

# AI를 활용한 중국시 연구 방안 탐색\*

— 《당시삼백수》와 AI 시를 대상으로 —

김 지 영\*\*

<目 次>

I. 서론	2. 당시와 인공지능 시 작품 비교 분석
II. 텍스트로서의 《당시삼백수》와 인공지능	IV. 중국시 연구에서 인공지능 활용방안 탐색
III. 《당시삼백수》와 인공지능 시 비교 분석	V. 결론
1. 《당시삼백수》의 시어분석	

## I. 서론

지금 우리사회에게 가장 이슈로 떠오른 화두는 단연 ‘4차 산업혁명’일 것이다. 4차 산업혁명은 “IT 기술 주도에 의한 혁신으로 모든 것이 연결되고 보다 지능적인 사회로의 진화”라고 정의되는데, 이를 이끌어갈 핵심 기술 가운데 하나가 인공지능(AI)이다. 최근 들어 빅 데이터에 기초한 AI의 알고리즘이 다양하게 활용되면서 문학작품 창작에서도 이를 활용한 시도가 다양하게 이뤄지고 있다. 이런 흐름에서 간과할 수 없는 것이 한시를 짓는 AI의 출현이다. 2013년 바이두(Baidu)에서 처음으로 모바일 앱(APP)

\* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구 자지원사업의 지원을 받아 수행된 연구임(NRF-2020S1A5A2A01042596).

\*\* 성결대학교 파이데이아학부 조교수

度秘(Dumi, 英文名:Duer)에 ‘爲你寫詩’ 기능을 출시하면서 일반인들도 詩題를 입력하면 쉽게 한시를 지을 수 있게 되었다. 제목 그대로 사람을 대신해 한시를 지어 주는 AI가 등장한 것이다. 심지어 2017년 6월에는 AI 샤오빙(小冰)이 쓴 현대시 시집인 《陽光失了玻璃窗》이 출간되어 사람들에게 “시란 단지 알고리즘에 불과한 것인가”라는 논쟁을 불러일으켰다.<sup>1)</sup>

이제 AI는 일상생활에서 뿐만 아니라 쉽게 대체하기 어려울 거라 예상했던 문학과 예술 분야에서도 활발히 활용되고 있다. 특히 시인이 아니면 짓기 어려웠던 중국 한시도 고전 시가라는 빅데이터 자료를 학습한 AI가 등장하여 한시를 쉽게 지을 수 있는 시대가 된 것이다. 인공지능이 생성한 시를 이전 시인이 지었던 한시와 같은 선상에서 볼 수 있는가 하는 문제는 앞으로 논의되어야 할 것이다. 그러나 한시를 짓는 AI가 출현한 시점에서 AI가 지은 시가 어느 정도의 수준이며, 중국시 연구에 AI를 활용할 수 있는 방안은 무엇인지 고찰해보는 것은 매우 의미 있고 시의성 있는 작업이라 생각한다. 필자는 인공지능이 중국시 연구에 도움을 줄 수 있을 것이라는 기대로 중국 고전시가 연구에 AI를 활용한 연구를 진행하고자 한다.

이러한 연구의 일환으로 당시선집인 《당시삼백수》와 AI 시를 대상으로 이를 비교 분석하여 연구를 진행하고자 한다. 《당시삼백수》는 대표적인 당시선집으로 양과 질에서 비교적 균형있게 정리되어 있어 AI 시와 함께 연구할 고전 텍스트로서 대표성을 지닌다. 본고에서는 먼저 텍스트로서의 《당시삼백수》와 인공지능에 대하여 살펴보고, 《당시삼백수》에서 추출한 주요 시어를 키워드로 당시와 인공지능의 시를 비교 분석하며, 더 나아가 중국시 연구에서 AI의 활용방안을 탐색해보고자 한다.

1) 〈詩歌只是一種算法? 人工智能寫詩出詩集〉, 華夏經緯網, 2017.06.

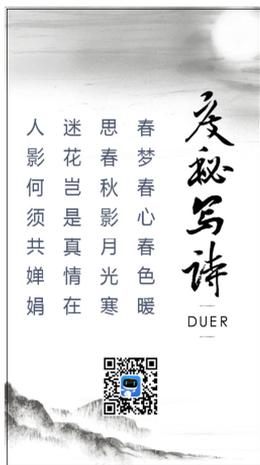
## II. 텍스트로서의 《당시삼백수》와 인공지능

AI를 활용한 중국 고전시 연구 방안을 고찰하기 위해 먼저 《당시삼백수》가 텍스트로서 가지는 의미와 인공지능의 작시 알고리즘의 원리에 대하여 살펴보기로 한다.

