

УДК 004.912

Г. Мацюк¹, Н. Кунанець²

¹(Тернопільський національний технічний університет імені Івана Пулюя)

²(Національний університет «Львівська політехніка»)

СТАНДАРТИ, ЩО РЕГЛАМЕНТУЮТЬ ФОРМАТ ПОДАННЯ ІНФОРМАЦІЙНО-ПОШУКОВОГО ТЕЗАУРУСА

Існує низка міжнародних та вітчизняних стандартів, що регламентують побудову інформаційно-пошукових тезаурусів.

Основними документами є наступні стандарти:

- міжнародні стандарти ISO (International Organization for Standardization) міжнародної організації зі стандартизації: *ISO 2788:1986. Documentation – Guidelines for the establishment and development of monolingual thesauri* (ISO 2788:1986 «Вказівки щодо створення і розвитку одномовного тезауруса») та *ISO 5964:1985. Documentation – Guidelines for the establishment and development of multilingual thesauri* (ISO 5964:1985 «Вказівки щодо створення і розвитку багатомовного тезауруса»). Стандарт ISO 2788:1986 залишався актуальним протягом 25 років до появи стандарту *ISO 25964-1:2011 Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval* (ISO 25964-1:2011 «Інформація та документація – Тезаурус та зв'язок з іншими словниками – Частина 1: Тезаурус для інформаційного пошуку») в 2011 році та його продовження *ISO 25964-2:2013 Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies* (ISO 25964-2:2013 «Інформація та документація – Тезаурус та зв'язок з іншими словниками – Частина 2: Зв'язок з іншими словниками»). Стандарт ISO 2788-1986 визначає тезаурус як набір термінів, пов'язаних між собою відповідними зв'язками (відношеннями). Структура багатомовних тезаурусів регламентується стандартом *ISO 5964-1985*. У ньому, крім зв'язків між термінами, є також зв'язки між еквівалентними термінами на різних мовах, серед яких існують такі типи зв'язків: – повна еквівалентність; – неповна еквівалентність (значення термінів не співпадають, але перетинаються); – часткова еквівалентність (значення одного терміну ширше, ніж значення іншого); – еквівалентність «один до багатьох» (значення одного терміну відповідає сукупності значень декількох термінів). Стандарт *ISO 13250* визначає тематичні карти, інформація подається у форматі xml з поясненнями до визначень xmt. У стандарті є ключові елементи та поняття.

- стандарти: *ГОСТ 7.25-2001* для одномовного тезауруса та стандарт *ГОСТ 7.24-2007* для багатомовного тезауруса. Ці стандарти були створені на базі вищеописаних стандартів ISO і, фактично, мало чим відрізняються від них. Вони встановлюють правила розробки, структуру, склад і форму подання інформаційних тезаурусів, орієнтованих на використання лексики російської мови.

- національні стандарти, що переважно базуються на міжнародних стандартах (ISO) цієї серії: *ДСТУ 4031-2001* (ISO 5964:1985) «Інформація та документація. Багатомовний тезаурус. Методика розроблення». Цей стандарт регламентує загальні вимоги до укладання багатомовного тезауруса, управлінські рішення у цій галузі, інші мовні проблеми тощо; *ДСТУ 4032-2001* (ISO 2788:1986) «Інформація і документація. Одномовний тезаурус. Методика розроблення». У стандарті наведено рекомендовані процедури керування складом словника, встановлено терміни індексування та складені терміни, визначено базові відношення в тезаурусі, розглянуто засоби встановлення та відтворення певних різновидів відношень між термінами індексування.