

# 웹에서 사용되는 데이터 신뢰성에 관한 연구

이석기\*, 김재중\*\*, 곽계달\*\*\*, 신승중\*\*\*\*

\*한양대학교 공학대학원 컴퓨터공학과

\*\*한양대학교 신뢰성분석연구센터

\*\*\*한양대학교 전기전자컴퓨터공학부

\*\*\*\*한세대학교 IT 학부

e-mail : expersin@hansei.ac.kr

## A Study on Data Reliability Used on the Internet

Seok-Ki Lee\*, Jae-Jung Kim\*\*, Kae-Dal Kwack\*\*\* and Seung-Jung Shin\*\*\*\*

\*Division of Electronics and Computer Engineering, Hanyang University

\*\*Reliability Analysis Research Center, Hanyang University

\*\*\* Dept. of Electronics and Computer Engineering, Hanyang University

\*\*\*\*Dept. of Information Technology, Hansei University

### 요 약

이 글은 신뢰성 문제와 관련된 현대 정보사회의 특성을 되짚어 본 다음, 정보보호와 신뢰성의 관계에 대해 살펴보고 웹 정보를 중심으로 신뢰성의 개념, 신뢰성 평가의 방법, 웹 사이트의 신뢰성 요인, 범주별 웹 사이트 신뢰도 등에 대해 이용자들을 대상으로 조사, 분석하고자 한다. 웹 사이트의 신뢰성 평가 항목과 조사 대상자의 표본 수가 제한되어 있다는 한계를 갖고 있으므로 이 부분은 향후 지속적인 조사와 연구를 통해 보완해 나갈 필요가 있다

### 1. 서론

현대사회에서 정보란 말만큼 흔하고, 어떤 다른 말이라도 붙여 쓰면 뜻이 통하게 되는 말은 드물 것이다. 이러한 현상은 심지어 현대사회가 정보사회라고 불릴 만큼 정보가 차지하는 비중이 크다는 것을 말해준다. 현재 인류가 1년 동안 생산해내는 정보량은 약 5~10EB(10<sup>15</sup>bytes)정도에 이를 것으로 추정되어 이를 텍스트로 환산하면 약 5~10 조 권 정도에 해당된다. 그러나 그 중의 극히 일부만 수집, 정리되어 재이용이 가능한 상태가 되고 있으며, 나머지 대부분의 정보는 생산과 동시에 사라지고 있는 가운데 증가 속도 또한 기하급수적임에 비해 우리 역사와 사회에 책임 있는 지식 정보의 유통은 홀대 받고 있다.

각종 디지털 기기로 다양화되었으나, 디지털 저장매체가 불안정하고 그 수명도 상대적으로 매우 짧다는 치명적인 문제점을 갖고 있어서 미래의 후손들이 현재 생산되고 있는 디지털 자료를 지금과 같이 활용할 수 있을 것인가에 대해서는 의문을 품지 않을 수 없다. 정보의 신뢰성에 대한 평가 책임은 전적으로 이용자들의 판단에 맡겨지고 있는 상황이며, 오프라인(off-line)정보에 비해 온라인(on-line)정보, 디지털 정보의 왜곡과 편향이 훨씬 더 심각한 상황이다. 따라서, 인터넷 이용자들이 웹 사이트의 신뢰성 문제를 어떻게 인식하고 있으며, 어떤 방식으로 평가하는지에 대한 실질적이고 포괄적인 조사와 연구가 필요할 것이다. 아울러 디지털 정보의 신뢰성을 이용자가 스스로 판단할 수 있도록 도움을 주기 위한 기준, 장치, 시스템, 제도의 마련도 함께 필요할 것이다.

