Institut(ter): FORCE Technology	**Aktivitetsplan (titel):** Produktdesign via virtuelt lyttepanel **Idéforslags titel på www.bedreinnovation.dk:** Produktdesign via virtuelt lyttepanel	**Aktivitetsplan nr.:** A11	FoU

1) Manchettekst (kort resumé)

Et virtuelt lyttepanel vil blive udviklet til estimering af subjektiv lydkvalitet indenfor kunde/produktspecifikke anvendelser. Ydelserne vil nedbringe tid og ressourcer anvendt på tests, hvilket muliggør bredere anvendelse og af flere virksomheder.

2) Aktiviteten kort (resumé)

FORCE Technology SenseLab vil gennem denne aktivitet udvikle services baseret på et virtuel lyttepanel (VLP). Dette betyder i praksis, at vi vil anvende data fra lytteforsøg (eksisterende databaser og nyindsamlede data) til at træne prædiktive modeller med brug af regression, Machine Learning og deep learning. Ydelserne vil være målrettet eksisterende kundesegmenter og SMV'er, men også nye segmenter der traditionelt set ikke har været i stand til at evaluere lydkvalitet ved brug af dyre og tidskrævende metoder baseret på lytteforsøg. Indenfor de første to år af aktiviteten forventer vi, at det virtuelle lyttepanel kan erstatte visse anvendelsesspecikke evalueringer. Det vil udgøre et supplement til de eksisterende lytteforsøgsbaserede lydkvalitetsevalueringer, som fortsat anses som værende mest troværdige i en årrække frem. Det virtuelle lyttepanel vil således udvide SenseLabs forretning og være til gavn for mindre virksomheder.

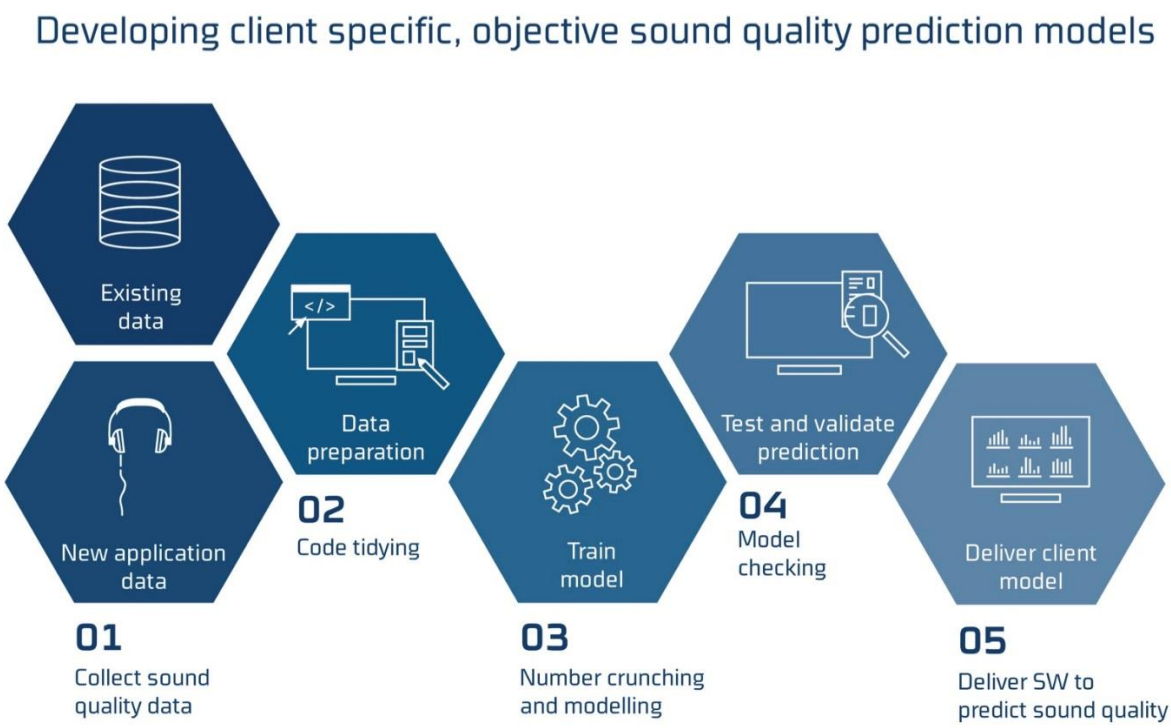
Aktiviteten vil udbygge de eksisterende kompetencer og databaser indenfor subjektiv evaluering af lyd, som er indsamlet via SenseLabOnline og LabView værktøjer de seneste ti år. Anvendelsesområdet vil omfatte det eksisterende marked indenfor høreapparater og højttalere, men også det nye og hastigt voksende marked indenfor Hearables[1] og intelligente hovedtelefoner, hvor danske og internationale virksomheder er aktive. Den nye tilgang vil desuden åbne metoderne op for andre virksomheder, der anvender lyd.

