V oblasti počítačového zpracování přirozené jazyka není slovotvorba v porovnání s (flektivní) morfologií dostatečně pokryta jazykovými zdroji. Již existující zdroje zachycující slovotvorbu se navíc liší v mnoha aspektech. V rámci této diplomové práce jsou popsány jak existující jazykové zdroje zachycující slovotvorbu napříč jazyky, tak sjednocení (harmonizace) jejich datových struktur a souborových formátů. První dvě kapitoly uvádí základní pojmy z oblasti slovotvorby a zároveň detailní přehled a kvantitativní i kvalitativní srovnání existujících jazykových zdrojů slovotvorby. Jádro diplomové práce tvoří popis harmonizačního procesu a jeho aplikace na vybrané zdroje. Jsou představena nejen kritéria výběru, ale také základní rozhodnutí týkající se harmonizačního procesu. Výsledné harmonizované zdroje reprezentují příbuzná slova jako zakořeněné stromy uložené ve sloupcovém souborovém formátu. Tato datová struktura a souborový formát aktuálně používá DeriNet 2.0. Navržená harmonizační procedura využívá řízené strojové učení a algoritmus hledající kostru v orientovaném grafu. Natrénovaný strojový model přiřazuje skóre každému slovotvornému vztahu a zmíněný algoritmus následně na jejich základě nalezne v každé slovotvorné rodině kostru orientovaného grafu, tj. strukturu zakořeněného stromu. Výsledná kolekce pokrývá 20 evropských jazyků, je publikována pod názvem ‚Universal Derivations' (UDer) a je volně dostupná.