《당시삼백수》는 역대의 시선집 가운데 가장 대표적인 당시선집이다. 《전당시》에는 4만 8900여 수나 되는 시가 수록되어 있어 사람들이 唐詩를 쉽게 익히기가 어려웠다. 이에 청대 孫洙(1711~1778)는 글방의 교본으로 唐詩를 배울 수 있도록 시의 양을 조절하고 작가와 작품도 고르게 엄선하여 이 책을 편찬하였다. 《당시삼백수》에 수록된 작가의 수는 77명이고, 시의 제목은 294편이며 실린 시는 연작시를 포함하여 320수이다. 이 책의 체례를 살펴보면 오언고시·칠언고시·오언율시·칠언율시·오언절구·칠언절구로 분류되어 있고, 작품 아래 評點과 간단한 주석을 달아놓았다. 내용 역시 광범위하여 서사시·서정시·염정시·변재시·회고시·영회시·기행시·민가시·송별시·증답시·영물시·유선시·은일시·전원시 등의 작품이 고르게 분포되어 있으며, 唐詩 작가도 문인뿐 아니라 제왕·사대부·승려·기녀·무명씨에 이르기까지 다양하게 수록해 놓았다. 많은 당시선집이 주로 전문적인 학술적 가치에 목적을 두고 만들어졌다면, 이 책은 唐詩 학습용 교과서로 편찬된 점이 폭넓게 애독되었던 이유이기도 하다. 이처럼《당시삼백수》는 시 형식에 따라 다양한 내용을 살펴볼 수 있도록 《전당시》에 실린 唐詩를 압축하여 균형 있게 선별했다는 점에서 AI 시와 비교할 고전 텍스트로서 충분한 가치가 있다고 하겠다.

컴퓨터로 한시를 짓게 하는 기술은 2013년 바이두에서 모바일앱을 통해 처음 선보였다. 이때 제공한 ‘爲你寫詩’ 기능은 이용자가 사진을 업로드 하면 그림의 내용에 따라 자동적으로 4구의 한시를 만드는 방식이었다. 2013년판 ‘爲你寫詩’의 핵심 알고리즘은 SMT(통계기반 기계번역, Statistical Machine Translation) 방식으로 1.0 버전이었다. SMT는 컴퓨터가 언어에

서 말뭉치(Corpus)를 학습하여 통계학적으로 번역하는 기술이다. 이 1.0버전에서는 한시에 운율과 압운을 만들 수 있고, 對仗도 맞출 수 있었으나 시의 내용이 뒤 구로 갈수록 주제에서 벗어나는 경우가 많았다. 반면 2016년에 나온 AI 度秘(Dumi)의 ‘爲你寫詩’ 기능은 2.0 버전의 NMT(인공신경망 기반 기계 번역, Neural Machine Translation) 방식으로 바뀌었다. NMT는 신경망 알고리즘으로 AI가 데이터 학습을 통해 문장 단위로 언어를 번역하는 딥러닝 기술이다. 따라서 본격적인 AI 시는 2016년에 나온 度秘의 ‘爲你寫詩’ 기능에서 시작되었다고 할 수 있다. 2.0 버전인 度秘는 이전의 통계기반 기계번역(SMT) 방식에서 인공신경망 기반 기계번역(NMT) 방식으로 알고리즘이 설정되어 1.0버전보다 생성된 한시가 더욱



정교해졌다. 2.0버전의 신경망 알고리즘은 이용자가 시제를 입력하면 AI가 생성될 내용을 먼저 예측하고, 각 시의 하위 주제를 예측하며 이어서 하위 주제별로 단어를 배치한다.<sup>2)</sup> 이렇게 지어진 시의 맨 앞구절에 시제가 그대로 나오는 것이 특징이었다. 예를 들면 AI 度秘에 <春思>라는 시제를 입력하면 “春夢春心春色暖, 思春秋影月光寒. 迷花豈是真情在, 人影何須共婵娟.”이라고 시를 생성하는데 첫 구와 둘째 구의 앞글자에 시제가 들어가는 게 특징이다. 왼쪽은 인공지능 度秘가 지은 시이다.

2016년에 중국에서 AI가 지은 시의 수준을 가늠할 수 있는 실험을 진행한 적이 있었다. 일반인과 중국문학 연구자를 포함한 40명의 평가자를 선정해, 고대 시인의 시 20수와 AI 시 20수를 무작위로 섞어 놓고 고대 시인이 쓴 시를 고르게 하였다. 그 결과 일반인 그룹에서는 50%만이 고대

2) Zhe Wang, Wei He, Hua Wu, Haiyang Wu, Wei Li, Haifeng Wang, Enhong Chen. <Chinese Poetry Generation with Planning based Neural Network>. <<University of Science and Technology of China. Hefei>>, 2016.

시인의 시를 골랐고, 전문가 그룹에서는 84%가 정확히 골라내었다. 이는 당시 AI 시가 일반인이 구분하기 어려울 정도의 수준은 되나, 아직 전문가를 속일 정도의 수준까지는 미치지 못했음을 말해 준다.<sup>3)</sup>

이후 칭화대학에서 인공지능 九歌(Jiuge)를 개발하여 수준이 더욱 향상된 인공지능 시를 지을 수 있게 되었다. 九歌는 자연어 처리 및 사회 인문학 컴퓨팅 연구소에서 개발한 인공지능 시 쓰기 시스템이다. 이 시스템은 최신 딥 러닝 기술을 채택하고, 시 생성을 위해 특별히 설계된 여러 모델을 결합하고, 예전 시인이 쓴 30만 수 이상의 시를 기반으로 훈련·학습하였다. 다른 시 생성 시스템과 달리 九歌는 다중 모드 입력, 다중 장르 및 다중 스타일, 인간-컴퓨터 상호 작용 생성 모드의 특성을 가지고 있다.<sup>4)</sup> 아래는 인공지능 九歌가 지은 시이다.



