남북정보기술 용어표준

특 집
04

목 차

1. 서 론
2. 용어의 일반적 표준화 원칙
3. 비슷한 혹은 반대 개념을 가지는 용어들의 대역어 구분
: 일반적 용어원칙 적용예
4. 결 론

최기선 · 변정용 · 신효식 · 옥철영
(한국과학기술원 · 동국대학교 · 전북대학교 · 울산대학교)

1. 서 론

다음은 정보기술용어 "access management service"를 북한용어로 정의한 것이다.

"사용자대행체와 통보이송대행체가 호상간의 접근관계를 수립하고 련관된 정보들을 관리할수 있도록 하는 봉사형태."

이것은 ISO 2382의 영어 정의문을 번역한 것인데, 의미가 잘 전달되는지 생각해보자. "대행체", "통보이송", "봉사"라는 단어가 정보기술용어로서 쓰이고 있다는 것을 알 수 있다.

"A service that enables a user agent and a message transfer agent to establish access to one another and to manage associated information."

다시 KS (한국산업표준)의 용어 (KS X001)규격과 TTA (한국정보통신기술협회) 용어표준에 따라서 영어 정의문을 번역하면 각각 다음과 같다.

"사용자 대리인과 메시지 전송 대리자가 서로의 접근관계를 수립하고 연관 정보를 관리할 수 있도록 하는 서비스 형태."

"사용자 에이전트와 메시지 전송 에이전트가 (아래는 위와 같음)"

"agent"가 북에서는 "대행체", 남에서는 "대리인", "대리자", "에이전트"의 세 가지로 쓰인다는 것을 알 수가 있다. "message"는 북에서는 "통보", 남에서는 "메시지"로 쓰이고, "transfer"는 북에서는 "이송", 남에서는 "전송"으로 쓰인다. TTA 용어사전에 의하면, "transmission", "transfer" 두 용어가 모두 "전송"으로 쓰이고, "migration"이 "이송"으로 쓰이고 있다. 그러면 transfer와 transmission이 같은 개념을 갖고 있는 것일까? "service"는 북에서는 "봉사", 남에서는 "서비스"를 음차하여 쓰고 있다.

일반적인 용어 표준화 원칙을 제2절에서 설명하고, 이를 정보기술용어 규격인 ISO 2382에 적용하면서 발생한 사항에 대하여 제3절에서 기술한 다음 결론으로 끝을 맺는다.

2. 용어의 일반적 표준화 원칙

2.1 용어는 되도록 쉬운 어휘를 사용하여 구성한다.

용어는 많은 경우에 형태적으로 두 개 이상의 어휘가 결합한 합성어, 혹은 파생 접사를 포함하는 파생어와 같은 복합어이거나, 기능어를 포함하는 구 형태이다. 용어를 형성하는데 사용된 성분 어휘들에 대해서 가능하면 일상적인 어휘 표현인 기본 어휘를 선택함으로써 궁극적으로 쉬운 용어를 표준안으로 택한다. 이러한 배경에는 용어를 학습하는 사람들의 단계별 수준에 용어의 난이도가 부응해야 한다는 인식이 깔려 있다. 예를 들어, 초등학교 교재에 등장하는 용어는 초등학생 어휘력 수준에서 개념적으로 이해될 수 있는 것이어야 한다. 단, 영어 용어의 어휘 자체가 매우 어려운 수준일 경우에는 무리하게 쉬운 용어로 바꾸지 않아도 된다.

예) "data object"의 우리말 대역어는 "자료 대상", "자료 객체" 모두 가능하지만 영어에서 "object"의 등급이 가장 쉬운 1등급이므로, 한국어에서도 쉬운 어휘가 선택되도록 한다. "대상"의 등급(1등급)이 "객체"의 등급(3등급)보다 낮은 쉬운 어휘이기 때문에, "data object"는 "자료 대상"이 된다.

2.2 외래어는 순우리말화 한다.

