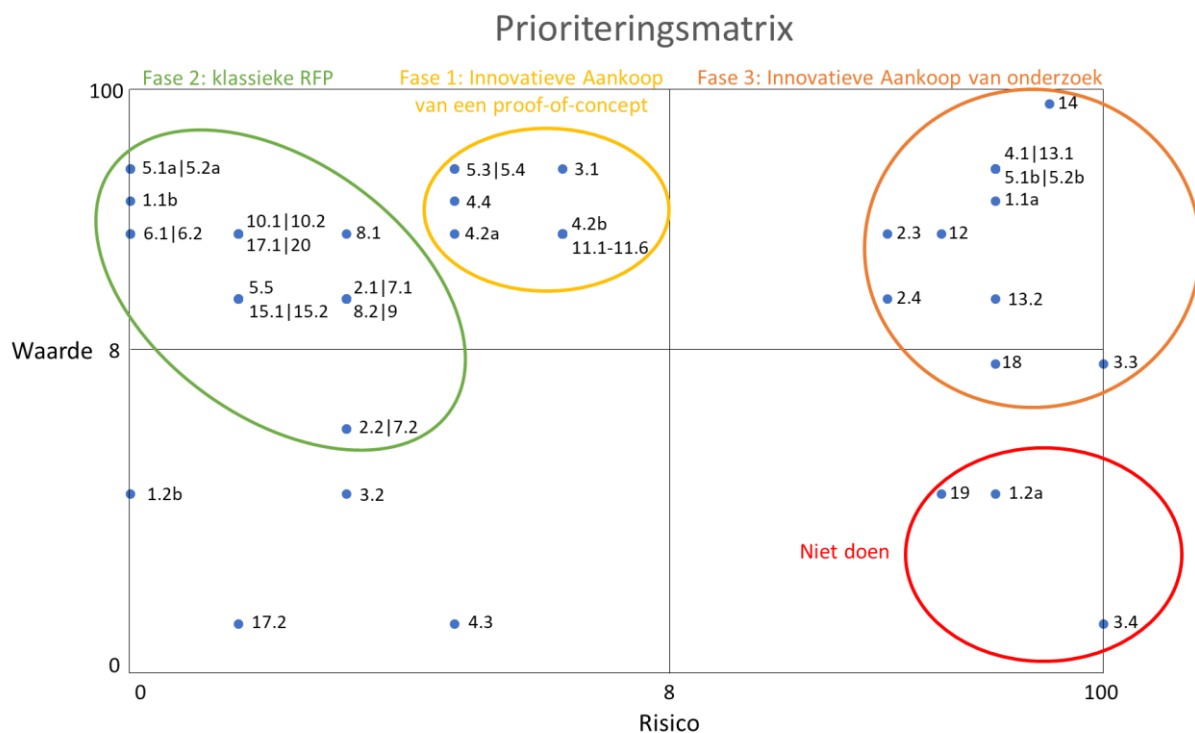
6.1 Fases en doelstellingen van het commercieel traject

Om de vooropgestelde ambities van dit project uit te bouwen zal het nodige budget beschikbaar gesteld moeten worden en zullen er heel wat inspanningen geleverd moeten worden. Tijdens de 1e fase wordt de focus het best gelegd op een precommercieel voortraject. Tijdens dit voortraject wordt met beperkte middelen een relevant prototype opgebouwd om de grootste risicobronnen te reduceren of zelfs volledig weg te werken. Na dit voortraject kan het Vlaams Parlement de gewenste functionaliteiten selecteren en overgaan tot fase 2, waarin een klassieke RFP wordt uitgeschreven en een geïntegreerde oplossing wordt ontwikkeld. Tot slot kan eventueel nog worden overgegaan tot de 3^e fase waarin de innovatieve aankoop van onderzoek wordt gedaan.

Zelfs na een succesvol voortraject is het aan te raden om een geïntegreerde oplossing op te bouwen aan de hand van een iteratieve aanpak. Steeds opnieuw dienen incrementele, haalbare stappen gezet te worden naar het einddoel.

In de volgende sectie worden de verschillende fases uitgelicht en de eraan verbonden use-cases en functionaliteiten worden aangehaald.

6.2 Prioriteiten van de geïntegreerde oplossing



Figuur 2 Prioritering van functionaliteiten

Figuur 2 weerspiegelt de prioritaire uitdagingen van de toepassing in een zgn. prioriteringsmatrix van waarde vs. risico. Op de verticale as bevindt zich de waarde of score van het innovatiepotentieel. De bepaling hiervan is gedaan aan de hand van de use-cases die voorgelegd zijn aan de gebruikers via planning poker zoals omschreven in sectie 4.3. Op de horizontale as bevindt zich het risico of de

risicoscore. De bepaling hiervan is gedaan aan de hand van de use-cases die voorgelegd zijn aan de industriespelers en kennisinstellingen via planning poker zoals omschreven in sectie 5.2.

Binnen deze matrix worden typisch een viertal zones onderscheiden. Linksboven liggen de kritische functionaliteiten van het eindproduct die een hoge toegevoegde waarde hebben voor de eindgebruiker maar waarvan het risico eerder beperkt is. Rechtsboven situeren zich de elementen met een hoge toegevoegde waarde voor de eindgebruiker en met bovendien een aanzienlijk risico. Linksonder liggen de details van het eindproduct waarvan de toegevoegde waarde eerder beperkt is en die daarenboven geen hoog risico met zich meedragen. Het laatste kwadrant bevat de zaken die weinig toegevoegde waarde leveren en toch een hoog risico inhouden. Deze zijn voor het eindproduct op dit moment helemaal niet belangrijk.

De hoogste prioriteiten op de roadmap van een prototype worden typisch gevormd door de elementen met hoge toegevoegde waarde voor de eindgebruiker, en met op relatief korte termijn goed te reduceren risico. Hoe hoger het risico van een facet, hoe moeilijker om dit risico op korte termijn afdoende te reduceren. Afhankelijk van de beschikbare middelen en de spreiding van de elementen op de grid kan een afbakening gemaakt worden van de minimale scope van het prototype.

6.2.1 Fase 1: Innovatieve Aankoop van een proof-of-concept

Tijdens het precommerciële ontwikkelingstraject dient de scope scherp gesteld te worden op maat van de gebruikersgroep om de hoge toegevoegde waarde effectief te realiseren en tegelijkertijd dienen de technologische, implementatietechnische risico's gereduceerd te worden.

In een eerste fase wordt een precommerciële aanbesteding gelanceerd die de focus legt op het realiseren van use-cases die door de huidige technologie ondersteund kunnen worden maar een zekere inspanning verlangen. Na het trainen en ontwikkelen van een prototype kan geëvalueerd worden welke van deze use-cases opgenomen dient te worden in de uiteindelijke RFP. Het gaat dus in essentie om het verlagen van de risico's omtrent deze use-cases.

Het gaat hier om de use-cases met een hoge toegevoegde waarde en een risicoscore tussen 3 en 8. In concreto draait het om het vlot introduceren van nieuwe sprekers en bijbehorende sprekersherkenning, het gebruik maken van een lexicon en het behandelen van verschillende accenten en spreekstijlen van de 124 parlementairen en 10 ministers. Dit alles omvat de volgende use-cases: 3.1, 4.2a, 4.2b, 4.4, 5.3, 5.4, 11.1-11.6.

Uiteraard zal ook use-case 1.1, de transcriptie van audio tot tekst in Word, moeten worden meegenomen in het prototype als centrale functionaliteit.

6.2.2 Fase 2: Klassieke RFP

In de tweede fase wordt overgegaan tot een klassieke RFP, waarin zowel de use-cases met een laag risico als de use-cases uit fase 1 waarvan het risico succesvol verlaagd is, in rekening worden gebracht. Bij aankoop ligt de focus op evalueren, integreren en deployen van de toepassing.

Het gaat hier om de use-cases met hoge toegevoegde waarde en een risicoscore lager dan 3, aangevuld met de succesvol ontwikkelde use-cases uit fase 1. In concreto draait het om een accuraatheidsaanduiding van woorden of zinsdelen, het gebruik maken van redactionele richtlijnen op woordniveau, het linken van tekst met bijbehorend audiofragment, het live aanpassen van de

transcriptie, de robuustheid van het systeem, ondertiteling van video op de website, de invloed van omgevingslawaaï, het meegeven van een vooraf gegenereerde tekst, integratie in de bestaande workflow en het omgaan met haperingen en stottingen. Dit alles omvat de volgende use-cases: 1.1b, 2.1, 2.2, 5.1a, 5.2a, 5.5, 6.1, 6.2, 7.1, 7.2, 8.1, 8.2, 9, 10.1, 10.2, 15.1, 15.2, 17.1, 20.

6.2.3 Fase 3: Innovatieve Aankoop van onderzoek

In een derde en laatste fase kan overgegaan worden tot de aankoop/medewerking aan onderzoek. Het draait hier om het verlagen van het risico op lange termijn. Dit enkel voor de use-cases met een voldoende hoge waarde vanuit gebruikersstandpunt. Zoals al eerder vermeld kan het archief met audio en bijbehorende verslagen hier van grote waarde zijn als trainingsmateriaal voor onderzoeksinstellingen.

De use-cases die hierin omvat zitten bevatten op enkele uitzondering na een component waarin de betekenis van taal een rol speelt. Het draait hier om het live herkennen van sprekers, de rol van de voorzitter herkennen, grammaticaal en semantisch correcte zinnen, gestructureerde samenvattingen en het verwijderen van nodeloze herhalingen.

De volgende use-cases zitten al in fase 1 en 2 vervat, maar kunnen kwalitatief versterkt worden door training en onderzoek. Het gaat om algemene transcriptie van spraak naar tekst en het verwerken van redactionele richtlijnen (hier vooral richtlijnen ivm stijl en herstructurering van zinnen).

Dit alles omvat de volgende use-cases: 1.1a, 2.3, 2.4, 3.3, 4.1, 5.1b, 5.2b, 12, 13.1, 13.2, 14, 18.

