

WHITE PAPER

Introduction to audio

Acoustics, speakers and audio terminology

September 2024

Table of Contents

1	Introduction	3
2	Audio frequency	3
	2.1 Audible frequencies	3
	2.2 Sampling frequency	3
	2.3 Frequency and wavelength	3
3	Acoustics and room dimensions	4
	3.1 Echoes	4
	3.2 The impact of room dimensions	4
	3.3 Professional solutions for neutral room acoustics	4
4	Measures of sound	5
	4.1 Human sound perception and phon	5
	4.2 Watts	6
	4.3 Decibels	7
	4.4 Sound pressure level	7
5	Dynamic range, compression and loudness	8
6	Speakers	9
	6.1 Polar response	9
	6.2 Speaker sensitivity	9
	6.3 Built-in digital signal processor	10
	6.4 Speaker types	10
	6.5 Placement of speakers	15
	6.6 AXIS Site Designer	16

1 Introduction

The audio quality that we can experience in a certain room is affected by a number of things, for example, the signal processing done on the audio, the quality of the speaker and its components, and the placement of the speaker. The properties of the room itself, such as reflection, absorption and diffusion, are also central. If you have ever been to a concert hall, you might have noticed that the ceiling and the walls had been adapted to optimize the audio experience.

This document provides an overview of basic audio terminology and of the properties that affect the audio quality in a room. It also presents a background on different speaker types and their optimal placement for an audio installation.

2 Audio frequency

2.1 Audible frequencies

The human ear is, in theory, able to perceive frequencies from 20 Hz to 20 kHz. The upper limit of 20 kHz is lowered with age but the high frequencies can still add "character" through overtones to audio with lower frequencies. Human speech, being complex with lots of harmonies, is scattered over frequencies from around 85 Hz (lowest for human male) to around 8 kHz (overtones for human female). In telephony, only the range of 300 Hz to 3.4 kHz is commonly used, and while it makes the voice audible, the audio will not be as clear as a full frequency range recorded voice.

2.2 Sampling frequency

The sampling frequency is the number of audio "snapshots" taken per second of the analog input audio, in order to digitally reconstruct it. In audio files and CDs, 44.1 kHz is a commonly used sampling frequency, thus using 44,100 samples per second. The sampling frequency must be at least twice as high as the highest input audio frequency that should be reconstructed.

2.3 Frequency and wavelength

There is a simple, inverse, relation between frequency (f , in Hz) and wavelength (λ , Greek letter lambda, in m):

$$\lambda = v/f$$

The wavelength is equal to the speed of sound ($v=340$ m/s in air) divided by the frequency. For quick conversion between wavelength and frequency, there are also online tools that can be used. To provide some examples of audio wavelengths: a frequency of 20 Hz corresponds to a wavelength of about 17 m (56 feet), while a higher frequency of 20 kHz corresponds to a shorter wavelength of about 1.7 cm (0.7 inches). Obviously, there is a wide spread of the wavelengths of audio that we can perceive.

3 Acoustics and room dimensions

3.1 Echoes

In a room that is completely empty, there will be reverb and/or delay in the sound. This is, of course, because all the flat surfaces are perfect for the audio waves to reflect against. If fabrics and uneven surfaces are added, such as sofas, curtains, and carpets, there will be less reverb, but the sound will also be perceived slightly less loud because of the absorption.

Sound waves are often reflected multiple times before reaching our ears. Knowing that the speed of sound in air is around 340 m/s (1020 feet/s), we can calculate the distance that an echo has travelled. If we hear the echo 0.25 s after the initial sound, for example, the sound has travelled around 85 m ($0.25 \text{ s} \times 340 \text{ m/s}$), or 255 feet. For each reflection, the audio fades a little bit until we cannot hear it anymore.

3.2 The impact of room dimensions

The size of the room has a large effect on the audio experience. With wavelengths up to 17 m (56 feet) for the lowest bass, audible sound waves in a small room will be reflected against the walls before the waves have properly developed. This results in resonances and associated standing waves, causing some frequencies to be amplified (higher volume), and others to be attenuated (lower volume). We need a rather large room to hear the bass without distortion.

The impact of resonances on the experienced audio quality increases with the sound volume. With higher volume, the reflections will interfere more with the sound from the source.

