東洋學 第84輯(2021年 7月) 檀國大學校 東洋學硏究院

디지털인문학시대『통합디지털한한대사전』표제어휘의 활용

- 워드프로세서용 인명 서명 변환기 개발 -

양 창 진*

국문초록

오늘날 모든 연구 저술 행위는 컴퓨터 워드프로세서로 이루어진다. 연구 결과로 발표되는 각종 저술과 논문에는 인명, 지명, 서명, 관직명 등 고유명칭이 수시로 등장한다. 이 중에서도 전통시대를 연구한 저술이나 논문들에는 인명과 서명들이 반드시 포함된다. 역사 연구는 '인간'에 대한 연구이자 인간들이 만든 '사건'과 '문헌'에 대한 연구이기 때문이다.

그런데 인명과 서명은 그 속성으로 인해 단순히 문자로만 표기되지 않고, 특수기호와 숫자가 추가적으로 다양하게 결합하여 표기된다. 연구자가 선호하는 표기 형태가 다르고, 학술지마다 표기 방식이 다르기 때문이다. 인명 표시의 대표적 유형으로는 ① 이순신(李舜臣, 1545~1598), ② 李舜臣(이순신, 1545~1598), ③ 이순신(1545~1598), ④ 李舜臣(1545~1598) 등이 있다. 서명 표기 사례로는 ①『三國史記(삼국사기)』, ②『삼국사기(三國史記)』, ③ ≪三國史記(삼국사기)≫, ④ ≪삼국사기(三國史記)≫, ⑤ ≪三國史記≫, ⑥ ≪삼국사기≫ 등이 있다.

위 사례는 빈번하게 등장하는 형식 중 일부이다. 문제는 표기 형태가 다양하기도 하지만 입력이 번거롭다는 점이다. 그리고 입력할 때 잘못된 자료를 참조하여 틀리게 입력하기도 한다. 이러한 불편함과 오류를 줄이기 위해 공신력 있는 사전의 어휘를 워드프로세서에 처음부터 내장하면 연구 효율 향상과 입력 정보 형식의 일관성 유지에도 도움이 될 것이다.

이 논문은『통합디지털한한대사전』의 인명, 서명 표제어휘들을 활용하여 워드프로세서의 고유명칭 입력 형식 변환기 개발 및 사전의 새로운 활용 사례를 제시하는데 목적이 있다.『통합디지털한한대사전』이 거의 완성된 지금 이 사전에 수록된 표제어를 활용하는 특별한 사례가 될 것으로 기대된다.

[주제어] 통합디지털한한대사전, 고유명칭, 서명(書名), 문서작성기, 온라인사전

목 차

Ⅰ. 서 론
Ⅱ. 연구 필요성 및 인명 형식 개발 사례
Ⅲ. 서명 입력 형식 변환기로의 확장
Ⅳ. 결론 및 향후 과제

* 한국학중앙연구원 수석연구원 / uridul@aks.ac.kr

Ⅰ. 서 론

4차 산업혁명의 물결이 세계를 휩쓸면서 새로운 정보와 지식도 전에 없을 정도로 대량 생산되고 있다. 그런데 지식은 정제되고 표준화된 것만이 오래도록 전승된다. 지식의 정제는 용어의 정제와 정교한 정의(definition)에서 시작된다. 따라서 지식의 불확실성이 난무할수록 정제된 용어를 체계적으로 정리한 '사전'의 중요성도 같이 커지고 있다.

그런데 인터넷이 우리 실생활의 기본 도구로 자리 잡으면서 각종 사전들도 인터넷으로 유통되고 있다. 대표적으로 민간 포털 네이버는 40여개 이상의 어학 사전과 5개 이상의 종합 백과사전의 검색 서비스를 무료로 제공하고 있다. 이렇듯 이용료 없는 온라인 사전이 일반화되면서 서적형 사전은 쇠퇴하고[1] 사전의 중요성도 예전만 못할 것이라 주장되기도 한다.[2]

이렇게 사전에 대한 인식이 달라진 이유는 두 가지를 들 수 있을 것이다. 첫째, 온라인 사전이 대세를 이루면서 서적 편찬 출판사들의 수익성 악화로 사전 편찬을 중지하여 주변에서 서적형 사전을 보기 힘든 상황 때문이다. 둘째, 민간 포털에서 정보를 쉽게 검색해 볼 수 있어서 정작 사전을 이용하면서도 사전을 이용하고 있다는 생각이 들지 않아 사전의 효용성이 피부에 와 닿지 않기 때문이다.

그러나 지금도 사전은 '여러 가지 방식으로' 계속 편찬되어 '여러 가지 방식으로' 이용되고 있다. 이렇게 '여러 가지 방식'으로 편찬되는 사전은 고유의 편찬 목적이 있다. 서적형 사전으로 출판하거나 온라인 서비스하는 것이 이러한 본래의 목적 중의 하나이다. 하지만 사전 편찬이 쇠퇴하는 현실에서 사전 편찬의 의의를 확장하고 사전의 활용성을 넓히기 위해서는 본래 의도했던 "편찬 목적 이외의 2차 활용방안"을 강구하는 것도 좋은 길이 될 것이다. 필자는 이 2차적 활용을 위한 사례 연구 대상으로 『통합디지털한한대사전』(이하 '통합대사전')을 선정하고, 이 사전에 수록된 표제어휘를 이용하여 워드프로세서의 고유명칭 입력 형식 변환기로 만드는 방안을 강구해 보았다.

Ⅱ. 연구 필요성 및 인명 변환기 개발 사례

1. 연구의 필요성

단국대학교 동양학연구원(이하 '연구원')이 1976년에 시작하여 2008년에 완간한 『한국한자어사전』 4책과 『한한대사전』 16책은 해방 후 한국 학계가 이룩한 대표적 역작이다. 그리고 연구원은 이 둘을 통합하고 더

1) 단국대학이 편찬한 세계 최대의 한자 사전인 『한한대사전(漢韓大辭典)』도 증보판은 서적형 편찬을 포기하고 『통합디지털한한대사전』이라는 명칭의 인터넷 사전으로 전환하고 있다. 김지영, 「『통합디지털한한대사전』의 DB구축과 온라인 사전편집기」, 『동양학』 54, 2013, 213~246쪽.

2) 정철, 「온라인 백과사전 만들기」, 『한국사전학』 11, 2008, 65~110쪽.

욱 발전시킨『통합디지털한한대사전』을 편찬하고 있다. 특히『통합대사전』은 편찬 방식도 정보화시대에 걸맞게 기존의 책 출판이 아니라 디지털 데이터 베이스 형식을 취함으로써 사전의 양과 질은 물론 이용의 편의성도 대폭 확장될 것으로 기대된다.

그런데 시대는 바야흐로 정보화시대를 지나 '디지털인문학'이 화두가 되는 시대로 접어들고 있다. 디지털을 바탕으로 한 정보 제작과 검색을 넘어 디지털로 표현하고 소통하는 디지털인문학적 연구 방법론[3]이 한국학 영역으로 확산되고 있다. 이러한 시대적 조류에 따라『통합대사전』도 기존의 사전 기능을 포함하면서 동시에 디지털인문학이 추구하는 방향으로 활용처를 확대할 필요성이 제기되고 있다. 현재 동양학연구원이 발간한『한국한자어사전』은 포털 네이버를 통해 서비스되고 있다. 이를 이어『통합대사전』도 인터넷 검색 서비스를 제공함과 아울러 새로운 영역으로 활용을 확장할 필요가 있다.

이 논문은『통합대사전』의 새로운 활용 방법을 탐색하기 위한 연구의 하나로서,『통합대사전』에 수록된 인명(人名), 서명(書名) 등의 고유명칭 표제어휘가 본래의 사전 편찬 목적을 넘어 다른 방식으로 활용되는 사례를 연구하는 데 목적이 있다. 요약하면『통합대사전』에 수록된 고유명칭 표제어휘들을 활용하여 워드프로세서의 고유명칭 표기 자동 변환기를 개발하는 방안을 연구하는 것이다. 그동안 워드프로세서에 수록되는 한자 사전과 관련해서는 주로 어떻게 하면 많은 단어를 수록하고 검색을 편리하게 하고 입력된 한자를 편리하게 이해할 것인가 하는 부분 즉, 콘텐츠에 집중되었다.[4] 그런데 워드프로세서는 콘텐츠 입력기이기 때문에, 하나의 단어를 다양한 조건에 따라 다양한 형식으로 쉽게 변환 입력시킬 수 있는 방법, 즉 콘텐츠의 입력 및 표시 방법 또한 매우 중요하다.