7 Voorwaarden tot succes

Best wordt de beoogde spraaktechnologie gradueel ingevoerd, om onder andere kwaliteit en snelheid van transcriptie te doen stijgen. Bovendien geeft dit het Vlaams Parlement de mogelijkheid te bekijken welke features een absolute meerwaarde bieden en ze allemaal te evalueren zonder al te hoge kosten op te lopen. Ten tweede geeft dit het Vlaams Parlement de tijd om het menselijke vertrouwen in de technologie op te bouwen. Zo kan de huidige organisatie gradueel mee evolueren met de technologie.

Deze aanpak zorgt ervoor dat de invoering van nieuwe technologie zich langzaam in productiviteitswinsten zal vertalen, en zeker niet toelaat om op korte termijn bijvoorbeeld het aantal redacteuren te verminderen.

Bijlage A: Technische omschrijving digitale opnames Vlaams Parlement

Analoge audio naar glasvezel omvormer : DHD

De basis van het systeem is de conversie van de analoge signalen die afkomstig zijn van de verschillende zalen in het Vlaams Parlement naar een digitaal audio signaal. Dit digitaal audio signaal is een optisch MADI signaal dat later wordt opgenomen op verschillende servers (loggers). De vermelde conversie vindt plaats op twee verschillende plaatsen in het Vlaams parlement.

- Server lokaal Vlaams Parlement : Hier komen alle publieke audio lijnen toe van de verschillende zalen
- Technisch (server)lokaal VP+3: Hier komen alle geheime audio lijnen toe van de verschillende zalen

Server Lokaal VP-2 - Publieke omvormer :

De publieke omvormer is van het merk DHD en converteert 16 analoge audio signalen via XLR aansluiting naar twee parallel geschakelde optische MADI signalen. Het eerste MADI signaal gaat naar Logger 1 en het tweede MADI signaal gaat naar Logger 2. Op deze manier is er steeds een dubbele en volledig gescheiden opname van dezelfde audio signalen. Hieronder een lijst van de ingangen en de desbetreffende signalen die momenteel zijn verbonden:

- Audio Ingang 1 : Lokaal 301 - Pieter Bruegel
- Audio Ingang 2 : Lokaal 322 - Jan Van Eyck
- Audio Ingang 3 : Lokaal 215 - Pieter Paul Rubens
- Audio Ingang 4 : Lokaal 201 - Antoon Van Dijck
- Audio Ingang 5 : Lokaal 206 - Jeroen Bosch
- Audio Ingang 6 : Lokaal 207 - James Ensor
- Audio Ingang 7 : Lokaal 212 - Valerius De Saedeleer
- Audio Ingang 8 : Lokaal 213 - Constant Permeke
- Audio Ingang 9 : Lokaal 219 - Quinten Metsijs
- Audio Ingang 10 : Lokaal 220 - Hans Memling
- Audio Ingang 11 : Lokaal 050 - Koepel 1
- Audio Ingang 12 : Lokaal 322 - Jan Van Eyck (Vertaling)
- Audio Ingang 13 : Lokaal 547 - De Schelp
- Technisch (server)lokaal VP+3-

Geheime omvormer :

De omvormer is van het merk DHD en converteert analoge audio signalen via XLR aansluiting naar twee parallel geschakelde optische MADI signalen. Het eerste MADI signaal gaat naar Logger 1 en het tweede MADI signaal gaat naar Logger 2. Op deze manier is er steeds een dubbele en volledig gescheiden opname van dezelfde audio signalen. Hieronder een lijst van de ingangen en de desbetreffende signalen die momenteel zijn verbonden :

- Audio Ingang 1 : Pieter Bruegel (Geheim)
- Audio Ingang 2 : Jan Van Eyck (Geheim)
- Audio Ingang 3 : Pieter Paul Rubens (Geheim)
- Audio Ingang 4 : Antoon Van Dijck (Geheim)

- Audio Ingang 5 : Jeroen Bosch (Geheim)
- Audio Ingang 6 : James Ensor (Geheim)
- Audio Ingang 7 : Valerius De Saedeleer (Geheim)
- Audio Ingang 8 : Constant Permeke (Geheim)
- Audio Ingang 9 : Quinten Metsijs (Geheim)
- Audio Ingang 10 : Hans Memling (Geheim)
- Audio Ingang 11 : Koepel 2

Opsomming inclusief serienummers gebruikte hardware:

Artikel	Benaming
DHD-05RM5200D-5063	Stagebox Frame
DHD-0RM5200D-5047	Power Supply Module 48V / 100W
DHD-05RM5200D-5422	MB/CR Controller, redundant
DHD-05RM5200D-5210	MB/CR/XD 4 Ch. Line In, 18dBu
DHD-05RM5200D-5063	Stagebox Frame 3U/19
DHD-0RM5200D-5047	Power Supply Module 48V / 100W
DHD-05RM5200D-5422	MB/CR Controller, redundant
DHD-05RM5200D-5210	MB/CR/XD 4 Ch. Line In, 18dBu
DHD-0RM5200D-5047	Power Supply Module 48V / 100W
DHD-05RM5200D-5422	MB/CR Controller, redundant
DHD-05RM5200D-5210	MB/CR/XD 4 Ch. Line In, 18dBu

Loggers

Het doel van de loggers is het opnemen van alle audio signalen afkomstig van beide omvormers. De twee loggers (HP DL380G6) zijn telkens voorzien van twee optische audio kaarten van het merk RME. Deze audio kaarten zullen rechtstreeks het optische MADI signaal opnemen dat afkomstig is van de DHD omvormers. Elke logger heeft een optisch MADI signaal dat afkomstig is van de publieke omvormer en een optisch MADI signaal dat afkomstig is van de geheime omvormer. Op deze wijze is een permanente opname gegarandeerd van zowel de publieke als geheime signalen op beide loggers.

De audio opname software is het logdepot systeem van Arbor en is ingesteld om de opnames te bewaren met compressiemethode WMA (Windows media audio 9.2) aan 48kbit/s mono (1 kanaal per zaal), 44kHz, 1-pass CBR.

Audio componenten loggers :

Eén optische MADI-kaart merk RME : voor de publieke kanalen

Eén optische MADI-kaart merk RME : voor de geheime kanalen

Gebruikte optische MADI kaarten, model: RME-HDSPMADI-E HDSP MADI

Fetchers + storage

Elke fetcher bestaat uit een HP-server met daaraan verbonden een storage kabinet Het doel van de fetchers is het opslaan van de opgenomen audio bestanden op langere termijn, in ons geval is dit momenteel 6 maanden.

Elk storage kabinet is geconfigureerd in RAID6. Hierdoor kunnen tot 2 harde schijven tegelijkertijd falen zonder een verlies aan data en audio opnames. Tevens door het ont dubbeln van het hele systeem zijn er twee storage cabinets die exacte dezelfde opnames bevatten waardoor een kans op verlies aan data en audio opnames praktisch onbestaande wordt.

Onderdelen gebruikt voor de storage:

Artikel	Benaming
PS00031305	DL360 E5420 2,5GHZ 1P 2GB SA-P400 BBWC
P300031305	700W REDUNDANT POWER SUPPLY
P300031305	146GB 10K RPM SFF SAS HOT PLUG HDD 2,5"
PS00031305	MSA60 SAS
PS00031305	500GB 7.2K rpm HP-SATA HDD MDL

Bijlage B: Redactieapplicatie Woordelijk Verslag

De redactieapplicatie van het Woordelijk Verslag ('Handelingen van vergaderingen') is webgebaseerd en beschikbaar op de Kiosk-pagina van het intranet van het Vlaams Parlement. Ze wordt zowel in de vergaderzaal gebruikt, om de gegevens van de vergadering te registreren, als in het redactielokaal, om het verslag te redigeren.



Bij het inloggen in de applicatie verschijnt het startvenster met een overzicht van vergaderingen. De gebruiker heeft de keuze tussen plenaire en commissievergaderingen. Standaard verschijnt een scherm met een overzicht van alle commissievergaderingen van de week.

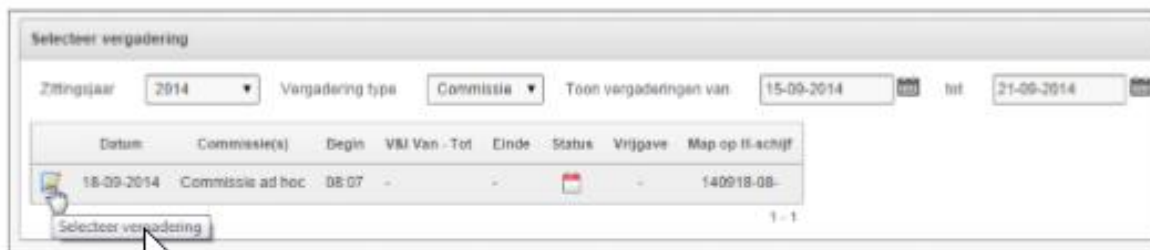
Week	14-20-2017	21-27-2017	28-03-2017	04-10-2017	11-17-2017	18-24-2017	25-31-2017
14-05-2017	Commissie voor Landbouw, Natuur en Milieu	14:00	-	-	-	-	17:00-18:00 uur
15-05-2017	Commissie voor Cultuur, Jeugd, Sport en Media	14:00	-	-	-	-	17:00-18:00 uur
16-05-2017	Commissie voor Economie, Werk, Sociale Economie, Innovatie en Antidiscriminatie	14:00	-	-	-	-	17:00-18:00 uur
17-05-2017	Commissie voor Onderwijs	14:00	-	-	-	-	17:00-18:00 uur
18-05-2017	Commissie voor Mobiliteit en Openbare Werken	13:30	-	-	-	-	17:00-18:00 uur
19-05-2017	Commissie voor Onderwijs	13:00	-	-	-	-	17:00-18:00 uur
20-05-2017	Commissie voor de opvolging van het Algemeenbeleid in Vlaanderen	13:00	-	-	-	-	17:00-18:00 uur
21-05-2017	Commissie voor Cultuur, Jeugd, Sport en Media	13:00	-	-	-	-	17:00-18:00 uur
22-05-2017	Commissie voor Economie, Werk, Sociale Economie, Innovatie en Antidiscriminatie	13:00	-	-	-	-	17:00-18:00 uur
23-05-2017	Commissie voor Wetgeving, Antidiscriminatie en Gender Kwestie	13:00	-	-	-	-	17:00-18:00 uur
24-05-2017	Commissie voor Landbouw, Natuur en Milieu	13:15	-	-	-	-	17:00-18:00 uur
25-05-2017	Commissie voor Streekl en de Vlaamse Rand	13:30	-	-	-	-	17:00-18:00 uur
26-05-2017	Commissie voor Landbouw, Natuur, Ruimtelijke Ordening, Energie en Omgeving	14:00	-	-	-	-	17:00-18:00 uur
27-05-2017	Commissie voor Wetgeving, Antidiscriminatie en Gender	14:00	-	-	-	-	17:00-18:00 uur

De commissies waarin er die week vragen om uitleg en/of interpellaties geagendeerd staan en waarvan er dus een Woordelijk Verslag wordt gemaakt, krijgen een volgnummer toegekend (dat bepaald wordt door de datum, het beginuur en de alfabetische volgorde van de diverse commissienamen).

