# Предложение

# за възлагане на дипломна работа

# към катедра “Информационни технологии”,

# ФМИ, СУ ”Св. Кл. Охридски”

**Дипломант: Марин Маринов Ножчев**

**Магистърска програма: Био- и медицинска информатика**

**Научен ръководител: Калин Георгиев, ФМИ**

**Тема на дипломната работа: *Анонимизация на клинична информация за пациенти***

**Анотация**

Наличието на биомедицински данни за пациенти е критично за изследванията в медицината. Основен източник на данни са болниците, но здравните заведения нямат право по закон да ги разпространяват. Данните могат да се използват за изследвания само ако се трансформират по начин, който не позволява информацията да се свърже с конкретни пациенти. Този процес се нарича „медицинска анонимизация“.

Основна пречка за пълната анонимизация са немаркираните лични данни особено в неструктурирани медицински текстове като анамнези и диагнози в амбулаторни листи. В тези документи често се споменават лични данни за пациентите дори извън съответните полета на формуляра. Например възрастта на пациента често се споменава в диагнозата, ако възрастта е релевантна към нея.

Целта на дипломната работа е анализ на state-of-the-art алгоритмите за анонимизация на клинични данни и реализация на алгоритъм, който маркира лични данни в свободен медицински текст. Софтуерът ще бъде по-гъвкав от много от съществуващите решения.

Задачи, произтичащи от целта:

1. Обзор на стандарти алгоритми за анонимизация и тяхната софтуерна реализация.
2. Създаване на софтуер за автоматично откриване на потенциално идентифициращи данни в клинично документи базиран на статистически анализ.

Ограничаващи/облекчаващи условия:

1. Наличието и качеството на експериментални данни ограничава качеството на заключенията на дипломната работа.
2. За част от експериментите с алгоритми за анонимизация ще се използват готови техни реализации с отворен код.

Срок за изпълнение: октомври 2013

Дата: 06.05.2013г. Заявител:

Научен ръководител: