

# Компьютер и язык

---

Леонид Лейбович Иомдин

Лаборатория компьютерной  
лингвистики Института проблем  
передачи информации РАН им.  
А.А.Харкевича

# Компьютер и язык

---

**Доклад посвящён задачам компьютерной обработки текстов на естественном языке.**

**Автор рассказывает об основных типах таких задач (автоматический перевод текстов с одних языков на другие, интеллектуальный поиск информации, общение с компьютерными системами на человеческом языке, анализ и порождение устной речи) и о современных методах их решения.**

**Подробно демонстрируется работа одной системы автоматического перевода, разрабатываемой в Институте проблем передачи информации с участием докладчика. Что должен знать компьютер, чтобы из английского выражения *What is your name* получить его русский эквивалент *Как тебя зовут?***

**Автор показывает, как развитие теоретической лингвистики и прогресс в компьютерной науке взаимно дополняют друг друга. Особое внимание уделяется нерешенным проблемам компьютерной лингвистики - науки, которая лежит в основе рассматриваемой области человеческой деятельности.**

# АВТОМАТИЧЕСКИЙ ПЕРЕВОД

---

Лингвистическая теория «Смысл  $\leftrightarrow$  Текст»  
Человеческий диалог и перевод

.



http://proling.iitp.ru/etap-bin/register.exe/etap/start.html



DAEMON Search



Ссылки



gmail



OSP



SDM



Admin



Intuit



Catalog



Diofant.ru



HPCU



HPCU.adm



Яндекс

>>



Система автоматического пе...



Домой



Веб-каналы



Печать



Страница



Сервис

>>

# Лаборатория компьютерной лингвистики



Введите в окошко текст для перевода:

Моих друзей зовут Светлана и Никита.

или задайте имя файла:

Обзор...

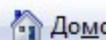


III



Интернет | Защищенный режим: выкл.

130%



Российская Академия наук  
Институт проблем передачи информации



## Лаборатория компьютерной лингвистики

**Исходный текст:** Моих друзей зовут Светлана и Никита.

**Перевод:** My friends' names are Svetlana and Nikita.



Система  
машинного  
перевода  
**ЭТАП-3**



http://proling.iitp.ru/etap-bin/register.exe/etap/start.html



DAEMON Search



Ссылки



gmail



OSP



SDM



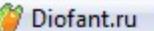
Admin



Intuit



Catalog



Diофант.ру



HPCU



HPCU.adm



Яндекс



Система автоматического пе...



Домой



Веб-каналы



Печать



Страница



Сервис



Российская Академия наук  
Институт проблем передачи информации

# Лаборатория компьютерной лингвистики

Введите в окошко текст для перевода:

Моих друзей зовут Светлана и Никита.  
Светлана и Никита зовут моих друзей.

или задайте имя файла:

Обзор...



Интернет | Защищенный режим: выкл.

160%



Российская Академия наук

Институт проблем передачи информации



# Лаборатория компьютерной лингвистики

**Исходный** Моих друзей зовут Светлана и Никита.

**текст:** Светлана и Никита зовут моих друзей.

**Перевод:** My friends' names are Svetlana and Nikita.

Svetlana and Nikita call my friends.



Система  
машинного  
перевода  
**ЭТАП-3**



**Исходный  
текст:**

Моих друзей зовут Светлана и Никита.

**Перевод:**

<S>

<W DOM="2" FEAT="А МН ВИН ОД" ID="1" LEMMA="МОЙ" LINK="опред">Моих</W>

<W DOM="3" FEAT="S МН МУЖ ВИН ОД" ID="2" LEMMA="ДРУГ" LINK="1-компл">друзей</W>

<W DOM="\_root" FEAT="V НЕСОВ ИЗЪЯВ НЕПРОШ МН 3-Л" ID="3" LEMMA="ЗВАТЬ">зовут</W>

<W DOM="3" FEAT="S ЕД ЖЕН ИМ ОД" ID="4" LEMMA="СВЕТЛАНА" LINK="2-компл">Светлана</W>