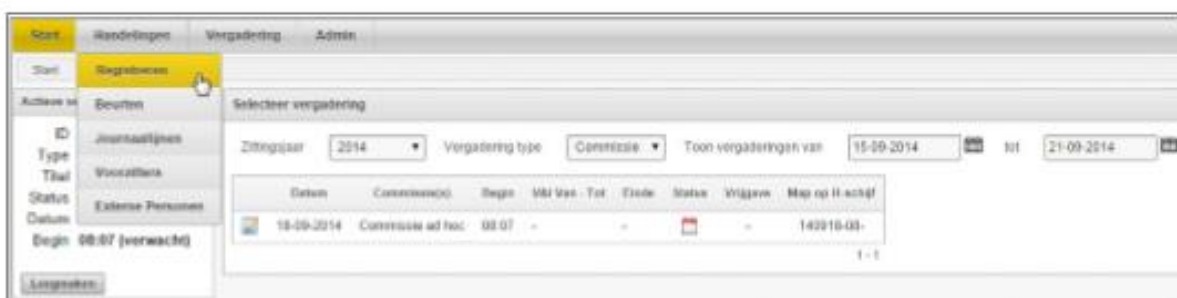
Registratie

In de vergaderzaal wordt er geregistreerd: **aanvullen**

Eerst selecteert de (eind)redacteur een vergadering door op het -icoontje te klikken.



Vervolgens opent hij de registratiemodule door in het menu *Handelingen* op *Registreren* te klikken.



Er verschijnt dan een 'tussenscherm', waarop een aantal gegevens van de betreffende vergadering zichtbaar worden: datum, beginuur, de vaste (commissie)voorzitter en vergaderzaal.



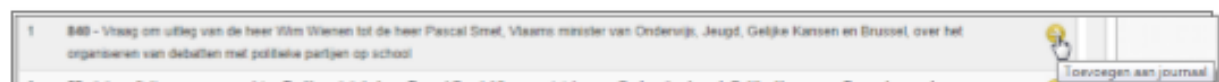

Als op dat moment al duidelijk is dat een andere persoon dan de vaste (commissie)voorzitter als voorzitter van de vergadering zal optreden, kan de (eind)redacteur dat al meteen invoeren, door op de knop **Wijzigen** onder de foto van de voorzitter te klikken en de juiste persoon te selecteren.

Om een eerste redactiebeurt voor een plenaire vergadering aan te maken, vult de eindredacteur de initialen van de eerste redacteur in. Vervolgens klikt hij op de knop **Vergadering Openen**. De eerste beurt start.

Om een eerste redactiebeurt voor een commissievergadering aan te maken, klikt de redacteur op de knop **V&S Sessie Aanvangen**. De eerste beurt start automatisch.

Vervolgens klikt de (eind)redacteur het agendapunt aan dat wordt behandeld.

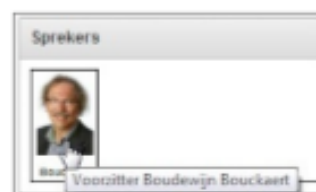
Als een agendapunt slechts uit één item bestaat, klikt hij op het pijltje ➡ achteraan om dat item toe te voegen aan het verslag.



Als een agendapunt uit meerdere samengevoegde items (meestal vragen om uitleg en en/of interpellaties of actuele vragen bestaat, klikt hij op het kettinkje 🔗 vooraan om ze toe te voegen aan het verslag.

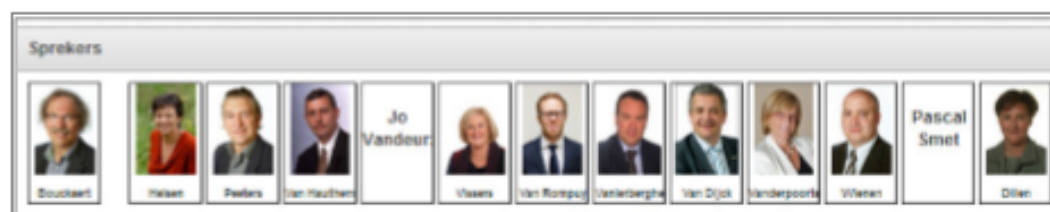


Telkens wanneer de (eind)redacteur een agendapunt toevoegt in de applicatie, wordt meteen ook automatisch de voorzitter als eerste spreker toegevoegd. Om hem later opnieuw als spreker toe te voegen, klikt de redacteur op de foto van de voorzitter, links op het scherm.



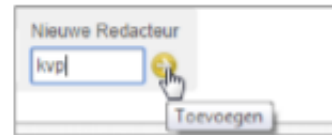
Om andere sprekers toe te voegen, klikt hij op de knop **Nieuwe Spreker** en selecteer hij de juiste persoon.

De twaalf meest recente sprekers komen met hun foto in een rijtje te staan naast de foto van de voorzitter. Die sprekers kunnen ook opnieuw worden ingevoerd door gewoon op hun foto te klikken.



Opm. Wanneer de persoon die als voorzitter optreedt, in eigen naam spreekt, wordt hij op dat moment als gewone spreker onder zijn eigen naam ingevoerd.

Om een nieuwe beurt aan te maken, vult de (eindredacteur voor) de volgende redacteur in het vak *Nieuwe Redacteur* de juiste redacteursinitialen in en klikt hij op het pijltje erachter.

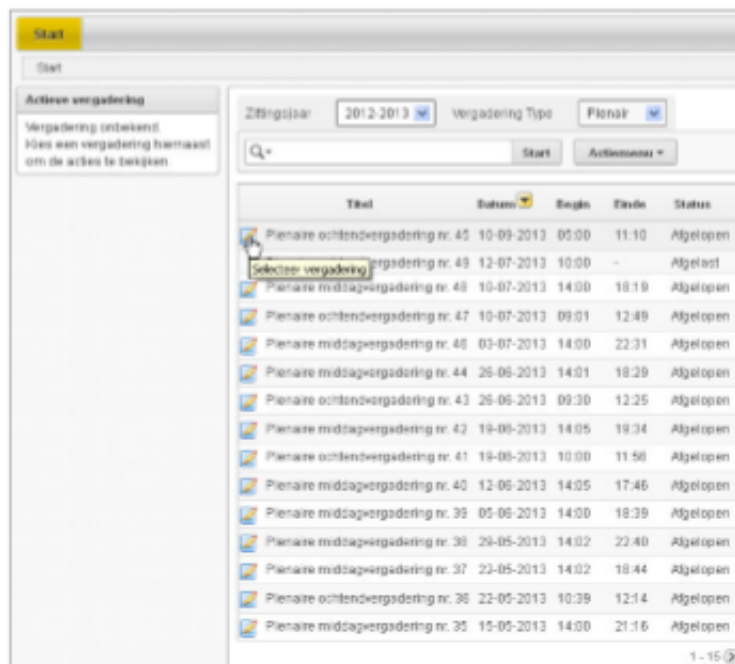


Om te stoppen met het registreren van een plenaire vergadering, klikt hij op de knop **Vergadering Beëindigen** en daarna op **Ok**. Om te stoppen met het registreren van een V&I-sessie in een commissievergadering, klikt hij op de knop **V&I Sessie Beëindigen** en daarna op **Ok**. Opmerking: Soms gebeurt het dat er binnen een vergadering twee of meer V&I-sessies zijn, bv. wanneer er een schorsing is of wanneer er andere agendapunten behandeld worden tussen twee V&I-sessies in.

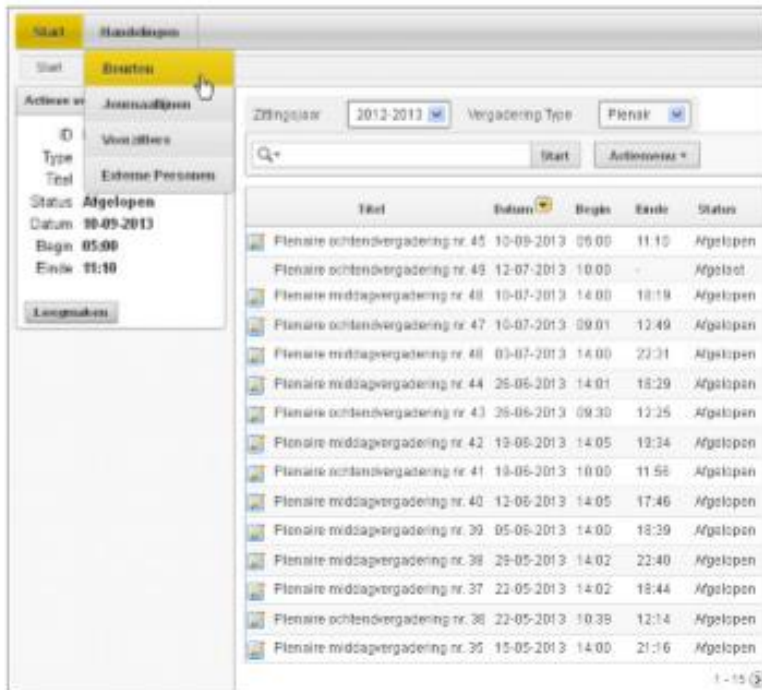
Redactie

Op basis van het stramien dat in de vergaderzaal al is aangemaakt, wordt in het redactielokaal het eigenlijke verslag geredigeerd en gereviseerd. De diverse redacteurs herbeluisteren het digitale audiobestand van de 'beurt(en)' die ze voor zichzelf in de vergaderzaal hebben aangemaakt en zetten, per onderwerp en per spreker, het gesproken woord om in een zo woordelijk mogelijk geredigeerd verslag.