이러한 인식에 근거하여 수행된 이 연구는 현재 흔글 워드프로세서에 내장된 인명 표기 변환기 개발 경험을 바탕으로 한 것이기 때문에 보다 현실적인 활용성이 있을 것으로 보인다. 대부분의 연구와 저술이 워드프로세서로 이루어지는 현실에서 이러한 기능을 개발한다는 것은 단지 컴퓨터 프로그램 개발을 넘어서 한국학과 동양학 연구 진흥에도 기여하게 될 것이다. 다른 한편으로는 앞서 말한 사전의 중요성을 다시 환기시키는 계기가 될 수도 있을 것이다.

이 연구의 대상인『통합대사전』은 크게 ① 표제자를 중심으로 ② 표제자의 뜻갈래, ③ 표제어휘, ④ 표제어휘의 뜻갈래 등으로 구성되어 있다.[5] 특히『통합대사전』에 수록된 약 50만개의 표제어휘는 다양한 활용가치를 지니고 있다고 할 수 있다. 더구나 이 표제어휘들은 출처를 증명할 수 있는 출전과 예문이 확실한 것들[6]이기 때문에 활용성뿐만 아니라 학문적 신뢰성도 높다. 그래서 필자는『통합대사전』에 수록된 표제어휘를 종합적으로 추출하여 새로운 방식으로 활용할 수 있는 방안을 강구해 보았다. 이러한 고민이 표제어휘를 워드프로세서에 탑재하여 인명, 서명 표기 형식을 자유롭게 변환시킬 수 있는 프로그램을 개발하는 아이디어로 귀결되었다.

---

3) 김현,「디지털 인문학과 고문헌 자료 연구」,『열상고전연구』50, 2016, 13~38쪽.
4) 김은희,「한글 워드프로세서 한자사전의 분석과 개선안 모색」,『중국어문학지』72, 2020, 271~295쪽.
5) 양창진,『통합디지털한한대사전』편찬 현황의 기계적 점검 방안 연구－표제자 및 표제어휘 통합 분석을 중심으로－』,『동양학』60, 2015, 195~218쪽.
6) 정재철,「『디지털한한대사전』의 표제어 구성에 대하여」,『동양학』50, 2011, 163~186쪽.

한국학 연구 성과로 발표되는 각종 저술과 논문에는 인명, 지명, 서명, 관직명, 사건명 등 각종 고유명칭이 수시로 등장한다. 이 중에서도 역사적으로 비중이 있는 '인명(人名)'과 '서명(書名)'의 출현 비중이 높다. 역사 연구는 곧 '인간'에 대한 연구이자 그들이 만들어 낸 '사건'과 '문헌'에 대한 연구이기 때문이다. 그런데 이러한 고유명칭들은 문서마다 다양한 형식으로 표기된다. 한자로 표기되기도 하고 한글과 한자가 병기되기도 한다. 더 나아가 해당 용어가 가지는 특별한 속성으로 인해 단순히 문자로만 표기되는 것이 아니라 특수기호와 숫자가 다양한 형태로 결합되어 표기된다. 연구자마다 선호하는 표기 방식이 다르고, 또 학술지마다 표기 방식이 다르게 집필 기준을 정하기 때문이다. 예를 들어 KCI에 등재된 여러 학술지에 등장하는 인명 정보 표기를 조사하여 이 중 대표적인 형태를 다음과 같이 정리해 보았다.

① 이순신(李舜臣, 1545~1598)
② 李舜臣(이순신, 1545~1598)
③ 이순신(1545~1598)
④ 李舜臣(1545~1598)
⑤ 이순신(李舜臣)
⑥ 李舜臣(이순신)

이외에도 소수이지만 '이순신(1545~1598)'과 같은 표기 유형도 있다. 예에서 보듯이 용어의 한글 표기에 한자와 다른 특수기호들이 추가되면 더욱 복잡해질 수 있다.

한편 인명 정보와 같이 표기 방식이 다양하게 나타나는 고유명칭에는 서명도 있다. 서명이 표기되는 유형을 조사하기 위해 민간 학술정보 서비스 DBPia와 '한국학술지 인용색인(Korea Citation Index, KCI)'에서 제공하는 한국연구재단 등재 학술지 논문들을 검색해 보았다.[7] 학술지들을 조사해 본 결과 아래와 같이 연구자들은 다양한 형식으로 서명을 표기함을 알 수 있었다.

①『삼국사기(三國史記)』
②『三國史記』
③『삼국사기』
④ ≪삼국사기(三國史記)≫
⑤ ≪三國史記≫
⑥ ≪삼국사기≫

현재 연구자들이 서명을 입력할 때는 먼저 ① 한글을 입력하고 ② 한자 변환 과정을 거친 후 ③ 특수기호 문자판을 열고 ④ '『'와 '≪' 같은 특수기호를 별도로 입력하는 과정을 거쳐야 한다. 인명의 경우, 인명과 생몰

7) http://DBPia.co.kr 및 https://www.kci.go.kr 참고.

년을 같이 표기할 때는 생몰년을 별도로 찾아서 입력하는 번거로운 과정을 거쳐야 한다. 이 과정에 잘못된 생몰년을 입력하는 경우도 있다.

예를 들어 세종 때 북방 영토를 개척한 김종서(金宗瑞)의 경우를 보면 논문에 따라 '김종서(1390~1453)'[8] 또는 '김종서(金宗瑞, 1383~1453)'[9]로 표기하기도 하는데, 한글과 한자 등이 조합되어 표기되는 형태가 다르다. 둘 중 하나는 오류일 수 있고 둘 다 오류일 수도 있다. 문서를 작성할 때 서로 다르게 기록된 인명사전이나 학술논문을 참고하여 입력하였기 때문에 이런 현상이 발생하는 것이다. 이것이 연구자들 사이에서 계속 반복되면 잘못된 자료가 지속적으로 양산되어 더 이상 손댈 수 없게 될 수도 있다. 이 점을 고려하면 이를 개선하기 위한 방안을 연구할 필요성이 대두된다고 할 수 있다.

그래서 이러한 고유명칭 표기를 오류 없이 간단히 표준화된 형태로 입력해 줄 수 있는 문서작성기 [워드프로세서. word processor] 내장용 프로그램을 개발하면 좋지 않을까 하는 아이디어에서 이 연구는 출발하였다. 그리고 이 연구가 현실성이 있는 것은 우리나라에서 인문학 연구자라면 거의 95% 이상이 ㈜한글과 컴퓨터에서 만든 워드프로세서를 이용하기 때문에, 이 워드프로세서에 이러한 기능을 탑재하면 연구자에게 부담이 되는 문제의 상당한 부분을 해결할 수 있을 것으로 판단되기 때문이다.

이러한 프로그램 개발을 위하여『통합대사전』에 수록된 표제어휘를 사용하는 것이다. 필자가 ㈜한글과컴퓨터와 협력하여 이러한 아이디어를 일차적으로 구현하여 적용한 결과가 흔글 워드프로세서에 탑재하여 보급되고 있는 '인명정보 형식 변환기'이다. 이하에서는 이 연구의 선행 사례로서 인명정보 형식 변환기 개발 사례를 자세하게 살펴보겠다.

2. 선행 사례

오늘날 학술 연구 성과를 문서화하는 문서작성 프로그램은 MS Word나 흔글 워드프로세서 등이 대표적이다. 그런데 모든 문자 생활이 워드프로세서에서 이루어지다 보니 한자 입력, 맞춤법 검사 등과 같이 인간의 지식에 의존하던 활동도 점차 정보 기술에 의존하게 되고 있다. 이러한 기술은 사전을 바탕으로 구현되어 오늘날 거의 모든 워드프로세서에 내장된다. 따라서 향후에도 많은 사전들이 문서작성 프로그램들과 직간접적으로 연계될 것이다. 이러한 상황에서 필자는 위에 제시한 다양한 유형의 인명 표기를 자동으로 변환하고 입력하는 방안을 실제로 연구하고 구현하였다.