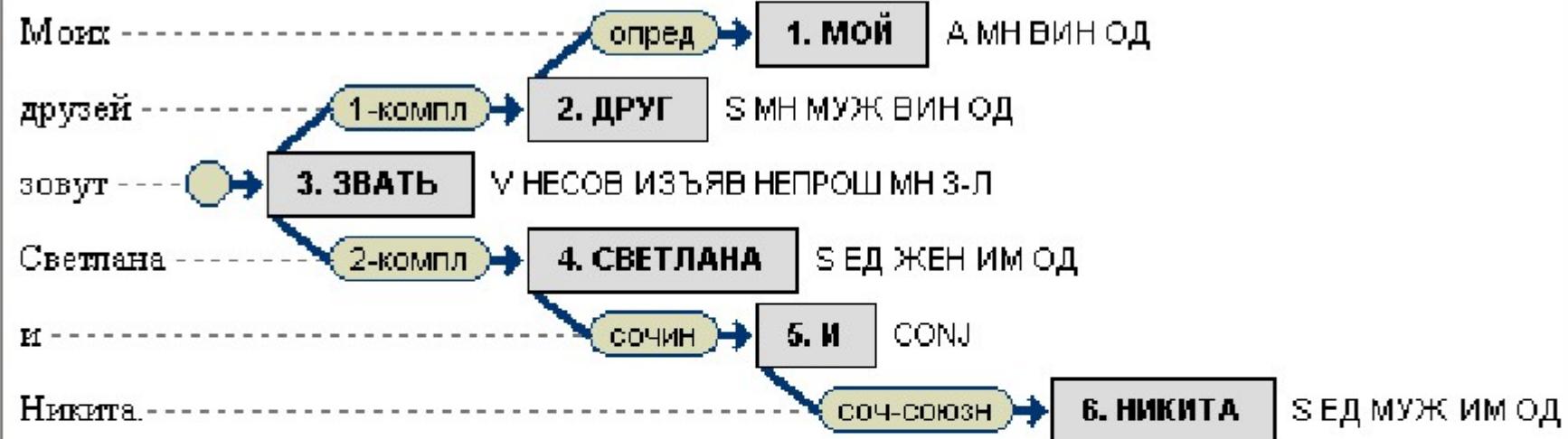
<W DOM="4" FEAT="CONJ" ID="5" LEMMA="И" LINK="сочин">и</W>

