

# Technical guide

Kantar podcast measurement system Norway

Change history:

<b>Version</b>	<b>Change</b>	<b>Date</b>	<b>Author</b>
0.1	Initial draft	07.08.2019	MC
0.2	Updated with GDPR	10.10.2019	MC
1.0	Misspelling and clarification	23.04.2020	MC

Contents

- 1. Introduction..... 3
  - 1.1 Audience..... 3
  - 1.2 Background..... 3
- 2. System overview ..... 4
  - Figure 2.2 Simplified platform overview ..... 4
- 3. Requirements ..... 5
  - 3.1 Requirements for podcast producers..... 5
  - 3.2 Requirements for podcast publishers ..... 5
  - 3.3 Log file required fields ..... 6
- 4. Metadata ..... 7
- 5. GDPR compliance ..... 8

# 1. Introduction

## 1.1 Audience

This documents primary audience is organizations or individuals who want to join the Norwegian podcast measurement and need to learn how to comply. This paper would also be of interest if you want to learn more about the methods used.

## 1.2 Background

In Q2 2019 Kantar won the contract for measuring podcast listening in Norway. The measurement itself is owned by a coalition of major broadcasters in Norway, including NRK, VGTV, Moderne Media, Bauer and P4.

Kantar is delivering the platform conforming to the IAB Podcast Measurement Guidelines\*. The platform is developed using Microsoft .Net and Azure.

\* [IAB Podcast Measurement Guidelines](#)

## 2. System overview

### 2.1 Short summary of functionality

Podcast publisher platforms normally produce standard log files which they upload regularly to Kantar. If the log file contains the necessary fields necessary to comply with the IAB-standard, Kantar will validate the date and by the end of each week make a summary for an official top list. This list is published to [podtoppen.no](http://podtoppen.no).

### 2.2 System architecture

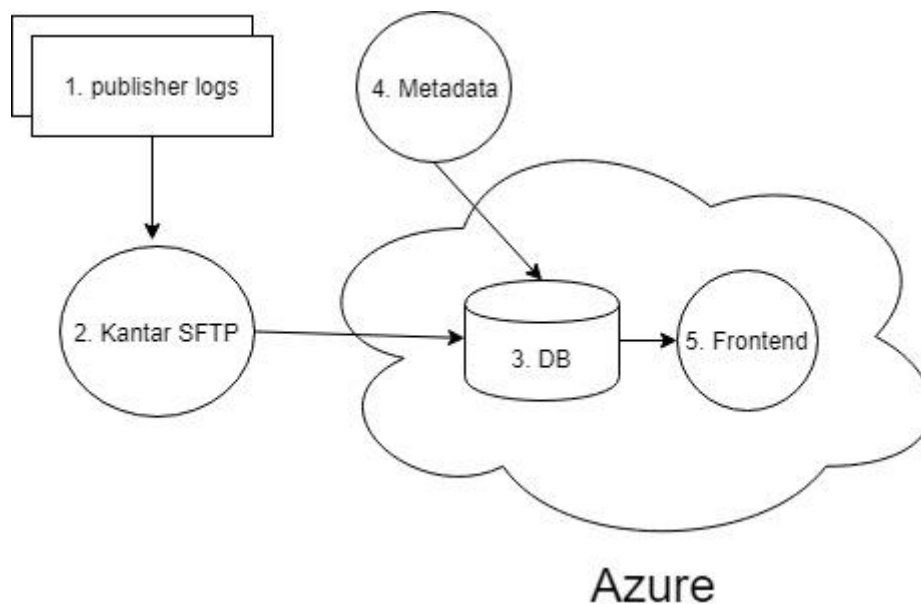


Figure 2.2 Simplified platform overview

## 3. Requirements

### 3.1 Requirements for podcast producers

For a podcast to be included in the measurement two criteria's needs to be fulfilled:

1. Someone, typically your publisher, needs to **register the podcast** on the Kantar podcast measurement maintenance page. This is done by registering the podcast RSS URL.
2. The publisher for the podcast need to **upload log files to Kantar** containing entries for your podcast.

### 3.2 Requirements for podcast publishers

The Measurement Steering Committee needs to approve all new publishers.

The actual method used in the measurement is based on Kantar filtering and validating HTTP GET request in the publisher server log. To set up scheduled deliveries of these log files you need to follow these steps:

#### **Step 1**

Publisher supplies Kantar with a sample log file. Kantar validates the sample log file and if necessary works with the publisher adjusting the log until format and content applies with Kantars platform. See log format description in chapter 3.3.

#### **Step 2**

Publisher registers the eligible podcasts in the Kantar measurement maintenance page. Credentials will be supplied individually from Kantar when this step is reached.

The reporting window is one week from Sunday 00:00. Therefore, you need to upload the last reporting day Sunday latest by 08:00 Monday morning (subject to change).

#### **Step 3**

Publisher configures an automated upload of log files to Kantar podcast measurement SFTP service. Credentials will be supplied individually from Kantar when this step is reached. Files are to be uploaded on a daily basis. The reporting window is one week from Sunday 00:00. Hence, the publisher needs to upload the server logs from the last reporting day Sunday by 08:00 Tuesday morning (subject to change).

### 3.3 Log file required fields

To comply with the IAB standards and the Norwegian measurement the publisher log file needs to contain a minimum of the following fields:

Field	Function
Timestamp	Full timestamp, including date and time, or just time.
Date	Date if above timestamp only contains time.
http status code	Used for filtering out HTTP 200, 206 and 304 requests.
http method	Used for filtering out HTTP GET requests.
IP address	Used to identify device. Represented either by a 4 octet IP-address or a unique hashed value. Kantar validating system will automatically hash full IP-addresses if not received hashed. To comply with GDPR.
Useragent	Used to identify device in combination with IP.
URL	Relative path for the podcast MP3 file.
Bytes	Used to identify size of the requested file. For HTTP status code 200, this is indicated by a singular value. For HTTP status code 206 this is indicated by a from-to value.

## 4. Metadata

Podcast metadata comprises of the podcast's underlying RSS feed, which must conform to the RSS standard for podcasts. The Kantar system will crawl the RSS URL and extract relevant data.

Necessary external metadata required is the podcast's bitrate. This is needed in order to comply to the IAB standard for validating log entries identified by HTTP status-code 206 (partial download). The IAB standard dictates that partial download requests must include data representing a minimum of 60 seconds of audio. In order to calculate correctly, the actual bitrate must be identified as *KiloBytes per Second* (kB/s).

However, in order to have a robust method, Kantar has established a four level hierarchy of bitrate assignment (priority order as follows):

1. Through the RSS feed: Publisher includes a <bitrate> tag in the RSS XML for the assigned podcast. Example: <bitrate>192</bitrate>
2. Through metadata registration: The user specifies the bitrate when registering the podcast RSS feed
3. Publisher specifies a default bitrate applying to all its podcasts
4. A global default measurement bitrate is set to 128kB/s

This is the bytes-to-seconds formula used for the calculation:

$$[\text{bytes}/(\text{bitrate}/8)]/1000 = \text{seconds}$$

## 5. GDPR compliance

IP-address is the only GDPR related data in the dataset. To comply to GDPR regulations, the following procedure is followed:

1. Publishers are to hash all IP-addresses in server logs prior to transferring them to Kantar.
2. Upload process is done on an HTTPS secured line.
3. If producer delivers non-hashed IP's, Kantar will hash the IP-addresses before saving any data in database.
4. The received log files are stored on a local Kantar server for 14 days and then automatically deleted permanently.

The combination of the hashed IP-address and the "User agent" is used to identify unique devices. This combination is not enough information to identify an individual.