단순히 음차표기 방식을 빌어 외래어를 우리 용어로 받아들이는 경향이 점점 높아지고 있다. 다른 용어에 포함되어 많이 사용되는 기본 용어까지 외래어로 사용되면서 용어에 대한 개념적 이해가 어렵게 되었다. 이러한 문제를 해결하기 위해서 가급적이면 이미 고착화된 외래어 용어들에 대해서도 우리말 용어를 제시하도록 한다. 외래어가 너무 보편적으로 많이 쓰여 굳어진 용어에 대해서는 경쟁적 우리말을 만들어 비우선어로 추가해 준다. 순우리말 표현을 우선적으로 받아 들이되, 표현에 있어서 파생력에 문제가 있거나 다른 어휘와 결합하여 새로운 용어를 만드는데 어색함이 있을 경우에는 한자어를 사용한다. 한자어를 사용할 경우에는 가급적이면 쉬운 한자어를 사용하도록 한다.

예) "cartridge"는 "카트리지", "교체통" 모두 가능하지만, 우리말 우선 원칙에 따라 "교체통"으로 한다.

예) "program"은 "프로그램"으로 굳어져 버렸기 때문에, 우리말 경쟁어 "실행순서"를 "프로그램(=실행순서)" 와 같이 비우선어로 추가한다.

2.3 외래어를 그대로 사용할 경우에는 외래어 음차표기 규정을 따른다.

전문용어의 우리말 표현에서 가능하면 우리말을 이용하는 것이 바람직하지만, 용어가 내포하는 개념을 명확하게 표현하기 위하여 외래어를 사용하여야 하는 경우도 있다. 따라서 외래어로 표현하는 경우는 국립국어원의 외래어 표기법을 따르거나, 전문용어언어공학 연구센터의 영어-우리말 음차표기 시스템의 결과를 따른다.

예) "program"은 "프로그램", "프로그람"으로 번역될 수 있지만 국립국어원의 외래어 표기법에 따라 "프로그램"을 대역어로 선정한다.

2.4 용어 내의 어휘는 일관성을 갖도록 한다.

우리말 용어의 성분 어휘에 대해서 개념상 동일한 원어 용어에 대해서 일관되게 단일한 우리말 대역어를 선택하도록 한다. 이들 기본 어휘들은 반복적으로 많은 용어에 성분 어휘로 사용되기 때문에 용어의 일관성을 획득하는데 파급효과가 크다. 기본 어휘들의 단일화를 통해서 영어 원어에 대한 우리말 용어의 일관성을 유지하도록 한다. 반대로 우리말 용어에 대한 영어 원어도 일관성을 유지하도록 한다.

예) "assignment"는 "대입", "할당", "대입", "지정" 등이 가능하지만, "배정"으로 선택하였다

고 하자. 그렇다면 의미상 문제가 없을 경우에 한하여 "assignment"가 들어간 모든 용어에서 "assignment"는 배정으로 번역되도록 한다. 즉, "assignment by name"은 "이름방식 배정", "assignment statement"는 "배정문", "credit/blame assignment"는 "신용/책임 배정"으로 한다.

2.5 말뭉치 빈도수가 높은 용어를 우선적으로 선택한다.

이 기준은 용어의 생명력과 관련이 있다. 현재 널리 사용되고 있는 용어를 표준으로 선정하려는 데 목적이 있다. 여러 가지 후보 대역어 중 말뭉치에서 가장 많이 나타나는 용어를 선정한다. 말뭉치의 크기가 작아서 통계적으로 의미 있는 빈도수를 구하기 어려운 경우에는 웹 검색 엔진에서의 빈도수를 이용한다.

예) "voice command"의 우리말 대역어로 "목소리 명령", "목소리 지령", "음성 명령" 등이 가능하지만 웹에서 가장 많이 검색된 "목소리 명령"을 대역어로 선정한다.

2.6 기존 용어 사전에 등재된 용어를 우선적으로 선택한다.