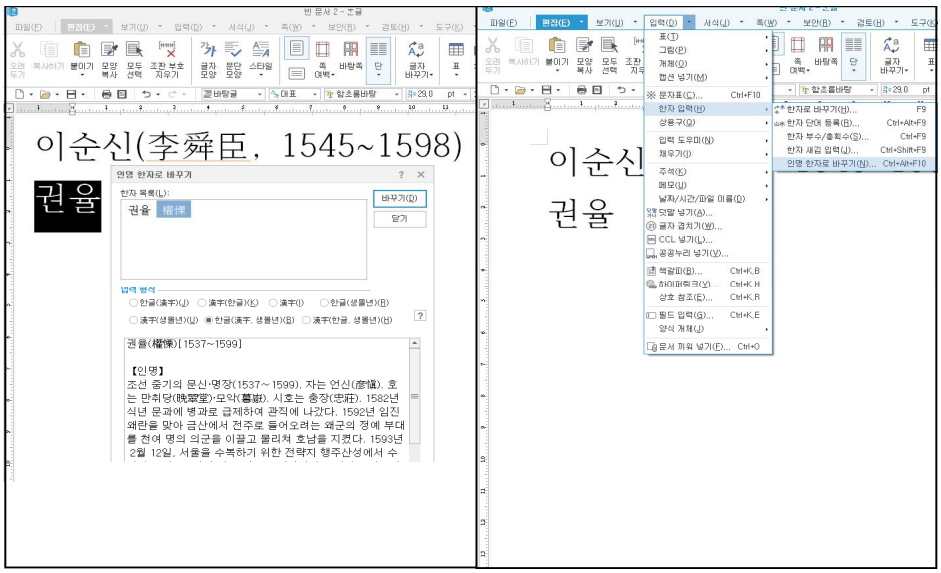
필자는 2006년부터 6년간 정보통신부가 주관하는 '국가DB 사업' 예산의 지원을 받아 한국역대인물종합정보DB(people.aks.ac.kr)를 개발하여 10여 년 이상을 운영하였다.[10] 이 DB에는 우리 역사에 족적을 남긴 인물에 대한 상세한 생애 정보를 담은 인물사전을 중심으로, 고려와 조선시대 과거 합격자 신상정보, 성씨와

8) 장인식,「영역 고시조에 나타난 번역상의 문제점: 이순신, 이조년, 김종서 시의 경우」,『번역학연구』5: 2, 2004, 145~162쪽.
9) 박대아,「세종과 한글을 활용한 지역 문화 정책 - 세종특별자치시의 사례를 중심으로」,『한국학연구』58, 2016, 5~29쪽 및『한국민족문화대백과사전』등.
10) 국가DB사업에 대해서는 양창진,「역사문화 콘텐츠 육성 정책의 회고와 전망 : 역사분야 국가DB를 중심으로」,『인문콘텐츠』35, 2014, 199~222쪽 참고.

본관정보 및 관직명 정보 등이 종합적으로 수록되어 있다. 그 결과 고조선 시대부터 현대에 이르기까지 약 3만 명의 인물정보와 조선시대 과거(문과, 무과, 생원진사시, 잡과) 합격자와 고려시대 과거 합격자 및 음관 정보 등을 통합하여 약 16만여 명의 인물 관련 정보를 서비스할 수 있게 되었다. 인물정보에서 서비스되는 자료 중 인명과 직접적으로 관련된 자료는 다음과 같다.[11]

◈ 인물 사전 : 29,235명
◈ 문과 급제자 : 15,150명
◈ 무과 급제자 : 25,901명
◈ 생원진사시 합격자 : 42,048명
◈ 잡과 합격자 : 7,750명
◈ 고려과거 급제자 : 1,893명
◈ 음관 : 7,793명

이 DB 구축 과정을 통해 수많은 인명 정보를 축적하였고 이를 2차적으로 활용할 방안을 찾던 중 ㈜한글과 컴퓨터와 협력하여 워드프로세서에 탑재할 수 있는 인명 정보 자동 변환기를 개발하였다. 그리고 변환에 참고하기 위해 기본으로 탑재한 인명 정보는 2만 5천 명 정도이다. 이 변환기는 현재 '흔글 2014', '흔글 2018' 워드프로세서 등에 탑재되어 시중에 판매되고 있다. 한국학 연구자라면 누구나 알고 있듯이 흔글 워드프로세서는 가장 대중적인 문서작성 프로그램이므로 이 변환기가 개발되어 활용되고 있다는 것은 큰 의미가



〈그림 1〉 흔글의 인명 정보 단축키 입력(좌)과 메뉴 입력(우)

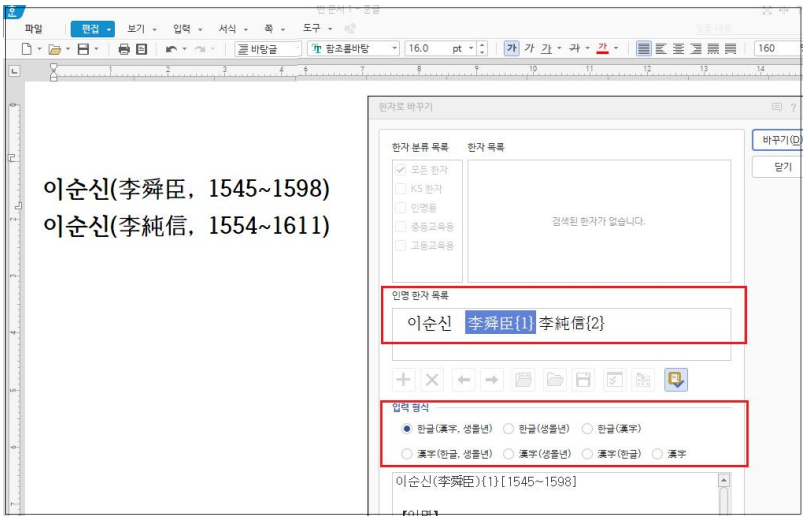
11) 「한국역대인물종합정보」(people.aks.ac.kr, 2021. 1. 1.)

있다고 할 수 있다.

2010년에 최초로 개발되어 흔글 2010SE 버전에 적용된 인명정보 입력기의 사용 방법은 다음과 같다. 이 워드프로세서를 실행하여 문서작성 화면이 나타나면 '이순신'을 입력한 후 단축키인 'Ctrl + Alt + F10' 키를 동시에 누르면 〈그림 1〉과 같은 화면이 나온다. 여기서 원하는 형태의 옵션을 선택하면 해당 옵션이 지정하는 형태로 인명 정보를 입력하여 준다. 단축기를 외우지 못하는 이용자들은 메뉴에서 '입력 → 한자입력 → 인명 한자로 바꾸기' 순으로 찾아가서 입력하는 것도 가능하다. 〈그림 1〉은 이 기능을 요약한 것이다.

흔글 워드프로세서에서 한글 ↔ 한자 변환 기능은 1990년대에 최초 개발 당시부터 'F9' 키나 '한자' 키에 배정되어 있었다. 메뉴를 차례대로 선택하여 입력할 수도 있으나 번거로워서 대부분은 단축키를 이용하여 입력한다. 오랫동안 이 방식에 적응된 연구자들은 이 키 [F9]에 매우 익숙하다. 그런데 인명 정보 변환 기능을 'F9' 키나 '한자' 키를 활용하지 않고 F10키를 이용한 것은 F9키와의 모든 조합 영역이 다음과 같이 한자 관련 기능 제공을 위해 사용 중이기 때문이다.