Figuren nedenfor illustrerer den overordnede proces, hvormed vi ønsker at udvikle de nye virtuelle lyttepanel services. Processen består af fem faser:

1. **Indsamling af lydkvalitetsdata**: Effektiv indsamling af subjective data (01a. Eksisterende databaser; 01b Nyindsamlede data).
2. **Klargøring af data**: Strukturering af subjektive data, optagelser, kalibrering, normalisering og equalisering af lydfiler som forberedelse til udtræk af lydkarakteristika og efterfølgende modeltræning.
3. **Træning af model**: Modeltræning og finjustering.
4. **Afprøvning og validering af prædiktionsmodellen**: Afprøvning af prædiktionsmodellens kvalitet op imod anden (uset) data med henblik på validering og forfinelse af modellen for at sikre bedst mulig præcision.

[1] Hearables er elektroniske *in-ear-devices* til forskellige formål som trådløs kommunikation, musik, medicinsk overvågning eller *fitness tracking*. Begrebet introduceres samtidig af Apple og af Nick Hunn i 2014, se sidstnævntes artikel "Hearables – The new Wearables", Wearable Technologies online (april, 2014)

5. **Levering af kundemodel**: Modellen bygges ind i et softwarebrugerinterface, der giver kunden mulighed for let at indsætte lydfiler og estimere den prædikterede subjektive lydkvalitet af deres produkttype. Det forventes, at softwaremodeller sælges på licensbasis.



Figur 1. Overblik over processen til at udvikle en anvendelsesspecifik Virtuelt Lyttepanel model.

3) Markedsbehov, erhvervs- og samfundsmæssige potentialer

There is fierce global competition between manufacturers of audio products. Asia, US and Europe have strong companies competing on technology and price. For example headphones/headsets segment is growing rapidly[2,3], and as a result of increased competition, there is a need to focus on development cost, development time (time to market) and competition parameters, such as sound quality. The rapid adoption of wireless headphone technology since 2017 has boosted this market and also introduced smart product categories including smart headphone and hearables[4], with the latter predicted to be the largest growing smart device segment[5]. The US 2017 *Over-the-Counter Hearing Aid Act*[6] has also reinvigorated and expanded the market potential for hearing aid manufacturers, targetted at a new/younger market segement with milder hearing losses, with users starting to employ these devices potentially ten years before traditional hearing aids[4].