기존의 정보통신 분야 사전에 등재여부를 이용하여 표준 용어를 선정한다. 기존의 사전에 등재된 경우는 그 용어가 널리 통용되고 있다는 사실을 의미하기 때문에 기존 사전에 등재 여부를 표준 용어의 기준으로 삼는다.

예) "word processor"는 "문서 편집기", "워드프로세서"로 번역될 수 있지만 "문서 편집기"가 다수의 사전에 등록되어 있기 때문에 "문서 편집기"를 대역어로 선정한다.

2.7 용어의 개념을 잘 표현하는 용어를 선택한다.

용어의 정의문 정보를 참조하여 용어의 개념을 잘 드러낼 수 있는 표현을 선택하도록 한다. 다만, 용어의 개념적 정의를 그대로 번역하여 우리말 대역어를 제시하는 방법을 택하기 보다는 용어를 확장적으로 해석함으로써 해당 개념을 유추할 수 있는지를 주요한 기준으로 삼는다.

(예) "function"은 우리말에서 "함수" 또는 "기능" 모두 가능하지만, "Boolean function"에서는 "함수"로 번역하는 것이 더 용어의 의미를 더 정확하게 전달한다. 따라서 일관성의 원칙에 위배되더라도 경우에 따라서는 한 개의 영어 단어가 여러 개의 우리말 단어로 변환될 수 있다.

2.8 기능어는 되도록 삭제한다.

용어의 언어적 경제성의 원칙에 따르면 용어는 가능하면 간략하여야 한다. 따라서 용어가 내포하는 개념을 충분히 전달할 수 있는 경우에는 용어에 사용되는 기능어를 삭제한다. 용어에서 많이 사용되는 기능어는 관형격 조사 "-적", "-의" 등이 있다.

예) "deductive inference"는 "연역적 추론", "연역 추론"으로 번역될 수 있지만 의미 전달에 애매성이 없기 때문에 용어의 경제성 원칙에 따라 "연역 추론"을 대역어로 선정한다.

2.9 한국어의 용어의 띄어 쓰기는 영어 단어 단위로 맞추어 준다.

우리말 용어는 띄어쓰기에 관하여 영어와 비교해서 비교적 자유롭지만 영어의 단어 단위로 띄어주면 용어의 구성 어휘를 쉽게 파악할 수 있어서 용어의 의미를 정확하게 전달할 수 있다. 또한 영어 단어 단위의 우리말 대역어의 일관성 검사를 자동으로 할 수 있는 장점이 있다. 그러나 우리말의 특성상 우리말 구성 어휘가 접미사 또는 접두사로 사용되어서, 띄어쓰기 원칙을 지키지 못하는 경우도 있다.

예) "data bank"는 "자료은행", "자료 은행"로 번역될 수 있지만 구성 어휘 별 띄어쓰기 원칙에 따라서 "자료 은행"을 대역어로 선정한다

3. 비슷한 혹은 반대 개념을 가지는 용어들의 대역어 구분: 일반적 용어원칙 적용예

다음 표는 비슷한 개념을 가지는 용어들의 목록이다. 의미가 비슷한 용어의 경우, 그 대역어들을 같은 용어로 한다면 한글 용어를 보고는 이들 사이의 구분이 되지 않게 된다. 그래서 이들 용어를 구분하기 위하여 되도록 다른 대역어를 사용하였다.

〈표 1〉 비슷한 개념을 가지는 용어들의 목록

	영어 단일어	대역어	영어 용어	용어 대역어
1	*delivery*	배달	*delivery notification service*	배달 통지 서비스
	transmission	전송	text transmission	본문 전송
	transfer	전달	message transfer system	통신문 전달 체계
2	character	문자	character set	문자 집합
	literal	글자값	integer literal	정수 글자값
	letter	글자	letter quality	글자 품질
	string	문자열,열	string literal	문자열 글자값
	word	단어,문서	numeric word	수 단어
3	deferred	지체	deferred maintenance	지체 유지보수
	delay	지연	delay element	지연 요소
4	description	서술	concept description	개념 서술
	specification	명세	specification language	명세 언어
5	fault	결함	fault trace	결함 추적
	failure	고장	failure access	고장 접근
	error	오류	error prediction	오류 예측
6	video	영상	inverse video	반전 영상
	image	화상	dynamic image	표시 화상
7	*instance*	실례	*negative instance*	부정 실례
	example	사례	example space	사례 공간

