Algorithm 1 The WimuQ source selection

Input: $\mathbf{Q}$ ▷ The SPARQL query
Output: $\mathbf{E}, \mathbf{H}, \mathbf{T}, \mathbf{N}$ ▷ Hash sets of relevant sources of SPARQL endpoints, HDT files, dataset with dereferenceable URIs, datadump with non-dereferenceable URIs, respectively

$\mathbf{U} = extractURIs(\mathbf{Q})$ ▷ Extract the URIs from the SPARQL query
for all $\mathbf{u} \in \mathbf{U}$ **do**
    $\mathbf{D} = \mathbf{D} \cup wimuLookup(\mathbf{u})$ ▷ Datasets from the WIMU
end for
for all $d \in \mathbf{D}$ **do** ▷ For each dataset
    if d is SPARQL endpoint **then**
        for all $t \in \mathbf{Q}$ **do** ▷ For each Triple pattern in query
            b = ASK(t,d) ▷ SPARQL ASK of t in d
            if b = true **then**
                E.add(d)
            end if
        end for
    end if
    if d is HDT file **then**
        H.add(d)
    end if
    if d is dataset with dereferenceable URIs **then**
        T.add(d)
    end if
    if d is datadump with non-dereferenceable URIs **then**
        N.add($RDFSlice$(d)) ▷ only add the relevant slice
    end if
end for
return $\mathbf{E}, \mathbf{H}, \mathbf{T}, \mathbf{N}$ ▷ final relevant sources set