[2] Global Audiophile Headphone Sales are 20M units in 2018 – estimated to grow to 36M unit in 2025: Global Audiophile Headphone Sales Market Report 2018, QY RESEARCH, June 2018, https://www.marketinsightsreports.com/reports/06196129 26/Global-Audiophile-Headphone-Sales-Market-Report-2018

[3] The global earphone and phone market was valued at $ 8.7 B in 2015 and is expect to reach $ 15.8 B in 2025: Earphones And Headphones Market Analysis By Product (In-Ear, Over-Ear), Grand View Research, 2017 https://www.grandviewresearch.com/industry-analysis/earphone-and-headphone-market

[4] https://www.nuheara.com/expanding-market-hearable-devices/

[5] https://emberify.com/blog/pixel-buds-airpods-and-the-future-of-the-hearable/

[6] http://www.hearingreview.com/2017/08/president-trump-signs-otc-hearing-aid-legislation-law/

Today, listening tests are reserved for the large corporations that can afford the time and resources to optimize sound quality through listening tests with test persons and facilities. Danish companies need the services that increase their competitiveness effectively. Additionally, predictive models of sound quality in telecommunication have been successfully stardardised, but only address sound quality in digital telephony[7,8].

Presently, clients use listening tests (internal or independent lab) only where critical, e.g. for marketing claims, standardization technology selection processes etc. Objective sound metrics are used widely, as always, but lack the perceptual details. Also when available, e.g. in telecommunications, standardized predictive models such as P.OLQA[8] or 3QUEST[7] are widely employed. However, such models are not existing for other fields and applications, such as hearing aids/hearables, wireless speaker technology, etc. This also means in practice that smaller companies and startups cannot afford to employ these methods – something that we would like to help change.

Our initial probing of the market interest through the bedreinnovation.dk process yielded positive comments and support towards the development of a *Virtual Listening Panel*, i.e. predictive models to estimate subjective sound quality. The interest came from the Danish hearing aid, headphone and speaker industries[9], aiming to apply a Virtual Listening Panel *more rapidly, frequently and cost effectively* compared to listening test[10]. A few references from this process are included as illustration:

"Applying perceptual evaluation through machine learning is a great idea, which will provide years of accumulated knowledge and data matriale with Senselab, a new value. This increases the possibility of frequent evaluation of sound quality through a development process to ensure proper course and short development time. At the same time, Senselab could be further strengthened, as a natural and important focal point, for the benefit of the audio industry.", **Bjarne Bjerre, Engineer, Acoustics/Electronics, Terma A/S**

"…[if] in the future we could potentially add machine learning as a test method, this would be attractive to many in the audio industry. In many places we often see a busy development process - both on budget and time. Rapid decisions must be made on a continuous basis, and in order to do so on an informed basis, these methods could be an excellent supplement.", **Michael Hoby, Senior Acoustic Researcher, Goertek Europe**

"Developing models for fast and frequent quantification of measurement results, using machinelearning, is an obvious opportunity to ensure that we (as industry) remain at the forefront of developments. Particularly the possibility of integrating these models into existing (and new) simulation tools is particularly valuable for SMEs; like ourselves.", **Tue Dissing, Founding Director, Lyd by Dissing ApS**

"…Using the listening panel is relatively time consuming, so developing a virtual panel based on machine learning will be a major advance especially in the early design phase and solid prototyping of our products.", **Ole Dyrlund, Director of Technology Integration, GN Hearing**

[7] ETSI 3QUEST: ETSI EG 202 396-3 https://www.etsi.org/deliver/etsi_eg/202300_202399/20239603/01.03.01_60/eg_20239603v010301p.pdf

[8] P.OLQA (or Perceptual Objective Listening Quality Analysis), also known as ITU-T Rec. P.863: https://www.itu.int/rec/T-REC-P.863-201803-I/en

[9] Comments from Oticon, Sennheiser Communication, Jabra, GN Hearing, Goertek, Terma, LOUDSOFT, EBB Consultant, Voetmann Akustik, Danish Standard, Google

[10] In the medium-term listening test data will continue to be considered the *"ground truth"* benchmark of subjective quality assessment.

