

## 과학기술 용어에 대한 용어 생명주기 고찰

-전자신문 말뭉치를 중심으로-

A Study on Term Life Cycle for Science & Technology Terms

-Focused on 'ETNEWS' Corpus-

정한민, 성원경  
한국과학기술정보연구원

Jung Han-Min, Sung Won-Kyung  
KISTI

### 요약

과학기술 분야에서의 기술 발전 속도에 맞추어 해당 분야의 용어 역시 생성과 소멸을 끊임없이 반복한다. 본 연구에서는 아직까지 체계적으로 다루어지고 있지 않은 용어 생명주기를 정의하고 다양한 관점으로 대용량 말뭉치에서 추출된 용어들을 분석함으로써 그 필요성을 주장하고자 한다. 용어 추출을 위한 대상으로서 12년 분량의 1,700여만 어절을 포함하는 전자신문 말뭉치를 선정하였는데, 그 이유는 여러 기간에 걸친 말뭉치이므로 용어 생명주기 추이를 파악하기 용이하며, 기술 발전이 빠른 대표적인 분야인 컴퓨터/IT/전기·전자를 포함하고 있기 때문이다. 본 연구를 통해 기간에 따른 용어 생명주기 변화, 용어 수명과 전문성과의 관계 등 여러 분석 결과를 얻었으며, 앞으로 일반 사전에의 등재 여부를 결정이나 유사 기술 간의 경쟁 구도를 분석 등 다양한 응용 분야에 활용될 것으로 기대한다.

### Abstract

Keeping pace with the speed of development of science & technology domain, the domain terms continuously repeat the step of creation and extinction. This study tries to define term life cycle and analyze extracted terms from a large corpus with the viewpoint. We chose 'ETNEWS' corpus which includes about 17 million Eojeols for 12 years because it is easy to inspect the transition of term life cycle and the corpus represents computer, IT, and electrotechnology domains. This study acquired several useful conclusions including the relation between specificity and life of terms. We expect that term life cycle will contribute to analyze the competition of similar technologies and determine which term be registered into general dictionary.

## I. 서론

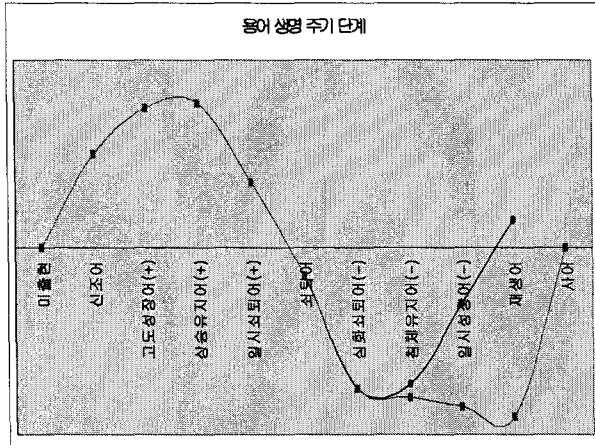
용어는 기술을 간접적으로 표현하므로 용어 수명 역시 해당 분야 발전 속도에 의존할 수밖에 없다. 과학기술 분야 중 특히 컴퓨터/IT/전기·전자 등의 경우 하루가 다르게 발전하는 기술 속도에 따라 새로운 용어가 끊임없이 생겨나고 소멸된다. 단적인 예로 2004년 전자신문 말뭉치에 출현한 복합어 중 약 44%가 그 이전 시기에 출현한 적이 없는 신조어이다. 분야의 기술 발전 속도가 빠를수록 용어의 수명이 짧아지는 경향이 있으므로 이러한 생명주기를 체계적으로 관리할 필요가 있다. 본 연구는 실제 사용하고 있는 대용량 말뭉치에서 용어를 추출하고 이것을 용어지배지수라는 측정자를 이용하여 용어 별 생명주기를 관찰하고자 한다. 특히, 용어 생명주기를 10단계로 세분화하여 이전 연구 [3]에서 깊이 다루지 못한 시간의 흐름에 따른 생명주기 추이를 좀 더 면밀하게 다룰 수 있게 되었다. 아직까지 세계적으로 용어 생명주기를 정의하고 분석하는 시도가 보이지 않는데, 다만 지식 관리 시스템 등에서 지식 생명

주기를 다루려는 시도 정도만 있었다 [4]. 본 연구에서는 용어 생명주기 단계를 세분화하여 말뭉치에서 추출된 용어에 적용하고 관찰 기간 길이, 기간 변화, 출현 빈도, 용어 수명 등 4가지 관점에서 용어 생명주기를 분석함으로써 다양한 결론에도 도달할 수 있도록 한다. 먼저 2장에서는 용어 생명주기를 정의하고, 3장에서 말뭉치 수집 및 용어 추출 방법을 기술한다. 4장에서 다양한 분석 결과를 보여줌으로써 용어 생명주기 관리의 필요성을 간접적으로 설명한다.