In small rooms at low frequencies, the room can be said to dominate the sound, whereas at higher frequencies, the speaker dominates the sound. For small rooms, the room transition frequency is often around 300 Hz. This is the frequency where the audio can be said to transcend from behaving like a wave to behaving like a ray.

3.3 Professional solutions for neutral room acoustics

In order to reduce annoying echoes in large or empty rooms, acoustic panels can be installed in the ceiling, on the walls, or both. The panels are made from sound-absorbing materials and create more neutral acoustics in spaces such as shopping malls, auditoriums, offices, and conference rooms. A similar effect can, however, be achieved by using curtains or other interior fabrics.

Acoustic panels are usually quite effective for frequencies above 300 Hz, while the absorption capabilities gradually decrease for lower frequencies.



Figure 1. Curtains and other pieces of fabric can significantly improve room acoustics.

4 Measures of sound

This section deals with human perception of sound, different measures of sound, and how these relate to each other.

4.1 Human sound perception and phon

Even though the human ear is sensitive to all frequencies between 20 Hz and 20 kHz, the sensitivity varies with the frequency. Sounds of a specific power will thus be perceived as having different loudness at different frequencies. The loudness unit "phon" takes this sensitivity into account and, for example, a sinusoidal tone of 50 phons is perceived as equally loud at all frequencies.

Figure 2 below shows equal-loudness curves. One line represents the sound level that must be used, in order for the sound to be perceived at the same volume for all frequencies. The different lines represent different phon values.

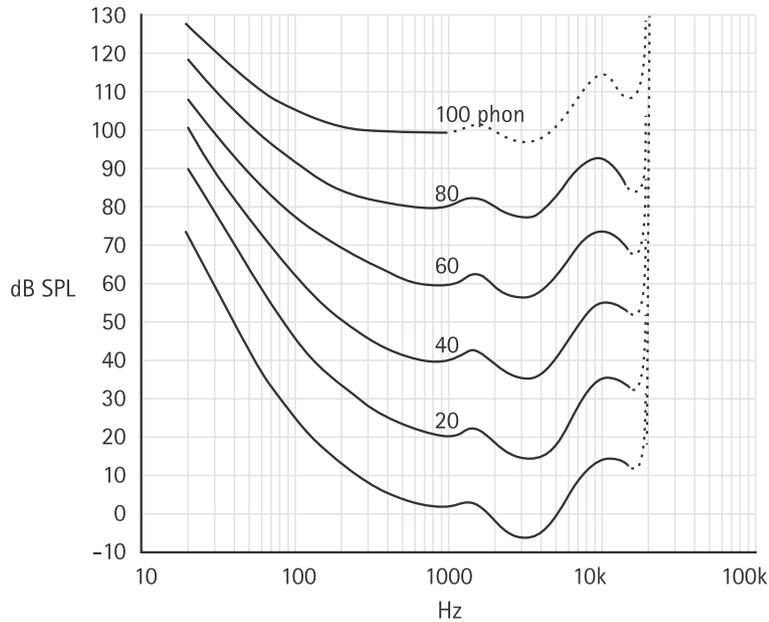


Figure 2. Sound pressure levels needed at different frequencies in order to make a sound perceived as equally loud over all frequencies. The curves are originally from the ISO standard ISO 226:2003.

It is evident from the curves that the sound level must be substantially higher at the lower frequencies in order to be perceived as equally loud as higher frequencies. This is because the human ear is less sensitive to lower frequencies. The minimum of the curves is placed around 2 - 5 kHz, meaning that this is the frequency range to which a human ear is most sensitive, and in which the ear can best decipher a conversation. It is also the frequency range of human speech.

4.2 Watts

The unit of power, watt (W), is familiar from various electrical components, such as light bulbs, laptop chargers, and speakers. The unit can, however, be used in different ways, and in audio terminology we come across varieties like instantaneous power, average power, RMS (root mean square) power, and peak power.

An amplifier might be constructed to be able to deliver 300 W over a very short period of time, such as when a drum, explosion, or any other audio with a short and loud transient, will be heard. This means that the instantaneous power will increase really fast from very low to very high. The same amplifier might, however, only be rated for 50 W continuous use, since continuous use will produce a lot more heat, which impacts both the electrical components and the amplifier's performance.

The human ear does not perceive a 10 W sound to be twice as loud as a 5 W sound. In fact, the sound power has to be 10 times higher (50 W) for the ear to perceive it to be twice as loud. This is where the decibel comes in.