F9 : 한글을 한자로 바꾸기
Ctrl + F9 : 부수 및 총획수로 한자 찾기
Ctrl + Alt + F9 : 한자 단어 등록
Shift + F9 : 한자 자전 보기
Alt + F9 : 여러 개의 한자를 한글로 바꾸기



〈그림 2〉 흔글2018의 개선된 인명 정보입력 화면

그래서 부득이하게 F10키와의 조합을 이용한 것이다. 그래서 흔글 워드프로세서에 이 기능이 있는지 모르는 이용자들도 많다. 하지만 그동안 개선 방안을 모색하였고 그 결과 불편한 사항이 2018년 업그레이드판에

서는 개선되었다. 즉 이용자들과 친숙한 F9키 또는 한자키 입력으로 위의 모든 형식을 입력할 수 있게 되었다. 다음 〈그림 2〉는 개선된 프로그램의 실행 화면이다.

아직 완벽한 것은 아니지만 훈글 2018 버전 이후의 워드프로세서에서 한자 변환 및 인명정보 변환과 관련된 키는 F9에 통합되어 이용이 편리해 졌다. 위와 같은 변환 프로그램을 구현하기 위해서는 인명사전DB를 먼저 만들어야 했다. 이를 위해 필자는 다음 〈표 1〉과 같은 형식으로 간단한 인명사전을 만들어 ㈜한글과 컴퓨터에 제공하였다.

〈표 1〉 인명사전 데이터 구조

항목	예시	비고
① 한글	이순신	
② 한자	李舜臣	
③ 동명	李舜臣{1}	동명이인 구분
④ 생몰	1545~1598	

이중 '③ 동명'은 한글 음이 같은 한자가 있을 경우 구분이 쉽도록 선택 화면에 표시해 주기 위해 마련된 것이다. 위의 그림을 보면 이순신을 누르고 F9키를 누르면 李舜臣{1}과 李純信{2}이 표시되고 둘 중에 어느 것을 선택하느냐에 따라 그 아래에 제시되는 내용이 변한다. 그리고 그 밑에는 간단한 인명사전 내용이 표시됨으로써 이용자가 선택하고자 하는 인물이 맞는지 확인하는데 도움을 줄 수 있다. 위의 프로그램 개발을 위해 작성된 데이터는 아래와 〈그림 3〉과 같다.

1	일련번호	한자	한글	동명	생몰년	사전내용
23081		崔漢文	최한문	최한문(崔漢文)	9999~9999	생몰년 미상. 조선 전기 문신. 본관은 강화(江華)이다.
23082		崔漢洙	최한수	최한수(崔漢洙)	1856~9999	1856년(철종 7)~미상. 조선 말기 유학자. 자는 광진(
23083		崔漢禎	최한정	최한정(崔漢禎)	1427~1486	1427년(세종 9)~1486년(성종 17). 조선 초기의 문신
23084		崔漢柱	최한주	최한주(崔漢柱)	9999~9999	생몰년 미상. 본관은 강릉(江陵).
23085		崔漢候	최한후	최한후(崔漢候)	9999~9999	생몰년 미상. 본관은 화순(和順). 자는 자방(子房), 호
23086		崔諴	최함	최함(崔諴)	1094~1160	1094년(선종 11)~1160년(의종 14). 고려의 문신. 본
23087		崔咸一	최함일	최함일(崔咸一)	9999~9999	생몰년 미상. 고려 후기의 문인.
23088		崔沆	최항	최항(崔沆){1}	972~1024	972(광종 23)~1024년(현종 15). 고려 초기의 문신.
23089		崔沆	최항	최항(崔沆){2}	1209~1257	1209년(희종 5)~1257년(고종 44). 고려무신정권의
23090		崔恒	최항	최항(崔恒){3}	1409~1474	1409년(태종 9)~1474년(성종 5). 조선 초기의 문신
23091		崔恒慶	최항경	최항경(崔恒慶)	1560~1638	1560년(명종 15)~1638년(인조 16). 본관은 영천(永)
23092		崔瀣	최해	최해(崔瀣){1}	1287~1340	1287년(충렬왕 13)~1340년(충혜왕 1). 고려의 문인
23093		崔海	최해	최해(崔海){2}	1895~1948	1895년(고종 32)~1948년. 독립운동가. 일명 해일(海
23094		崔海斗	최해두	최해두(崔海斗)	1668~1740	1668년(현종 9)~1740년(영조 16) 11월 11일. 조선
23095		崔海山	최해산	최해산(崔海山)	1380~1443	1380년(우왕 6)~1443년(세종 25). 조선 초기 화포(火
23096		崔海雲	최해운	최해운(崔海雲)	1682~1764	1682년(숙종 8)~1764년(영조 40). 조선 후기의 무신

〈그림 3〉 훈글2018에 탑재된 인명 정보 데이터

Ⅲ. 서명 입력 형식 변환기로의 확장

1. 서명 표기의 유형

이하에서는 위에서 고찰한 인명 정보 자동변환 프로그램을 개선하여 서명을 다양한 형태로 변환하여 입력하는 기능으로 확장하는 프로그램 개발을 염두에 두고 그 방안을 고찰해 보겠다. 한국학 관련 저술에서 가장 많이 나타나는 고유명칭이 인명 다음으로 서명이기 때문이다. 이를 위해 우선 서명이 나타나는 유형을 조사해 보았다. 이를 위해 한국에서 산출되는 학술논문들이 종합적으로 집적되어 있는 한국학술지인용색인(KCI)에서 논문 제목에 표기된 '삼국사기'를 검색하였다. 다음은 이 검색 결과 중 일부이다.

김일권, 「『삼국사기』「신라본기」의 별자리목록과 천문성변기록 분류 연구」, 『한국고대사탐구』 32, 2019, 567-743.
김성한, 「『三國史記』 百濟本紀의 溫祚傳承과 沸流傳承」, 『역사와 담론』 73, 2015, 117-153.
김일권, 「≪삼국사기≫ 일식기록의 한중 사료 대조와 일식상황 비교」, 『신라사학보』 37, 2016, 133-250.
김영일, 「〈〈삼국사기 지리지〉〉의 지명 고찰」, 『한글』, 257, 2002, 103-128.

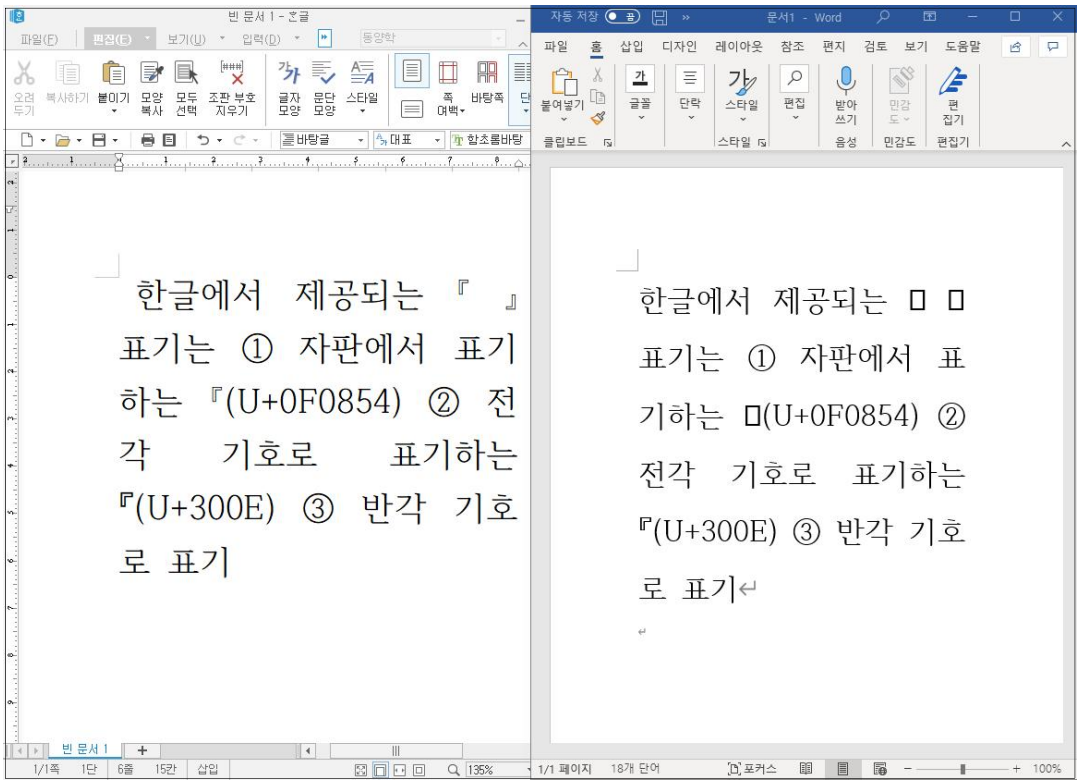
검색된 논문 제목에 표시된 서명의 유형이 다양하기는 하지만, 최근에는 『삼국사기』, 『三國史記』, ≪삼국사기≫, ≪三國史記≫로 표기 형태가 수렴되어 가는 듯 보인다. 즉 서명은 '『 』'나 '≪ ≫'로 감싸는 형태가 주를 이루고 있고 간혹 자판에 있는 '〈 〉'를 두 개 겹친 '〈〈'와 '〉〉'로 감싸는 경우도 보인다.

