LabTwin®

WHITE PAPER

Non-Functional Specifications for LabTwin Voice Assistant

Table of Contents

INTRODUCTION

Non-functional specifications for LabTwin's voice assistant in different lab environments.

This document outlines the non-functional specifications for **LabTwin's voice assistant** in different lab environments. The primary goal of this document is to report the accuracy for different scenarios that simulate lab environments and the risk of unintentional records of data. To report these, LabTwin conducted a series of tests.

SCENARIOS

Four different lab environments with varying background noise.

The scenarios simulated 4 different lab environments with varying background noise. These environments include a fume hood lab, an open lab, an analysis lab, and a vivarium. To ensure consistency in the scenarios, a sound level meter was installed to maintain a consistent dB level of background noise and speaker during testing. Additionally, a speaker was installed in the testing room to simulate background noise from different labs.

Figure 1: Tested environments.

Analysis Lab 50dB

Open Lab 60dB

Fumehood Lab 60dB

Vivarium 65dB

TESTING PARAMETERS

Measuring the accuracy of the LabTwin voice assistant

The tests included different parameters to measure the accuracy of the LabTwin voice assistant. These parameters include word error rate of speech-to-text services (before LabTwin's manual transcription review service) accidental triggers of the voice assistant, and the percentage of other voices being transcribed unintentionally (before LabTwin's manual transcription review service). The testing was conducted using 2 different hardware combinations, including the LabTwin recommended headset (Aftershokz OpenComm + iPhone 13 Pro / Google Pixel 6a).

TESTING PARAMETER 1

Accuracy

Test procedure

The accuracy parameter was tested using 5 test sentences with 3 repetitions each, resulting in a total of 120 utterances per hardware combination and lab environment. The speaker for this parameter is a female native English speaker.

Test sentences:

- The media is yellow so I might need to passage the cells today afternoon
- I have 5 mL of buffer solution
- The temperature of the reaction is currently 25 degrees Celsius
- The concentration of the protein sample is 2.5 milligrams per milliliter
- The weight of the sample is 5 grams

Results

LabTwin could transcribe speech in every lab environment with an accuracy above 96%, with little difference between the hardware or the levels of the background noise. The transcription was especially accurate in the fume hood lab environment with an accuracy of 98,7% with the Google Pixel and 99,4% with the iPhone.

The below infographic illustrates how accurately the LabTwin voice assistant can transcribe speech in the different simulated lab environments. The results are summarized in the below table (Table 1).

The speech-to-text accuracy of the voice assistant was found to be between 96% to 99.4% depending on the lab environment and hardware combination.

Table 1: Accuracy of LabTwin speech-to-text.

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
iPhone 13 Pro + Aftershokz OpenComm	98%	96%	99.4%	96.7%
Google Pixel 6a + Aftershokz OpenComm	96.7%	98%	98.7%	98.7%

TESTING PARAMETER 2A

Accidental trigger – same speaker

Test procedure

The wake word accidental triggers of the voice assistant was tested using 5 test sentences including similar sounding utterances to “Hey LabTwin” with 3 repetitions each, resulting in a total of 120 utterances per hardware combination and lab environment. The speaker for this parameter is a female native English speaker.

This test aims to show how often the user could accidentally trigger the LabTwin voice assistant by using similar sounding utterances to our branded wake word “Hey LabTwin”.

Test sentences:

- The hair lab treatment left my hair feeling soft, smooth, and shiny.
- The high lab workload has caused some researchers to feel overwhelmed and stressed.
- The new lab tin containers are designed to be more durable and long-lasting.
- The hay lab twins were born on the same day and are identical in every way.

Results

The LabTwin user pronouncing sentences with very similar sounding utterances that our branded wake word “Hey LabTwin” would accidentally trigger a recording in 5 to 15% of the cases. This can be prevented by turning off the wake word activation when not actively using the app for an extended period, or selecting a different wake word from our provided list which would be distinct from the user general vocabulary.

The below infographic illustrates the percentage of how many utterances accidentally triggered the LabTwin voice assistant. The results are summarized in the below table (Table 2).

The accidental trigger rate of the user was found to be between 5% to 15% depending on the lab environment and hardware combination.

Table 2: Percentage of utterances that accidentally triggered the LabTwin voice assistant.