4.3 Decibels

Because sound is perceived non-linearly, it is best measured and described using the non-linear unit decibel (dB). A doubling (measured in W) of the sound power equals to a 3 dB increase, and a doubling of the loudness equals a 10 dB increase. Figure 3 shows familiar sound sources and their power levels in dB.

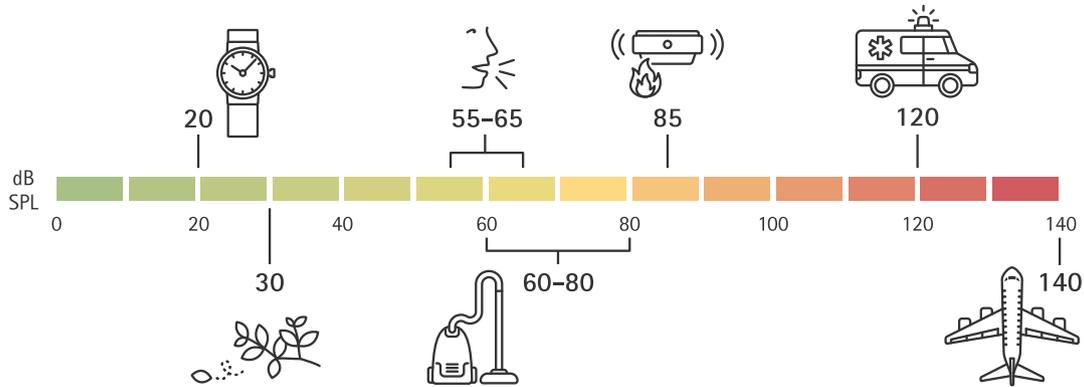


Figure 3. Approximate sound levels, in decibel, from familiar audio sources.

A sound pressure level given in the weighted dBA scale has been compensated for the human ear's frequency-dependent perception of sound, as discussed in section 4.1. Using the unweighted dB scale, a 100 dB level at 100 Hz will, for example, be perceived to have a loudness equal to only 80 dB at 1 kHz, while 100 dBA will be perceived as equally loud at all frequencies.

The decibel unit is often referring to a relative change in loudness. For expressing an absolute value, dB SPL should be used. A value of 0 dB SPL is the softest sound that the human ear can perceive.

4.4 Sound pressure level

Sound pressure level (SPL) is the RMS value of the instantaneous sound pressures measured, in dB, over a specified period of time. SPL is not a constant average value of loudness but rather an average of the short peak values.

An SPL value given for a speaker is assumed to be measured for a 1 kHz tone at a distance of 1 m, if nothing else is stated.

The sound pressure level of an audio source decreases with the distance from the source. Defined to start at 0 dB at 1 m from the source, the SPL decreases by 6 dB with each doubling of the distance from

the source, as illustrated in Figure 4. However, for more detailed information about the sound levels of a certain speaker, we need to look at its polar response as exemplified in section 6.1.

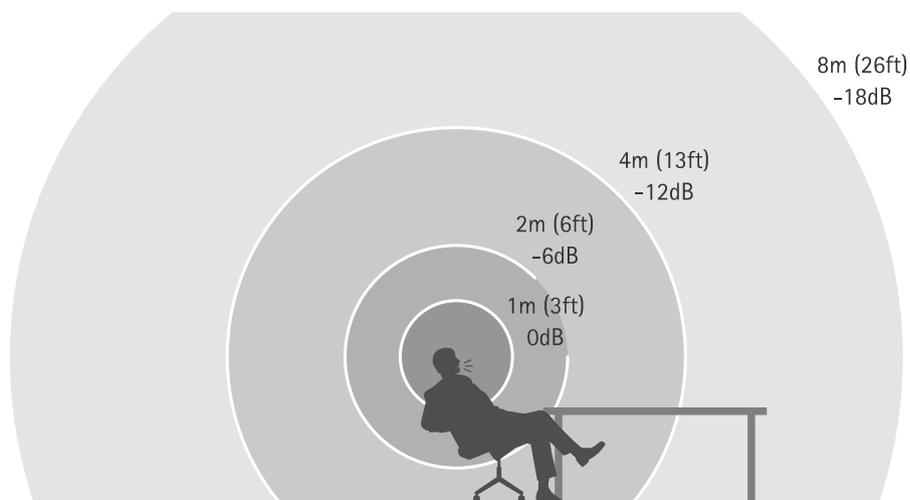


Figure 4. The sound pressure level from an audio source decreases by 6 dB with each doubling of the distance from the source.

