Appendix - Effect of cumulatively removing properties

Table 1: Effect of cumulatively removing properties, trigram similarity threshold $\sigma = 0.835$, Reuters-21578, $d = 2$, Prefix: dbpedia-owl.

<table>
<thead>
<tr>
<th>Property</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>owner</td>
<td>0.675</td>
</tr>
<tr>
<td>isPartOf</td>
<td>0.672</td>
</tr>
<tr>
<td>countySeat</td>
<td>0.670</td>
</tr>
<tr>
<td>city</td>
<td>0.666</td>
</tr>
</tbody>
</table>

OwningCompany, party, hometown, religion, memberOfParliament, occupation, influencedBy, locationCity, type, deputy, aircraftHelicopter, spouse, usingCountry, architect, knownFor, place, residence, governor, industry, timeZone, influenced, league, division, deathPlace, team, populationPlace, keyPerson, bandMember, operatedBy, officialLanguage, broadcastNetwork, march, education, season, regionalLanguage, battle, aircraftRecon, stylisticOrigin, militaryUnit, predecessor, wikiPageDisambiguates, leaderFunction, youthWing, derivative, fourthCommander, nationality, leaderParty, languageRegulator, leftTributary, parentCompany, language, country, militaryBranch, mayor, athletics, languageFamily, formerBandMember, largestCity, programmeFormat, state, location, territory, musicSubgenre, affiliation, aircraftTransport, associatedBand, garrison, primeMinister, part, otherParty, restingPlace, locationCountry, currency, aircraftElectronic, foundationPlace, related, neighboringMunicipality, award, relation, chairman, foundedBy, distributingLabel, isPartOfMilitaryConflict, colour, secondCommander, commander, ideology, lieutenant, formerTeam, broadcastArea, product, commandStructure, child, ground, routeStart, headquarter, spokenIn, profession, musicFusionGenre, recordLabel, anthem, successor, notableCommander, regionServed, governingBody, operator, service, hubAirport, militaryRank, aircraftTrainer, president

parent, jurisdiction, vicePresident, district, genre, pictureFormat, capital, birthPlace, region, governmentType, leaderName, leader, ethnicGroup, targetAirport, campus, monarch, thirdCommander, subsidiary, twinCity, almaMater, instrument, internationalAffiliation

| managerClub                           | 0.661    |
| manager                               | 0.663    |
| sisterStation                         | 0.664    |
Appendix - Algorithm for expansion policy

**Algorithm 1**: Search candidates for named entities and expansion policy

**Data**: $N = \{N_1, N_2 \ldots N_n\}$ sorted in ascending order of their string length, trigram similarity threshold $\sigma$

**Result**: set of candidates $C$

begin

heuristicExpansion ← $\emptyset$, $C$ ← $\emptyset$;

for $N_i \in N$ do

label ← string($N_i$);

tmp ← label;

expansion ← false;

for key ∈ heuristicExpansion do

if key contains label then

if tmp.length $>$ key.length $\&\&$ tmp $\neq$ label then

    tmp ← key;

    expansion ← true;

if tmp.length $<$ key.length $\&\&$ tmp $==$ label then

    tmp ← key;

    expansion ← true;

label ← tmp;

if ¬expansion then

    heuristicExpansion ← label $\cup$ heuristicExpansion

end

$C$ ← $C$ $\cup$ searchCandidates(label,$\sigma$);

end
Appendix - Figures of experimental results

Figure 1: Accuracy of AGDISTIS on the RSS 500 corpus.
Figure 2: Accuracy of AGDISTIS on the news.de corpus.
Figure 3: Comparison of the average runtime of AGDISTIS and AIDA on Reuters-21578 corpus with respect to the number of entities per sentence. AGDISTIS is clearly more time-efficient than AIDA.