"A virtual listening panel will have many advantages for Danish consultants / audio system developers: Fast and cheaper development of speakers, headphones and test equipment; Reliable and objective way to convince foreign customers of the quality of the above products, especially with OEM customers; Better opportunities to launch the products in a competitive market; Highlight Denmark's already good reputation for acoustics worldwide.", **Peter Larsen, President, LOUDSOFT Ltd**

New digital services within prediction of basic sound quality parameters can ensure rapid and inexpensive screening, so the sound quality of an upcoming product meets expectations. This will be an attractive tool for both large and small businesses as part of an efficient and quality-oriented development process, and in keeping with FORSK2025, pointing to "***Endvidere er der behov for fokus på skalérbare metoder, værktøjer og teknikker, der understøtter udvikling af systemer med henblik på hurtig og prisbillig udvikling af nye produkter,…***" (p29). The last ten years FORCE Technology SenseLab collected large amounts of data from listening tests with both trained expert listeners and consumers. At the same time tools in Machine Learning matured and were made available. In particular, data collected by the trained panel, which is structured in Lydhjulet[11] will give a decisive advantage to develop reliable prediction models for Danish industry. This is in line with with FORSK2025, pointing to ***"… behov for forskning, der afdækker hvordan: forudsigelser om nye data kan laves ved brug af modeller baseret på tidligere data (machine learning, dynamiske systemer og stokastiske systemer).****"* (p29).

Denmark has a strong position in the global market in audio products (especially speakers, speakerphones, headsets, hearing aids). In 2016, an industry analysis[12] conducted by the Innovation Network Danish Sound showed that the Danish audio industry had grown by 28 percent from 2012 to 2016 and had a total turnover of nearly 36 billion kr. in 2016. In count, 68 percent of the companies are categorized as SMEs and this indicates a growth potential in the segment where new innovative solutions for product development are vital for success. We expect 40 SME's to request the new VLP services within the first five years. There are further growth indications in the industry through the recent trend of foreign investments into the Danish industry (Dynaudio, TC Electronics, Libratone, Goertek, B&O). Additionally the developments in smart headphones will have a positive impact on companies such as GN Hearing/ JABRA, AiAiAi, Sennhesier Communication and the newly defined hearables cateogory leading to new growth in hearing aid companies (Oticon, Widex and GN Hearing).

The services within prediction models for sound quality will be relevant for developers of audio equipment, especially headset/headphone manufacturers, as well as companies that develop less compact speakers. The benefits can be used at an early stage of the design for both setting goals and testing in the development process in terms of component selection, signal processor setting etc. SenseLab's experience is that typically SMEs prioritise other activities above listening tests, in part due to the high required time and cost investement. The VLP should lower this entry barrier.

4) Vidensprednin og inddragelse

In order to ensure that this activity is focused on the market needs, we will be interacting with both potential Danish and International companies including startups, SMEs and larger corporations. Firstly an Advisory Board will be established with a combination of experts as well as potential client representatives. We will then also engage with the potential clients to learn more about how they would

11 http://assets.madebydelta.com/docs/share/Akustik/TEKnotat_Beskrivelse_af_lyden_fra_audioprodukter_DK.pdf
12 https://danishsound.org/Publications/DSN_2018_Industry%20Analysis%202017.pdf

ideally envisage such a VLP service to look like in detail. This will be done by a combination of meetings, group seminars/workshops and so forth.

Furthermore, we will invite companies to actively participate in pilot projects, by bringing real product usecases to be the subjective of the VLP development thought the 2 year project. This will allow clients to interact and comment on the protocol as it is developed, as well as shaping the final services. Example of potential partners might include:

GN Audio (JABRA), Sennheiser Communications, Goertek Denmark, Point Source, LOUDSOFT ltd, Lyd by Dissing ApS, GN Hearing, Dynaudio, B&O, AiAiAi, Terma A/S

We will also share the project developments with a wider audience through interaction with the broader community using the Danish Sound Network, organising Danish Audio Club meetings and hosting a yearly International meeting of SenseCamp. We will furthermore publish our engieering and scientific achievements through at least three conference or journal articles. Lastly, as soon as we have concrete services to offer, these will be published via our corporate web pages.

5) Konkrete aktiviteter

Main activities

The primary project aim is to build application specific audio quality prediction models for clients, which should - for these specific applications - replace the use of a subjective listening test. This will be achieved by collected data from existing databases in SenseLab as new datasets collected from the clients product category, comprising of recording audio samples from the product category (existing, prototype and new products), collecting subjective datasets through listening tests using novel techniques for rapid data collection. From this the audio features will be extracted, which will be used as an input in the modelling process. This process will aim to outline a model that can be used to predict the subjective data from the audio files/extracted audio features. Establishing the meaningful set of audio features will be vital at the generic level (loudness, distortion, etc.) and an application specific levels (e.g. spatial sound: envelopment, source width, etc.). Research and collaboration will be required to identify the best suited modelling approaches for our data type, beyond the commonly known approaches and will require iteration and testing to select the best approach.

Collaboration (with companies/potential clients):

- Offer interested SMEs & corporations the possibility to participate in demonstrator projects. Actively invite them to submit their new/existing products to be used for creating sound quality models. Potential markets include the fields of hearing aids, over-the-counter (OTC) hearables, headphones, wireless speaker technology;
- Provide companies with results from pilot listening tests, using new and efficient methods, as well as the outcome of the pilot modelling work, i.e. sound quality models.

Collaboration (with universities and expert organisations):

- Collaborate and repatriate specialist knowledge through secondment to various institutions: both university and expert/commercial orgnisations (see section 7). This will comprise of participating in training courses or short visits (max 1 week duration) to learn about methods and technniques in the various activity areas, where new knowledge is needed.

Data and data collection (**Fejl! Henvisningskilde ikke fundet.**: 01)

In order to be able to predict the subjective sound quality and to model these characteristics, we need data. In SenseLab we have collected sound files and related subjective scores for over a decade using SenseLabOnline. Some of this data can be re-used to either build models and/or to validate the accuracy of models. This task is about organising the existing database and learning how it can be re-used. Additionally for new clients we need to establish a new and efficient process to collect new data including a standardized approach to recording of audio stimuli, definition of a a novel and rapid subjective test method for input into the modelling process.

- Analyse and structure the existing SenseLabOnline and LabView audio quality database for usage within the project – for modelling and validation, where appropriate;
- Seek common denominator information, where possible, regarding most pertinent perceptual attributes and/or perceptual dimensions within each domain of application;
- Rapid experimental methods: Selection, definition and implementation of efficient and robust experiemental methods and designs for rapid data collection, e.g. using the experimental design/response surface methods;
- Establish and standardise a common recording technique for headphones, hearables, hearing aids (using high resolution head and torso simulator) and loudspeakers (using higher order ambisonic spatial microphone approaches);
- Record test stimuli and collect of new sound quality data (preferenece, mean opinion score and attribute based data) for client specific application areas for the development of application specific sound quality models.

Feature extraction, modelling and validation (**Fejl! Henvisningskilde ikke fundet.**: 02 - 05)
The core of this project is to establish a way to predict subjective scores based on extracted audio features from recording of client technology. Steps 02-04 are related to developing a process for preprocessing the collected subjective and audio data for subsequent modelling. Once the audio recordings are collected, we need to identify a meaningful set of audio features (02) and metrics that describe the nature of the audio well, such this this information can be used to estimate the subjective scores through modelling. Phase three is about finding appropriate modelling methods (e.g. partial least squares regression, neural networks, deep learning, etc.) for these data types. It will be an iterative process of testing and refinement between stages 02, 03 and 04 to find suitable audio feature extraction and modelling approaches that provide sufficient prediction accuracy and robustness. The validation phase (04) is vital to testing how well the model predictions work against previously unseen data (by the model) – this will give a measure of the prediction accuracy in the case of a *new prototype*. Lastly, we hope to be able to package and deliver model to the client is a user friendly manner. This will require software and an associated process for recording new audio samples, which can be read by the software, which then display the estimate audio quality of these new audio samples, compared to earlier recordings.

- Establishment of a core audio feature extraction toolset for analysis of recorded audio samples (02)
- Establish protocol to equalise, normalise and calibrate audio recordings (02)
- Research into families of modelling and prediction methods that may be appropriate for predicting subjective scores from recorded audio samples and associated extracted audio features. Select the most potential candidates (03)
- Research, development and testing of methods for creating objective predictive models from audio quality from audio (03)
- Testing of modelling techniques for different applications using different pilot data sets, with validation tests against other validation datasets (SenseLabOnline database or other datasets) (04);

- Input of audio files for analysis (05)
- Running prediction model interations (05)
- Output and plotting of modelled audio quality predictions results (05)

Risks/uncertainties

This is an ambitious project and comes with associated limitations and risks. As a whole the project objective are achievable within a two year period. Our aim is merely to develop predictive model of audio quality for client specific and limited applications. Furthermore in a two year period we will be highly reliant upon access and usage of existing toolboxes for the various stages of the work, i.e. psychoacoustic metrics, audio feature extraction tools, etc. There is thus an uncertainty as the quality level that can be achieved given these known hurdles in terms of level of quality, robustness and prediction accuracy of models which can be achieved.

6) Nyhedsværdi og ambitionsniveau

The project will develop several new services to compliment the existing SenseLab business of professional listening and AV testing services, consultancy and software tools. The new service will build upon and compliment the 10 year experience of SenseLab including:

- Service 1: Rapid approach to data collection using listening tests → quick screening listening test service. Expected ready for market by end of year one.
- Service 2: Application specific model development process. Expected ready for market by end of year two.
- Service 3: Licencing of application specific models. Service concept expected ready for market by end of year two.

Companies are seeking fast and reliable ways to evaluate sound quality without the use of test subjects, and especially in the development phase is the time factor is crucial. With the development of new services in the form of tools, in an objective way to describe the sound quality, it would also ensure more uniform quality across development teams and products. Companies, for example, working with brand identity in the form of a measurable sound profile, could benefit from the new models. Benefits in using data for more advanced and efficient listen trial designs (active learning) can contribute to greater data reliability and faster and cheaper test process, which will be beneficial for SMEs. FORCE Technology SenseLab has unique conditions for creating new services. As a leading independent testing lab in subjective evaluation of sound quality continuously collected large amounts of data that can be used to develop models. In addition to new services, it also creates a continued stronger data amount in Denmark in sound quality that helps to strengthen the Danish sound environment and clusters. The market does not even have a platform or another community that similarly will be able to collect large amounts of structured data. Therefore it is not considered likely that the type of services can be developed by the market, so it will benefit and equal access for industry.

7) Vidensamarbejde og -hjemtagning

The project will work with a network of institutions that have specific domain expertise. The majority of these have been selected from our network of partners, who can provide deep domain specific knowledge to enable the project to succeed within a short period of time. Knowledge is primarily to be gained in a collaborative manner through secondments or inviting experts to visit for max 1 week

- **Advanced statistics, experimental design** (**Fejl! Henvisningskilde ikke fundet.**: 01):

- Anne Hasted (QI statistics, UK) – Professional consultancy on design of experiments (DoE) for efficient listening tests.

- **Psychoacoustics models** (**Fejl! Henvisningskilde ikke fundet.**: 02-03)**:**
  - Prof. Ville Pulkki (Aalto University, Finland) – Collaboration regarding psychoacoustics metrics for sound quality and associated tool boxes.
    Aim to collaborate &/or host a MSc thesis student
- **Modelling, machine learning, data mining, adaptive learning** (**Fejl! Henvisningskilde ikke fundet.**: 03):
  - DTU Compute (Technical University of Denmark) - Collaboration of machine/deep learning approaches and adaptive learning techniques.
    Aim to collaborate &/or host a MSc thesis student

The activity will also maintain close relationship with DTU (Prof. Per Brockhoff), Nofima (Prof. Tormod Næs) and TU München.