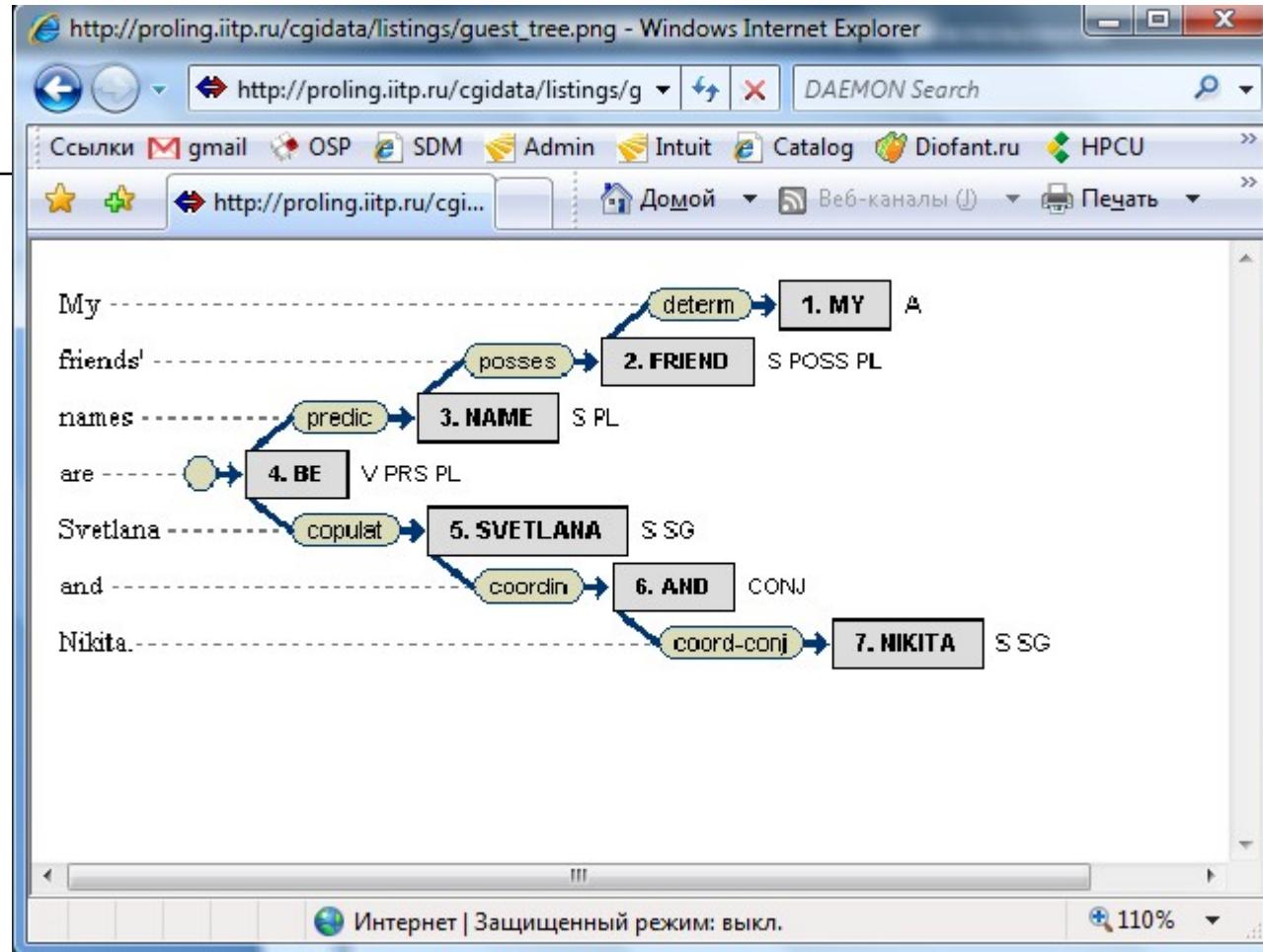
<W DOM="5" FEAT="S ЕД МУЖ ИМ ОД" ID="6" LEMMA="НИКИТА" LINK="соч-союзн">Никита</W>.

</S>



Система  
машинного  
перевода  
**ЭТАП-3**





# ИНТЕЛЛЕКТУАЛЬНЫЙ ПОИСК ИНФОРМАЦИИ

---

Информационный поиск и  
перифразирование

Распознавание чувств и настроений  
(sentiment recognition)

.

# ОБЩЕНИЕ С КОМПЬЮТЕРНЫМИ СИСТЕМАМИ НА ЧЕЛОВЕЧЕСКОМ ЯЗЫКЕ

---

Какова средняя зарплата  
сотрудников отдела № 7?

20391 рубль в месяц

.

# АНАЛИЗ И ПОРОЖДЕНИЕ УСТНОЙ РЕЧИ

---

Автоматический анализ речи

Автоматический синтез речи

.