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
iPhone 13 Pro + Aftershokz OpenComm	10%	10%	15%	5%
Google Pixel 6a + Aftershokz OpenComm	5%	5%	10%	5%

TESTING PARAMETER 2B

Accidental trigger – different speaker

Test procedure

The accidental trigger of the LabTwin voice assistant by a different speaker were tested using 20 wake word utterances for each hardware combination and lab environment. These activations were conducted at 4 different directions of the secondary speaker in relation to the primary speaker (left, right, face-to-face, back-to-back) and 3 different distances

- Distance 1: 60 cm (1 ft 11 in)
- Distance 2: 120 cm (3 ft 11 in)
- Distance 3: 180 cm (5 ft 11 in)

Resulting in a total of 1920 utterances. You can view the set up of the testing in Figure 2.

The Primary Speaker is our LabTwin user who wears the Aftershokz OpenComm Headset and runs LabTwin mobile app on iPhone 13 Pro or Google Pixel 6a. The secondary speaker is someone in close proximity, that is trying to activate LabTwin voice assistant by using the wake word.

Results

A colleague **using the same "Hey LabTwin" wake word in close proximity**, within two meters in front or on each side of the user, **could unintentionally trigger a recording** in the user app. Interestingly, speaking from behind the user significantly reduces the likelihood of triggering a recording. To avoid such accidental activations, we recommend that users working in close proximity align on different wake words.

The below infographics (Figures 3 and 4) display the percentage of how often LabTwin was triggered by the secondary speaker using the wake word "Hey LabTwin" from the different angles (front, left, right, back) for the two tested hardware in the different lab environments. The tables summarizing the results can be found in the appendix (Tables 3 and 4).

Figure 2: Set up of directional testing with two speakers.

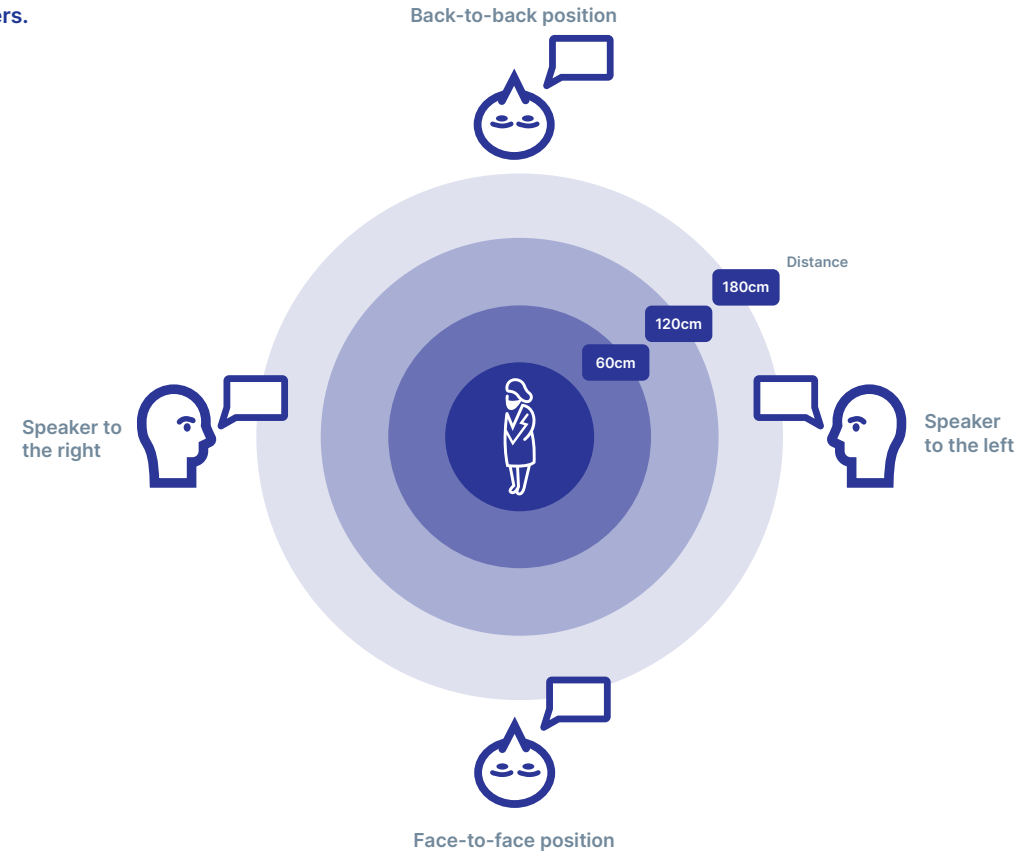
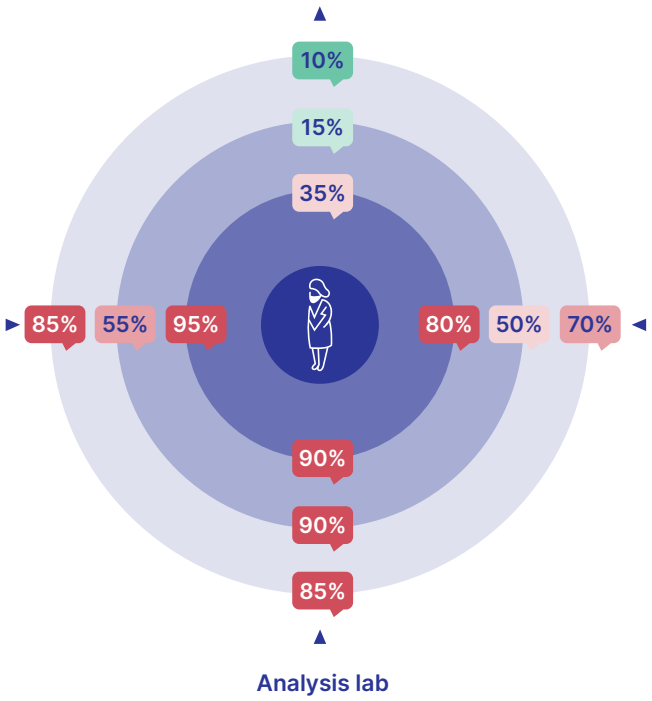
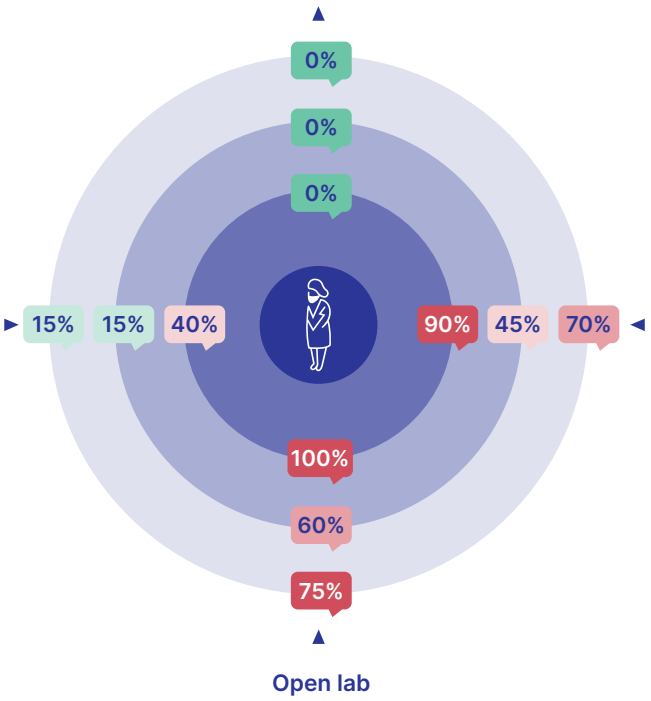