2. 서명 변환기 기능 고찰

위에서 고찰한 서명 표기 유형 변환을 위해서는 다음 두 가지 기능을 갖추어야 한다. 첫째 대부분이 많이 사용하는 기호로 서명을 둘러쌀 수 있도록 기본(default) 기호를 제공해야 한다. 위에 조사한 바에 의하면 『 』로 둘러싸는 경우가 가장 많았다.[12] 그런데 필자의 분석에 따르면 흔글 워드프로세서에서 제공되는 『 』 표기는 ① 자판에서 표기하는 『(U+0F0854) ② 전각 기호로 표기하는 『(U+300E) ③ 반각 기호로 표기하는 『(U+0F0854) 등 여러 가지가 있다. 이 기호들의 유니코드 번호는 서로 다르지만, 문서로 인쇄된 후에는 구분이 매우 힘들다. 그런데 외국의 학자들과 문서를 교환하거나 국제학술대회에 논문을 제출할 경우 문제가 발생할 수 있다. 즉, 문서 교환을 위해 흔글 문서를 MS Word 등으로 변환하면 다음 〈그림 4〉에서 보듯이 자판으로 입력한 '『'는 모두 다 '□' 모양으로 변형된다. 이 문제는 국제 표준 즉 유니코드에 규정된 전각 기

12) 흔글에서 제공되는 『 』 기호 입력 방법은 ① 기본 자판에서 입력, ② 문자표에서 전각기호로 입력 ③ 반각기호로 입력하는 방법 등 여러 가지가 있다. 그런데 이것들은 문서로 인쇄한 후에는 육안으로는 구분이 매우 힘들다.

호로 입력하게 하면 된다. 즉, 전각 기호 '『'를 본 논문이 목표로 하는 입력기에서 선택할 수 있는 기본 목록에 포함하면 될 것이다. 이 경우 인쇄할 때 전각 기호 주변 여백이 더 넓어 보이는 문제가 있을 수 있다. 이는 편집 단계에서 문서 편집진이 자체 규정 기호로 일괄 변환하면 될 것이다. 사실 이제는 모든 워드프로세서에서 유니코드가 거의 기본으로 사용되고 있다.



〈그림 4〉 흔글2010(좌)의 "『"를 MS Word(우)로 복사한 사례

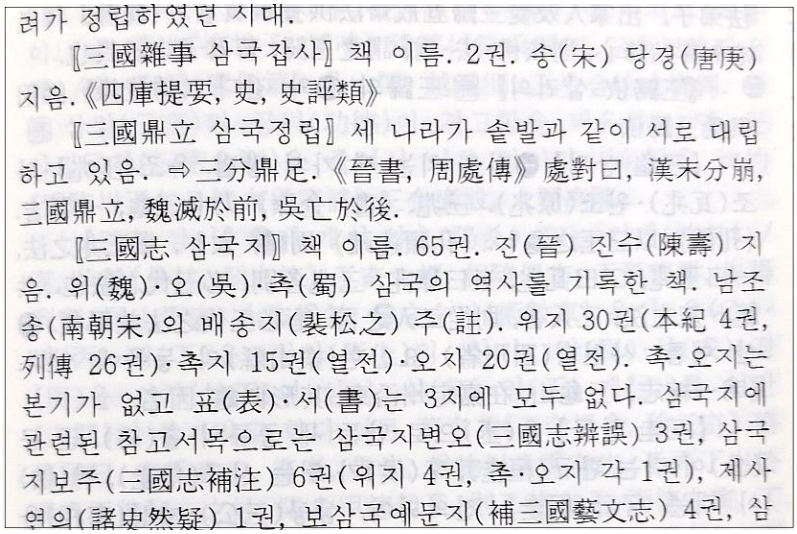
둘째, 디폴트로 제공되는 기호 이외의 기호를 이용자가 등록할 수 있어야 한다. 이 기능이 제공되면 흔글에서만 제공하는 『 』도 등록할 수 있을 것이다. 셋째, 선택한 기호 및 한글과 한자를 선택적으로 조합하여 서명 표기를 할 수 있어야 한다. 이것은 이미 인명 정보 변환기에서 구현한 바 있다. 그리고 네 번째, 마지막에 사용된 유형의 서명 표기는 '상태를 기억'하도록 하고 다음부터는 같은 방식으로 선택하지 않아도 바로 입력할 수 있도록 해야 한다.[13]

13) 이전 상태 기억 기능은 흔글에서 한자 입력 기능 등에 구현되어 있다. 한글로 단어를 입력하고 한자키를 누른 후, 제시되는 한자 목록에서 원하는 단어를 목록의 앞으로 이동시키면, 다음부터는 해당 단어가 가장 앞에 위치하여 있어 바로 입력할 수 있는 것도 이전 상태 기억 기능이다.

3. 서명 변환기의 자료 구조

서명 변환기 개발을 위해서는 기초 자료로 한국 고전 명칭들이 다수 정리되어 있는 서명 사전이 필요하다. 그리고 이 서명 사전 구축을 위한 기초 자료로는 단국대학교 동양학연구원이 편찬하고 있는『통합디지털한한대사전』에 수록된 표제어휘가 적절하다고 판단된다. 그 이유는 다음 두 가지가 있다. 첫째,『통합대사전』에 수록된 표제어휘들은 전부가 한글·한자가 병기되어 있다. 이것은 한글로 음을 입력하고 이에 해당되는 한자를 대치시키기 쉬운 구조의 자료를 만드는 토대가 된다. 한국에서는 동형이음 한자가 많이 사용되기 때문에 한글과 한자가 병기된 자료는 매우 유용하게 활용될 수 있다. 둘째, 표제어휘 마다 간단한 설명이 있어서 입력하고자 하는 한자 단어를 구분하기 쉽게 한다. 이런 이유로『통합대사전』의 표제어휘는 본 연구를 위한 적절한 대상이 될 수 있다.

대략 파악한 바에 의하면『통합대사전』의 표제자는 오차를 감안하더라도 5만여 자를 상회하며, 이에 딸린 표제어휘는 50만여 개로 추정된다. 그리고 이들 어휘 중에서 인명은 2만여 개, 서명은 5천여 개 정도로 알려진다. 원래 이전의『한한대사전』에는 아래의 그림처럼 서명들이 중국 고전 위주로 수록되어 있었다.


려가 정립하였던 시대.
〖三國雜事 삼국잡사〗 책 이름. 2권. 송(宋) 당경(唐庚) 지음.《四庫提要, 史, 史評類》
〖三國鼎立 삼국정립〗 세 나라가 솥발과 같이 서로 대립하고 있음. ⇨三分鼎足.《晉書, 周處傳》處對曰, 漢末分崩, 三國鼎立, 魏滅於前, 吳亡於後.
〖三國志 삼국지〗 책 이름. 65권. 진(晉) 진수(陳壽) 지음. 위(魏)·오(吳)·촉(蜀) 삼국의 역사를 기록한 책. 남조송(南朝宋)의 배송지(裴松之) 주(註). 위지 30권(本紀 4권, 列傳 26권)·촉지 15권(열전)·오지 20권(열전). 촉·오지는 본기가 없고 표(表)·서(書)는 3지에 모두 없다. 삼국지에 관련된 참고서목으로는 삼국지변오(三國志辨誤) 3권, 삼국지보주(三國志補注) 6권(위지 4권, 촉·오지 각 1권), 제사연의(諸史然疑) 1권, 보삼국예문지(補三國藝文志) 4권, 삼

『한한대사전』의 "三" 표제어휘 부분

그런데 이것이『통합대사전』으로 확장되면서 한국 고전들이 상당 부분 추가 수록된 것으로 확인된다. 다음의 표는 추가된 표제어휘의 예이다.