## II. 용어 생명주기

용어 생명주기는 용어지배지수 (TDV; Term Dominance Value)를 이용하여 모델링한다[3] [5]. 용어지배지수는 일정 기간 내에서의 관찰을 통해 이전 기간에 비해 관찰되는 용어가 어떤 추이를 가지는 지를 알려준다. 추이를 결정하는 요소로는 연도 별 말뭉치 크기를 고려하여 정규화된 용어 출현 빈

도 (NTF; Normalized Term Frequency), 관찰 기간 내에서의 NTF 평균 (ANTF; Average NTF), 기간 별 가중치 (PW; Period Weight), 시점 빈도 (PF; Period Frequency)가 있다.



▶▶ 그림 1. 용어 생명주기 단계와 TDV (기준선을 중심으로 TDV가 양수와 음수로 나누어진다.)

본 논문에서는 이에 더해 용어 생명주기 단계를 두 가지 기준을 이용하여 구분하고자 한다. 먼저 이전 기간과 상관없이 현재 기간에서의 절대적인 TDV 값에 따라 긍정적 단계와 부정적 단계로 나눈다. 긍정적 단계는 TDV가 양수일 때인 신조어, 고도성장어(+), 상승유지어(+), 일시쇠퇴어(+), 재생어 단계를 포함하며, 부정적 단계는 TDV가 음수 또는 0을 가질 때인 쇠퇴어, 심화쇠퇴어(-), 침체유지어(-), 일시성장어(-), 사어 단계를 포함한다. 긍정적 단계는 기간 가중치를 고려하여 관찰 기간 중 현재 시점에 가까울 때의 정규화된 용어 빈도가 상대적으로 높은 경우이므로 비록 일시적으로 TDV 값이 하락을 하더라도 용어 쓰임의 지속 시기가 어느 정도 기간 동안 유지될 가능성이 크다.

나머지 하나의 기준은 TDV의 상대적 추이 변화에 따라 구분하는 것으로 활성 단계와 위축 단계가 있다. 활성 단계는 현재 TDV가 이전 TDV에 비해 상승하는 경우에 해당하는 신조어, 고도성장어(+), 상승유지어(+), 일시상승어(-), 재생어 단계를 포함하며, 위축 단계는 현재 TDV가 이전 TDV에 비해 하락하는 (즉, 정규화된 용어 출현 빈도가 상대적으로 감소하는) 쇠퇴어, 심화쇠퇴어(-), 침체유지어(-), 일시쇠퇴어(+), 사어 단계를 포함한다. 활성 단계는 용어의 쓰임이 직전 기간 보다 상대적으로 증가했다는 것을 의미하며 그 연속성에 주목할 필요가 있다.

### III. 말뭉치 수집 및 용어 추출

용어 추출을 위한 말뭉치 선정을 위해서는 먼저 TDV를 결정하는 요소들을 고려해야 한다. 관찰 기간이 충분하도록 수집할 수 있는 기간이 긴 말뭉치를 수집 대상으로 선정해야 한다. 또한, 용어 생명주기를 잘 반영할 수 있는, 용어의 생성과 소멸이 뚜렷한 분야의 말뭉치를 대상으로 하는 것이 분석에 유리하다. 이러한 의미에서 과학기술 말뭉치(특히, 컴퓨터/IT/전기·전자 분야)가 선호되는데, 그 이유는 과학기술 용어가 타 분야의 그것에 비해 생명주기가 비교적 짧기 때문이다. 용어 생명주기는 분야의 발전 속도와 관련이 있는데, 인문사회 분야를 비롯한 타 분야에 비해 과학기술 분야의 발전 속도가 상당히 빠르기 때문에 이러한 기술을 대표하는 용어 또한 빠르게 움직이게 된다. 본 연구에서는 1994년부터 2005년까지 12년간의 전자신문 말뭉치를 수집하고 용어를 추출하여 이를 토대로 용어 생명주기를 살펴보고자 한다.

