ProApprox is a query processor over probabilistic trees that represents a first step towards building a fully-fledged probabilistic semistructured data management system. It relies on a generalization of the different uncertain data models in XML proposed in the literature and allows for efficient data querying with a subset of the XPath query language, through techniques of exact calculations or efficient approximations of the result.

**Probabilistic Data**

- **Boolean query languages on trees:**
  - **Tree Pattern Queries**
    - A
    - B
    - C
    - D
  - **Tree Pattern Queries with joins**
    - A
    - B
    - C
    - D

- **Processing the Query:**
  - **Exact Computation:**
    - Use the $w_i$'s of the mappings to run exact computation (whenever possible).
  - **Additive Approximation:**
    - Draw values for the $w_i$'s of the mappings. Evaluate the mappings. Use this process $n$ times running Additive App.
  - **Multiplicative Approximation:**
    - Evaluate Multiplicative App. formula using mapping sequences for draws and evaluation.

- **What we aim for:** A probabilistic DBMS using XML technology capable of efficiently querying discrete probabilistic data models.
- **And...** deal with aggregate queries. The result of a query that makes use of aggregate functions is a set of possible values (for each possible document), each with its probability.
- **Also...** move to a distributed framework to manage probabilistic data, i.e., in an open file sharing environment or in the case of data integration.
- **Update operations also belong to future implementation perspectives.**