다음 표는 반대 개념을 가지는 용어들의 목록이다. 반대 개념을 가지는 용어들은 용어의 일관성을 위하여 대역어를 일관성 있게 부착하였다.

〈표 2〉 반대 개념을 가지는 용어 목록

	영어 단일어	대역어	영어 용어	용어 대역어
1	row	가로줄	tape row	테이프(=띠) 가로줄
	column	세로줄	*card column*	카드(=딱지) 세로줄
2	upload	올려싣다	upload	올려싣다
	download	내려싣다	download	내려싣다
3	*uplink*	위로연결	uplink	위로연결
	downlink	내림연결	downlink	내림연결
4	supercomputer	대형컴퓨터	supercomputer	대형컴퓨터
	minicomputer	소형컴퓨터	minicomputer	소형컴퓨터
	microcomputer	초소형컴퓨터	microcomputer	초소형컴퓨터
5	foreground	전경	foreground	전경
	background	배경	background	배경

4. 결 론

이 연구에서 한편으로는 전문용어의 표준화 지침을 마련하고, 다른 한편으로는 구체적으로 ISO 2382 정보기술분야 용어를 중심으로 남북한 용어 표준으로서 통일안을 마련하고자 하였다. 엄격히 말해서 본 연구의 범위는 용어 표준화 지침을 마련하여 남북 통일안에 대비한 남한측 용어 표준안을 제시하는 것이었다. 다행히도 본 연구 수행 중에 남북 학자들이 만나서 그 동안의 연구 성과를 토대로 상당 부분 남북 표준안에 대한 합의를 도출할 수 있었다.

이와 같은 연구 결과가 현실적인 성과를 거두기 위해서는 여러 가지 측면에서 고려할 점들이 있다. 용어 표준의 공감대는 형성되었지만, 현재의 정치 사회적 환경의 차이로 말미암아 언어 습관이 완전히 동일하다고는 할 수 없기 때문에 어느 정도 현실적인 고려는 불가피하다. 남한의 '상호'라는 어휘가 북한에서는 '호상'이라고 표현되며, 남한의 '지시' 대신에 북한에서는 '지령'이 자연스럽게 사용되고 있는 예가 이에 해당한다. 이런 사회 이념적 색체가 가미된 용어에 대해서는 무조건 통일안을 만들기 보다는 상대방 안을 인정해야 하는 경우라고 할 수 있다.

기존 우리말 용어들의 분석을 통해서 표준화 지침을 마련하였지만 모든 용어에 대해서 이러한 지침을 충분히 적용했다고는 볼 수 없다. 굳어진 외래어 등이 이에 해당하는데 이는 표준안을 도출하기 위해서는 "단칼에 전부를 해결하기"보다는 점진적인 추진전략이 필요함을 의미한다. 사실 현재의 표준화 지침을 통한 용어의 표준안은 현 상황에서 가능한 최대한의 작업결과 일뿐, 향후에는 현재 미흡했던 부분에 대해 계속적으로 표준화를 확대 적용해야 하며, 다른 한편으로는 마구 쏟아지는 신조용어들에 대해서 현 표준화 방향에서 일관성 있게 적용해야 하는 양방향의 과제가 남아 있다.

북한에서는 통일안이 규격화를 통해서 강제력을 갖고 산업현장에서 실천될 수 있는 반면에 남한에서는 규격화가 된다고 하더라도 사회적 파급효과는 훨씬 낮다고 할 수 있다. 강제력이 약하기 때문이다. 따라서, 본 연구의 성과물이 규격화되고 교육 및 산업 현장에서 실천될 수 있도록 하기 위해서는 각 정부기관, 학회간의 연계를 활발히 하는 전략적인 노력이 필요하다.