[표 1] 수집·정제된 전자신문 말뭉치

연도	어절 수
1994	38,103
1995	126,104
1996	109,963
1997	135,008
1998	175,312
1999	228,976
2000	435,337
2001	1,219,308
2002	2,092,567
2003	1,609,785
2004	4,944,764
2005	6,374,469
합계	17,489,696

#### 1. 전자신문 말뭉치 수집

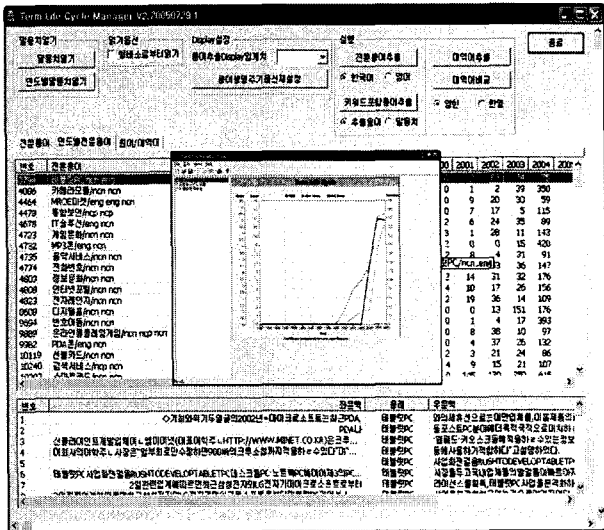
전자신문 말뭉치 수집을 위해 웹 로봇을 이용하여 전자신문 웹 사이트 (<http://www.etnews.co.kr/>)로부터 웹 문서를 수집하였다. 수집된 웹 문서에서 불필요한 HTML 태그 정보를 제거하여 텍스트 파일 형식의 원문을 추출하였다. 이를 통해 얻어진 전자신문 말뭉치 크기는 연도 별로 다음과 같다.

2005년의 경우 1994년에 비해 12년 동안 약 170배 정도 커졌다는 사실을 통해 정보의 증가 속도에 맞추어 전자신문 말뭉치 역시 그 크기가 꾸준히 증가했음을 알 수 있다.

#### 2. 과학기술 용어 추출

본 연구에서 사용한 과학기술 분야 용어 추출기 [1]는 형태소 분석기와 불용어 처리기를 포함하고 있으며, 여러 연도에 걸친 말뭉치로부터 용어를 추출할 수 있다. 또한 용어 생명주기 기법을 적용하여 각 용어에 대한 생명주기 단계를 결정해

주는 기능을 가지고 있다. 불용어 처리는 추출된 용어에 과학 기술 용어로서 부적절한 일반 명사 또는 고유 명사를 포함하는 경우를 필터링하기 위함이며, 형태소 분석은 명사 상단어 구 및 복합 명사의 추출을 가능하게 한다. 연도별로 추출된 용어 정보는 TDV로 표현되며 그 결과는 그림 2의 용어 생명주기 그래프를 통해 시각화된다.



▶▶ 그림 2. 용어 추출기와 용어 생명주기 그래프 화면 예

#### IV. 용어 생명주기 분석

용어 생명주기 분석은 전 장에서 추출한 용어를 대상으로 이루어진다. 본 장에서는 용어 생명주기 분석을 관찰 기간 크기, 관찰 기간 변화, 출현 빈도, 용어 수명 등 4가지 관점에서 수행한다.

##### 1. 관찰 기간 길이에 따른 용어 생명주기 분석

관찰 기간 길이에 따라 용어 생명주기 단계가 어떻게 변하는 지 살펴보기 위해 용어 생명주기 단계 중 출현 빈도가 상대적으로 높은 고도성장어(+), 신조어, 쇠퇴어, 재생어, 일시성장어(-), 심화쇠퇴어(-) 단계를 선택했다. 용어 생명주기의 거의 모든 단계를 거치는 경우에는 그림 1과 같이 신조어, 고도성장어(+), 상승유지어(+), 일시쇠퇴어(+), 쇠퇴어, 심화쇠퇴어(-), 일시성장어(-), 재생어 단계를 차례로 거치게 된다. 이러한 생명주기 변화 추이는 관찰 기간 길이에 따라 그 급격함에서 차이가 생긴다. 예를 들어, 고도성장어(+) 단계의 경우에는 관찰 기간이 3 (년)이나 4 (년)인 경우에는 일시쇠퇴어(-) 단계로의 이동이 가장 많은데 비해 2 (년)인 경우에는 쇠퇴어 단계로의 이동이 가장 많다. 관찰 기간이 짧을수록 용어 출현 빈도 변화에 민감하게 반응하는데, 그러한 이유로 관찰 기간이 짧은 경