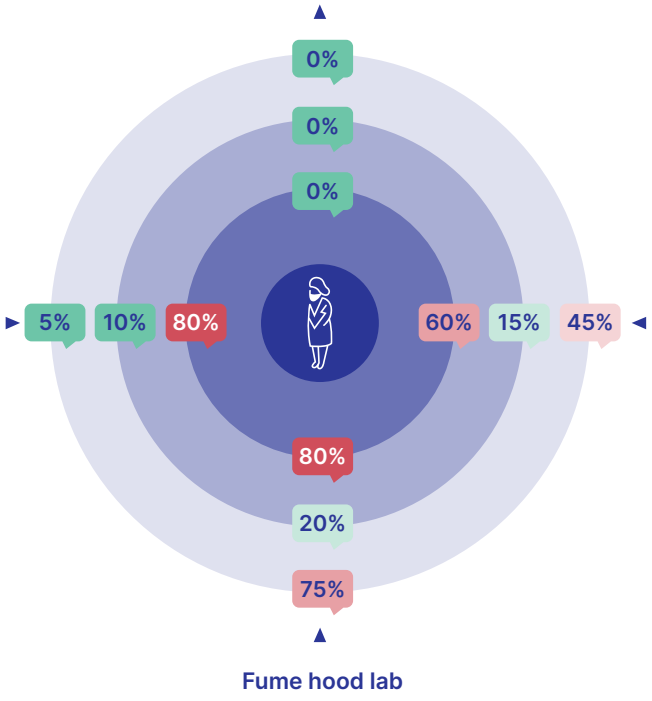
Figure 3: Percentage of how often LabTwin voice assistant was triggered by secondary speaker for iPhone 13 Pro + Aftershokz OpenComm.



