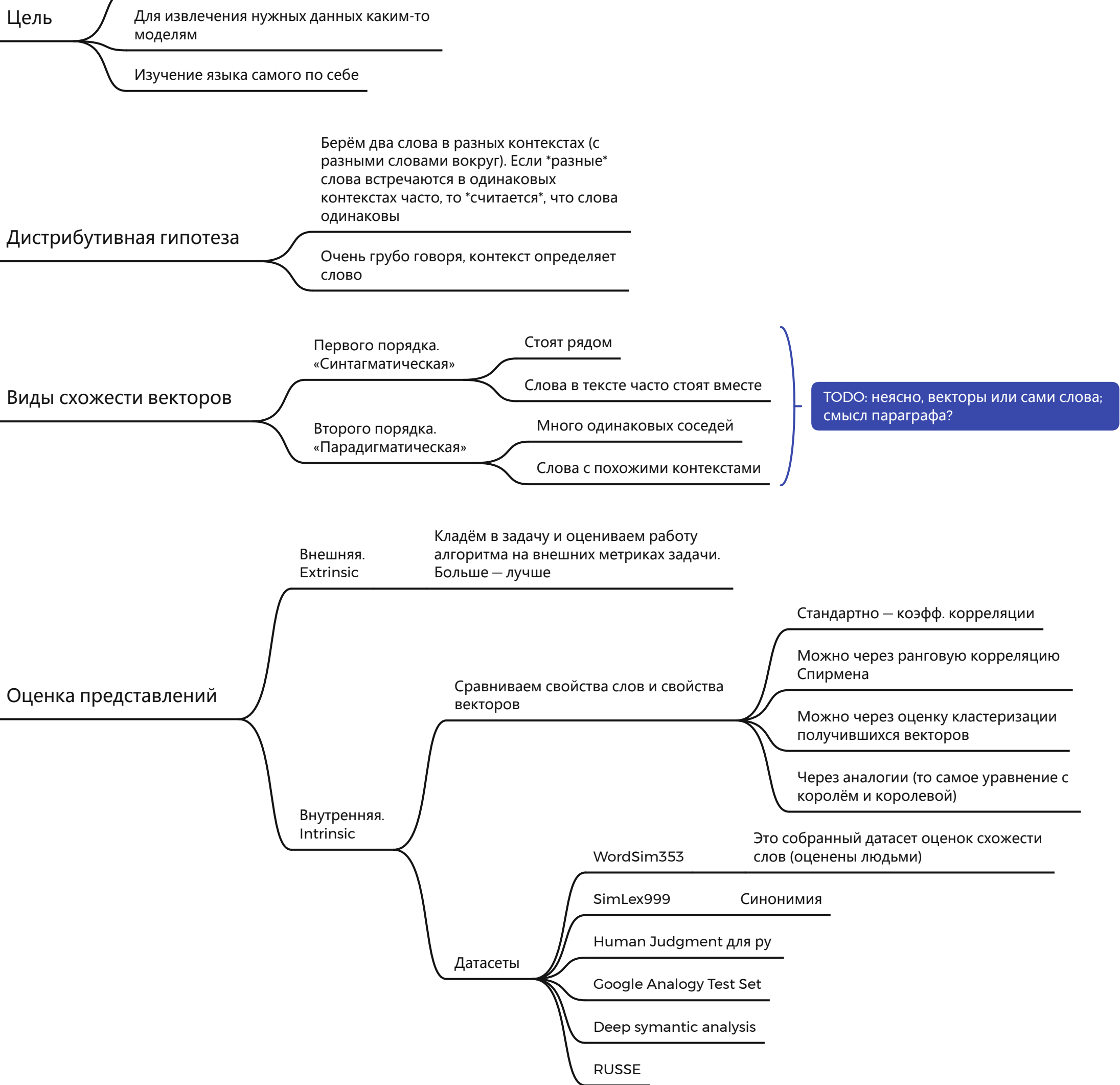
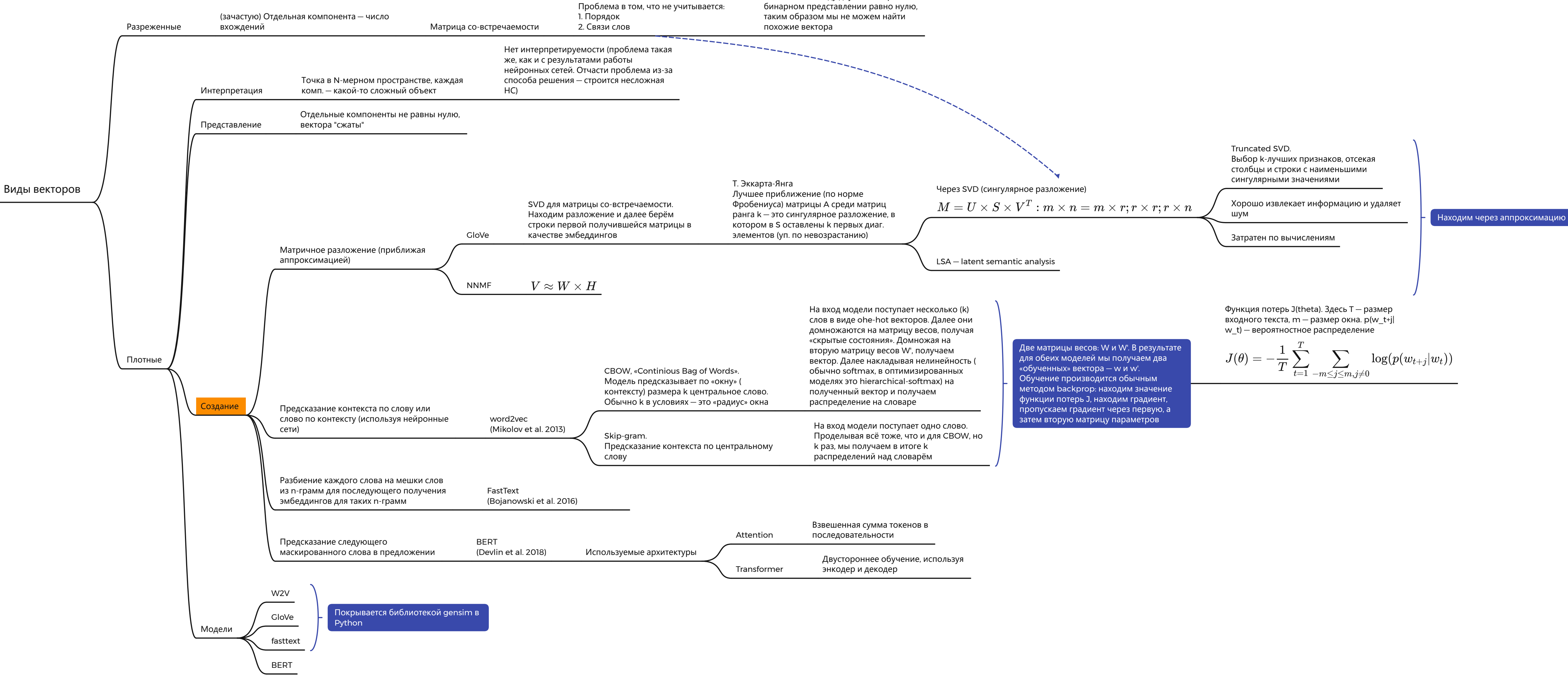


Векторная семантика



TODO: неясно, векторы или сами слова; смысл параграфа?



Через SVD (сингулярное разложение)

$$M = U \times S \times V^T : m \times n = m \times r; r \times r; r \times n$$

LSA — latent semantic analysis

- Truncated SVD. Выбор k-лучших признаков, отсекая столбцы и строки с наименьшими сингулярными значениями
- Хорошо извлекает информацию и удаляет шум
- Затратен по вычислениям

Находим через аппроксимацию

Две матрицы весов. W и W'. В результате для обеих моделей мы получаем два «обученных» вектора — w и w'. Обучение производится обычным методом backprop, находим значение функции потерь J, находим градиент, пропускаем градиент через первую, а затем вторую матрицу параметров

Функция потерь J(theta). Здесь T — размер входного текста, m — размер окна. p(w_{t+j}|w_t) — вероятностное распределение

$$J(\theta) = -\frac{1}{T} \sum_{t=1}^T \sum_{-m \leq j \leq m, j \neq 0} \log(p(w_{t+j}|w_t))$$