우 용어 생명주기 단계가 급격하게 움직인다. 반면, 관찰 기간이 긴 경우에는 다음 단계로 이동할 때 보다 많은 단계를 거치면서 완만한 흐름을 가진다 (예를 들어, 관찰 기간이 4 (년)인 경우에 소멸 후 3년 간 심화쇠퇴어(-) 단계를 거친다). 고도성장어(+)의 경우에는 다음 단계로서 고도성장어(+), 상승유지어(+), 일시쇠퇴어(-)를 가질 때 보다 완만한 용어 생명주기의 흐름이라 볼 수 있다. 그렇지만 바로 쇠퇴어로 이동하는 경우에 그 흐름은 급격히 바뀌게 되므로 용어 생명주기 관찰이 불안정할 수밖에 없지만 대신 용어의 쇠퇴와 소멸을 빨리 파악할 수 있다. 즉 기술 발전 속도에 맞추어 관찰 기간을 설정하는 것이 필요하다.

[표 2] 관찰 기간 길이에 따른 용어 생명주기 변화 추이

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율	
2	고도성장어(+)	고도성장어(+)	7	6.14% (7/114)	
		상승유지어(+)	1	0.88% (1/114)	
		쇠퇴어	78	68.42% (78/114)	
3	고도성장어(+)	고도성장어(+)	23	15.44% (23/149)	
		상승유지어(+)	1	0.67% (1/149)	
		쇠퇴어	36	24.16% (36/149)	
4	고도성장어(+)	고도성장어(+)	49	27.07% (49/181)	
		상승유지어(+)	0	0.00% (0/181)	
		쇠퇴어	14	7.74% (14/181)	
			일시쇠퇴어(+)	118	65.19% (118/181)

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율
2	신조어	쇠퇴어	5	55.56% (5/9)
		일시쇠퇴어(+)	4	44.44% (4/9)
3	신조어	고도성장어(+)	1	12.50% (1/8)
		일시쇠퇴어(+)	7	87.50% (7/8)
4	신조어	고도성장어(+)	2	25.00% (2/8)
		일시쇠퇴어(+)	6	75.00% (6/8)

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율	
2	쇠퇴어	심화쇠퇴어(-)	40	22.35% (40/179)	
		일시성장어(-)	87	48.60% (87/179)	
3	쇠퇴어	재생어	52	29.05% (52/179)	
		심화쇠퇴어(-)	70	50.73% (70/138)	
4	쇠퇴어	일시성장어(-)	48	34.78% (48/138)	
		재생어	20	14.49% (20/138)	
4	쇠퇴어	심화쇠퇴어(-)	116	69.46% (116/167)	
		일시성장어(-)	38	22.76% (38/167)	
		재생어	12	7.19% (12/167)	
			침체유지어(-)	1	5.99% (1/167)

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율
2	재생어	고도성장어(+)	20	7.69% (20/260)
		상승유지어(+)	1	0.38% (1/260)
		쇠퇴어	186	71.54% (186/260)
		일시쇠퇴어(+)	53	30.39% (53/260)
3	재생어	고도성장어(+)	49	36.57% (49/134)
		상승유지어(+)	2	1.49% (2/134)
		쇠퇴어	35	26.12% (35/134)
		일시쇠퇴어(+)	48	35.82% (48/134)
4	재생어	고도성장어(+)	20	35.71% (20/56)
		쇠퇴어	15	26.79% (15/56)
		일시쇠퇴어(+)	21	37.50% (21/56)

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율
2	일시성장어(-)	심화쇠퇴어(-)	23	24.47% (23/94)
		일시성장어(-)	35	37.23% (35/94)
		재생어	36	38.30% (36/94)
3	일시성장어(-)	심화쇠퇴어(-)	23	19.33% (23/119)
		일시성장어(-)	40	33.61% (40/119)
		재생어	55	46.22% (55/119)
4	일시성장어(-)	심화쇠퇴어(-)	18	26.47% (18/68)
		일시성장어(-)	27	39.71% (27/68)
		재생어	23	33.82% (23/68)

관찰 기간	기간1 (2000~2004)	기간2 (2001~2005)	출현 빈도	출현 비율
2	심화쇠퇴어(-)	심화쇠퇴어(-)	5	8.07% (5/62)
		일시성장어(-)	35	56.45% (35/62)
		재생어	22	35.48% (22/62)
3	심화쇠퇴어(-)	심화쇠퇴어(-)	5	4.63% (5/108)
		일시성장어(-)	66	61.11% (66/108)
		재생어	37	34.26% (37/108)
4	심화쇠퇴어(-)	심화쇠퇴어(-)	19	19.00% (19/100)
		일시성장어(-)	60	60.00% (60/100)
		재생어	21	21.00% (21/100)

## 2. 기간 변화에 따른 용어 생명주기 분석

용어 별로 생명주기 변화 추이를 살펴본다면 용어마다 편차가 심해서 생명주기 흐름을 파악하기 어렵다. 이에 본 절에서는 용어 집합을 선정하고 해당 용어 집합의 생명주기 변화 추이를 집단적으로 관찰함으로써 분야 수준에서 용어 쓰임의 성숙 및 소멸 단계로의 이동 여부를 파악하고자 한다. 용어 집합은 기간1 (2000~2004)과 기간2 (2001~2005) 모두에 대해 관찰 기간 (4년) 동안 50회 이상 출현한 용어 781개를 대상으로 하였다.

표 3을 보면, 기간1에서 기간2로 이동할 때 생명주기 단계가 부정적 단계 (두 번째 표 참조)로, 그리고 위축 단계 (세 번째 표 참조)로 이동하는 경향을 발견할 수 있다. 용어 출현 빈도를 다르게 하더라도 이 결과는 동일하다 (4.3절의 표 4와 표

5 참조). 반면, 기간 2에서만 용어 출현 빈도 50이상을 가지는 용어 집합 (네 번째 표 참조)의 경우 긍정적 단계와 활성 단계 비율이 기간 1에서부터 유지되어 온 용어 집합 (첫 번째 표 참조)에 비해 20% 정도 높게 나타남을 알 수 있는데, 이 결과는 용어 생명주기 상에서의 성숙도가 낮다 (즉, 신생 용어 집합)는 것을 의미한다.

[표 3] 기간 변화에 따른 용어 생명주기 변화 추이 (관찰 기간 내 용어 출현 빈도 50이상)

관찰 기간	생명주기 단계	기간1 (2000~2004)		기간2 (2001~2005)	
		출현 빈도	출현 비율	출현 빈도	출현 비율
4	고도성장어(+)	181	23.18%	94	12.04%
	상승유지어(+)	2	0.26%	1	0.13%
	쇠퇴어	167	21.38%	148	18.95%
	신조어	8	1.02%	0	0.00%
	심화쇠퇴어(-)	100	12.80%	153	19.59%
	일시성장어(-)	68	8.71%	125	16.01%
	일시쇠퇴어(+)	198	25.35%	203	25.99%
	재생어	57	7.30%	56	7.17%
	침체유지어(-)	0	0.00%	1	0.13%
	합계	781	100.00%	781	100.00%

기간	단계 유형	비율
기간1 (2000~2004)	긍정적 단계	57.11%
	부정적 단계	42.89%
기간2 (2001~2005)	긍정적 단계	45.33%
	부정적 단계	54.67%

기간	단계 유형	비율
기간1 (2000~2004)	활성 단계	40.46%
	위축 단계	59.54%
기간2 (2001~2005)	활성 단계	35.34%
	위축 단계	64.66%

생명주기 단계	출현 빈도	긍정적 단계	활성 단계
고도성장어(+)	155	0	0
사어	0		
상승유지어(+)	0	0	0
쇠퇴어	41		
신조어	7	0	0
심화쇠퇴어(-)	24		
일시성장어(-)	41		0
일시쇠퇴어(+)	134	0	
재생어	57	0	0
침체유지어(-)	0		
합계	459	76.91% (353/459)	56.64% (260/459)

### 3. 출현 빈도에 따른 용어 생명주기 분석

용어 집합의 평균 출현 빈도에 따라 용어 생명주기의 변화 속도가 달라진다. 출현 빈도가 낮은 경우에는 용어의 쓰임이 상대적으로 활발하지 못하다는 의미이므로 이러한 용어가 지속적으로 살아남을 확률 역시 떨어질 수밖에 없다.