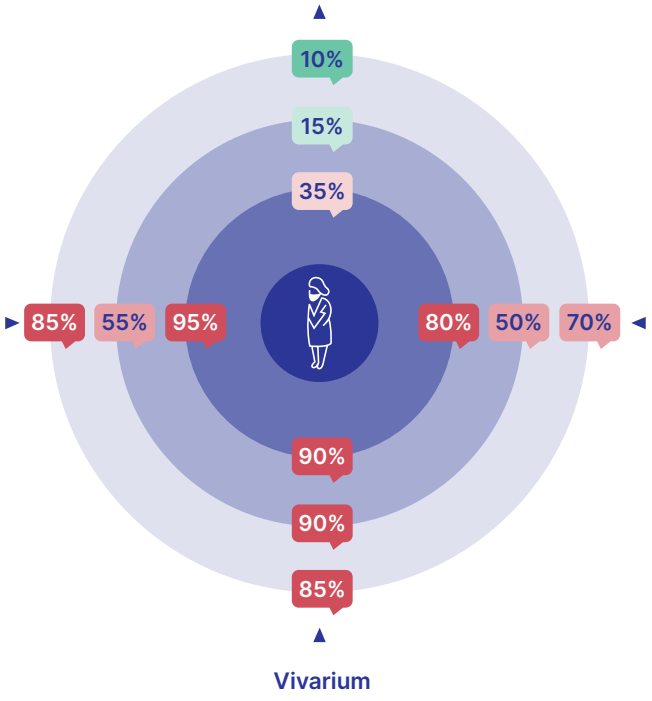
Analysis lab



Open lab



Fume hood lab

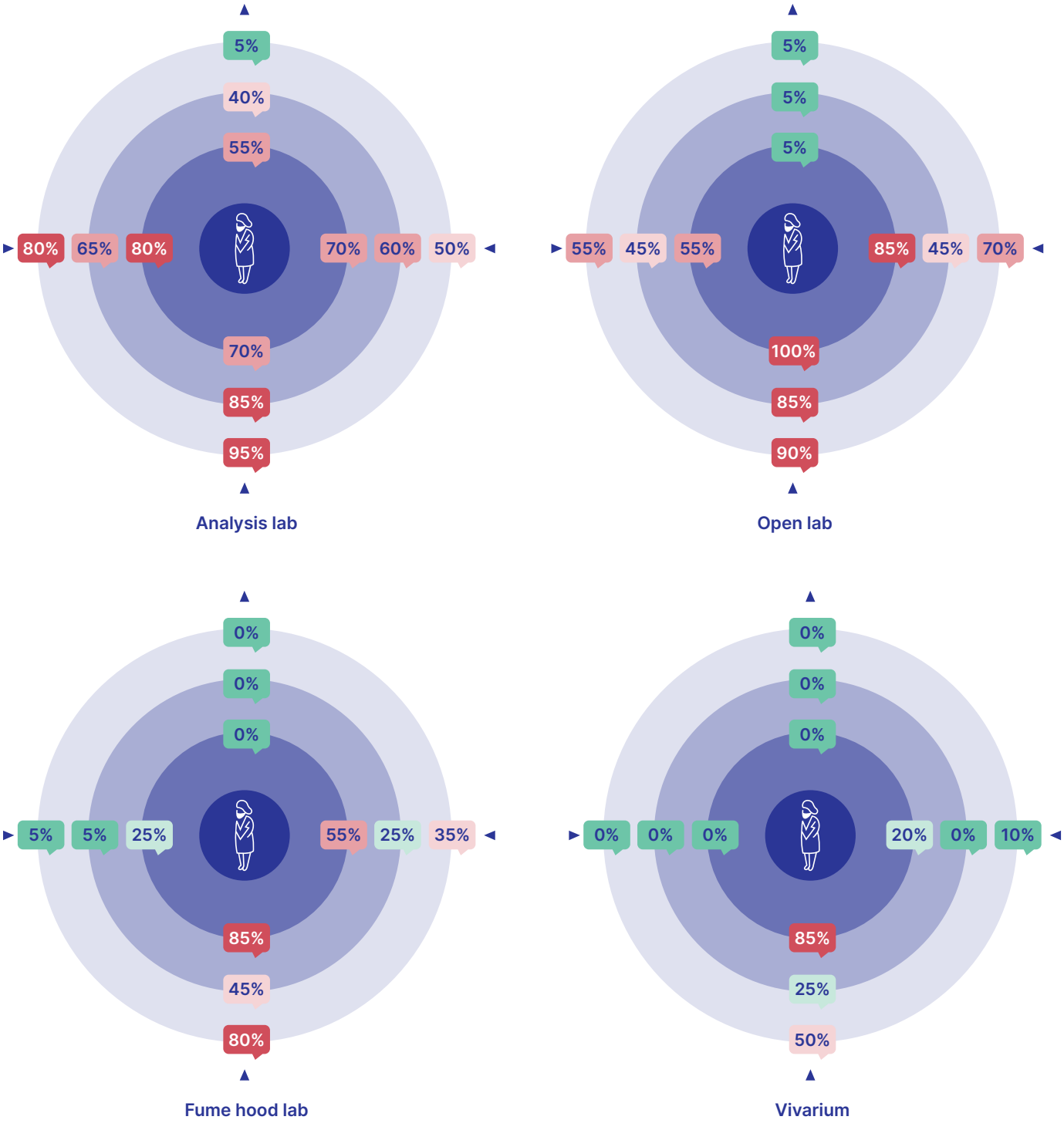


Vivarium

► Direction of view of the second speaker

% Frequency percentage at which LabTwin transcribed the voice of the second speaker at the respective distance

Figure 4: Percentage of how often LabTwin voice assistant was triggered by secondary speaker for Google Pixel 6a + Aftershokz OpenComm



TESTING PARAMETER 3

Other voices

Test procedure

The test parameter of other voices being transcribed unintentionally was tested using 5 test sentences with 3 repetitions each at 4 different directions of the secondary speaker in relation to the primary speaker (left, right, face-to-face, back-to-back) and 3 different distances

- Distance 1: 60 cm (1 ft 11 in)
- Distance 2: 120 cm (3 ft 11 in)
- Distance 3: 180 cm (5 ft 11 in)

This resulted in a total of 1440 utterances per hardware combination and lab environment.

The Primary Speaker is our LabTwin user who wears the Aftershokz OpenComm Headset and runs LabTwin mobile app on iPhone 13 Pro or Google Pixel 6a. The secondary speaker is someone in close proximity, that is speaking at the same time.

Test sentences:

Utterance 1
- Speaker 1: the weight of the sample is 5 grams
- Speaker 2: the pH of the solution measures 7.2

Utterance 2
- Speaker 1: the concentration of the protein sample is 2.5 milligrams per milliliter
- Speaker 2: the melting point of the compound is 142 degrees Celsius

Utterance 3
- Speaker 1: the density of the liquid is 1.02 grams per milliliter
- Speaker 2: the reaction was initiated at 3:15 PM

Utterance 4
- Speaker 1: the duration of the incubation period is 2 hours
- Speaker 2: I centrifuged samples to separate the different layers in the mixture

Utterance 5
- Speaker 1: I incubated the samples at 37°C for 30 minutes
- Speaker 2: I used a western blot to separate and identify proteins

