

Extracting relational data from Wikipedia

Chris Perivolaropoulos

Tuesday 9q May 2016

Contents

I	αμποτράκι	2
II	Introduction	4
III	Wikipediabase	5
IV	WikipediaMirror	6
V	References	7

Part I

αμπστρακτ

MiT InfoLab's START (SynTactic Analysis using Reversible Transformations) is the world's first web-based question answering system. For accessing most data sources it takes advantage of Omnibase, the "*virtual database*" providing uniform access to multiple sources on the web. We developed WikipediaBase to provide an Omnibase-like way for START to access unstructured and semi-structured information in Wikipedia. As part of this goal we also created wikipedia-mirror, a program to create Wikipedia clones that can be run locally, to provide control and unrestricted access to the wikipedia data set without depending on or abusing wikipedia.org.

Part II

Introduction

Part III

Wikipediabase

Part IV

WikipediaMirror

Part V

References