

NISTEPにおける大学・公的機関名辞書の整備と名寄せプログラムの開発 —より精確な研究機関同定(名寄せ)を目指して—

文部科学省 科学技術・学術政策研究所 科学技術予測・政策基盤調査研究センター

小野寺夏生, 伊神正貫

要旨

科学技術・学術政策研究所(NISTEP)では、我が国における研究開発動向の機関レベル、組織レベルでの分析のための基盤として、大学・公的機関名辞書の整備と、データベース中の英語機関名データをこの辞書中の機関に名寄せするプログラムの開発を進めている。本報告書は、この機関名辞書と名寄せプログラムを活用したい人たちのために、それらの特徴、機能、利用の際に留意すること等を述べる。機関名辞書は、研究開発を行う国内機関約16,000とそれらの主要な下部組織4,000の基本的情報を収録する。名寄せに用いるための種々の英語揺らぎ名を収録していること、代表機関と下部組織の間に階層関係を付けていること、非現存機関とその継承機関の間に関係を付けていることなどが特徴である。名寄せプログラムにおいてはこれらの特徴を活かし、類似英語名を持つ機関の識別、下部組織の同定、変遷前後の機関の識別に工夫を凝らしている。

Construction of the Dictionary of Names of Universities and Public Organizations and development of the Program for Organization Name Disambiguation in NISTEP: Toward more precise and accurate research organization name disambiguation

ONODERA Natsuo and IGAMI Masatsura

Center for S&T Foresight and Indicators, National Institute of Science and Technology Policy (NISTEP)

ABSTRACT

National Institute of Science and Technology Policy (NISTEP) has been promoting construction of the Dictionary of Names of Universities and Public Organizations and development of the Program for Organization Name Disambiguation, which enable to identify an organization name description (in English) to an organization included in the Dictionary. NISTEP aims that the Dictionary and the Program serve as the infrastructure to analyze the current state and trend of research and development in Japan in micro- and meso-level. This report describes their characteristics and functions, and also issues that should be noted at their use. The Dictionary includes the basic information on about 16,000 principal organizations in Japan and their about 4,000 main subsidiary organizations. Its features are many variant organization names for disambiguation, hierarchical relation among organizations, and association of closed organizations with their succeeding ones. The Name Disambiguation Program devises precise and accurate identification of organizations, particularly in discrimination between organizations with similar names, identification of adequate subsidiary organizations, and discrimination between succeeded and succeeding organizations.