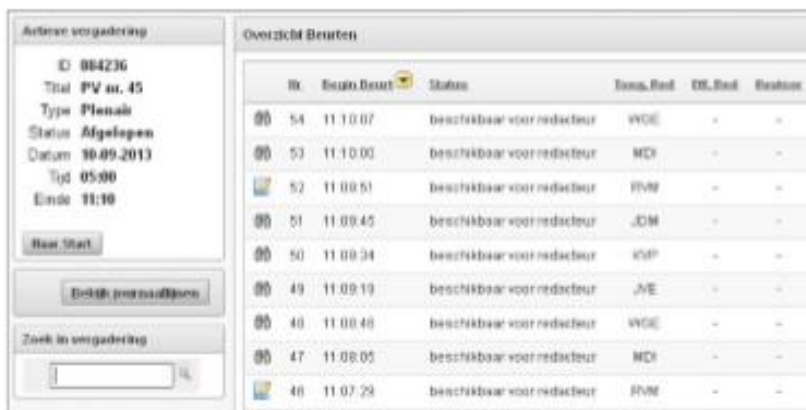
Eerst selecteert de redacteur een vergadering door op het -icoontje te klikken.



Vervolgens opent hij de redactiemodule door in het menu *Handelingen op Beurten* te klikken.

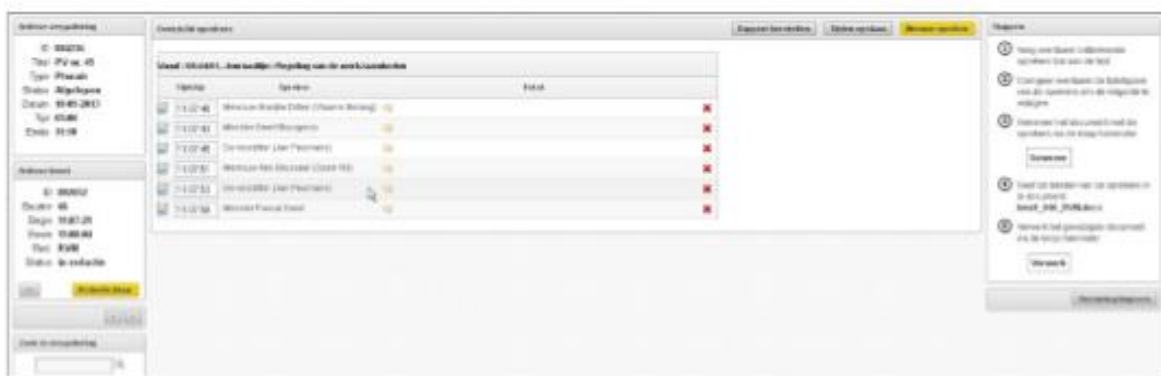


Er verschijnt dan een venster met een overzicht van de beurten van de desbetreffende vergadering.



Om een beurt te selecteren klikt de redacteur op het -icoontje (*Beurt bewerken*).

Er verschijnt een venster met alle gegevens van de desbetreffende beurt.



Links op het scherm, onder het gegevensblok van de actieve vergadering, staat nu ook een gegevensblok van de actieve beurt. In dat gegevensblok staan o.a. het beurtnummer, de redacteursinitialen en de status van de beurt vermeld. Verder staan in dat gegevensblok ook nog een knop **Redactie klaar** om de redactie van de beurt af te sluiten en een knop  om terug te keren naar het venster met het overzicht van de beurten.

Centraal op het scherm staat het overzicht van de sprekers van die beurt. Achter de namen van de sprekers staan (lege) tekstballonnetjes . Nadat een beurt volledig is geredigeerd en verwerkt, komen op die plaats de teksten van de sprekers te staan.

In het overzicht van de sprekers kan de redacteur sprekerstijden aanpassen, sprekers verwijderen, nieuwe sprekers toevoegen en bestaande sprekers door een andere spreker vervangen.

Het stappenblok rechts op het scherm laat de redacteur toe om, met een druk op de knop **Genereren**, vanuit de applicatie een Word-document te genereren. Er wordt een Word-document geopend met daarin reeds de titels van de journaallijnen en de namen van de sprekers (in beveiligde secties). In dat document wordt de eigenlijke tekst van het verslag geredigeerd. Wanneer de redacteur daarmee klaar is, slaat hij het Word-document op, sluit hij het af en gaat hij terug naar zijn beurt in de applicatie om het Word-document te verwerken, door op de knop **Verwerken** te drukken.

In de applicatie worden de (lege) tekstballonnetjes  achter de namen van de sprekers nu vervangen door de teksten van de sprekers.

Overzicht sprekers Rapport herstellen Tijden opmaken Beurten spreken

Vaard | 890481, Jozefia | Regeling van de werkzaamheden

Tijdstip	Spreker	Tekst
11:07:40	Mevrouw Marijke Dillen (Maatsch. Belang)	Over de looflijken viel ook een theorie schiet af wat heeft een beurt. We zijn erom overtuigd dat we in een situatie van zelfstandigheid, of dat we materialen of instrumenten, geboden of bevestigd zijn, de meest belangrijke plaats hebben om de samenwerking die we doorlopen van vorige generatie, te herkennen en te begrijpen te herkennen voor het toekomst.
11:07:43	Minister Geert Bourgeois	Doe zijn we het dat schiet een een.
11:07:45	De voorzitter (Jan Peumans)	Mevrouw Bourgeois heeft het woord.
11:07:51	Mevrouw Ann Brussel (Open Vld)	Ik vind het wel belangrijk dat we ons bezig houden met.

Actieve beurt

ID 892052

Beurt nr 46

Begin 11:07:29

Einde 11:08:04

Red. RVM

Status in redactie

 **Redactie klaar**

Om de redactie van een beurt af te sluiten en de beurt beschikbaar te maken voor revisie, klikt de redacteur links op het scherm, in het gegevensblok van de actieve beurt, op de knop **Redactie klaar**. Nu kan hij een nieuwe beurt selecteren om te redigeren.

Met het oog op de revisie van de beurten die door de redacteur afgewerkt zijn, maakt ook de eindredacteur gebruik van de redactiemodule in de applicatie, door in het menu *Handelingen op Beurten* te klikken.

In het venster met een overzicht van de beurten van de desbetreffende vergadering ziet de eindredacteur welke beurten beschikbaar zijn voor revisie.

Vlaams Parlement

Redactie Handelingen

Commissievergadering

Start Handelingen Vergadering Bestanden Admin

Actieve vergadering

ID 1137488
Titel CV nr. 248 - WON
Type Commissie
Status Bezig
Datum 18-05-2017
Tijd 09:56
Zaal Hans Memling
H-map 170518-10-won

Naar Start

Bekijk journaallijnen

Zoek in vergadering

Overzicht Beurten

Nr	Begin Beurt	Status	Toeg. Red.	Eff. Red.	Revisor	Beurt bevat
00 12	11:05:40	beschikbaar voor redacteur	BWA			8 spreker(s)
00 11	11:00:43	beschikbaar voor redacteur	BWA			3 spreker(s)
00 10	10:55:46	beschikbaar voor redacteur	KAWAL			12 spreker(s)
00 9	10:48:36	in redactie	KAWAL	KAWAL		6 spreker(s)
00 8	10:43:40	beschikbaar voor redacteur	RVM			9 spreker(s)
00 7	10:38:44	beschikbaar voor redacteur	RVM			3 spreker(s)
00 6	10:33:48	beschikbaar voor redacteur	LDY			3 spreker(s)
00 5	10:28:49	beschikbaar voor redacteur	LDY			6 spreker(s)
00 4	10:21:48	in redactie	BWA	BWA		9 spreker(s)
00 3	10:14:48	beschikbaar voor revisor	KAWAL	KAWAL		5 spreker(s)
Beurt bekijken 07		in redactie	RVM	RVM		3 spreker(s)
00 1	10:00:56	in redactie	LDY	LDY		4 spreker(s)

1 - 12

Om een beurt te selecteren klikt de eindredacteur op het -icoontje (*Beurt bekijken*).

Ook nu weer verschijnt een venster met alle gegevens van de desbetreffende beurt, én de door de redacteur uitgeschreven tekst.