이를 표 4에서 확인할 수 있는데, 표 3보다 사어의 비율이 높으며, 기간 2에서 부정적 단계나 위축 단계로 진입하는 비율이 10% 정도 높게 나타난다. 고도성장어(+)의 비율 역시 상당한 차이를 보임을 알 수 있다. 표 5는 평균 용어 출현 빈도가 표 3과 표 4의 중간에 해당하는 경우인데, 용어 생명주기 변화 속도 역시 그 중간에 속함을 알 수 있다.

[표 4] 기간 변화에 따른 용어 생명주기 변화 추이 (관찰 기간 내 용어 출현 빈도 10~19)

관찰 기간	생명주기 단계	기간1 (2000~2004)		기간2 (2001~2005)	
		출현 빈도	출현 비율	출현 빈도	출현 비율
4	고도성장어(+)	134	10.33%	17	1.31%
	사어	3	0.23%	12	0.93%
	상승유지어(+)	7	0.54%	0	0.00%
	쇠퇴어	281	21.67%	333	25.67%
	신조어	174	13.42%	0	0.00%
	심화쇠퇴어(-)	189	14.57%	291	22.44%
	일시성장어(-)	121	9.33%	204	15.73%
	일시쇠퇴어(+)	282	21.74%	352	27.14%
	재생어	105	8.10%	81	6.25%
	침체유지어(-)	1	0.08%	7	0.54%
	합계	1,297	100.00%	1,297	100.00%

기간	단계 유형	비율
기간1 (2000~2004)	긍정적 단계	54.12%
	부정적 단계	45.88%
기간2 (2001~2005)	긍정적 단계	34.70%
	부정적 단계	65.30%

기간	단계 유형	비율
기간1 (2000~2004)	활성 단계	41.71%
	위축 단계	58.29%
기간2 (2001~2005)	활성 단계	23.28%
	위축 단계	76.72%

[표 5] 기간 변화에 따른 용어 생명주기 변화 추이 (관찰 기간 내 용어 출현 빈도 20~49)

관찰 기간	생명주기 단계	기간1 (2000~2004)		기간2 (2001~2005)	
		출현 빈도	출현 비율	출현 빈도	출현 비율
4	고도성장어(+)	109	13.99%	40	5.13%
	사어	0	0.00%	1	0.13%
	상승유지어(+)	1	0.13%	0	0.00%
	쇠퇴어	169	21.69%	186	23.88%
	신조어	38	4.88%	0	0.00%
	심화쇠퇴어(-)	111	14.25%	164	21.05%
	일시성장어(-)	82	10.53%	128	16.43%
	일시쇠퇴어(+)	200	25.67%	190	24.39%
	재생어	68	8.73%	68	8.73%
	침체유지어(-)	1	0.13%	2	0.26%
	합계	779	100.00%	779	100.00%

기간	단계 유형	비율
기간1 (2000~2004)	긍정적 단계	53.40%
	부정적 단계	46.60%
기간2 (2001~2005)	긍정적 단계	38.25%
	부정적 단계	61.75%

기간	단계 유형	비율
기간1 (2000~2004)	활성 단계	38.25%
	위축 단계	61.75%
기간2 (2001~2005)	활성 단계	30.30%
	위축 단계	69.70%

### 4. 용어 수명과 전문성

전문용어의 전문성을 측정하는 방법으로 시소러스 상의 계층 관계를 이용한 시도가 있으나 [2] [6], 용어 생명주기 관점에서 용어의 전문성을 측정하는 연구는 아직까지 없는 실정이다. 과학기술 발전 속도는 해당 분야 전문용어 수명의 길이에 영향을 미친다. 표 6과 표 7을 보면 그 차이를 알 수 있는데, 전자는 10여 년 동안 계속 유지되는 용어 집합으로서 컴퓨터와 IT 분야에서 범용으로 사용되는 용어이다. 비록 이들 용어가 신규로 출현할 당시에는 전문성을 가진 용어였을지라도 시간이 흐르면서 전문성을 잃고 일상적으로 사용되는 쓰임으로 변화되었거나 지속적으로 발전하는 기술과 관련된 것으로 보인다. 반면 후자는 시대적인 유행에 따라 잠시 출현하고 소멸되었거나, 살아남지 못한 기술과 관련된 용어를 포함한다.