5 Dynamic range, compression and loudness

The recording has a large dynamic range, meaning that there are large differences between the quietest and the loudest part.



Figure 5. Visualization of a large dynamic range recording, without compression.

The quietest parts become louder, while the loud parts either stay the same or become less loud. The differences between peaks and dips are smaller, which makes us perceive this recording as louder. As can be seen in Figure 6, the dynamic range is decreased.



Figure 6. Visualization of the same recording as above, now after compression.

Compression of dynamic range is often applied in audio systems for restaurants, retail, and similar public environments that play background music at a relatively low volume. Apart from making the volume more constant, the compression also makes the quieter parts of the audio more audible over ambient noise.

6 Speakers

A speaker can have different physical shapes depending on its purpose. The component that distributes the audio, the speaker driver, is usually cone-shaped but can have other form factors if it should reconstruct high frequencies. Some speakers have a very narrow direction of sound in order to achieve a high sound pressure in one direction. Others are made to have as wide spread of the sound as possible. A speaker's ability to reconstruct audio signal is dependent on the frequency of the audio signal.

6.1 Polar response

The polar diagram in Figure 8 shows how different frequencies spread out differently from a generic example speaker, placed in the center of the diagram. It shows that lower frequencies have a wide spread (even behind the speaker, at 180 degrees) while higher frequencies are more directional.

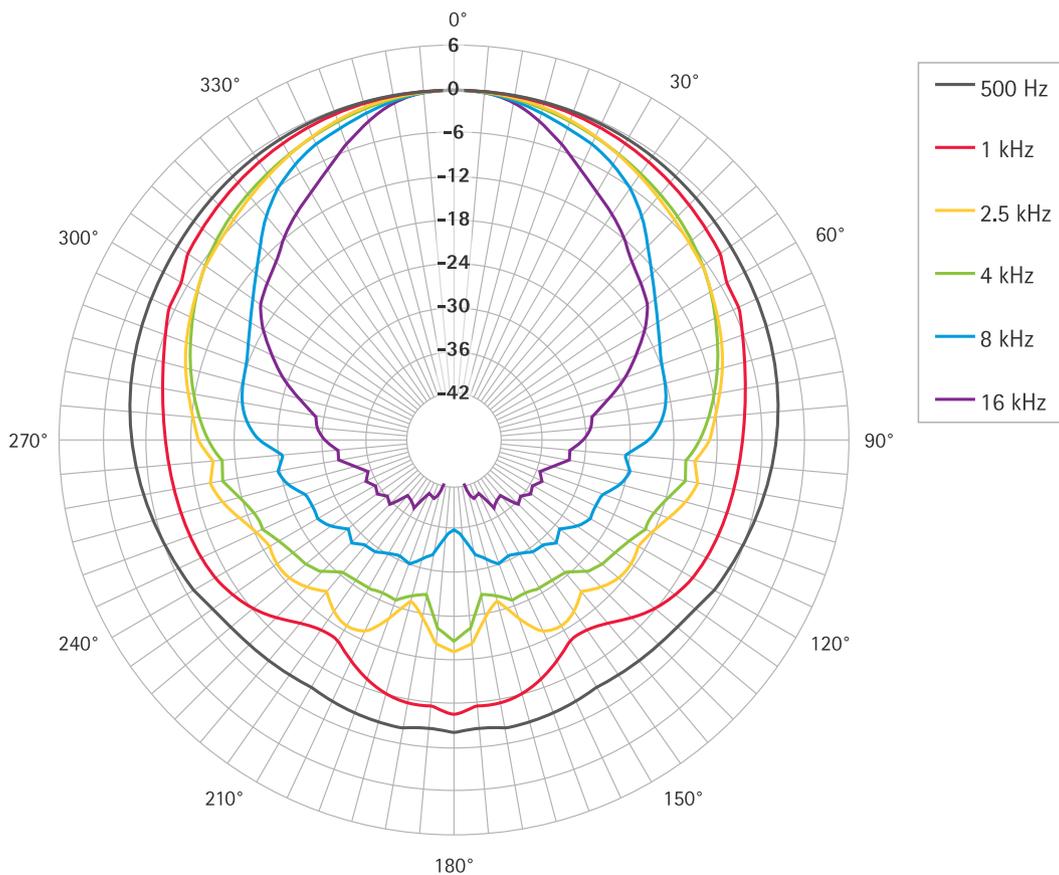


Figure 7. A polar diagram showing the spread from a generic example speaker (which was located in the center of the diagram). The lower frequencies have a wider spread (even behind the speaker, at 180 degrees) while higher frequencies are more directional.

