



The Road Runner Project Towards Automatic Data Extraction

Mirror Site at Roma Tre: <http://www.dia.uniroma3.it/db/roadRunner/>
Mirror Site at Università della Basilicata: <http://www.dia.unibas.it/db/roadRunner/>

Adobe Acrobat

正在等待转换...

ROAD RUNNER HOME PAGE

[Road Runner](#)

- [Publications](#)
- [Project Members](#)

Description:

Road Runner is a combined project of the [Database Group](#) of **Università di Roma Tre** and of the [Database Group](#) of **Università della Basilicata**. The project investigates techniques for extracting data from HTML sites through the use of automatically generated wrappers. In fact, many Web-based applications today use wrappers to extract data from HTML pages. These wrappers, however, are usually coded by hand, and therefore their generation and maintenance are difficult and labor intensive. To automate the wrapper generation and the data extraction process, the Road Runner project aims at developing original techniques to automatically generate wrappers.

A wrapper generation system has been implemented in a working prototype, which has been used to conduct a number of experiments on real-life data-intensive Web sites. These experiments confirm the feasibility of the approach and. The system prototype has been implemented in [Java](#).

On-Line Resources:

- [Publications](#)
Papers, Technical Reports and unpublished manuscripts related to the project;
- [Experimental Results](#)
Some of the wrappers that have been automatically generated by our system prototype
- [Downloads](#)
We have released under GPL the source code of the roadRunner system.

*This page is maintained by [Gianni Mecca](#) and [Paolo Meriardo](#)
Road Runner. [geococcyx californianus](#)*