[표 6] 기간0 (1995~1999)과 기간2 (2001~2005)에서 공유되는 용어 (기간0 내 용어 출현 빈도 50 이상)

용어	용어 생명주기 단계
중계유선	심화쇠퇴어(-)
정보통신기기	일시성장어(-)
정보문화	일시쇠퇴어(+)
위성방송	심화쇠퇴어(-)
디지털TV	쇠퇴어
전자문서	일시성장어(-)
이동통신	일시성장어(-)
개인정보	일시쇠퇴어(+)
이동전화	일시성장어(-)
실리콘밸리	일시성장어(-)
정보가전	재생어
비메모리	심화쇠퇴어(-)
컴퓨터프로그램	고도성장어(+)
컬러TV	심화쇠퇴어(-)
전기통신	심화쇠퇴어(-)
PC통신	일시성장어(-)
메모리반도체	일시성장어(-)
전자상거래	일시성장어(-)
정보통신망	심화쇠퇴어(-)
초고속망	일시성장어(-)
정보통신기술	일시쇠퇴어(+)
CD롬	심화쇠퇴어(-)
정보통신	심화쇠퇴어(-)
지식정보	재생어
디지털방송	쇠퇴어
전산시스템	일시성장어(-)
컴퓨터바이러스	재생어
통신서비스	일시성장어(-)
정보기술	일시성장어(-)
D램	심화쇠퇴어(-)
...	...
초고속정보통신망	일시성장어(-)

현 빈도 등 다양한 요소에 의해 같은 용어일지라도 여러 형태로 된다. 본 논문에서는 여러 요소들의 변화시키면서 나타난 현상들을 분석함으로써 용어 생명주기의 의미와 활용 가능성을 살펴보았다. 특히, 12년 분량의 1,700여만 어절을 포함하는 대용량 말뭉치를 이용해서 그 분석의 신뢰성을 높였다는 데 의의가 있다.

향후 본 연구는 상기에서 언급한 바와 같이 용어의 전문성을 측정하거나, 일반 사전에의 등재 여부를 결정하거나, 또는 유사 기술 간의 경쟁 구도를 분석하는 등의 다양한 응용 분야에 초점을 맞출 예정이다.

■ 참고 문헌 ■

- [1] 구희관, 정한민, 이병희, 성원경, 전문용어사전 구축을 위한 전문 용어 추출 및 순위화, 한국정보처리학회 춘계학술대회, 2005.
- [2] 구희관, 정한민, 이병희, 성원경, 수식어구를 이용한 한국어 용어의 전문성 측정, 한국컴퓨터종합학술대회, 2005.
- [3] 정한민, 구희관, 이병희, 성원경, 효율적인 자원 운영을 위한 전문용어 생명주기 관리 연구, 한국컴퓨터종합학술대회, 2005.
- [4] J. Birkinshaw and T. Sheehan, Managing the Knowledge Life Cycle, In Journal of MIT Sloan Management Review 44(1), 2002.
- [5] H. Jung, H. Koo, B. Lee, and W. Sung, Toward Managing the Life Cycle of Terms Using Term Dominance Trend, In Proceedings of Pacific Association for Computational Linguistics, 2005.
- [6] P. Ryu, Determining the Specificity of Terms using Compositional and Contextual Information, In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, 2004.

[표 7] 표 4의 기간2 (2001~2005)에서 사이어가 된 용어 (색칠된 부분은 추출 오류)

용어	용어 생명주기 단계
HDR서비스	사이어
파운데이션인큐베이터	사이어
V커머스	사이어
콘텐츠디바이스	사이어
네트워크경매	사이어
스마트패널	사이어
사이버은어	사이어
E북서	사이어
오레일리	사이어
MAN솔루션	사이어
보복전쟁이	사이어
콤포물	사이어

V. 결론

본 연구를 통해서 과학기술 분야에서의 용어 생명주기가 의미하는 바를 살펴보았다. 용어 생명주기는 관찰 기간, 용어 출