6.2 Speaker sensitivity

A speaker's sensitivity is its ability to reproduce sound when fed a certain power. Determining the sensitivity is usually done by feeding an audio signal of 1 W (typically at 1 kHz) and then measuring the sound pressure level in dB SPL at a 1 m distance. Common values for speakers are around 85 - 92 dB SPL. The higher the sensitivity, the louder the sound will be from the speaker when fed a certain power.

For analogue speakers, the sensitivity of the speaker is usually an indicator of the quality of the speaker. Lower sensitivity indicates a less powerful magnet and/or a smaller and cheaper coil. Therefore, in regards to audio quality a 10-inch speaker is not necessarily better than an 8-inch speaker.

However, for digital speakers, the amplifier is incorporated into the speaker. The measurement of the speaker's sensitivity is not vital for determining the speaker's quality.

6.3 Built-in digital signal processor

All Axis speakers have integrated amplifier and digital signal processor (DSP) for preconfigured sound quality. These ensure that the speakers can be used by anyone without needing audio expert to produce good sound. DSP analyzes and processes audio signals to improve speech intelligibility.

With a built-in DSP, Axis speakers filters background noise away and balances the audio frequency to enhance tone quality. It also compresses the dynamic range of an audio signal. An audio signal will often have peaks and troughs in volume, and dynamic range control can balance these to make sure that sound is broadcast at the ideal volume for listeners.

DSP compensates for quiet sounds that are less perceptible to the human ear at low volume. It boosts the frequency of such sounds to make sure that the listener doesn't miss anything. Furthermore, it processes, stores, and transmits audio digitally from the source to the speaker. This improves the sound quality and maintains signal strength, ensuring that the sound is well optimized for the speakers. The sound profiles for background music and voice are predefined so you don't have to manually control the audio quality.

6.4 Speaker types

Form factors, sound pressures, and mounting possibilities vary – some speaker types are optimal for conveying clear and audible announcements in noisy outdoor areas, while others work better in small spaces.



Figure 8. Axis speakers.

6.4.1 The hi-fi speaker

In hi-fi equipment, so-called '2-way' or '3-way' speakers are common. These speakers use several different speaker drivers, in order to accurately reproduce as many frequencies between 20 Hz and 20 kHz as possible. One driver might be responsible for reproducing sound up to 500 Hz, a second one for frequencies from 500 Hz to 9 kHz, and a third for frequencies above 9 kHz. These border frequencies are called 'crossover frequencies'. A hi-fi speaker is designed to reproduce audio very accurately at high loudness.

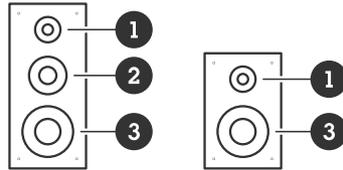


Figure 9. Hi-fi speakers. Treble driver (1), Mid-range driver (2) and Bass driver (3).

6.4.2 The horn speaker

The horn speaker has a completely different usage than a hi-fi speaker, and should not cover a large frequency range. Its purpose is instead to maximize the loudness of those frequencies to which the human ear is the most sensitive, so that the speaker can convey a message (a human voice or a siren, for example) as clearly as possible. The horn directs all sound in one direction, which further enhances its sound pressure.



Figure 10. Network horn speaker

6.4.3 Multipurpose speaker

Multipurpose speakers are easy to integrate and have all-in-one solutions you can use for live or prerecorded voice messages to give safety instructions or warn intruders off. You can also use a multipurpose speaker to play background music. Axis portfolio includes various multipurpose speakers:

6.4.3.1 The cabinet speaker

An Axis network cabinet speaker provides a medium sound pressure level. It can be used in most indoor areas, but is less optimal in very noisy environments. It can also be used semi-outdoors, which means it can be mounted below a roof that protects it from heavy rain. The cabinet speaker can be mounted horizontally or vertically, on a wall, in a ceiling, or with a pendant kit.