Methods for subjective testing and predictive modelling within telecommunication applications have been standardized in ITU-T and ITU-R where we have been active. We will continue to monitor these groups to stay abreast of developments as well as to gauge whether there is an opportune moment to rejoin these standardization activities with active contributions.

Additional knowledge sharing is expected on general modelling methods during the EU activity under Maria Skłodowska-Curie ITN RealVision project[13] where SenseLab is hosting a PhD student throughout the three year activity (2018-2020).

As this is an ambitious project, it could be forseen that the final results of the project may require further development and refinement of the services. We will thus consider - during the project period - manners in which to continue into commercial refinement, potentially through funded national or EU collaborations, e.g. EU H2020, Eurostars, Innnovation Fondation projects, etc.

8) Sammenhæng med instituttets strategi og afsæt i instituttets ressourcer

FORCE Technology's SenseLab has successfully served the Danish and global audio, hearing aid and audio-visual markets with sound quality evaluation services and research for over a decade. The strategic corporate aim is to continue to grow the SenseLab activity by exploiting the opportunities that lie within digitizing the services based, in part, on the extensive expertise and data accumulated in SenseLab over a decade. The activity links to FORCE Technology Strategy2020+ by both developing af Teknologisk Service 2.0 based on *Digitalisering 2.0* and hereby opening a new set of *Målgrupper 2.0*. The new digital services will reach out to both existing and new clients within the Danish Audio industry and aim to deliver new technological services that compliment the existing listening tests services with rapid and cost effective predictive models, and by this to companies with low capabilities themselves.

The Virtual Listening Panel services will enable the clients to perform sound quality evaluations when-ever needed in their product development cycle, attracting SMEs who have not previously been able to afford the time and expenditure for subjective listening tests and also new clients who are merely seeking to evaluate sound quality, but without listening tests. The success of this project should provide a complimentary growth path to the existing SenseLab business.

This activity is primarily linked to FORCEs strategic area: *Design and Development*.

[13] www.realvision-itn.eu/

9) Tidsplan og milepæle

Year 1

Knowledge collaboration and competence building

1.1 2 visits to or from collaborative partners.
1.2 1-2 meetings and/or visits to universities.
1.3 4-7 -visits or meetings with potential Danish and International clients.

Development of services

1.4 Develop standardised framework (methodology and experimental design) for rapid and efficient listening tests.
1.5 Define standardised framework for recording and pre-processing of audio files.
1.6 1-2 demonstrator projects.

Knowledge disseminations

1.7 2 Danish Audio Club meetings.
1.8 1 knowledge sharing event organised with > 50 Danish and international participants.
1.9 Participation in 1-2 relevant conferences.
1.10 Establishment of an Advisory Board.
1.11 Advisory board meeting(s).

Year 2

Knowledge collaboration and competence building

2.1 2 visits to or from collaborative partners.
2.2 1-2 meetings and/or visits to universities.
2.3 4-7 visits or meetings with potential Danish and international clients.

Development of services

2.4 1-2 client demonstrator projects.
2.5 New service for prediction of headphone/hearables/loudspeakers[14] preference.
2.6 Service ready for developing client specific models.

Knowledge disseminations

2.7 2 Danish Audio Club meeting.
2.8 1 knowledge sharing event organised with > 50 Danish and international participants.
2.9 Participation in 1-2 relevant conferences.
2.10 Publish at least 2 conference or journal papers.
2.12 Advisory Board meeting(s).

[14] The domain of application will be studied and selected based on potential client interest and available technologies.