**ABSTRACT:**

Distribution system operators will gather an increasing amount of data from the lower level of the grid to be able to cope with several challenges, such as an increase of distributed energy production, local storage, and electric vehicles. Most DSOs are not yet prepared for collecting, storing and processing these large amounts of data. This paper introduces a method for designing, validating and implementing such a system. The most important aspect of this method is the fact that the domain expert is able to validate the constructed conceptual model, before it is used to create a working system. This validation step adds to the quality of the model, and therefore the resulting system. We have applied our method to a use case where a DSO researcher wants to answer questions such as "Can we recognize which appliances are present in households?" and "How can we cluster similar households?". Initial results show that a clear distinction can be made between data that is easily handled by flat files or a single database server, and high volume, rapidly changing data that is much more suitable to be stored in a distributed database.