Vlaams Parlement

Redactie Handelingen

Commissievergadering

Actieve vergadering

ID 1137488
Titel CV nr. 248 - WON
Type Commissie
Status Bezig
Datum 18-05-2017
Tijd 09:56
Zaal Hans Memling
H-map 170518-10-won

Naar Start

Bekijk journaallijnen

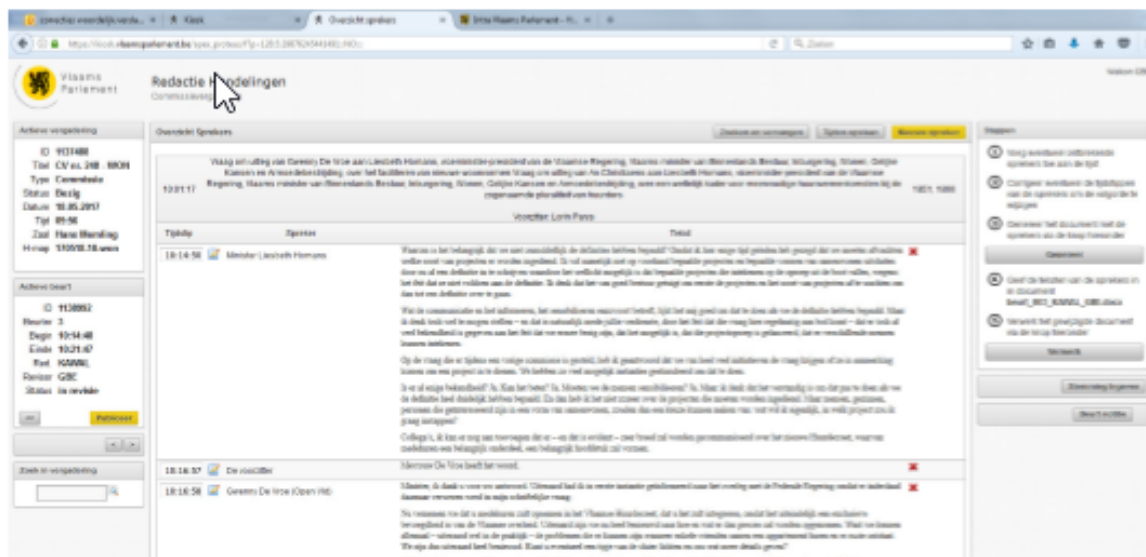
Zoek in vergadering

Overzicht Beurten

Nr	Begin Beurt	Status	Toeg. Red.	Eff. Red.	Revisor	Beurt bevat
00 12	11:05:40	beschikbaar voor redacteur	BWA			8 spreker(s)
00 11	11:00:43	beschikbaar voor redacteur	BWA			3 spreker(s)
00 10	10:55:46	beschikbaar voor redacteur	KAWAL			12 spreker(s)
00 9	10:48:36	in redactie	KAWAL	KAWAL		6 spreker(s)
00 8	10:43:40	beschikbaar voor redacteur	RVM			9 spreker(s)
00 7	10:38:44	beschikbaar voor redacteur	RVM			3 spreker(s)
00 6	10:33:48	beschikbaar voor redacteur	LDY			3 spreker(s)
00 5	10:28:49	beschikbaar voor redacteur	LDY			6 spreker(s)
00 4	10:21:48	in redactie	BWA	BWA		9 spreker(s)
00 3	10:14:48	beschikbaar voor revisor	KAWAL	KAWAL		5 spreker(s)
Beurt bekijken 07		in redactie	RVM	RVM		3 spreker(s)
00 1	10:00:56	in redactie	LDY	LDY		4 spreker(s)

1 - 12

De eindredacteur klikt vervolgens op de knop **Begin revisie**



Links op het scherm, onder het gegevensblok van de actieve vergadering, staat nu ook een gegevensblok van de actieve beurt. In dat gegevensblok staan o.a. het beurtnummer, de redacteursinitialen, de initialen van de eindredacteur en de status van de beurt vermeld. Verder staan in dat gegevensblok ook nog een knop **Publiceer** om de eindredactie van de beurt af te sluiten en een knop **Terug** om terug te keren naar het venster met het overzicht van de beurten.

Centraal op het scherm staat het overzicht van de sprekers van die beurt. Achter de namen van de sprekers staat nu de tekst ingevuld.

In het overzicht van de sprekers kan de eindredacteur sprekerstijden aanpassen, sprekers verwijderen, nieuwe sprekers toevoegen en bestaande sprekers door een andere spreker vervangen. Bovendien kan hij journaallijnen toevoegen en wijzigen.

Op dezelfde manier als de redacteur kan de eindredacteur opnieuw een Word-document genereren. In document kan hij de betreffende beurt reviseren. Wanneer hij daarmee klaar is, slaat hij het gewijzigde Word-document op, sluit hij het af en gaat hij terug naar de beurt in de applicatie om het Word-document te verwerken, door op de knop **Verwerk** te drukken.

Om de revisie van een beurt af te sluiten en de beurt beschikbaar te maken op de website, klikt de eindredacteur links op het scherm, in het gegevensblok van de actieve beurt, op de knop **Publiceer**.

In het beurtenoverzicht van de betreffende vergadering krijgt de beurt nu de status 'beschikbaar voor publicatie'.

https://kiosk.vlaamsparlement.be/apex_primeus/f?p=328:2:29876245441491::NO::5::&cs=3:ld9dD4bUAEchngRtHTudyKCE&

Vlaams Parlement
Redactie Handelingen
Commissievergadering

Start Handelingen Vergadering Bestanden Admin

Actieve vergadering

ID 1137488
Titel CV nr. 248 - WON
Type Commissie
Status Afgelopen
Datum 18-05-2017
Tijd 09:56
Einde 11:45
Zaal Hans Memling
H-map 170518-10-won

Naar Start

Bekijk journaallijsten

Zoek in vergadering

Overzicht Beurten

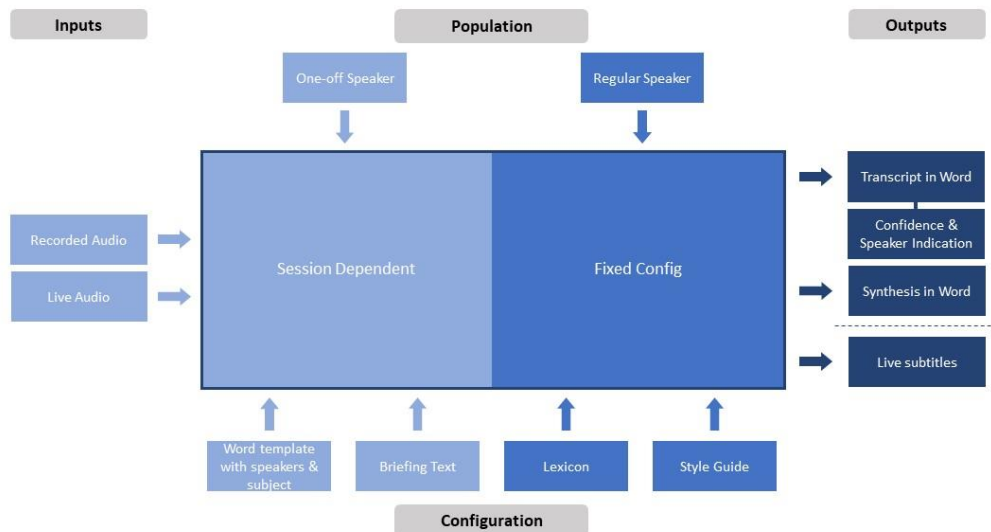
Nr	Begin Beurt	Status	Toeg. Red.	Ent. Red.	Revisor	Beurt bevat
00 12	11:05:40	beschikbaar voor redacteur	BWA			8 spreker(s)
00 11	11:00:43	in redactie	BWA	BWA		3 spreker(s)
00 10	10:55:46	beschikbaar voor redacteur	KAWAL			12 spreker(s)
00 9	10:48:36	in redactie	KAWAL	KAWAL		6 spreker(s)
00 8	10:43:40	beschikbaar voor redacteur	RVM			9 spreker(s)
00 7	10:38:44	in redactie	RVM	RVM		3 spreker(s)
00 6	10:33:48	beschikbaar voor redacteur	LDY			3 spreker(s)
00 5	10:28:49	in redactie	LDY	LDY		6 spreker(s)
00 4	10:21:48	beschikbaar voor revisor	BWA	BWA		9 spreker(s)
00 3	10:14:48	beschikbaar voor publicatie	KAWAL	KAWAL	GBE	5 spreker(s)
00 2	10:07:57	in revisie	RVM	RVM	BBO	3 spreker(s)
00 1	10:00:56	beschikbaar voor publicatie	LDY	LDY	BBO	4 spreker(s)

1 - 12

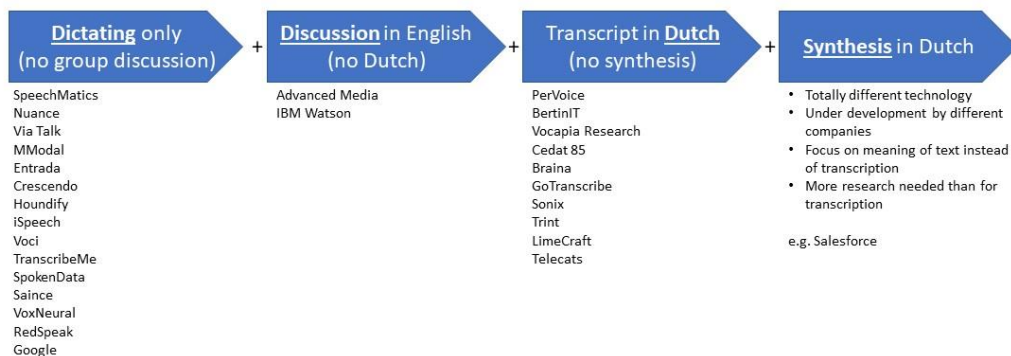
Op dat moment verschijnt de tekst van de betreffende beurt op de website van het Vlaams Parlement.

Bijlage C: State of the Art analyse

There are 2 types of outputs: the transcriptions whose accuracy is vital and the subtitles which need to be generated live



Most applications suited to transcribe group conversations support transcription of Dutch audio





Country: Italy
Email: info@pervoice.it
Telephone: +39 0461383960
URL: <http://www.pervoice.com/en/>

PerVoice provides enabling technologies and services for **reporting activities** for creating the official transcripts in Law Courts, Assemblies or Conferences. PerVoice solutions include:

- **Flyscribe**, an ASR cloud service that lets professionals create text files from recorded speech. Today, Flyscribe is used by a wide variety of clients working in the transcription of speech both for public administration and business uses.

- **PASS** (Pervoice AudioText Synchro Suite), a suite of software tools developed around Audioma™ Box or Flyscribe that lets you manage the entire transcription workflow. Users correct transcriptions in Microsoft Word® while listening to the synchronized audio file. The user can then store the synchronized audio and text permanently in PDF format.