# ОБРАТНАЯ СВЯЗЬ ТЕОРИИ И ПРАКТИКИ

---

**«Развитие теоретической  
лингвистики и прогресс в  
компьютерной науке взаимно  
дополняют друг друга»**

**Пример: длительные  
обстоятельства в русском языке**

# Длительное синтаксическое отношение

---

*Поезд стоял [X] там всего минуту [Y].  
Он работал год.*

*Больному надо вылежать дня три.  
Он посидел часик и ушел.*

# Нерешенные проблемы компьютерной лингвистики

---

## Модели управления предикатных слов

▪

# Словарные статьи русского компьютерного словаря

---

## СМОТРЕТЬ

POR:V

//СОВ: ПОСМОТРЕТЬ

SYNT:ТРАНЗИТ,ДАТПОС,\*ДО-КОНЕЧН

DES:'ДЕЙСТВИЕ','ФАКТ','АБСТРАКТ'

D2.1:ВИН

D2.2:КОНЕЧН

D2.3:QU

D2.3:ЧТОБЫ

SYN:ГЛЯДЕТЬ LF1\_CIN

\_REAL1-M:ВИДЕТЬ

\_SYN1:ВЗИРАТЬ/ГЛАЗЕТЬ/ТАРАЩИТЬСЯ

\_SING:ВЗГЛЯД

\_MAGN:ВНИМАТЕЛЬНО/ПРИСТАЛЬНО

\_ANTIMAGN:БЕГЛО

TRAF:ОБСТ-ТАВТ.10

LR:ВЗГЛЯД

TRAF:1-КОМПЛ.11

TRAF:1-КОМПЛ.51

TRAF:НЕАКТ-КОМПЛ.13

TRAF:1-КОМПЛ.28

TRAF:1-КОМПЛ.27

\*\*\*\*\*

ZONE:EN

TRANS:LOOK1

# Словарные статьи русского компьютерного словаря

---

## ЗАГЛЯДЫВАТЬ

POR:V АЛ+ЛВ

//COB: ЗАГЛЯНУТЬ

SYNT:БЕНЕФ,\*ДО-КОНЕЧН

DES:'ФАКТ','АБСТРАКТ','ДЕЙСТВИЕ'

D2.1:КОНЕЧН

\_S0:ЗАГЛЯДЫВАНИЕ

TRAF:HEAKT-КОМПЛ.12

\*\*\*\*\*

ZONE:EN

TRANS:PEEP2

REG:RA-TRADUCT2.D0 ЗАГЛЯДЫВАТЬ К КОМУ-ЛИБО -> CALL IN ON SMB

TAKE:X

N:01

CHECK

1.1 DOM-LEXR(X,Z,1-КОМПЛ,К)

1.2 DOM-EQU(Z,Z1,ПРЕДЛ,'ЧЕЛОВЕК','ЖИВОТНОЕ')

DO

1 ZAMAUZ:X(CALL\$IN)

2 ZAMAUZ:Z(ON1)

3 IZOT:(X,Z,2-КОМПЛ)-(X,Z,ADVERB)

TRAF:EXPANS.64

LR:ЗАГЛЯДЫВАНИЕ

# Словарные статьи русского компьютерного словаря

---

## **ВЗГЛЯДЫВАТЬ**

**POR:V АЛ**

**//СОВ: ВЗГЛЯНУТЬ**

**SYNT:БЕНЕФ,\*ДО-КОНЕЧН**

**DES:'ДЕЙСТВИЕ','ФАКТ','АБСТРАКТ'**

**D1.1:'ЛИЦО'**

**D2.1:КОНЕЧН**

**\_S0:ВЗГЛЯД**

**TRAF:HEAKT-КОМПЛ.12**

\*\*\*\*\*

**ZONE:EN**

**TRANS:GLANCE2**

**TRAF:EXPANS.64**

**LR:ВЗГЛЯД**

**TRAF:EXPANS.93**

**LR:HA1**

# Предлоги, задающие конечную точку движения

---

В1

COMMENT: "УПРАВЛЯЕТ ВИНИТЕЛЬНЫМ ПАДЕЖОМ"

EXAMPLE: "ВОШЕЛ В ДОМ"

POR:PR ЛИ

SYNT: ВРЕМ, КОЛИЧ-ПР, ОТРЫВ-НИ, ПЕРВООБР, КОНЕЧН, ВВОД

D2.1: ВИН

ВНУТРЬ1

COMMENT: "ПРЕДЛОГ, УПРАВЛЯЮЩИЙ РОД. ПАДЕЖОМ"

