

# ПРЕДСТАВЛЕНИЕ ПАТЕНТНО-КОНЪЮНКТУРНЫХ ДАННЫХ С ПОМОЩЬЮ АЛГЕБРЫ КОНЕЧНЫХ ПРЕДИКАТОВ

Король О.И., Шаронова Н.В.

*Национальный технический университет*

*"Харьковский политехнический институт", г. Харьков*

В работе рассмотрены вопросы представления и классификации неструктурированных патентно-конъюнктурных (ПКД) и ассоциативных данных (ПАД).

При организации поиска ПКД и ПАД, на вход поступает информация об объекте исследования из разных ресурсов, распределенных в сети. На выходе получаем множество материалов, объединенных в логическую последовательность в индивидуальном подходе поиска и обработки. Этот процесс основан на механизме адаптации, который организован в виде навигационных правил.

Предложенные средства основаны на использовании метода компараторной идентификации для разбиения на классы эквивалентности и связывания в гиперструктуру документов, отобранных в результате запроса к базе ПКД.

Используем теорию алгебры конечных предикатов для представления онтологий, по которым будем осуществлять поиск, а также для представления ограничительных текстовых полей в виде:  $F(x_1, x_2, \dots, x_n) = 1/0$ .

На вход системы подается множество сигналов (тексты патентных документов, информация о научных конференциях, выставках и т.д.):  $x_1, x_2, \dots, x_n$ , которые берутся из конечных множеств  $X_1, X_2, \dots, X_n$ , причем  $x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$ . На выход поступает множество элементов вида  $y_1, y_2, \dots, y_n$ , причем  $y_1 \in Y_1, y_2 \in Y_2, \dots, y_n \in Y_n$ . Элементы  $y_1, y_2, \dots, y_n$  однозначно зависят от сигналов  $x_1, x_2, \dots, x_n$ , что показывает на существование функций  $y_1 = f_1(x_1), y_2 = f_2(x_2), \dots, y_n = f_n(x_n)$ , каждая из которых представляет собой сюръекцию, отображающую множество  $X_i$  на множество  $Y_i, i \in \{1, 2, \dots, n\}$ .

Для каждой пары  $x$  и  $y$  существует точно определенное значение соответствия данного понятия проблемной области и рассматриваемого документа, которое выражено предикатом релевантности  $R$ .

Нижеследующие формулы представляют закономерности структурирования патентной информации через предикат  $R$ , объективно определяемый классификатором при проведении поиска:

$$Vx(x) = (y \in Y) (R(x, y) \sim R(x, y)),$$

$$Vr(y) = (x \in X) (R(x, y) \sim R(x, r)).$$