PerVoice prepares language models for every language it transcribes, including pronunciation and spelling variations for languages like English. PerVoice provides speaker independent language models, which means that users do not need to "train" the system. PerVoice can also prepare language models for specific customer environments to help recognize professional jargon and technical words. The PerVoice Speech-to-Text engine can transcribe virtually any language.

Great speech-to-text results

- Accurately recognize up to 95% of spoken words in optimal conditions.

Freedom to choose

- Choose the languages you want from the many existing ones.
- Work with us to create new language models.
- The PerVoice Speech-To-Text engine uses multicore parallel processing to make the recognition process fast and efficient.
- Speaker independence means there is no need to train the system.
- Add new words on-the-fly.

Get a bigger return from your audio data

- Analyze what customers say about you. (Speech Analytics)
- Find who is talking about you on TV. (Broadcast Monitoring)
- Make Audio/Visual archives easy to consult. (Media Indexing)

Speed up your transcription process.

- Get automatic drafts of speech. (Reporting)
- Create subtitles and closed captions efficiently. (Subtitling)



Country: France
Email: contact form on site
Telephone: +33 (0)1 39 30 60 00
URL: <https://www.bertin-it.com/en/vecsyt/>

Expert in Vocal Technologies, Vecsys offers software and services dedicated to the enhancement of multilingual speech content in audio and video sources and telephone interactions.

- Speech-to-Text Transcription & Synchronization
- Word-level time-coding
- Speaker segmentation
- Several "speed/accuracy" levels offered as options (from "fastest" to "most accurate")
- Speed up to 0.4 to 0.25xRT, i.e a 20 minutes program transcribed in 5 minutes
- Multiple languages covered: English, French, Spanish, Arabic, Russian, Chinese, ...

TRANSFORM SPEECH INTO TEXT AND MAKE YOUR RECORDING VALUABLE

- Full-text indexing
- Enhanced searchability
- Powerful audio-mining

EXPLORE AUDIO & VIDEO SOURCES AND IMPROVE YOUR BUSINESS

- Broadcast news & Web TV
- Conferences & Meetings
- Telecommunications

Services: corpus collection, models training and performance monitoring

Tools: Adaptation of the lexicon and language model, MediaSpeech Management Platform



Country: France
Email: sales@vocapia.com
Telephone: /
URL: <http://www.vocapia.com/>

Vocapia Research develops leading-edge, multilingual speech processing technologies. These technologies enable large vocabulary continuous speech recognition, automatic audio segmentation, language identification, speaker diarization and audio-text synchronization. Vocapia's VoxSigma™ speech-to-text software suite delivers state-of-the-art performance in many languages for a variety of audio data types, including broadcast data, parliamentary hearings and conversational data.

We provide solutions and expertise for core speech processing technologies in many languages. For example, speech to text transcription is available for the Arabic, Dutch, English, Finnish, French, German, Greek, Italian, Latvian, Lithuanian, Mandarin, Polish, Portuguese, Romanian, Russian, Spanish and Turkish languages, with several others under development. Our language identification module identifies the spoken language from a set of 50 languages, and clients can create models for their desired language set. We also work with our clients to adapt, tune or create specific models or systems tailored to their application needs.

Automatic on-the-fly adaptation allows the user to provide texts related to the audio document being processed, what can be considered topic/domain adaptation. These accompanying texts serve to increase the lexical coverage of the speech-to-text system and to adapt the language model to the specific domain of the audio document with the aim of improving the transcription accuracy.



Country: Italy
Email: cedat85@cedat85.com
Telephone: /
URL: <http://www.cedat85.com/en/>

We are not just a technology company, nor just a service provider. We are an integrated award-winning technology company and service provider specialized in developing and implementing the world's most powerful and customizable speech-to-text technology solutions for the private and public sector. Our proprietary +VOCE engine represents the most advanced, customizable, and complete system for the real-time voice-independent transcription and indexing of any multimedia content containing speech.

Cedat 85 has revolutionized the transcription, archival, and indexing process of content from political meetings and events.

What is it?

DIGITAL4DEMOCRACY satisfies the ever-growing demand from citizens for political transparency, going far beyond simple streaming or recording of government meetings and representing a leap forward in the transparency and accessibility capabilities of governments.

DIGITAL4DEMOCRACY is a highly powerful, advanced institutional communication solution developed by Cedat 85 Group that facilitates open government and real digital democracy through the automatic, real-time transcription and simultaneous archiving of government speeches and meetings, making them accessible via easily indexed and "searchable" digital content (accessible live as video is being recorded), and making it easier for citizens to take an active part in democratic life.

DIGITAL4DEMOCRACY reduces the amount of time needed for drafting minutes from public meetings by transcribing them in real time and synchronizing audio and text, ensuring that all content is immediately catalogued so users can run instant, intuitive searches, just like online search engines, including directly on public administration websites.

DIGITAL4DEMOCRACY in Action

DIGITAL4DEMOCRACY has been used by national and local governments across Italy—including the Italian Chamber of Deputies and more than 200 Italian cities—to improve their efficiency and transparency, as well as improving citizens trust in government.

Cedat 85's software solution avoids the need for repeat dictation, directly transcribing the voice of each member of parliament ("voice independent" technology) and starting a new paragraph every time the speaker changes. This halved the time needed for transcription, and with 95% accuracy, meaning members of staff now only intervene to correct potential transcription errors caused by noise interference or speech impediments.

Cedat 85's technology also immediately indexes audio and video content, word by word, so that citizens can quickly find any clip in the archive by running a simple Google-like search.

Braina

Country: India
Email: [contact form on site](#)
Telephone: /
URL: <https://www.brainasoft.com/braina/speech-to-text.html>

Braina Pro is the world's best speech recognition program that allows you to easily and accurately dictate (speech to text) in over 100 languages of the world, update social network status, play songs & videos, search the web, open programs & websites, find information and much more. You can use your voice to dictate text to your Windows computer, automate processes and improve your personal and business productivity. For example, you can dictate text to fill out online forms; or you can dictate text to a word-processing program, such as Microsoft Word quickly and accurately. Apart from converting live microphone input to text, Braina can also transcribe recorded audio files to text.

Supports microphones built into many laptops. You can even use Braina for Android app and turn your Android device into an external wireless microphone over a WIFI network. Using Android app, you can speak into your phone or tablet to dictate text to your PC!

Braina Pro works out of the box without the need of any training. It can accurately convert most of the accents and it can be used by multiple users in real time without creating or switching any voice profiles. Braina's voice recognition works even in noisy environment.

Braina can also be customized to recognize custom words, create canned answers and templates. You can teach uncommon names of people, technical jargon, address to Braina etc. Braina has ability to recognize unusual vocabulary and can already understand most of the medical, legal and scientific terms.

Go Transcribe

Country: /
Email: info@go-transcribe.com
Telephone:
URL: <https://go-transcribe.com/>

Advanced cloud based transcription service powered by artificial intelligence. Automatically converting voice into text saves time, money and effort.

Get more work done in less time

Simply upload an audio file onto our platform using any device and we transcribe it automatically into a document for you.

Fast text based editing

Use our online editor to review the transcription. Playback the original file and jump to specific parts of the audio by clicking on words on the transcription. Our unique editor allows you search your audio and make amendments to the transcription where required.

Manage your transcriptions

Download your transcription to several popular formats including Microsoft Word and PDF. With our advanced packages, you can assign transcriptions to your assistants for them to review and amend.

- Take time control of your time, spend more time on more important activities and leave the transcribing to us.
- Get meeting minutes automatically converted into text without having to write long notes.
- No need to spend money on recording devices. Use your mobile phone or any pre-existing dictation products to record your audio.
- Upload your file onto our cloud based service from any devices including mobile phones and Tablets.
- You can work from anywhere at any time making it easy and convenient for you with no complicated software to install.
- Receive the text file within minutes of uploading to help get the most out of your time.
- Setup workflows between users and assistants to allow your assistant to review the transcription.
- Always have the latest version of the product at no extra cost.



Country: USA
Email: wecare@sonix.ai
Telephone: /
URL: <https://sonix.ai/>

Sonix transcribes your recordings so you and your team can be more productive, less stressed, and just a little bit happier.

EASY UPLOADING

Our simple drag-and-drop interface lets you securely upload your audio and video files in record time.

FAST TRANSCRIPTION

Our automated algorithms will transcribe in less time than the length of your recording.

EMAIL NOTIFICATIONS

Leave and grab some coffee as we transcribe. You'll get an email when we are finished.

SYNCHRONIZED AUDIO & TEXT

Click on any part of the transcript and play the audio from that exact moment.

CATEGORIZE SPEAKERS

Hit enter and you can separate who said what and select from a dropdown list of names.

HIGHLIGHTING WITH TIMESTAMPS

Mark those important conversational moments and know exactly when something was said. Perfect for producers and audio engineers.

ORGANIZE WITH FOLDERS

Easily categorize your many files with a simple folder structure.

SHARE A READ-ONLY VERSION

Let others see your transcript with a simple link. You are the only one who can edit.

EXPORT TRANSCRIPT

Your transcript is available in many file formats to better suit your workflow.



Country: UK
Email: support@trint.com
Telephone: /
URL: <https://trint.com/>

We go beyond automated transcription to provide the world's most innovative platform for searching, editing and getting the most out of your audio and video content.

Speech-to-text technology is really smart, yet it can't produce perfect transcripts.

With Trint's ingenious editing platform, you'll get the fastest turnaround time possible for reaching perfect audio/video transcripts.

FAST AUTO-TRANSCRIBING

Just drop your audio and video files* into Trint and you'll get your transcripts in less time than the length of the original recording.

QUICK SEARCH & EASY CORRECTING

The amazing Trint Editor marries a text editor to an audio/video player. Click on a word and hear your original audio, making it easier than ever to quickly search, verify and (if necessary) correct your content.

OPTIMIZING YOUR CONTENT

Use Trint's innovative features to make the most of your content for collaborating, sharing or exporting... like you've never seen before!