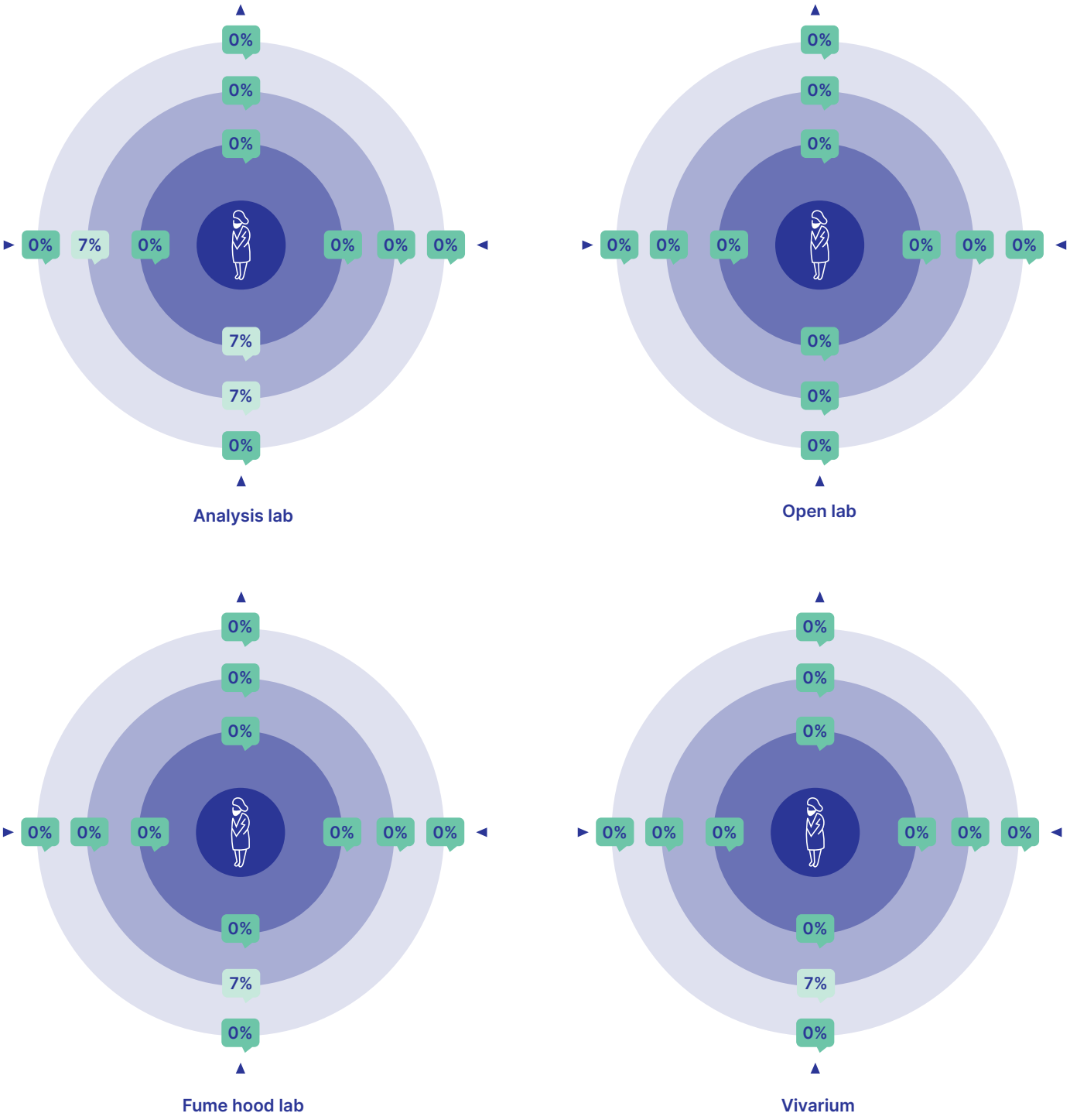
Results

In a shared workspace, the likelihood of **cross-contamination between colleagues' recordings is extremely low**. The test results indicate that the voice of the secondary speaker is only detected when in a face-to-face interaction or within a proximity of less than 60cm from the LabTwin user, and even then, the probability is just 7%. To minimize any potential interference, we recommend avoiding these relatively uncommon situations or waiting for the other person to finish speaking before initiating a recording.

The Figures 5 and 6 display the percentage of how often a cross-contamination was detected from the different angles (front, left, right, back) for the two tested hardware in the different lab environments. The tables (Table 5 and 6) summarizing the results can be found in the appendix.

In order to mitigate the risk of accidental triggers we provide different wake words that users can set when they work in close proximity.