Figure 11. A cabinet speaker

6.4.3.2 The ceiling speaker

An Axis network ceiling speaker provides a medium sound pressure level and should be used in less noisy indoor or outdoor areas, such as hospitals, retail stores, or office buildings. It can be mounted in a drop ceiling where it will be very discreet and physically well-integrated.



Figure 12. A ceiling speaker

6.4.3.3 The Pendant Speaker

An Axis network pendant speaker has a medium sound pressure level and is suitable for less noisy indoor areas with high ceilings. It comes in two sizes, and the cable length can be adjusted to fit any high ceiling.



Figure 13. A pendant speaker.

6.4.3.4 The mini speaker

An Axis network mini speaker provides a low sound pressure level and should be used in quieter indoor areas. It is small and discreet and fits into small spaces or corridors, where it can be surface mounted on a wall or ceiling. It has a wide audio coverage which means that you need fewer speakers. The mini speaker has a built in PIR sensor for motion detection, which can be set up so that the speaker automatically plays an audio message when someone is approaching.



Figure 14. A mini speaker.

6.4.3.5 The Sound Projector

An Axis network sound projector has a high sound pressure level and natural, rich sound. This means that a message can be conveyed as clearly as possible, but that background music will sound good too. A sound projector can be used in outdoor installations or noisy indoor areas and can be mounted on a pole, wall, or ceiling. It can be installed in easy-to-reach locations where the risk for vandalism is higher

– the sound projector is vandal-resistant and also has a sleek, minimalistic design that easily blends into the environment.



Figure 15. A sound projector.

6.5 Placement of speakers

There are many possible ways to place the speakers. The general rule is to, if possible, always point the sound along the room. That is, if you have a rectangular room, try to place the speakers on the short walls pointing out along the longer walls. This will let the sound spread as far as possible before being reflected on the walls. However, it is not recommended to place a speaker in a corner, since that would unevenly amplify the bass sound.

6.5.1 The cluster placement

If you prioritize simple and low-cost installation, you can install the speakers in clusters. This will minimize cabling, but might not be the best way to get a good spread of the sound.

6.5.2 The wall placement

If the room dimensions allow, and you do not mind the extra cabling, a wall placement solution will probably spread the sound better. With the same number of speakers as in the cluster placement example, the installation might look like the below figure. If the room is large, however, the reach of the speakers might be too short.

6.5.3 The ceiling placement

If the room has a drop ceiling, or if it is possible to install built-in ceiling speakers, a ceiling placement can be a discreet solution. However, this placement is very sensitive to the ceiling height. The lower the ceiling, the more speakers you need in order to cover a certain area.

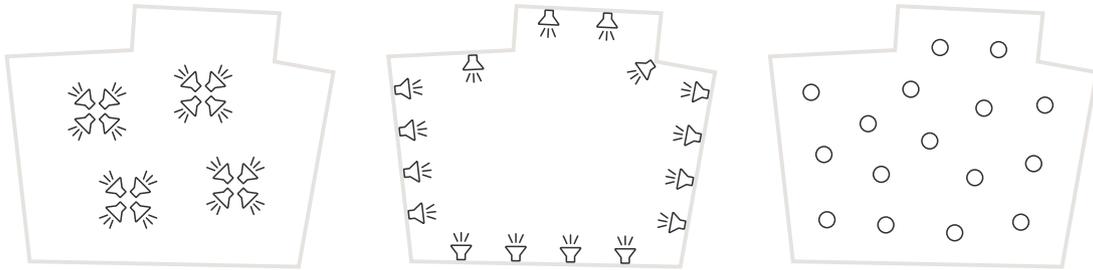


Figure 16. Cluster, wall, and ceiling placements of speakers.

6.6 AXIS Site Designer

AXIS Site Designer (<https://sitedesigner.axis.com>) is a helpful online tool for planning and designing an audio installation (as well as a video installation), including which speakers to use, how many speakers are needed, their optimal placement, and so on, with regard to the conditions at the site.

About Axis Communications

Axis enables a smarter and safer world by creating solutions for improving security and business performance. As a network technology company and industry leader, Axis offers solutions in video surveillance, access control, intercom, and audio systems. They are enhanced by intelligent analytics applications and supported by high-quality training.

Axis has around 4,000 dedicated employees in over 50 countries and collaborates with technology and system integration partners worldwide to deliver customer solutions. Axis was founded in 1984, and the headquarters are in Lund, Sweden