INSTANT TIMING

Select a quote within your trinted transcript and instantly Trint tells you its precise time in the recording and its precise length. A huge time-saver for broadcasters and podcasters.



Country: Belgium
Email: form on site
Telephone: (+32) 09 330 13 23
URL: <http://limecraft.com/>

Limecraft Transcriber uses Speech-To-Text technologies to assist editorial staff and archivists to transcribe and pre-cut factual or fictional digital content. During the transcription, the user has access to correct the result at anytime, and then prepare scenes into shots that can be directly imported into any video editing software. Transcriber provides output in various formats, including PDF, CSV or any common video editing suites (Avid, Premiere, FC) and is currently available in English (American and British), French, German, Spanish and Dutch.

We analyse the audio in your media and recognize all speech with a high accuracy. This reduces the time it takes to create an high quality transcript by many orders of magnitude!

Every single word in your transcript is linked to a position in your video. This means you can create subclips by simply selecting the spoken text in your transcript! These subclips can be used to create a first pre-cut which can be played out or exported to your edit suite.

Transcriber cuts the timeline into distinct speech fragments and renders a timeline with fragments grouped per speaker.

Once you have created a transcript you can search for subclips where a certain text is spoken out loud. You can also search all subclips spoken by a given speaker.



Country: The Netherlands
Email: form on site
Telephone: +31 (0)53 488 99 00
URL: <https://www.telecats.nl/en/>

Telecats has ensured that plenary debates of the Dutch Parliament come with subtitles. Anyone can watch these debates online. Voice recognition software ensures that the text of the acts is synchronized using the voice of the politicians, so that there is an auto-generated subtitles. The text of the subtitles in the videos is equal to the uncorrected report, such as that by the Clerk's Office was created and on the website of the Dutch Parliament. It is now possible to search a specific spoken word and view that specific clip from the debate. This makes the search possibilities enormous. Through the joint offer of voice and text, its easier to follow the debates, for example, for people who have trouble with the Dutch language. The subtitles ensures that debates can be followed without the sound. For example on displays at stations, in canteens and in government buildings. With the push of a button the subtitles can be switch on or off.



Country: Japan
 Email: form on site
 Telephone: /
 URL: <https://www.advanced-media.co.jp/english/solution/conferenceproceedings/>

Preparing minutes for meetings and conferences is time consuming. Outsourcing costs money. Solutions from Advanced Media help solve issues related to the preparation of minutes and transcriptions.

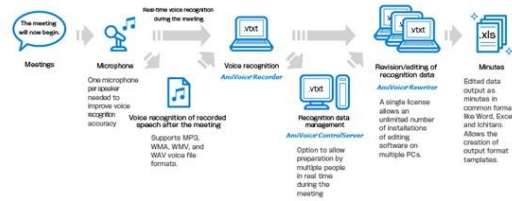
Feature 1

Voice recognition engine customized for customers

Converts the speech of unspecified individuals to text with a high level of precision, creates a dedicated dictionary based on a customer's past minutes data, identifies proper nouns (such as people and locations), and even deals with different dialects.

Feature 2

Simple revisions and editing of only those areas affected by recognition errors



Country: USA
 Email: form on site
 Telephone: /
 URL: <https://www.ibm.com/watson/services/speech-to-text/>

Powerful real-time speech recognition

Automatically transcribe audio from 7 languages in real-time. Rapidly identify and transcribe what is being discussed, even from lower quality audio, across a variety of audio formats and programming interfaces (HTTP REST, Websocket, Asynchronous HTTP)

Highly accurate speech engine

Customize your model to improve accuracy for language and content you care most about, such as product names, sensitive subjects or names of individuals. Recognizes different speakers in your audio Spot specified keywords in real-time with high accuracy and confidence

Built to support various use cases

Transcribe audio for various use cases ranging from real-time transcription for audio from a microphone, to analyzing 1000s of audio recording from your call center to provide meaningful analytics



Country: UK
Email: info@speechmatics.com
Telephone: +44 (0)1223 794 497
URL: <https://www.speechmatics.com>

Speechmatics' versatile automatic speech recognition technology, based on decades of research and experience in neural networks, is enabling world-leading companies to power a speech-enabled future.

AL, our language building framework, uses cutting-edge ASR techniques and machine learning. AL can build any language in the world in as little as a week.

A pioneering new theory of how all languages work, AL's unique framework identifies patterns in global languages, and uses them to automate and generalise tasks across all new language builds.

Requiring minimal data and previous language knowledge, AL's smart algorithms and cross-build learning make it a fast, accurate and dynamic ASR solution.

We currently support 28 languages, and with our language-building framework we have the ability to develop new languages at a rapid rate.

Our unique deep neural network (DNN) approach provides high accuracy levels across many use cases.

Results

- via email txt or JSON
- Confidence scores for words
- Time-stamping of words
- Speaker change and labelling

Markets the Speechmatics private deployment for pre-recorded files is suitable for: Interviews, Podcasts, Search Engine Optimisation, Captioning, Analytics Discoverability

Our technology can be deployed in the cloud, privately (including on-premises) or embedded on-device.



Country: USA
Email: form on site
Telephone: 781-565-5000
URL: <https://www.nuance.com/>

If you're a mobile professional—or anyone on the go—who relies on a digital voice recorder or smartphone to capture notes and memos, use Dragon's robust transcription features to turn your recordings into text quickly, easily and accurately. Dragon can even transcribe another single speaker's voice from pre-recorded audio files or from podcasts to maximize flexibility and drive productivity.

Many corporations, government agencies and other organizations face significant documentation requirements as part of conducting business. Inefficient or cumbersome processes too often result in missed deadlines, inaccurate documentation and too much time and money focused on administrative work. Join the thousands of organizations that have deployed Dragon across the enterprise to improve service and documentation and reduce reliance on outside transcription services. Along with the Nuance Management Center (NMC), Dragon can easily be customized and managed enterprise-wide.

Since Dragon is speaker-dependent software, it's not designed for use with multiple speakers. Thousands of customers still successfully use Dragon to transcribe interviews or lectures using a basic technique called "voice writing" or "parroting."

- Record a lecture using a digital recorder or mobile device and download the audio file to your PC or Mac.
- Activate your Dragon microphone, listen to the recording through your Dragon headset and repeat the recorded text as you hear it.
- Speaking the text aloud in your own voice enables Dragon to accurately transcribe the audio using the Dragon profile tuned to your voice.
- Dragon turns your voice into text as quickly as you can speak the words, so there's no need to constantly rewind the audio while you try to type out the corresponding text.



Country: China

Email: sales@corbell.com.sg

Telephone: /

URL: <http://sg.penpower.net/product.asp?sn=701>

ViaTalk turns voice into text. Talk to your Smartphone, and it types up to 200 words in a minute for you. It's easy to use. Just connect your Smartphone with your PC with ViaTalk, open the APP and start talking. Support Office software, email, Skype, Facebook and more. ViaTalk is also an excellent translation tool. Let you edit the translated result on PC easily.

- Input up to 200 words each minute by talking to Smartphone
- Recognize and translate multiple languages
- Support PC application and type for your on social network
- Add various symbols by talk or typing easily
- Apply with your smartphone
- PC Requirements : Windows 8 / 7 / Vista / XP, USB port
- Mobile device : iPhone 4S or above model with iOS 7 or above, Android 2.3 or above
- Dimensions : 56 x 27 x 12 mm
- Weight : 30g
- 3.5mm audio cable length : 60cm
- In the Box : Audio signal transmission USB cable, Lapel microphone, Quick guide, Registration card
- Available in Black color only



Country: USA

Email: [contact form on site](#)

Telephone: 800-233-3030

URL: <https://mmodal.com/>

Our cloud-based clinical documentation solutions connect workflows for efficient and accurate medical speech recognition, medical transcription, CDI, and coding. We facilitate physician-patient relationships by making it easy for doctors to capture the patient story, and by delivering information when it is needed most – at the point of care.

Is the goal of EHR utilization slowing you down and restricting the patient narrative? Our front-end medical speech recognition solutions provide the flexibility, convenience and efficiency you need to document patient encounters completely and quickly. With our superior speech-based documentation solutions, you can eliminate the cost and delay of transcription while improving the quality of clinical documentation and the user experience.



Country: USA
Email: info@entradahealth.com
Telephone: 615-750-2077
URL: <http://www.entradahealth.com/>

Entrada Rhythm enables caregivers to capture relevant and robust patient information at the point-of-care utilizing speech, no matter when and where care occurs. Entrada Rhythm is a cloud-based mobile dictation solution that provides multiple workflows for documentation, including editing options that can be changed throughout the day. Caregivers dictate the different sections of the patient chart from anywhere with a simple tap on your device's screen.

Entrada offers one solution for every documentation workflow...

- Mobile Scribe offers full EHR template support via remote assistance
- Entrada Rhythm offers both front-end speech and back-end transcription within a single mobile workflow
- Entrada offers other speech-driven workflows, including traditional transcription options



Country:
Email: support@crescendo.com
Telephone: +1 (450) 973-8029
URL: <http://www.crescendo.com/index.html>

Crescendo Speech Processing has revolutionized the way healthcare facilities work. A truly professional solution, the application can be customized with distinct Lexicons for a variety of languages and medical specialties and loaded with terminology specific to them.

FRONT-END SPEECH RECOGNITION

Physicians wishing to create their own reports on-the-fly are presented with Crescendo Speech Processing Front-End. The physician will have the ability to dictate, self-correct, sign and release reports at their own pace.

BACK-END SPEECH RECOGNITION

For those wishing to simply dictate, Crescendo Speech Processing Back-End can be implemented. This option produces text while being fully transparent to the user. The report is then sent to transcription for correction and quality assurance. Crescendo even allows the user to switch from one mode (back-end, front-end) to the other while logged in whereas other systems would force users to log out, saving valuable time.

NO TRAINING REQUIRED!

Crescendo Speech Processing eliminates the need to train the application to the user's voice, it learns "on the job" with remarkable accuracy. The system quickly and seamlessly adapts to individual voice characteristics while dictation habits remain unchanged.