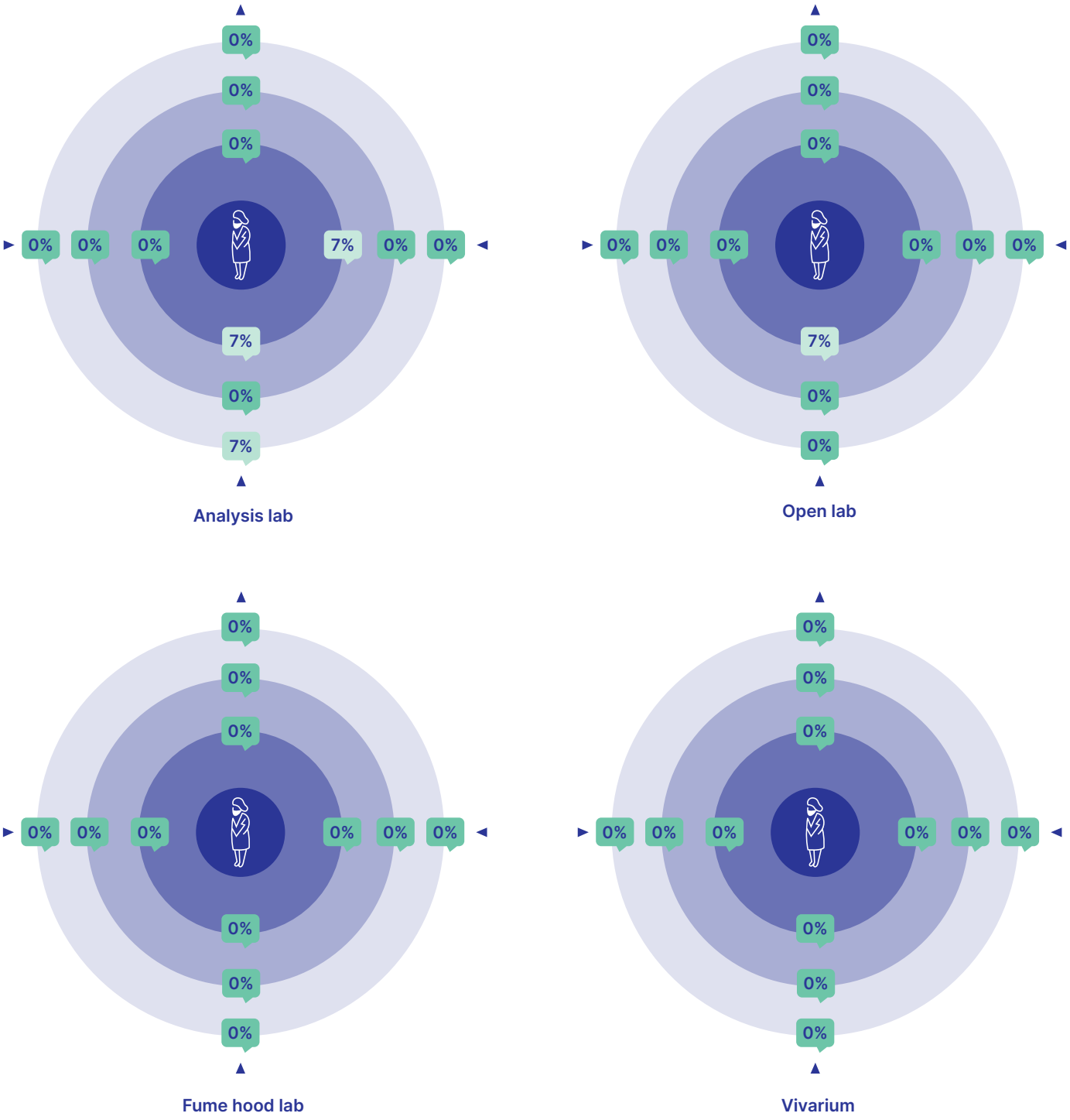
Figure 5: Percentage of how often LabTwin transcribed voice of secondary speaker for iPhone 13 Pro + Aftershokz OpenComm.



► Direction of view of the second speaker

% Frequency percentage at which LabTwin transcribed the voice of the second speaker at the respective distance

Figure 6: Percentage of how often LabTwin transcribed the voice of secondary speaker for Google Pixel 6a + Aftershokz OpenComm



► Direction of view of the second speaker

% Frequency percentage at which LabTwin transcribed the voice of the second speaker at the respective distance

CONCLUSION

96% to 99.4% speech-to-text accuracy

Based on the results of the tests conducted by LabTwin, it can be concluded that the LabTwin voice assistant performs accurately in different lab environments with varying levels of background noise. The speech-to-text accuracy of the voice assistant was found to be **between 96% to 99.4%** depending on the lab environment and hardware combination.

In terms of accidental triggers of the voice assistant by the same user, the test results show that the voice assistant is not highly susceptible to false activations. The accidental trigger rate was found to be **between 5% to 15%** depending on the lab environment and hardware combination.

Furthermore, it appears that LabTwin has almost no difficulty filtering out the voice of the secondary speaker in most cases. In particular, there were no instances where the secondary speaker's voice was transcribed at a distance of 180 cm, regardless of the direction or lab environment. Additionally, even at a distance of 120 cm or closer, LabTwin only transcribed the secondary speaker's voice accidentally in a few cases, and mostly in the face-to-face direction.

However, LabTwin has the most problems with accidental triggers caused by other speakers. Based on the testing conducted, the results show that accidental trigger of the LabTwin voice assistant by a different speaker can occur, especially at close distances and in certain directions. The highest percentage of accidental triggers occurred at a distance of 60cm and facing the primary speaker in the analysis lab and vivarium. The lowest percentage of accidental triggers occurred when the secondary speaker was at a distance of 180cm and facing away from the primary speaker in the open lab. It is worth noting that accidental triggers occurred more frequently in the analysis lab and vivarium environments compared to the open lab and fumehood lab environments. In order to mitigate the risk of accidental triggers we provide different wake words that users can set when they work in close proximity.

Overall, while the LabTwin voice assistant performs well in terms of speech-to-text accuracy and accidental triggers, care must be taken to minimize the risk of unintentional activation by other speakers.

