

ANALYSIS OF DATA QUALITY ISSUES IN HETEROGENEOUS ENVIRONMENT

SAPNA PUJARA¹ & KANWAL GARG²

¹Research Scholar, DCSA, Kurukshetra University, Kurukshetra, India

²Assistant Professor, DCSA, Kurukshetra University, Kurukshetra, India

ABSTRACT

Amid the growth of distributed computing, the collection and maintenance of data from diverse and heterogeneous sources is the major necessity of each organization. Organizations need to exchange and share the information across the globe. In such a scenario, the databases are intrinsically scattered and managed by different people with different objectives, which intensifies the diversity between the technologies and standards being used at each site. Such Heterogeneity is actually attributed to syntactic and semantic differences while describing same real world entity in various data sources. Moreover, such heterogeneity leads to some security & privacy issues. In this keynote paper an in-depth investigation of primary issues (syntactic, semantic), succeeding issues (preprocessing, integration & transformation) and consequential issues of privacy in such heterogeneous environment.

KEYWORDS: Data Quality, Heterogeneous Environment