Kareng Spärck Jones, "A Statistical Interpretation of Term Specificity and Its Application in Retrieval" (1972)
In this case of precise exploration of information retrieval, Jones Spärck explains a more quantitative view of specificity as opposed to purely semantic, claiming that a term's discriminating power is determined by its usage prevalence across collections. She posits the now widespread count-based weighting IDF, stating that matches on rarer terms should be classified as more important than matches on common terms. Using the three test corpora Cranfield, INSPEC, and Keen, she demonstrates that blindly removing high frequent terms increases recall, and that simple frequency-based weighting greatly improves precision and recall. The elegance of mathematics is shown to trump the sophistication of language processing, leading to the aforementioned foundation of TF IDF scoring relied on by modern web search engines.

Edgar F. Codd, "A Relational Model of Large Shared Data Banks" (1970)
Codd's Turing award paper suggests data be stored as relations which are sets of tuples defined over named domains blurring the logical structure with physical storage regions. After critiquing the most common hierarchical and network systems concerning their ordering, indexing, and access path dependency dominated for: order, Codd added his first relational innovations, primary key and foreign key, as well as projection, join, and normalization (the employee tree turned into flat tables). The model's ease is exemplified by tables like supply while first order predicate calculus permits symmetric, declarative queries. Codd remarks that this abstraction guarantees independence and consistency of data, a claim that was later achieved in System R, Ingres, SQL, and almost all commercial RDBMS.

Discussion question:
Both papers substitute for ad hoc data handling with abstraction defensible on rigorous mathematical grounds (IDF weighting; relational algebra).
>What steps can current practitioners take to extend this legacy---such as formalizing trust signals in ranking or integrity constraints for distributed databases---to deal with the scale and complexity of contemporary information systems?