- **LabTwin performs accurately in loud lab environments.**
- **The voices of others are never transcribed from a 180cm distance.**
- **False activation by the user are not highly susceptible (<15%).**
- **The accidental triggering by others can be effectively resolved by utilizing various wake words available in our collection.**

Appendix

Table 3: Percentage of how often LabTwin voice assistant was triggered by secondary speaker for iPhone 13 Pro + Aftershokz OpenComm.

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
→ **Right - 60cm**	95%	40%	80%	95%
→ **Right - 120cm**	55%	15%	10%	55%
→ **Right - 180cm**	85%	15%	5%	85%
← **Left - 60cm**	80%	90%	60%	80%
← **Left - 120cm**	50%	45%	15%	50%
← **Left - 180cm**	70%	70%	45%	70%
→← **Face-to-face - 60cm**	90%	100%	80%	90%
→← **Face-to-face - 120cm**	90%	60%	20%	90%
→← **Face-to-face - 180cm**	85%	75%	75%	85%
←→ **Back-to-back - 60cm**	35%	0%	0%	35%
←→ **Back-to-back - 120cm**	15%	0%	0%	15%
←→ **Back-to-back - 180cm**	10%	0%	0%	10%

Table 4: Percentage of how often LabTwin voice assistant was triggered by secondary speaker for Google Pixel 6a + Aftershokz OpenComm

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
→ Right - 60cm	80%	55%	25%	0%
→ Right - 120cm	65%	45%	5%	0%
→ Right - 180cm	80%	55%	5%	0%
← Left - 60cm	70%	85%	55%	20%
← Left - 120cm	60%	45%	25%	0%
← Left - 180cm	50%	70%	35%	10%
→← Face-to-face - 60cm	70%	100%	85%	85%
→← Face-to-face - 120cm	85%	85%	45%	25%
→← Face-to-face - 180cm	95%	90%	80%	50%
←→ Back-to-back - 60cm	55%	5%	0%	0%
←→ Back-to-back - 120cm	40%	5%	0%	0%
←→ Back-to-back - 180cm	5%	5%	0%	0%

Table 5: Percentage of how often LabTwin transcribed the voice of secondary speaker for iPhone 13 Pro + Aftershokz OpenComm.

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
→ Right - 60cm	0%	0%	0%	0%
→ Right - 120cm	7%	0%	0%	0%
→ Right - 180cm	0%	0%	0%	0%
← Left - 60cm	0%	0%	0%	0%
← Left - 120cm	0%	0%	0%	0%
← Left - 180cm	0%	0%	0%	0%
→← Face-to-face - 60cm	7%	0%	0%	0%
→← Face-to-face - 120cm	7%	0%	7%	7%
→← Face-to-face - 180cm	0%	0%	0%	0%
←→ Back-to-back - 60cm	0%	0%	0%	0%
←→ Back-to-back - 120cm	0%	0%	0%	0%
←→ Back-to-back - 180cm	0%	0%	0%	0%

Table 6: Percentage of how often LabTwin transcribed the voice of secondary speaker for Google Pixel 6a + Aftershokz OpenComm.

	Analysis Lab	Open Lab	Fumehood Lab	Vivarium
→ Right - 60cm	0%	0%	0%	0%
→ Right - 120cm	0%	0%	0%	0%
→ Right - 180cm	0%	0%	0%	0%
← Left - 60cm	7%	0%	0%	0%
← Left - 120cm	0%	0%	0%	0%
← Left - 180cm	0%	0%	0%	0%
→← Face-to-face - 60cm	7%	7%	0%	0%
→← Face-to-face - 120cm	0%	0%	0%	0%
→← Face-to-face - 180cm	7%	0%	0%	0%
←→ Back-to-back - 60cm	0%	0%	0%	0%
←→ Back-to-back - 120cm	0%	0%	0%	0%
←→ Back-to-back - 180cm	0%	0%	0%	0%

LabTwin creates the next generation of integrated, connected smart digital lab tools to save scientists time while increasing data quality and interoperability.

Who We Are

LabTwin

LabTwin GmbH is an AI software company on a mission to create the next generation of integrated, connected smart digital lab tools. Their leading voice-powered digital lab assistant enables hands-free data capture and exchange with informatics systems at the point of experimentation. Both user- and data-centered, LabTwin saves scientists time while increasing data quality and interoperability. LabTwin is a Germany-based company led by an experienced, highly skilled management team and backed by Sartorius and BCG DV. For more information visit: **www.labtwin.com**