

<http://dx.doi.org/10.17703/JCCT.2024.10.5.135>

JCCT 2024-9-16

정보시각화 표현의 진실과 오류에 관한 연구

Study on Truth and Error of Information Visualization Expression

김성곤*

Sungkon Kim*

요약 정보시각화는 목적에 따라서 정보전달형, 정보묘사형, 그리고 유희형 표현으로 나눈다. 이 중에서, 데이터의 테이블 형식과 내용을 파악한 후, 데이터 일정 부분을 수사학적으로 강조하여 사용자에게 전달하는 정보묘사형 표현에는 거짓이 아닌 오류가 있다. 이때 사용된 시각적 수사학표현은 사용자의 내용전달에 도움을 주지만, 사용된 데이터 그래픽은 일부 오류가 되는 딜레마가 있다. 본 연구에서는 역설법, 점층법, 생략법, 제유법, 운율법, 열거법 그리고 은유법 등의 수사학적 묘사표현에서 생기는 데이터 수치와 다른 그래픽 면적, 선택적 데이터 일부만 보여주기, 부적절한 그리드 사용, 기하학적 단순 묘사의 오해, 그리고 부적절한 그래픽 블록화 등과 같은 오류에 관해 논하였다.

주요어 : 정보시각화, 오류, 데이터 시각화, 수사학적 시각표현

Abstract Information visualization is divided into delivery, description, and beautiful fun. The dilemma is delivered to the user by emphasizing a certain part of the data as a visual rhetorical expression after grasping the table format and content of the data. The visual rhetoric representation used at this time helps users deliver content, but the data graphics used have a dilemma that becomes some false. In this study, errors that occur in rhetorical descriptions such as paradox, gradation, omission, synecdoche, meter, enumeration, and metaphor were discussed. Errors include graphic areas that differ from data values, showing only selective data, inappropriate use of grids, misunderstanding simple geometric depictions, and inappropriate graphic blocking.

Key words : Information Visualization, Error, Data Visualization, Rhetorical Visual Presentation

1. 서론

데이터를 그래픽 표현의 다양한 요소로 사용자에게 일정한 내용을 전달하는 것이 정보시각화 목적이다. 정보시각화는 목적에 따라서 3가지의 표현 방식으로 나뉜다. 첫째, 데이터 내용을 사용자에게 정확하게 전달하는 것이 목적인 정보전달(傳達)형 표현, 둘째, 사용자에게 특별한 목적을 가지고 데이터 내용을 여러 시각적 효과로 효율적으로 전달하는 정보묘사(描寫)형 표현[1], 마지막으로 데이터 내용을 예술적 혹은 풍자적으로 전달

하는 정보 유희(儒戲)형 표현이 있다.

정보시각화 오류의 딜레마는 정보묘사형 표현에서 생긴다. 정보전달형 표현은 주로 과학 실험보고 형식에서 사용한다. 여러 다이어그램을 사용하여 데이터 내용을 정확하게 표현하면 된다. 작은 공간에 많은 데이터를 동시에 표현하는 것이 중요하지만, 또한, 사용자의 인지력을 고려해야 한다[2]. 정보 유희형 표현은 광고나 예술전시회에서 사용하는 정보시각화다. 사용자에게 데이터 내용을 정확히 전달하는 목적보다는 특정 내용의 흥미 유발과 관심 유도 혹은 데이터를 활용한 미적인

*정회원, 울산대학교 디지털콘텐츠디자인학과 교수
접수일: 2024년 6월 10일, 수정완료일: 2024년 7월 10일
게재확정일: 2024년 9월 1일

Received: June 10, 2024 / Revised: July 10, 2024

Accepted: September 1, 2024

*Corresponding Author: sungkon@ulsan.ac.kr

Dept. of Digital Content Design, Univ. of Ulsan, Korea

창의적 표현에 중점을 둔다.

II. 정보시각화 진실과 오류의 딜레마

정보묘사형 표현은 데이터의 모든 값을 그래프로 정확히 전달하기보다는 정보시각화 개발 목적에 따라 의도된 데이터 일정 부분 내용만 강조하여 정확히 전달한다. 이를 위해, 정보시각화 그래픽에 표현된 일정 내용은 사실이 아닌 왜곡 혹은 부정확한 그래픽 표현을 사용한다. 이때, 의도된 왜곡표현을 위해 소설에서 사용되는 역설법, 은유법, 제유법 등의 수사학적 표현과 유사한 시각적 수사학표현이 사용된다. 그러나 이때 사용된 시각적 수사학표현은 사용자의 데이터 내용 이해에 도움을 주지만, 사용된 그래픽은 데이터 내용을 일부 부정확하게 표현한다. 내용 오류와 내용 이해 사이에서 시각적 수사학표현의 사용 여부와 표현 방식의 선택은 정보시각화 개발자의 딜레마이다.

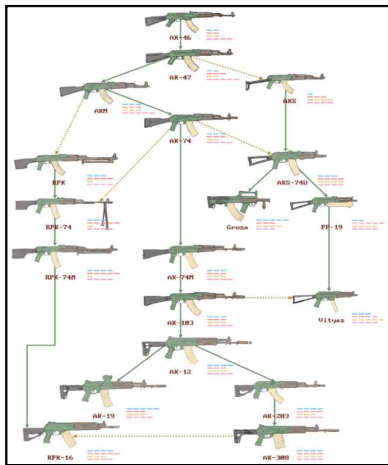


그림 1. AK47 변천사 (김정엽, 2022)
Figure 1. The History of AK47 (Jung-yeop Kim, 2022)

그림 1은 AK47 소총이 제작된 이후, 시간의 흐름에 따라서 새롭게 개조된 AK47 소총의 변화를 직관적으로 알 수 있게 만든 정보시각화이다. 각 총기 부품 부위에 따른 차이를 색상의 차이로 표현하여, 개조된 각각의 AK47 모델이 어떠한 모델에서 업그레이드되었는지를 알 수 있다. 또한, 모델별 정확도, 안정성, 착용감, 그리고 총기 화력의 정도를 비교할 수 있다.

그림 2는 금붕어의 품종 개량이 이루어진 진화 과정을 보여준다. 금붕어가 개량되어 명명된 이름, 품종 개

량이 이루어진 국가, 그리고 머리, 지느러미, 꼬리, 색상 등의 개량된 부위의 차이점을 단순하게 색상과 기하학적 도형으로 표현하였다. 또한, 각각의 품종 개량된 금붕어의 유사 아종의 형태를 부가적으로 알 수 있다.

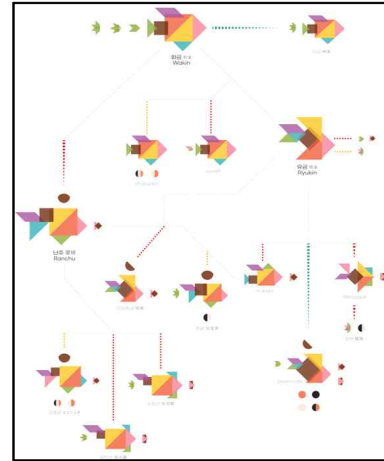


그림 2. 금붕어 품종 개량 (임수연, 2022)
Figure 2. The improvement of goldfish (Soo-yeon Lim, 2022)

그림 1의 AK47 소총과 그림 2의 금붕어에 관한 정보시각화에서 차이점은 그래픽의 수사학적 단순화 사용 여부이다. AK47 소총은 각 부위를 사실적으로 묘사해서 부위별 개조 형태를 비교하였고, 금붕어는 단순한 기하학적 형태와 색상을 사용한 추상적 묘사로 부위별 개조 형태를 비교하였다. 두 그림이 다르게 묘사된 이유는 다음과 같다. 대규모 생산 공정이 필요한 AK47 소총은 개조된 형태 변화는 다양하지 않다. 그러나 금붕어의 유전적 돌연변이의 형태는 매우 다양하다. 이런 이유로, 금붕어의 모든 형태를 사실적으로 묘사할 수 없다. 여러 금붕어 중에서 대표적 형태 표현만 강조하여 단순하게 표현했다. 수학에서 정수와 실수의 차이와 유사하다. 여러 사물의 개조를 시각화하면서 개발자는 사실적 묘사와 단순화 정도의 사이에 딜레마가 있다.

그림 3은 무역동반자 관계도를 표현한 정보시각화이다. 브라질, 미국, 중국, 그리고 호주가 어느 나라와 어떠한 상품으로 수출입 관계를 얼마만큼 가지고 있는가에 관한 시각화이다. 메르카도르 도법과 구드도법을 절충한 비교적 사실적 세계지도를 사용하고 있다.

그림 4는 전 세계 성인 에이즈 환자 분포도를 세계 지도를 통해 보여주는 정보시각화이다. 실제 대륙의 면적과는 다르게 에이즈 환자의 수에 따라 임의로 대륙의

크기를 조정하여 표현하였다.



그림 3. 세계 무역 동반자 관계도
 Figure 3. World trade partnership chart (Raconteur The Times)

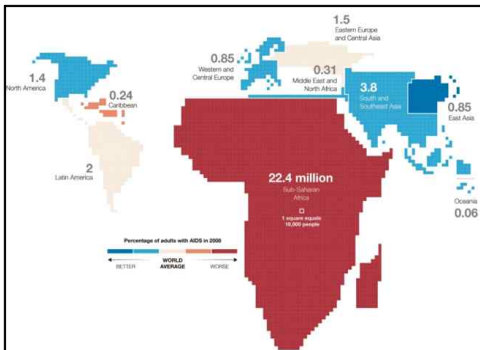


그림 4. 세계 에이즈 환자 분포도
 Figure 4. AIDS in the world (www.datavizproject.com)

그림 3.의 무역동반자 관계도와 그림 4.의 세계 에이즈 환자 분포도에 관한 정보시각화에서 차이점은 그래픽의 거짓된 역설적 표현 사용 여부이다. 두 그림이 다르게 묘사된 이유는 다음과 같다. 그림 3.은 컨테이너 무역선과 수화물 항공기를 사용한 지역 간의 운송량과 운송 거리의 내용이 포함된다. 지리적 사실적 표현이 중요하다. 그림 4.는 에이즈 환자의 대륙별 분포에 관한 내용이다. 에이즈 환자의 대륙 간 이동이나, 정확한 대륙 지역의 특정 지역 표현이 필요하지 않다. 대륙의 대략적 모습과 위치만 표현하면 된다. 중요한 것은 대륙별 분포 환자 수이다. 정보시각화에 세계지도를 표현하면서, 메르카도르 도법처럼 정확한 지도를 그릴 것인가 아니면, 지도 모습을 왜곡한 역설적 지도를 그릴 것인가에 대한 딜레마가 있다.

오류(誤謬)의 용어 사전적 정의는 ‘행위자가 의도하지 않았고, 어떤 기준에 맞지 않는 것으로 시스템이 이

해할 수 있는 한계를 넘어서는 것’이다. 오류는 ‘기준’과 ‘의도성’과 연관이 있다는 것을 말해준다. 시각화의 정보묘사 오류는 사전적 오류와 다르게 일반적인 기준에 맞지 않는 의도된 행위이다. 정보시각화에서 의도된 표현오류에는 다음과 같다[3]. 불필요한 그래픽, 설명글의 부적절한 위치, 부적절한 색상 및 명도 사용, 데이터 수치와 다른 그래픽 면적, 부적절한 배경 그래픽 사용 등의 사용자 중심의 디자인 표현으로 부적합한 ‘형식(形式)오류’와 부적절한 그래픽 블록화, X축과 Y축 그리드의 부적절한 사용, 의미 없는 그리드 사용, 그래프 비율 왜곡 등의 그리드와 관계 표현의 부정확한 사용에 기인하는 ‘묘사(描寫)오류’, 그리고 적은 데이터 자료 비교, 선택적 데이터 일부만 보여주기, 규칙 없는 나열 등의 어떤 기준에 맞지 않게 표현되어, 실질적 데이터 의미와 다르게 이해되는 ‘기준(基準)오류’ 등의 세 가지 범주가 있다. 이러한 오류는 수사학적 시각표현을 위하여 의식적으로 의도되어 표현된다.

III. 수사학적 묘사표현의 진실과 오류

사람을 설득하기 위한 언어기법을 연구하는 학문이 수사(修辭)학이다. 수사학은 문장을 장식하는 수단뿐만 아니라, 현대에는 정확한 전달과 설득을 위한 수단으로 사용된다. 이와 유사하게, 한정된 그래픽 공간 안에서 데이터 내용을 효율적으로 전달하여 사용자를 쉽게 이해시키기 위해 수사학적 정보묘사가 가능하다. 본 연구에서 수사학적 정보묘사 사례와 이와 관련된 의도된 오류에 대해 논한다[4]. 언어기법의 수사학에는 영탄법, 직유법, 경구법, 인용법, 대유법, 대구법, 제유법, 환유법, 점강법, 점층법, 은유법, 중의법, 도치법, 열거법, 의인법, 설의법, 풍유법, 반어법, 역설법, 문답법, 생략법, 그리고 비약법 등이 있다. 그러나 모든 수사학적 기법이 정보시각화의 정보묘사에 사용되지 않는다. 수천 개의 정보시각화를 조사하였고, 그중에서 정보묘사가 가능한 수사학적 기법에 관하여 사례를 논한다.

1. 역설법

역설(逆說, paradox)법은 단순히 보면 모순되어 보이지만, 표면적 내용을 떠나 깊이 생각하면, 근거가 확실한 진실을 담고 있는 표현을 뜻한다. 그림 5.는 인구 이동에 따른 지역별 인구수를 1960년대부터 2005년까지 5

년 연차적으로 보여준다. 1960년에는 국내의 인구 분포가 한국의 지도와 유사하지만, 2005년에 와서는 수도권 중심으로 집중되었다. 특히, 전라남북도의 인구가 많이 감소하였다. 한국 지도 모습이 비록 변경되었지만, 상하 좌우의 위치를 짐작하여 각 지역을 알 수 있다. 위의 그림에서 지도의 지역 위치는 모든 사람이 짐작할 수 있지만, 만약 좀 더 많은 데이터가 포함된 정보묘사일 경우 사용자는 혼돈이 올 것이다. 사용자가 이해할 수 있는 적절함의 한계의 정도는 개발자가 정해야 하는 딜레마이다.

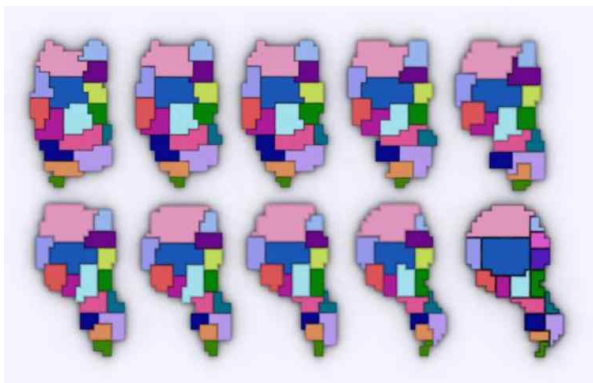


그림 5. 연도별 인구 이동 수 (박근영, 2006)
Figure 5. Population movement by year (Geun-young Park, 2006)

2. 점층 및 점강법

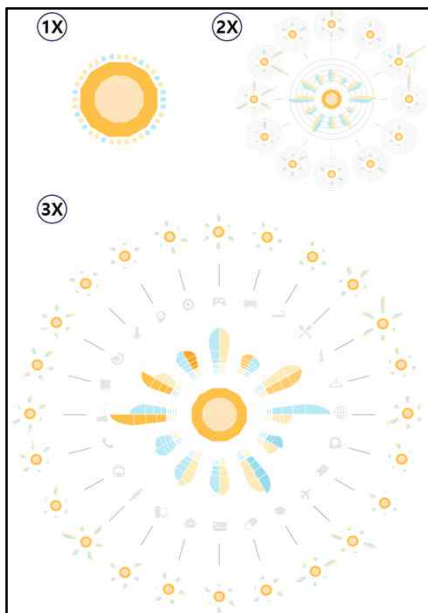


그림 6. 동적 주식 정보 (정보민, 2020)
Figure 6. Dynamic Stock Information (Jung-Bo Min, 2020)

점층(漸層, gradation)법은 내용의 비중 혹은 강도를 점차 넓히거나 높여 표현하는 기법이다. 작고 약하고 좁은 것에서 크고 강하고 넓은 것으로 확대해가는 표현 방법이다. 그림 6은 꽃잎 모양의 그래프와 색상의 변화로 주식 데이터를 파악할 수 있는 동적 정보시각화이다. 1X 배율일 경우에는 주제별 종목의 오르내림을 큰 원 주위의 색상으로만 알 수 있다. 2X 배율일 경우에는 최대 12개까지 사용자가 즐겨보기 한 주제별 각 해당 종목의 오르내림을 그래프와 색상으로 알 수 있다. 3X 배율일 경우에는 현재 주가, 시가 총액, EPS, PER, 영업이익이익률의 변화에 대한 정보를 자세히 볼 수 있다. 줌인 아웃을 사용하여 내용을 축소 확대할 수 있는 점강/점층법을 사용하여 정보를 묘사하였다. 1X 배율일 경우에는 적은 정보의 묘사로 사용자가 픽토그램 상징성 이해의 오류가 존재한다. 3X 배율일 경우에는 많은 정보의 묘사로 사용자가 혼란스러워한다. 사용자에게 각 배율 단계별 최적의 정보량을 어떻게 설정하는 것은 개발자의 딜레마이다.

3. 생략법

덜 중요한 부분을 생략하고, 나머지는 상상에 맡기는 수사법이 생략(省略, Omission)법이다. 어떤 특정 부분의 말이 생략되어도 뜻의 내용이 간결해서 여운과 함축을 지니게 하는 방법이다.



그림 7. 서울 간략 지도(장근형, 박병규, 2018)
Figure 7. The Simple Map of Seoul (Geun-hyung Jang Byeong-gyu Park, 2018)

그림 7은 기하학적으로 단순화된 서울을 표현한다. 한강을 중심으로 25개 구와 주요 범례, 도로, 지하철 노선, 그리고 산림지역을 보여준다. 서울을 처음 방문하는 외국인에게 서울의 주요 범례의 위치를 찾고, 그리고

주요 도로와 지하철의 노선을 파악하기 위함이다. 그림에서는 한강과 도로 모양은 단순 직선이 아닌 곡선이다. 그러나 그림 7.과 같이 단순화 과정을 통해 만든 간략 지도는 서울 지형을 기억하는 데 도움이 될 것이다. 곡선이 포함된 정확한 지형의 사실적 묘사의 장점과 기하학적 선으로 그려진 단순 묘사의 장점에서 개발자의 선택 딜레마가 존재한다.

4. 제유법 및 환유법

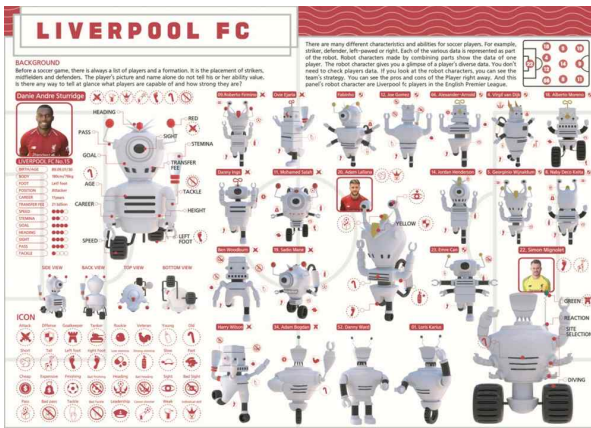


그림 8. 리버풀 FC 캐릭터 (박민정, 2018)
 Figure 8. Liverpool FC characters (Min-jeong, Park 2018)

사물 일부분이나 특징으로 사물 전체를 표현하는 수사법이 제유(提喻, synecdoche)법이다. 유사 수사법으로 대유(代喻)법, 환유(換喻)법이 있다. 그림 8.은 리버풀 FC 선수들을 캐릭터로 묘사했다. 축구 선수들은 각각의 포지션과 장단점이 있다. 선수에 따라서 스트라이커, 수비수, 원발, 신입, 나이, 키, 지구력, 현황판단, 통솔력, 패스 능력 그리고 연봉 등이 다르다. 이러한 선수들의 능력 및 개성을 로봇의 신체 일부와 연계하여 표현하였다. 그림 8.에서 로봇 하나는 신체 부분이 모두 모여 표현된 고유 선수의 캐릭터 묘사이다. 캐릭터를 통하여 선수들의 포지션이나 장단점의 정도를 대략 알 수 있다. 그리고 팀별 로봇 캐릭터 나열을 보면, 대략 팀의 전력을 알 수 있다. 이러한 캐릭터 로봇을 이용하여 팀의 작전 전략을 수립할 수 있다. 제유법에서 부분 묘사는 기하학적 형태 크기나 색상으로 그 데이터의 정도 값을 표현한다. 그러나 때로는 사용자가 그 정도 값을 완벽하게 인식하지 못하는 경우가 있다. 묘사의 함축적 의미 표현과 실제 수치의 구체적 표현 사이에서 개발자는 딜레마를 가진다.

5. 운율법

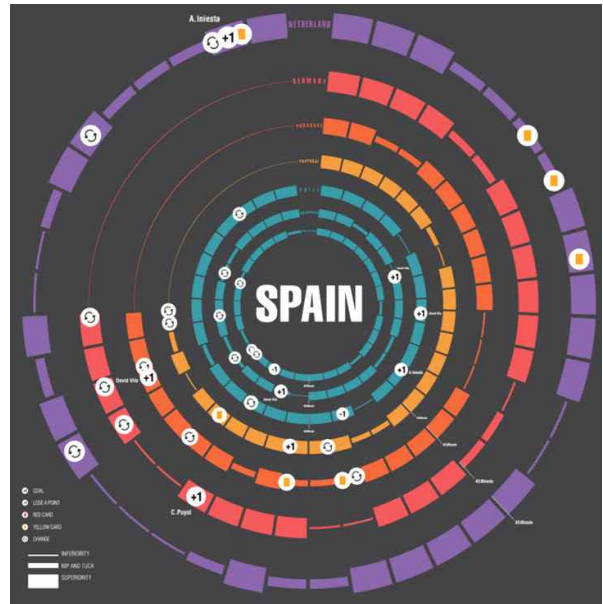


그림 9. 2010년 스페인 월드컵 경기 결과 (이현규, 2013)
 Figure 9. 2010 Spain World Cup (Hyunkyu Lee, 2013)

운문을 이루고 있는 소리의 반복적 요소의 양식을 운율(韻律, meter)법이라 말한다. 문장이나 글을 일정한 규칙을 가지고 반복 나열하여 내용을 비교 분석하거나, 전체의 뜻을 강조하는 방법이다. 두운(頭韻)법, 요운(腰韻)법, 그리고 각운(脚韻)법 등이 유사 수사법이다. 그림 9.는 2010년 스페인의 일곱 경기의 연장 포함한 모든 경기의 내용을 보여준다. 가장 안쪽부터 바깥까지 각각의 원형 띠가 32강부터 결승까지의 경기 내용을 알려준다. 원형에서 각 마디는 3분을 의미하며, 전반전 15마디, 후반전 15마디, 그리고 연장전 전후반 10마디로 구성된다. 마디의 두께는 공의 점유율, 마디 위의 원 안의 픽토그램들은 코너킥, 프리킥 그리고 득점의 시기를 보여준다. 다른 월드컵 연도 혹은 다른 국가를 선택 가능한 동적 시각화로 제작되어서, 경기 내용을 비교 분석할 수 있다. 운율법은 같은 반복적 항목이 일정하게 존재하는 경우, 항목의 규칙적 배열을 통하여 데이터의 변화를 비교할 수 있다. 그러나 많은 데이터는 일정하게 동일 반복하지 않는 경우가 많다. 이로 인해, 불필요한 그래픽, 부적절한 그래픽 블록화, 그래프 비율 왜곡 등의 관계 표현의 부정확한 사용에 기인한 딜레마가 존재한다.

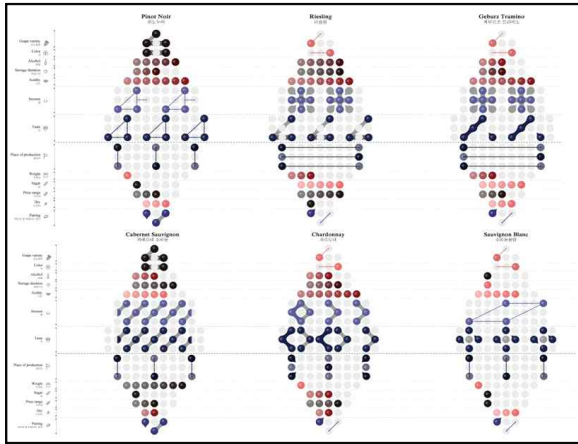


그림 10. 포도주 종류별 특성 (김수현, 2021)
Figure 10. Attribute by Wine types (Soo-hyun Kim 2021)

6. 열거법

어떤 분류나 계통상 비슷하거나 같은 맥락에 있는 것을 나열하는 수사법을 열거(列擧, enumeration)법이라 한다. 이때, 열거된 각각의 말은 의미상 연관성을 가지고 있고, 동격이며, 그리고 유사한 문장 성분으로 이어져야 한다. 그림 10. 여러 포도주 품종의 특징을 보여주는 정보시각화이다. 포도송이와 유사하게 생긴 각각의 원들은 위에서부터 포도품종, 색상, 알코올, 저장 기간, 산도, 향, 맛, 생산지, 바디감, 당, 가격분포, 그리고 드라이 정도를 보여준다. 추상적인 상징과 색상으로 여러 포도주의 내용을 비교할 수 있다. 열거법에서 보여주는 상징들의 하나의 데이터 세트는 하나 혹은 소수의 몇 개만 보여 줄 경우는 의미가 없고, 여섯 개 이상의 데이터 세트를 나열할 때만 정보묘사의 효과가 있다. 그러나 이러한 데이터 세트는 기하학적 모양의 추상화된 상징이라서 데이터의 수치를 정확히 표현하기 힘들다. 데이터 수치의 표현과 수치를 추상적으로 표현한 상징 사이에서 개발자는 선택의 딜레마를 가진다.

7. 은유법

원관념은 숨기고 보조관념만으로 대상을 설명 혹은 그 특징을 묘사하는 표현법이 은유(隱喩, metaphor)법이다. 원관념과 비유되는 보조관념을 같은 것으로 보므로 ‘A(원관념)는 B(보조관념)다’의 형태로 묘사된다. 제주도 200여 개의 오름 중에서 16개의 주요 오름을 표현하였다. 정보시각화에는 오름의 대략 사실적 모습과 오름을 보기에 가장 아름다운 시간의 정보(태양의 위치와 구름의 색상), 오름의 높이, 등반 난이도, 그리고 주요

관찰 포인트 정보가 은유적으로 묘사되어있다. 은유법은 데이터 수치의 표현과 수치를 추상적으로 표현한 상징 사이에서 선택의 딜레마를 가진다.



그림 11. 제주오름 (김지영, 2017)
Figure 11. Jeju Oreum (Ji-young Kim, 2017)

IV. 결 론

클라우드, 데이터마이닝, 인공지능 등 데이터 네트워크 기술 발전으로 정보 소비자의 패턴이 데이터 수리적 표현에서 민감도와 흥미 유발이 높은 데이터 묘사적 표현으로 변화되었다[5]. 이를 위해, 소비자를 위하여 전달하고자 하는 수리적 데이터는 테마가 있고, 스토리가 있으며, 그리고 수사적으로 묘사된 정보시각화로 재가공되고 있다. 이러한 변화로, 정보시각화는 과학 보고서에 주로 사용되는 정보전달형 표현, 광고나 신문 매체에서 여러 시각적 효과로 의도된 스토리를 전달하는 정보묘사형 표현, 창조적 예술 가치와 재미 중심의 정보유희형 표현으로 형식이 구체화 되었다. 본 연구에서는 특정 목적을 위해 수사학적 수법을 적용한 정보묘사형 표현에서 흔히 발생하는 의도된 오류 중심으로 논했다.

정보묘사형 표현은 데이터의 테이블 형식과 내용을 파악한 후, 데이터 일정 부분을 일정한 형식으로 강조하여 사용자에게 전달한다. 이를 위해, 설득의 언어학에서 사용되는 수사법과 유사한 형식을 사용한다. 이때 사용된 시각적 수사학표현은 사용자의 내용전달에 도움을 주지만, 사용된 그래픽은 데이터 내용을 일부 거짓되게 표현한다. 데이터 내용전달의 효율과 데이터 내용표현의 오류 사이에서 생기는 개발자 딜레마가 있다. 수사학적 묘사표현에서, 역설법은 데이터 수치와 다른 그래픽 면적, 의미 없는 그리드 사용, 그리고 선택적 데이터 일부만 보여주기에, 점층법은 줌인 아웃으로 인

한 너무 적거나 혹은 너무 많은 정보량과 급격한 내용 변화에 따른 그리드의 부적절한 사용에서, 생략법은 정확한 사실적 묘사의 장점과 기하학적 단순 묘사에서, 제유법은 묘사의 함축적 의미 표현과 실제 수치의 구체적 표현 사이에서, 운율법은 불필요한 그래픽, 부적절한 그래픽 블록화, 그래프 비율 왜곡에서, 그리고 열거법과 은유법은 데이터 수치의 표현과 수치를 추상적으로 표현한 상징 사이에 딜레마가 있다. 이러한 진실과 오류 사이의 딜레마는 개발자가 데이터 테이블 형식과 내용에 따라서 고려되어야 한다.

References

- [1] Donna M. Wing, Wall Street Journal Infographic Guide, insights, pp.32-33, 2014.
- [2] Won-Jo Lee, “A Study on Word Cloud Techniques for Analysis of Unstructured Text Data”, The Journal of the Convergence Culture Technology (JCCT) Vol. 6, No. 4, pp.715-720, 2020.
- [3] Sungkon Kim, “The Error and Integrity of Information Visualization Design”, Journal of Digital Design, Vol. 15, No. 1, pp.65-76, 2015.
- [4] Sungkon Kim, “Digital Information Visualization Methods with Rhetorical Presentation - Focus on Digital Visual Element, Visualization form, Rhetorical Effect”, Journal of Digital Design, Vol. 12, No. 1, pp.21-32, 2012.
- [5] Fayyad, Usama, Information Visualization in Data Mining and Knowledge Discovery, Morgan Kaufmann Publisher, 2002.

※ 이 논문은 2022년 울산대학교 연구비에 의하여 연구되었음