표제자	표제어휘	구분	비고
三	三國史節要	서명	추가
	三國史略	서명	추가
	三國遺事	서명	추가
	三國志	서명	
	三國史記	서명	추가
	柳道三	인명	
	尹重三	인명	

이 중 '표제어휘'와 '뜻갈래[정의]' 부분을 기본으로 하여 다음의 〈표 2〉와 같이 자료 구조를 구상하였다. 이 자료 구조는 위에서 살핀 인명 정보 변환기를 위한 자료 구조와 거의 유사하다. 이렇게 하면 5천여 개 레코드를 가진 서명 사전을 구축할 수 있다. 이 후에도 이 서명 목록은 계속 확대하여 한국학 연구자들에게 알려진 주요 문헌을 상당 부분 포함하게 되면 유용한 입력 도구가 될 것이다.[14]

〈표 2〉 서명 사전 자료 구조

항목	예시	비고
① 한글	국사	
② 한자	國史	
③ 동명	國史{1}	한글음 동명의 문헌 구분
④ 간행년	545	
⑤ 의미	신라 진흥왕 때 편찬된 신라의 역사서	

4. 고유명칭 한자 어휘 통합 변환기로의 확장

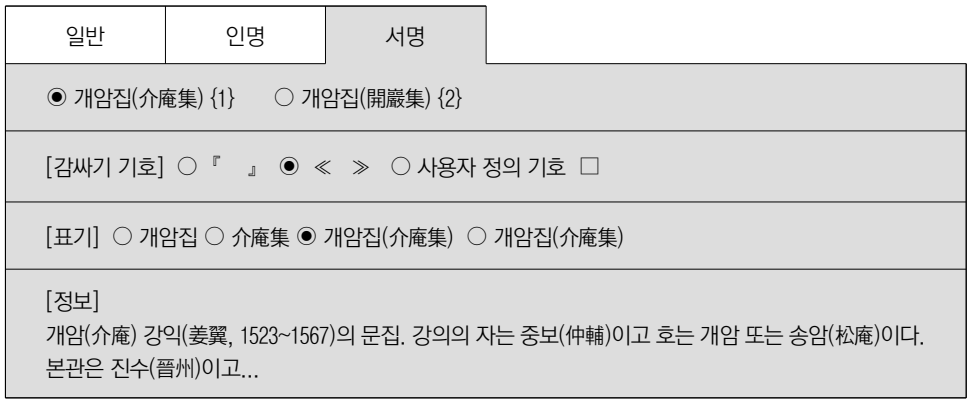
고유명칭 한자 통합 변환기의 목적은 기존의 ① 한글 ↔ 한자 변환 기능에 ② 한글 인명 → '다양한 인명 표기' ③ 한글 서명 → '다양한 서명 표기'로의 변환 기능을 'F9' 또는 '한자' 키 하나에 통합하기 위한 것이다. 이 기능을 구현하기 위해 일반 한자 용어와 인명, 서명 사전을 하나로 통합한 DB의 자료 구조는 다음 〈표 3〉과 같다. 이 자료 구조는 앞서 살핀 서명 사전의 자료 구조에 용어를 구분하기 위한 '구분' 필드만 추가하면 된다.

14) 한국에는 『문헌비고』와 『동국문헌비고』와 같이 동일 서명이 약간 다르게 불리는 경우가 있다. 이 경우 각각을 별도의 레코드로 등록하면 될 것이다.

〈표 3〉 고유명칭 통합 사전의 자료 구조

번호	1	2	비고
① 한글	국사	이순신	
② 한자	國史	李舜臣	
③ 동명	國史{1}	李舜臣{1}	한글음 동명의 문헌 구분
④ 연도	545	1545~1598	서명 간행년, 인물 생몰년
⑤ 의미	신라 진흥왕 때 편찬된 신라의 역사서	조선 중기 임진왜란 때 활약한 장군	
⑥ 구분	서명	인명	인명, 서명, 일반 용어 구분

다음 〈그림 5〉는 인명, 서명 입력기가 하나로 통합된 한자 단어 입력기가 개발되었을 경우 예상되는 입력창의 모습이다.



〈그림 5〉 고유명칭 형식 선택 입력창 예시

문서를 작성할 때 명사를 입력하고 변환하는 예상 시나리오는 다음과 같이 정리될 수 있을 것이다.

① '개암집'을 입력하고 F9키나 '한자' 키를 누른다.
② 단어 구분 필드를 판단하여 '인명'이나 '서명'에 없으면 일반 한자 입력창을 보여 준다.
③ 인명에 있으면 인명탭을, 서명에 있으면 서명탭을 활성화 한다.
④ 둘 다 있으면 인명탭과 서명탭을 둘 다 활성화하여 이용자가 선택할 수 있도록 한다. 그러나 인명과 서명이 같은 경우는 매우 드물다.
⑤ 서명에 있을 경우 다시 철자가 같은 서명이 있는지 판단한다.
⑥ 같은 음의 서명이 있으면 해당 문헌명들을 {1}, {2}를 붙여 나열하고 없으면 하나만 제시한다.
⑦ 두 개의 문헌이 있을 경우 맨 아래의 [정보]의 내용을 참고하여 원하는 문헌을 선택한다.

⑧ 원하는 감싸기 기호 라디오 버튼(◉)을 선택한다.
⑨ 원하는 형식의 서명 형식을 '표기' 라디오 버튼으로 선택한다.
※ 이상의 과정을 통해 한 번 설정된 표기 형식은 다시 변경하지 않는 한 그대로 유지되도록 한다.

위의 〈그림 5〉에서 ◉로 선택된 것을 그대로 입력하면 '≪개암집(介庵集)≫'으로 표시된다. 간혹 [개암집(三國史記)]처럼 위의 위의 서명 표기 사례에 없는 ' ['와 같은 기호들이 필요한 경우에는 '사용자 정의 기호'에 한 번 입력하여 프로그램에 저장해 두고, 다음부터는 별도의 입력 없이 계속 사용하도록 한다. 사용자 정의 기호는 하나를 입력하면 자동으로 해당 기호와 쌍을 이루는 기호가 입력되도록 하면 되기 때문에 첫 글자 하나만 입력하도록 한다. 아니면 처음부터 이런 기호들의 목록을 프로그램에 내장하고 아래 그림처럼 콤보 박스 목록으로 제공하여 선택하도록 하는 것도 좋은 방법이 될 것이다. 워드프로세서에서 내장하여 제공되면 좋을 것으로 판단되는 기호들로는 다음과 같은 것들이 있다.

〈표 4〉 서명 자동입력 용 쌍기호

번호	쌍 기호	첫 자의 유니코드
1	『, 』	U+300E
2	「, 」	U+300C
3	[,]	U+3014
4	≪, ≫	U+226A
5	《, 》	U+300A
6	〈, 〉	U+3008

