

Доклад на тему «Создание онтологической базы знаний рекрутинговой системы»

Подготовил :

Нифталиев В.Э.

Руководитель:

Привалов М.В.

Постановка проблемы

При составлении резюме, соискатели в большинстве случаев используют либо слишком детализированный способ описаний своих навыков и их уровня, либо наоборот поверхностный, что усложняет поиск оптимальных кандидатов для работодателя. Так же отсутствие единой стандартной формы резюме затрудняет процесс первичного анализа резюме.

Цели и задачи

- Цель
 - разработка онтологической базы знаний рекрутинговой системы, содержащей словарь терминов предметной области и логические связи, которые описывают, как эти термины соотносятся между собой.
- Задачи
 - определение терминов предметной области и взаимосвязей между ними;
 - выбор метода создания онтологии;
 - создание модели предметной области;
 - определение классов, иерархии классов и их свойств;
 - разработка онтологической базы знаний.

Что такое онтология?

- 1) Онтология – явная спецификация концептуализации.
- 2) Онтология - описание декларативных знаний, сделанное в виде классов с отношением иерархии между ними.

Обоснование онтологического подхода

Причины использования онтологий:

- возможность повторного использования знаний в предметной области;
- для того чтобы сделать допущения в предметной области явными;
- для отделения знаний в предметной области от оперативных знаний;
- для анализа знаний в предметной области.

Использование онтологий в рекрутинговой системе

- Определение словаря терминов предметной области;
- Согласование концептуальных описаний навыков соискателя;
- Создание иерархии навыков и определение взаимосвязей между ними.

Методологии построения онтологий

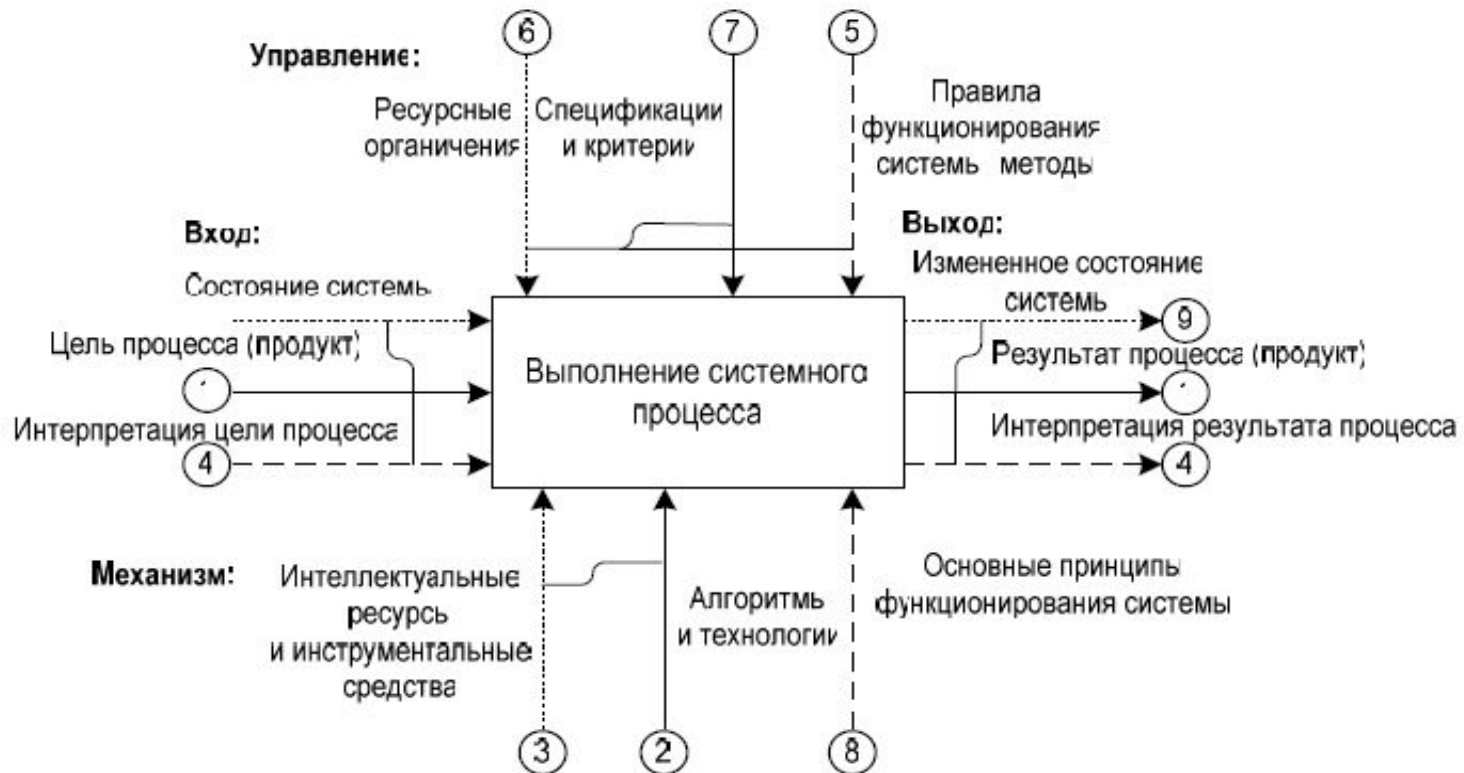
Для решения поставленной задачи предлагается исследовать применение следующих способов построения онтологий:

- Системно-когнитивный
- Иерархический

Системно-когнитивный подход

Данная методология предназначена для анализа и принятия решений в плохо определенных ситуациях. Она основана на моделировании субъективных представлений экспертов о ситуации. Задачей когнитивного моделирования является выявление и описание абстрактных или существующих коррелируемых явлений, событий, происходящих на объекте контроля и управления

Системно-когнитивный подход



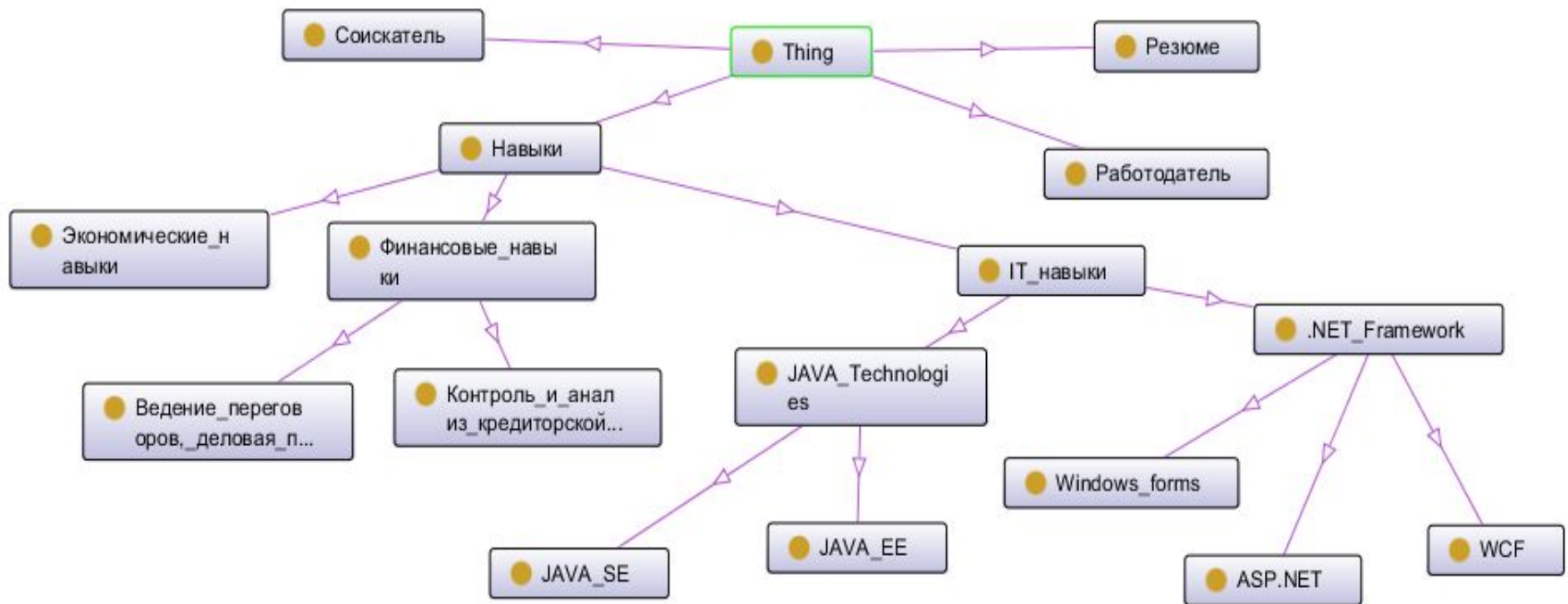
Интерпретация узла знакового ориентированного графа в методологии IDEF0

Иерархический подход

Иерархический подход включает следующие этапы:

- определение классов в онтологии;
- расположение классов в таксономическую иерархию (подкласс - надкласс);
- определение свойств классов и описание допускаемых значений свойств;
- заполнение значений слотов экземпляров.

Пример разработанной онтологии



Часть онтологии рекрутинговой системы, разработанной с помощью Protégé 4.2

Задача оптимального поиска кандидатов

Дано:

Максимальный уровень владения навыком:

$$X_{max} \in R$$

Вакансия – набор навыков и их уровень, необходимых для работы:

$$C = \{x_1, \dots, x_n\}, \text{ где } x_i \in [0, X_{max}], i, n \in N \text{ уровень } i\text{-го навыка}$$

Важность навыка в данной вакансии:

$$K = \{k_1, \dots, k_n\}, \text{ где } k_i \in R \text{ важность } i\text{-го навыка}$$

Исходный список кандидатов:

$$P = \{p_1, \dots, p_m\}, \text{ где } m \in N \text{ количество кандидатов}$$

Кандидат – набор навыков и их уровень, которыми он владеет:

$$P_j = \{p_{1j}, \dots, p_{nj}\}, \text{ где } x_i \in [0, X_{max}], i, n \in N$$

Возможное ограничение поиска в рамках каждого уровня владения навыком в меньшую и большую стороны:

$$l_1, l_2 \in R$$

Расстояние от кандидата до вакансии (Евклидово расстояние с ограничением):

$$\rho(P_j, C) = \sqrt{\sum_{i=0}^n k_i^2 * (x_i - y_{ij})^2} \text{ , если } \forall j, y_{ij} \in [x_i - l_1, x_i + l_2]$$

$$\rho(P_j, C) = \infty \text{ если } \exists j, y_{ij} \notin [x_i - l_1, x_i + l_2]$$

Найти:

$$\operatorname{argmin}_j \rho(P_j, C)$$

Заключение

Использование онтологического подхода для построения общей концептуальной модели рекрутинговой системы позволяет формировать общие понятийные пространства, обеспечивающие адекватное понимание информации, используемой в предметной области, и таким образом устранить семантическую неопределенность.

Спасибо за внимание

Вопросы?