Country: USA
Email: [contact form on site](#)
Telephone:
URL: <https://www.houndify.com/>

Understand the meaning, not just the words. The **Speech-to-Meaning™** engine delivers unprecedented speed and accuracy and can be integrated with mobile applications, cloud software, connected devices and, ultimately, the Internet of Things.

Integrate a powerful, flexible and scalable voice-enabled conversational intelligence into your product, service or business and instantly simplify your user interface. Our technology combines both Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU) into one engine, resulting in dramatically increased speed & accuracy.

The Houndify Platform offers 125+ Domains of Understanding, Knowledge Graphs and redistribution rights from content providers - over 4x more Domains than Siri, Google and Cortana. Your product, platform or service can be powered by Houndify.

Custom (Private) Domains
Create a custom domain for your business.

Custom Trigger Phrase
Make your product respond to a custom trigger phrase of your choice, similar to "Ok Hound..."

Custom Recognition Models
Create custom recognition models for noise, acoustic conditions, or unusual vocabulary.

Product Integration Support
Receive support with Integration of Houndify SDKs and Domains in your products.



Country: USA
Email: sales@ispeech.org
Telephone: /
URL: <https://www.ispeech.org/text-to-speech.tts.saas.api>

Speak any message and iSpeech Dictation™ will put it into text format. Dictate using BlackBerry Messenger (BBM), text (SMS), email, or voice notes into text and send.

The app's human-quality speech recognition is brought to you by iSpeech®, the creator of DriveSafe.ly®, award-winning leader in texting while driving applications.

Speak any phrase or message and iSpeech Dictation™ will translate it into text. Talk and type.

- Translate any message into text by speaking it
- After speaking the message, the text will appear in the text box
- Use to send BlackBerry Messenger (BBM), SMS, emails, voice notes



Country: USA
Email: info@vocitec.com
Telephone: 412-621-9310
URL: <https://vocitec.com/what-we-do>

Voci develops state-of-the-art voice transcription and speech analytics solutions for our clients and partners.

We take 100% of your audio and video recordings and convert them into searchable text that can be viewed alongside email, social, and chat messages for a global view of your data.

Our transcriptions are guaranteed to be 80% accurate out-of-the box, but we don't stop there—we also include detailed metadata covering everything from call duration to speaker emotion and generate these results quickly and accurately.

We also work with you to develop custom implementations, and continually develop new ways to integrate with your existing infrastructure to give you an all-in-one research platform that delivers valuable insights and radically improves business outcomes.

Voci provides voice data solutions for customer experience, call center operations, compliance & risk management, eDiscovery, and visual voicemail.



Country: USA
Email: sales@transcribeme.com
Telephone: +1-800-275-5513
URL: <https://transcribeme.com/>

The Enterprise Speech Recognition (ESR) uses state-of-the-art speech recognition technologies and expertise to provide enterprise with the highest accuracy automated transcriptions available on the market today.

By customizing the ESR platform to each client's use case, we are able to improve overall speech to text accuracy by as much as 10% to 20%, when compared to the industry's best, non-customized speech recognition services.

ESR phonetics-based keyword spotting is designed to accurately identify and analyze words or phrases instantaneously, and provide Hot Alerts automatically to management personnel or enterprise systems for resolution.

Typical keyword or key-phrase spotting use cases include topic detection, risk mitigation, fraud detection, agent training and monitoring, and customer satisfaction.

- Highest accuracy automated speech to text solutions for specific enterprise use cases.
- Identify key customer demographics, including age, gender and accents.
- Quickly identify topics and mitigate risk with real-time keyword spotting.
- Easily integrate text and metadata with any analytics software tools and WFO platforms.
- Hybrid automated speech to text and human transcription provides flexible enterprise solutions.
- Best-in-class voice biometrics to authenticate callers for fraud detection.



Country: Czech Republic
Email: info@spokendata.com
Telephone: +420-603-171-863
URL: <https://www.spokendata.com/>

Your Speech-to-Text all in Cloud.

Select the desired technology from speech to text, voice activity detection, speaker segmentation or text to audio alignment. You are notified by email when the automatic processing is done.

- Manage your team of transcribers using tags and categories.
- Help them with transcription by automatic voice-to-text technology.
- Integrate SpokenData into your application via our REST API.
- We adapt the voice-to-text on your data domain to maximize the transcript accuracy and lower your labor costs.
- Track your transcribers.
- Browse through transcriber statistics.
- Suitable for: small teams of professional transcribers, academy/business projects, ...



Country: USA
Email: sales@saince.com
Telephone: +1-678-967-3400
URL: <https://www.saince.com/transcription-services/>

Saince's revolutionary medical dictation software allows us to offer our customers unmatched flexibility. Physicians can use toll-free numbers, voice recorders, laptop or PC integrated mics, or our iPhone or Android apps to record their medical dictations. With revolutionary voice technology integration, dictations using our mobile apps are clearer and sharper than those made over phone lines!

Verbatim's powerful voice transcription software ensures that you receive accurate medical transcriptions every time. Less time spent correcting errors made by medical speech recognition software means less editing means more productivity.

Verbatim is designed for humans, not robots. We made our state of the art dictation service for physicians to use in the real life situations and complications that arise every day in the medical field. Urgently need to attend a patient? Not in the mood to spend time completing a medical coding project? With Verbatim, you can save the incomplete report as a draft to be completed later, or with the click of a button instantly send your voice file out for transcription. Your audio voice file and verbatim transcription draft immediately transferred to a trained medical transcriptionist who will send the completed report back to you shortly. Verbatim is the best medical dictation software available on the market and our integrated transcription technology is a first in the industry.

Verbatim is fully a cloud based digital dictation service so you do not run any risk of technology obsolescence, and there is no need to upgrade any software or pay any maintenance fees.



Country: USA
Email: form on site
Telephone: /
URL: <https://www.voxneural.com/>

Looking for a simple Voice Recognition API platform?
Look no further.

Let's face it - audio data is a pain in the ear canal to analyse. We are Voxneural. We convert voice to text using deep learning neural networks.

We convert most audio files. We'll process them asynchronously, convert them to searchable text and send them right back to you. It's that simple. We can even take it further and draw real-time insights from this text using Vox Analytics.

The result? Invaluable business intelligence. Sign up now and we'll send you instructions to hook up to our API.



Country: The Netherlands
Email: form on site
Telephone: /
URL: <http://www.redspeak.nl/>

RedSpeak allows you to create letters, e-mails and to create reports using only your voice. Where and whenever you want. Behind the wheel (hands free) in the car, at the Office, abroad or at home. You speak your text (digital dictation) on your Smartphone and before you know it your voice is converted to digital text by the high-security RedSpeak server. The server is equipped with high-end business speech recognition, so you are assured of the highest degree of recognition of no less than 99% accurate. With RedSpeak, you can be up to 95% more productive processing texts, because now you can spend your time optimally by creating texts anywhere. With RedSpeak you get the most out of your working day.

RedSpeak is built so you are able to work smarter on any device at any time. That is why RedSpeak is web-based. You have access to RedSpeak with the operating systems: Windows, OSX, iOS, Android and Linux. In addition the interface automatically resizes to the size of the device that you are using so you always get the best user experience.

To further increase efficiency in your organization, RedSpeak's collaboration portal allows for flexible configurations. Support staff can see the dictations from multiple authors for further processing. This means the workload can be distributed, leading to increased efficiency.



Google Cloud Platform

Google Cloud Speech API enables developers to convert audio to text by applying powerful neural network models in an easy to use API. The API recognizes over 110 languages and variants, to support your global user base. You can transcribe the text of users dictating to an application's microphone, enable command-and-control through voice, or transcribe audio files, among many other use cases. Recognize audio uploaded in the request, and integrate with your audio storage on Google Cloud Storage, by using the same technology Google uses to power its own products.

Automatic Speech Recognition

Automatic Speech Recognition (ASR) powered by deep learning neural networking to power your applications like voice search or speech transcription.

Global Vocabulary

Recognizes over 110 languages and variants with an extensive vocabulary.

Streaming Recognition

Returns recognition results while the user is still speaking.

Word Hints

Speech recognition can be customized to a specific context by providing a set of words and phrases that are likely to be spoken. Especially useful for adding custom words and names to the vocabulary and in voice-control use cases.

Real-time or Pre-recorded Audio Support

Audio input can be captured by an application's microphone or sent from a pre-recorded audio file. Multiple audio encodings are supported, including FLAC, AMR, PCMU and Linear-16.

Noise Robustness

Handles noisy audio from many environments without requiring additional noise cancellation.

Inappropriate Content Filtering

Filter inappropriate content in text results for some languages.

Integrated API

Audio files can be uploaded in the request or integrated with Google Cloud Storage.

Country: USA

Email: form on site

Telephone: /

URL: <https://cloud.google.com/speech/>



Country: The Netherlands

Email: form on site

Telephone: 1-800-667-6389

URL: <https://www.salesforce.com/eu/>

Product URL: <https://www.salesforce.com/blog/2017/05/ai-salesforce-research-text-summarization.html>

In 2017, the average person is expected to spend 12 hours and 7 minutes every day consuming some form of media and that number will only go up from here. Connected devices are ubiquitous with everything from Alexa and our thermostats to our cars and cell phones creating new information at a dizzying pace. Today the struggle is not getting access to information, it is keeping up with the influx of news, social media, email, and texts. But what if you could get accurate and succinct summaries of the key points of every email you receive or news article you scroll by?

Salesforce Research is tackling this exact challenge and today we're excited to announce two new breakthroughs in natural language processing towards the goal of automatically summarizing a long text and serving up coherent, digestible highlights that help you stay informed in a fraction of the time. Text summarization is a very tough challenge, especially for longer texts such as news articles, and the work we are doing at Salesforce Research is pushing the state of the art. I'm honored to work with Caime Xiong and Richard Socher to introduce a more contextual word generation model and a new way of training summarization models with reinforcement learning (RL). Together, these AI improvements enable algorithms to create easy-to-read, highly accurate multi-sentence summaries of long text, significantly improving on previous results.