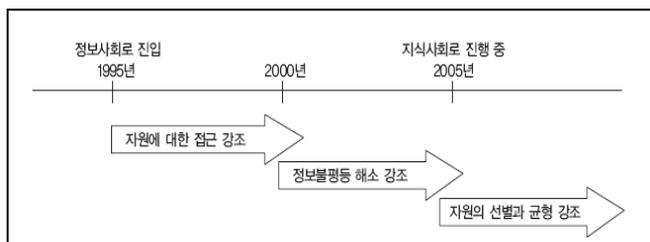


Fig.1 우리 사회 정보문화의 변화

정보사회에서는 컴퓨터의 발전으로 종이를 대신할 수 있는 각종 전산 매체가 개발되어 많은 양의 정보를 저장, 이용할 수 있게 되었을 뿐 아니라 인터넷 등 대중통신수단의 발달로 정보를 신속, 저렴하게 광범위한 사람들에게 전달할 수 있게 되었다. 지식정보의 저장 매체가 종이 중심에서

### 2. 본론

#### 1) 정보신뢰의 중요성

현대사회는 정보과다 사회이다. 이 ‘정보과다’란 말을 대하는 방식에 있어서는 두 가지 대립되는 흐름이 존재하고 있다. 하나는 ‘정보의 바다’라는 말로 대표되는 긍정적인 입장이고, 다른 하나는 ‘정보의 쓰레기’로 표현되는 다소 부정적인 입장이 그것이다.

컴퓨터의 정보처리 기술과 데이터 통신기술의 발달에 근거한 ‘정보의 바다’ 편에 설 경우, 누구나 원하기만 하면 어떤 정보든지 구할 수 있게 될 것이다. 그러나 이러한 정보는 정보생산자의 필요에 의해 생산되며, 특히 자본주의 사회에서는 이윤을 남길 수 있는 지식과 정보가 중심이 되어 적극적으로 생산, 유통된다. 따라서 가치 있는 모든 정보가

생산되는 것은 아니며, 모든 정보가 양적, 질적으로 같은 비중으로 유통되는 것은 더욱 아니다. 모든 지식과 정보는 생산과 유통, 축적 과정에서 어떤 형식으로든지 선택되며 변환 또는 왜곡될 소지를 충분히 안고 있는 것이다.

반면에 ‘정보의 쓰레기’ 라는 비판적인 입장을 취할 경우 현대사회는 극히 일부의 사람들이 자신들의 필요에 의해 만들어 놓은 정보, 원한다면 볼 수 있는 그런 사회가 되어 버린다. 범람하는 정보의 홍수 속에서 정작 우리가 필요로 하는 정보는 극히 일부에 지나지 않으며 진실하며 유효한 정보를 가려내기가 더욱 어려워졌다는 것이다.

현대사회의 지식, 정보현상을 이해하는데 있어서 정보의 과다 문제와 맞물려 고려해야 할 것은 ‘정보의 왜곡현상’에 관한 문제이며 이미 우리사회 곳곳에서 정보의 왜곡현상은 감지되고 있다. 따라서 지식 정보자원에 대한 접근을 제공하는 것도 중요하지만, 질 높은 자원을 하나하나 선정하고 안내하는 역할에 대한 중요성이 증가하고 있다.

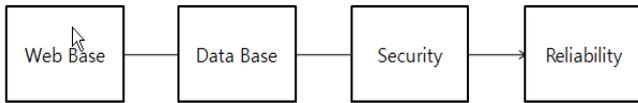


Fig.3 현대사회에서 정보신뢰의 중요성

3. 웹에서 사용되는 데이터 신뢰성 평가 연구

신뢰성의 개념과 관련된 포괄적인 연구로 Fogg 와 Tseng의 연구를 들 수 있다. 신뢰성이란 믿을 수 있음을 의미한다. 즉 신뢰성 있는 사람은 믿을 수 있는 사람을 말하며, 신뢰성 있는 정보란 믿을 수 있는 정보를 말하는 것이다. 신뢰성과 관련된 문제에 있어서 웹은 상당히 예외적인 정보원으로서, 경우에 따라 웹은 매우 믿을만한 정보원이 되기도 하며, 반대로 가장 믿기 어려운 정보원이 되기도 한다. 그 이유는 정보의 신뢰성에 대한 평가 책임은 전적으로 이용자들의 판단에 맡겨지고 있는 상황이기 때문이다. 따라서 이용자들이 웹 사이트의 신뢰성을 어떻게 평가하는지를 실제 이용자들을 대상으로 조사, 분석하고자 한다.

- 1) 웹 사이트 이용자들의 행태에 대해 분석하는 작업이다. 웹 사이트 접속 경로, 인터넷 사용 빈도, 이용자 정보의 제공 여부, 인터넷으로 물건을 구매해 본 경험, 약관 또는 개인정보 보호 정책을 읽어보는 빈도, 인터넷 검색 엔진에 정보를 등록해 본 경험 등이 포함된다.
- 2) 웹 사이트의 유형에 따른 신뢰도 조사이다. 인터넷 사이트를 사이트의 유형과 주요 제공 서비스내용에 따라 포털, 전자상거래, 뉴스, 학술 DB, 여행, 금융기관, 정부기관과 같이 7 개의 범주로 나눈 다음, 사람들이 자주 이용하는 사이트와 순위정보 사이트의 정보를 이용하여 각 범주별로 7 개씩의 웹 사이트를 선정하여 신뢰도를 조사한다.
- 3) 이용자들의 신뢰성 판단에 영향을 미친 요인을 분석하는 작업이다. 웹 사이트의 어떤 요소가 이용자들의 신뢰도 판단에 영향을 미치는지에 대한 조사, 분석 작업을 진행한다. 신뢰 요인을 유형화하여 해당 사이트를 방문하도록 결정하게 된 중요성의 정도를 분석한다.

조사는 예비조사와 본 조사로 나누어 진행하고 위의 세 가지 항목을 토대로 하여 설문조사를 실시한다. 각 항목에 대한 응답은 5 점 리커트 척도법(5-point Likert-type)을 이용하고 조사를 실시한 후 분석을 통하여 신뢰성에 대한 조사결과를 도식화 한다.

성별	카운트	CumCnt	백분율	CumPct
남성	32	32	53.33	53.33
여성	28	60	46.67	100.00
N=	60			

항목 1	카운트	CumCnt	백분율	CumPct
만족	17	17	28.33	28.33
매우 만족	16	33	26.67	55.00
매우 불만족	7	40	11.67	66.67
불만족	12	52	20.00	86.67
불만족	8	60	13.33	100.00
N=	60			

항목 2	카운트	CumCnt	백분율	CumPct
만족	12	12	20.00	20.00
매우 만족	18	30	30.00	50.00
매우 불만족	9	39	15.00	65.00
불만족	14	53	23.33	88.33
불만족	7	60	11.67	100.00
N=	60			

항목 3	카운트	CumCnt	백분율	CumPct
만족	28	28	46.67	46.67
매우 만족	8	36	13.33	60.00
매우 불만족	7	43	11.67	71.67
불만족	6	49	10.00	81.67
불만족	11	60	18.33	100.00
N=	60			

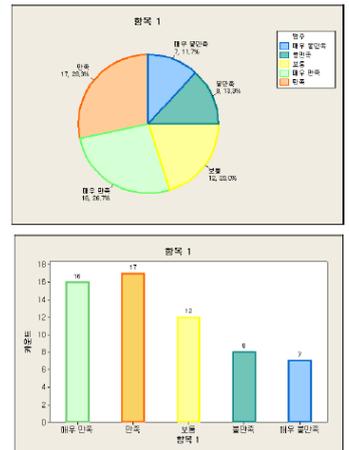


Fig.4 결과분석 예시

4. 결론

웹에서 사용되는 데이터 신뢰성 평가 연구를 통해 디지털 정보의 신뢰성을 이용자가 스스로 판단할 수 있도록 도움을 주기 위한 시스템이나 제도가 필요할 것으로 생각된다. 웹 사이트의 신뢰성 판단에 영향을 미치는 요인들에 대한 포괄적인 조사와 분석이 필요함에도 불구하고, 웹 사이트의 신뢰성 평가 항목과 조사 대상자의 표본 수가 제한되어 있다는 한계를 갖고 있으므로 이 부분은 향후 지속적인 조사와 연구를 통해 보완해 나갈 필요가 있다.

REFERENCES

- [1] 김영기(2007), KADO 이슈리포트 2007 년 10 월 통권 47 호
- [2] 최두진 등(2006), 유비쿼터스 사회의 지식, 참여, 그리고 생활, 한국정보문화진흥원 연구보고서(06-18)
- [3] Ronald B. Standler(2004), Evaluation Credibility of Information on the Internet, www.rbs0.com/credible.pdf
- [4] http://www.socialresearchmethods.net/ (2006)
- [5] Stanford web Credibility Research, http://credibility.stanford.edu
- [6] Peter Honeyman, Interdependence of Reliability and Security