산업자원부 기술표준원 산하의 전문용어전문기술위원회인 ISO/TC 37 한국전문위원회와 우리나라의 각 용어표준화 및 학회의 공조를 통해서 용어 표준화 지침을 검토하고 표준안 도출 작업을 수행하였으며, 그러한 바탕 위에서 남북의 합의안을 도출하였다. 이러한 단계적 성과를 발판으로 해서 남북 산업 전반의 표준화를 지속적으로 추진할 수 있는 계기가 마련될 것이라고 생각한다.

감사의 글

이 연구는 산업자원부 지원 "남북한 IT 기본 규격용어 표준화 과제" (2004.4~2005.3)의 결과로서 그 보고서에서 발췌한 것이다.

저자약력

최 기 선

1978년 서울대학교 자연과학대학 수학과 졸업 (학사)
1980년 한국과학기술원 전산학과 졸업 (석사)
1986년 한국과학기술원 전산학과 졸업 (박사)
1987년~1988년 일본 NEC C&C 정보연구소 연구원
1988년~현재 한국과학기술원 전산학과 교수
1997년~1998년 미국 스탠포드대학 CSLI 객원교수
2002년~2003년 일본 NHK 방송기술연구소 초빙연구원
2006년-현재 한국인지과학회 회장, ISO/TC37 부의장, 시맨틱웹첨단연구센터소장
2003년-현재 국가지정 언어자원특수소재은행장
http://bola.kaist.ac.kr
2002년-현재 ISO/TC37/SC4 언어자원관리표준 Secretary, TermNet 회장
2000년-현재 ACM TALIP, IJCPOL 편집위원, IAMT council member
1998년-현재 전문용어언어공학연구센터
http://korterm.or.kr/
관심분야 : 온톨로지, 텍스트마이닝, 인공두뇌, 지식획득, 창의계산론, 언어공학, 시맨틱웹
이 메 일 : kschoi@cs.kaist.ac.kr

변 정 용

1980년 2월 동국대학교 컴퓨터공학과(학사)
1983년 2월 동국대학교 컴퓨터공학과(석사)
1994년 2월 홍익대학교 전자계산학과(박사)
1982년~1988년 한국전자통신원 선임연구원
1988년-현재 동국대학교 컴퓨터멀티미디어학과 교수
1986년-현재 한국정보과학회 종신회원
1998년-현재 한국정보처리학회 종신회원
관심분야 : 한글공학, 데이터베이스, 웹서비스
이 메 일 : byunjy@dongguk.ac.kr

신 효 식

1980년 2월 전북대학교 독어교육과 (학사)
1984년 2월 서울대학교 독어독문과 (석사)
1991년 8월 독일 Regensburg 대학 언어학과 (Dr. Phil.)
2002년 8월 연세대학교 컴퓨터시스템학과 (석사)
2000년 L&H Korea 책임연구원
2001년~2002년 한국과학기술원 전산학과 초빙교수
2003년~2005년 한국과학기술원 전문용어언어공학연구센터 선임연구원
2006년-현재 전북대학교/군산대학교 시간강사
관심분야: 전문용어, 통사 의미론, 온톨로지, 디지털 언어자원

옥 철 영

1982년 서울대학교 컴퓨터공학과(학사)
1984년 서울대학교 컴퓨터공학과(석사)
1993년 서울대학교 컴퓨터공학과(박사)
1994년 러시아 TOMSK 공과대학 교환교수
1996년 영국 GLASGOW 대학교 객원교수
1984년-현재 울산대학교 컴퓨터정보통신공학부 교수
관심분야 : 한국어정보처리, 의미분별, 온톨로지, 지식베이스, 기계학습, 문서분류
이 메 일 : okcy@ulsan.ac.kr