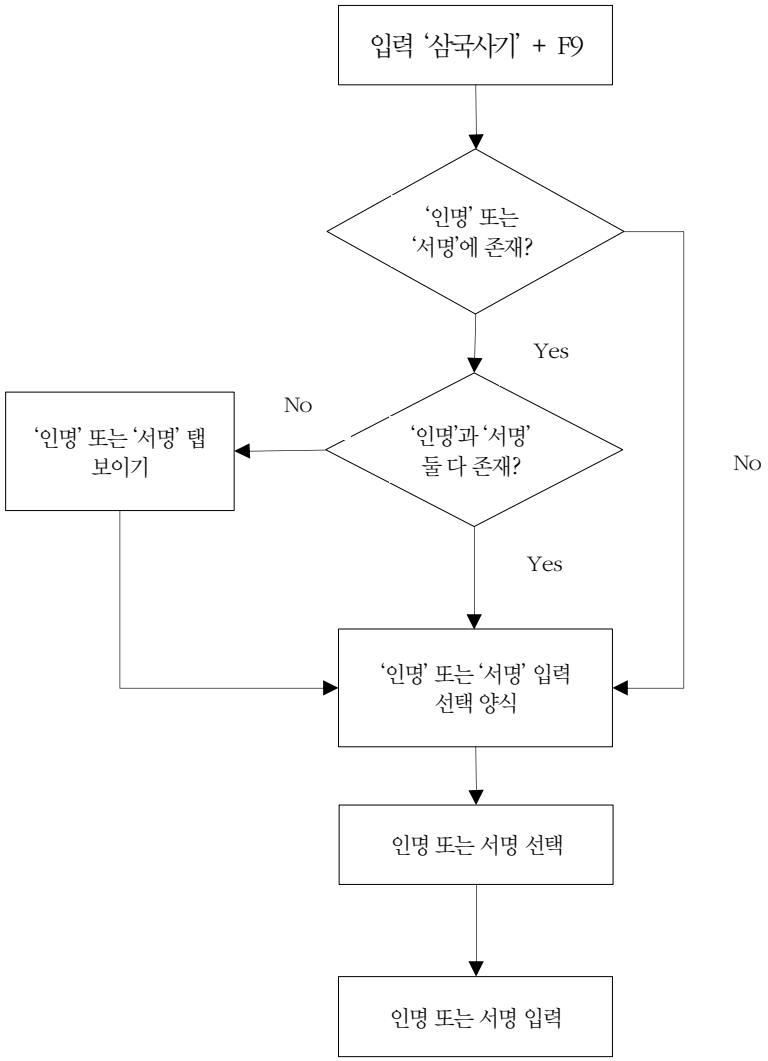
〈표 5〉 콤보 박스(▼)로 쌍기호 선택

번호	쌍 기호 선택 ▼
1	『, 』
2	「, 」
3	[,]
4	≪, ≫
5	《, 》
6	〈, 〉

이렇게 하면 '한자' 키 한두 번의 조작으로 인명, 서명을 원하는 형식으로 입력할 수 있다. 다음의 표는 이 기능을 통해 표시할 수 있는 서명 표시의 일부 유형이다.

한글	한자	한자병기
『삼국사기』	『三國史記』	『삼국사기(三國史記)』
[삼국사기]	[三國史記]	[삼국사기(三國史記)]
≪삼국사기≫	≪三國史記≫	≪삼국사기(三國史記)≫
《삼국사기》	《三國史記》	《삼국사기(三國史記)》

이상의 내용이 입력되고 변환되는 과정을 순서도로 표시하면 다음 〈그림 6〉과 같다.



〈그림 6〉 인명, 서명 형식 변환 프로그램 순서도

5. 부가적 기대효과

(1) 내용 태깅의 자동화

최근의 문서는 인공지능 번역 및 문서 내용의 기계적 판별이 가능하도록 XML(eXtensible Mark-up Language) 형식으로 제작되고 교환되고 있다. 그리고 이를 위해 각종 내용 태깅을 상세하게 하는 데까지 발

전하고 있다. 즉 서론, 본론. 결론 등과 같이 문서의 구조와 형태의 파악을 위한 '구조 태깅(Structure Tagging)' 뿐만 아니라 문장의 요소를 구성하는 각종 인명, 지명, 서명, 관직명, 사건명 등에 대한 '내용 태깅(Content Tagging)'을 하고 있다. 흔글 HWP 문서도 내부적으로는 구조가 XML 형식으로 태깅된 문서로 저장된다. 따라서 문서를 처음 입력할 때부터 이상에서 논의된 방식으로 입력을 하면 인명과 서명의 내용 태깅은 자동으로 이루어질 수 있다. 한편 자동 태깅에서 중요한 것은 어휘의 중의성을 해소하여 정확한 매칭을 하는 것이다.[15] 아래의 예시를 보면 '인종(仁宗, 1515~1545)'을 입력하면 같은 이름을 가진 고려의 왕이 아니라 조선의 왕을 의미하는 태깅을 하고 『한국민족문화대백과사전』에 있는 '인종(仁宗)' 항목과 링크하는 URL도 삽입할 수 있을 것이다.[16] 1990년대에 『창작기계』라는 SF 소설에서 구상된 기계적 글쓰기[17] 및 인공지능적(AI) 내용 분석도 결국은 이같은 정확한 내용 태깅이 뒷받침되어야 실현 가능할 것이다.

〈?xml version="1.0" encoding="utf-8"?〉
〈묵재일기〉
〈메타데이터〉
　　〈자료 ID="AKS-2014-KFR-1230007"〉
　　〈자료명〉1552년 7월 1일 오늘은 인종(仁宗)의 국기일(國忌日)이다.〈/자료명〉
　　〈시대〉조선〈/시대〉
　　〈유형〉일기〈/유형〉
　　〈집필자〉묵재(默齋) 이문건(李文楗)〈/집필자〉
〈/메타데이터〉
〈본문〉
　　〈원문〉
　嘉靖 三十一年 壬子 孟秋 七月 大 戊申 初一日辛巳, 霧, 午晴炎, 夕有雷聲 ○ 今日乃仁廟國忌 ○ 堂 齋素 服松末 ○ 下見兒等, 朝食而返, 困眼 ○ 兩兒習業 ○ 玄風倅朴承 任送還破笠補造者 熅物也給之
　　〈/원문〉
　　〈번역〉
　가정(嘉靖) 31년(1552) 임자 맹추(孟秋) 7월 큰달 무신. 1일 신사, 안개 끼다가 낮에 맑고 덥더니, 저녁 때는 천둥소리가 남. 1552년 7월 1일 오늘은 〈인명 url= "http://encykorea.aks.ac.kr/Contents/Item/E0047034"〉 인종(仁宗, 1515~1545) 〈/인명〉의 국기일(國忌日)이다. ○ 당(堂)에 머물렀다. 재계하고 소식(素食)했다. 솔가루 〔松末〕 를 먹었다. ○ 내려가 아이들을 보고, 아침밥을 먹고 돌아와 피곤하여 잤다. ○ 두 아이가 글공부를 했다. ○ 현풍(玄風)의 〈인명 URL="http://encykorea.aks.ac.kr/Contents/Item/E0004695"〉곽재우(郭再祐, 1552~1617)〈/인명〉가 깨진 갓을 수선하여 돌려보냈는데, 온(熅)의 것이므로 그에게 주었다.
　　〈/번역〉
〈/본문〉
〈/묵재일기〉

〈XML 문서의 구조 및 내용 태깅 사례〉

15) 이용구 · 정영미, 「사전 정보를 이용한 단어 중의성 해소 모형에 관한 실험적 연구」, 『정보관리학회지』 24: 1, 2007, 321~342쪽 참고.
16) 인터넷으로 서비스될 『통합대사전』의 표제어를 활용하면 해당 URL을 그대로 활용할 수 있을 것이다. 민간포털이나 위키백과의 데이터를 활용할 경우는, 표제어와 URL을 연계시키는 별도의 작업이 필요할 수 있다.
17) 조형래, 「워드프로세서 · 글쓰기 · 문학, 1980 · 1990 : 한국의 초기 워드프로세서 보급과 작가의 글쓰기 Ⅰ」, 『동악어문학』 73, 2017, 57~93쪽.

(2) 콘텐츠 자동 연계

서적으로 편찬된 사전 콘텐츠에는 서적이라는 물리적 한계로 인해 담을 수 있는 정보의 양에 한계가 있다. 사전이 각 분야, 주제별로 편찬되는 것도 이러한 이유에서 기인한 측면이 있다. 그러나 콘텐츠들이 전자화되면서 담을 수 있는 양이 무한대로 늘게 되었다. 방대한 분량의『조선왕조실록』이 DVD 한 장에 수록되었고, 더 발전하여 인터넷으로 서비스되는 상황은 이같은 현실을 상징적으로 보여 준다. 이 경우 중요한 것은 홍수처럼 범람하는 정보 속에서 보다 연관성 높은 개별 콘텐츠를 연계해 주는 것이다. 사실 인터넷이 정보화 시대를 상징하게 한 것도 하이퍼텍스트(Hyper Text)에 기반한 정보 연계 기능이다. 이 기능으로 인해 꼬리에 꼬리를 물고 마우스 클릭만으로 정보의 형태에 관계없이 원하는 정보에 접근할 수 있게 된 것이다. 아래 〈그림 7〉에서 보듯이 임진왜란을 설명하는 콘텐츠에서 인명사전에 있는 '원균(元均, 1540~1597)'을 연계함으로써 내용을 보다 정확하고 입체적으로 이해할 수 있다.