EXAMPLE: "Я ЗАГЛЯНУЛ ВНУТРЬ КОМНАТЫ"

POR:PR ЛИ

SYNT: \*Н-МЕСТ, КОНЕЧН

D2.1: РОД

ДО1

COMMENT: "ПРЕДЛОГ"

EXAMPLE: "ДО ДОМА, ДО ПЯТНИЦЫ, ДО ДЕСЯТИ"

POR:PR ЛК

SYNT: ВРЕМ, КОНЕЧН, ОТРЫВ-НИ, ПЕРВООБР, ПРИБЛ

D2.1: РОД

ЗА1

COMMENT: "ПРЕДЛОГ, УПРАВЛЯЮЩИЙ ВИНИТЕЛЬНЫМ ПАДЕЖОМ"

EXAMPLE: "ЗА ОВРАГ, ЗА МЕНЯ, ЗА ПЯТЬДЕСЯТ"

POR:PR ЛИ

SYNT: КОНЕЧН, ОТРЫВ-НИ, ВРЕМ, ПЕРВООБР

D2.1: ВИН

# Предлоги, задающие конечную точку движения

---

К

POR:PR ЛИ  
SYNT:КОНЕЧН,ВРЕМ,ОТРЫВ-НИ,ВВОД,ПЕРВООБР  
D2.1:ДАТ

НА1

COMMENT:"ПРЕДЛОГ, УПРАВЛЯЮЩИЙ ВИНИТЕЛЬНЫМ ПАДЕЖОМ"  
EXAMPLE:"НА УКРАИНУ, НА ПЯТЬ БОЛЬШЕ"  
POR:PR НП  
SYNT:КОНЕЧН,ОТРЫВ-НИ,ВВОД,КОЛИЧ-ПР,ПЕРВООБР  
D2.1:ВИН

ПОВЕРХ1

COMMENT:"ПРЕДЛОГ"  
EXAMPLE:"ВЗГЛЯНУЛ ПОВЕРХ ОЧКОВ"  
POR:PR ЛИ  
SYNT:/\*Н-МЕСТ,КОНЕЧН,ЛОКАТ  
D2.1:РОД

ПОД1

COMMENT:"ПРЕДЛОГ, УПРАВЛЯЮЩИЙ ВИН. ПАДЕЖОМ"  
EXAMPLE:"ПОД ДВЕРЬ, ПОД ПЯТЬДЕСЯТ, ПОДСТРАИВАТЬСЯ ПОД НАЧАЛЬСТВО"  
POR:PR ЛК  
SYNT:КОНЕЧН,ОТРЫВ-НИ,ПЕРВООБР  
D2.1:ВИН

# Предлоги, задающие конечную точку движения

---

## ПОПЕРЕК2

COMMENT: "ПРЕДЛОГ"

EXAMPLE: "БРЕВНО ПОПЕРЕК ДОРОГИ"

POR:PR КЛ

SYNT:ЛОКАТ,КОНЕЧН

D2.1:РОД

## ПОСЕРЕДИНЕ2

COMMENT: "ПРЕДЛОГ"

EXAMPLE: "ПОСЕРЕДИНЕ ДВОРА"

POR:PR ЛИ

SYNT: \*Н-МЕСТ,ЛОКАТ,КОНЕЧН

D2.1:РОД

## ПОСРЕДИ

POR:PR ЛИ

SYNT: \*Н-МЕСТ,ЛОКАТ,КОНЕЧН

D2.1:РОД

# Языковой материал

---

*Идти в город, на гору, за ворота, под стол, до  
поворота*

*Смотреть на стол, в глаза, за горизонт, под шкаф,  
\*до вершины*

*Заглянуть в щелочку, под шкаф, \*до шкафа*

*Взглянуть на рисунок, в зеркало, \*до зеркала*