〈그림 7〉 Naver에서 웹 콘텐츠와 '인명'이 연계된 예시

그런데 이것을 구현하기 위해서는 〈a href="URL"〉용어〈/a〉 형식의 HTML 태그를 수동으로 입력해야 한다. 이것은 문서의 크기가 클수록 연관된 내용이 많을수록 매우 많은 시간이 소요되는 고된 노동이기도 하다. 하지만 이상에서 논의된 방식으로 문서를 만들면 특정 인명과 서명에 대해서는 자동으로 인명사전 또는 서명사전과 연계하도록 할 수 있다. 위에 제시된 XML 문서는 "인종"을 "인종(仁宗, 1515~1545)"으로 입력할 때 이때 자동적으로 인명사전에 있는 URL을 같이 삽입할 수 있도록 내부적으로 태깅하는 것을 예상한 것이다. 이 자동 태깅과 자동 연계 부분은 이 논문의 후속 연구에서 논의되고 프로그램으로 개발되어야 할 부분으로 남겨 둔다.

Ⅳ. 결론 및 향후 과제

이상에서 고찰한 내용은 주로 '한국학 분야'의 인명과 서명으로 분류되는 표제어휘를 활용하는 부분이었다. 그런데 필자가 통합대사전의 표제어휘를 조사해 보니 중국의 인명과 서명이 상대적으로 더 많았다. 한국학 연구 분야에 따라 중국의 인명과 문헌이 자주 나오는 분야도 있지만, 한국을 대표하는 사전의 위상을 확보하기 위해서는 한국관련 인명과 서명을 대대적으로 수집 정리하고 이를 워드프로세서에 포함하면 한국학은 물론 중국학 연구에도 유용한 도구가 될 수 있을 것으로 보인다.

이상에서 통합대사전의 표제어휘를 활용하여 워드프로세서용 고유명칭 표기 형식 변환기를 개발하는 방안을 살펴보았다. 단순히 이론적으로만 검토한 것이 아니라 실제로 인명정보 표기 형식 변환기를 개발하여 흔글 워드프로세서에 탑재하고 보급한 경험을 바탕으로, 서명 변환 기능까지 확장한 것이므로 연구의 실천적 활용도도 클 것으로 판단된다. 워드프로세서라는 거대한 프로그램도 결국은 이러한 작은 사전 프로그램들이 모여서 형태를 갖추고 각종 편의 기능을 제공한다.

이 연구는 단순히 고유명칭 입력 형식 변환기 개발이라는 측면에서만 유용한 것은 아니다. 최근 온라인 사전의 범람으로 인해 존립 위기를 겪는 각종 사전들의 '2차적 활용방안'을 실천적으로 모색해 본 사례라는 점에서도 의의를 지닐 수 있을 것이다. 사전은 당대 지식의 정수가 체계적으로 정리된 집합체이다. 그러나 정보화의 물결로 인해 과거와 달리 서적형 사전은 쇠퇴하고 있다. 사전을 편찬하던 대형 출판사들이 사전 개정 회수를 현저히 줄이는 것만 보아도 확인할 수 있고, 서적형 사전을 출판하더라도 경제적 수익 압박이 덜한 국립국어원. 고려대학교 민족문화연구원, 한국학중앙연구원 등 일부 연구 기관들만 편찬하는 것에서 알 수 있다.

그럼에도 불구하고 많은 학회, 연구자들이 크고 작은 분야별 사전들을 지속적으로 편찬하고 있다. 그리고 이러한 사전 편찬 연구들은 사전의 기본 뼈대를 구성하는 전문 지식을 기초로 정보 기술 및 유관학문과 연계하는 융합학문으로 발전하고 있다. 이 연구가 『통합디지털한한대사전』을 비롯하여 동양학연구원에서 편찬했거나 편찬하고 있는 사전들이 본래의 편찬 의도를 넘어서서 디지털인문학적인 재활용 사례 탐구에 하나의 방향을 제시하는 한편, 한국학 연구자들이 오랜 기간의 연구 성과를 집약하여 편찬하는 사전들이 정보화 시대에 다양한 방식으로 새롭게 탄생되는 계기가 될 수 있기를 기대한다.

〈참고문헌〉

김은희,「한글 워드프로세서 한자사전의 분석과 개선안 모색」,『중국어문학지』72, 2020.
김지영,「『통합디지털한한대사전』의 DB구축과 온라인 사전편집기」,『동양학』54, 2013.
김현,「디지털 인문학과 고문헌 자료 연구」,『열상고전연구』50, 2016.
박대아,「세종과 한글을 활용한 지역 문화 정책－세종특별자치시의 사례를 중심으로」,『한국학연구』58, 2016.
양창진,「역사문화 콘텐츠 육성 정책의 회고와 전망 : 역사분야 국가DB를 중심으로」,『인문콘텐츠』35, 2014.
양창진,『통합디지털한한대사전』편찬 현황의 기계적 점검 방안 연구 － 표제자 및 표제어휘 통합 분석을 중심으로 －』『동양학』60, 2015.
이용구 · 정영미,「사전 정보를 이용한 단어 중의성 해소 모형에 관한 실험적 연구」,『정보관리학회지』24: 1, 2007.
장인식,「영역 고시조에 나타난 번역상의 문제점: 이순신, 이조년, 김종서 시의 경우」,『번역학연구』5: 2, 2004.
정재철,「『디지털한한대사전』의 표제어 구성에 대하여」,『동양학』50, 2011.
정철,「온라인 백과사전 만들기」.『한국사전학』11, 2008.
조형래,「워드프로세서 · 글쓰기 · 문학, 1980 · 1990 : 한국의 초기 워드프로세서 보급과 작가의 글쓰기 Ⅰ」,『동악어문학』73, 2017.

네이버 지식백과(https://terms.naver.com)
한국학술지인용색인(http://www.kci.go.kr)
한국민족문화대백과사전(http://encykorea.aks.ac.kr)
한국역대인물종합정보(http://people.aks.ac.kr)
DB피아(http://www.dbpia.co.kr)

* 이 논문은 2021년 5월 27일에 투고되어,
2021년 6월 11일에 심사위원을 확정하고,
2021년 6월 30일까지 심사하고,
2021년 7월 7일에 게재가 확정되었음.

▌Abstract ▐

Study on the Using Headwords of *The Great Chinese-Korean Digital Dictionary* in the Age of Digital Humanities

Yang, Chang-jin*

Today, all research writing is done with a computer word processor. There are many proper names, such as names of people, place names, book titles, and official titles, in writings and papers published as a result of research. Among them, names and book titles must be included in the study of the traditional period. This is because these studies are the study of humans as well as the events and documents created by humans.

However, the names of people and book titles are not recorded only with alphabetic characters, but are indicated by combining various special symbols and numbers. It is also because the researcher's preferred input format is different, and the notation format is different for each journal.

Examples of the book titles are ① *『三國史記(Samguk sagi)』*, ② *『Samguk sagi(三國史記)』*, ③ *≪三國史記(Samguk sagi)≫*, ④ *≪Samguk sagi(三國史記)≫*, ⑤ *≪三國史記≫*, ⑥ *≪Samguk sagi≫*.

The above examples are some of the frequently appearing forms. The problem is that the notation form is varied and the input is complicated. And if you refer to the wrong data, you may enter it incorrectly. In order to reduce such inconvenience and errors, embedding the vocabulary of a reliable dictionary into the word processor program will help improve research efficiency and maintain consistency of information format.

The purpose of this thesis is to develop a proper name format converter for a word processor and to present a new use case of a the headwords of *The Great Chinese-Korean Digital Dictionary*.

[Keywords] *The Great Chinese-Korean Digital Dictionary*(統合漢韓大辭典), Proper Name, Book Title, Word Processor, On-line Dictionary

* Chief Researcher, The Academy of Korean Studies