2024년 3월 5일자 <九歌 인공지능 시 쓰기 앱 공식 웹사이트 다운로드 (《九歌》人工智能诗歌写作app官网下载)>에 소개된 내용을 살펴보면 다음과 같은 내용이 있다: “九歌는 AI 기술을 기반으로 한 시 창작 도우미로,

3) <自動寫詩 PK 古代詩人：百度‘為你寫詩’技術深度揭秘> 百度NLP, 2016.3.14.

4) 칭화대학 인공지능 九歌 시스템 연구소 소개서.

풍부하고 다양한 시 콘텐츠를 제공할 수 있습니다. 또한 이 소프트웨어는 사용자가 이해하고 배울 수 있는 풍부하고 포괄적인 시 콘텐츠를 제공하여 사용자가 시 창작 능력을 지속적으로 향상시킬 수 있도록 돕습니다.”<sup>5)</sup> 이처럼 키워드만 입력하면 시를 창작할 수 있는 인공지능 시인인 九歌는 사용자들이 한시를 이해하고 배울 수 있도록 하는 목적으로 계속 알고리즘을 발전시켜나가고 있다. 지금 사용되고 있는 九歌는 앞에서 보았던 것보다 조금 더 다양한 기능을 갖추고 있다. 아래는 인공지능 九歌의 최신 버전이다.



인공지능 九歌가 처음 나왔을 때 시를 생성하는 모드는 絶句, 风格绝句, 藏头诗, 集句诗, 律诗, 词로 총 6개였으나, 지금은 생성 모드가 絶句, 风格绝句, 藏头诗, 集句诗, 律诗, 词, 对对子, 类文句搜奇로 8개로 늘어나 더욱

- 5) 珞珞, <<九歌>>人工智能诗歌写作app官网下载>, “九歌, 也可以被叫做九歌AI诗人, 这是一款基于AI技术打造的诗歌创作助手, 它能够为你提供丰富多样的诗词内容, 只需要输入相应的关键词就可以快速生成, 满足你与众不同的诗词创作需求. 软件还提供丰富全面的诗词内容供用户了解和学习, 帮助用户不断提升自己的诗词创作能力. 软件的界面简洁明了, 操作简单易用, 让用户能够轻松上手. 它还提供实时在线客服和社区交流功能, 方便用户随时获取帮助和与其他用户交流心得.”(2024.3.15)

다양한 시를 생성할 수 있게 되었다. 이처럼 인공지능 九歌의 알고리즘은 계속 향상되고 있음을 알 수 있다.

본고에서는 당시를 대표할 수 있는 《당시삼백수》와 한시 짓는 AI ‘度秘’와 ‘九歌’를 텍스트로 고전 한시와 인공지능 시 작품을 비교 분석하고자 한다.

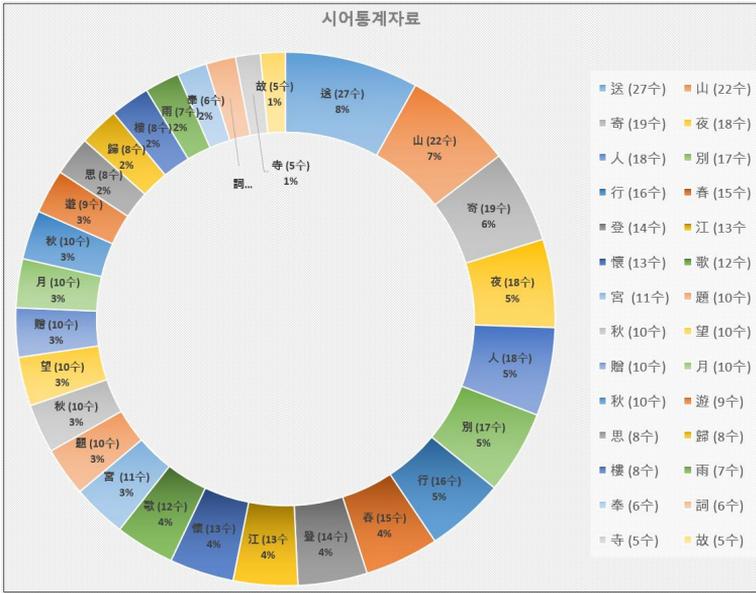
### Ⅲ. 《당시삼백수》와 인공지능 시 비교 분석

#### 1. 《당시삼백수》의 시어분석

《당시삼백수》와 AI 시 비교 분석을 위해서는 먼저 어떤 주제로 작품을 선별할 것인지에 대한 기준을 만들 필요가 있다. 필자는 이를 위해 《당시삼백수》를 텍스트 마이닝해서 주요 시어를 추출하고 이를 시각화하였다. 먼저 《당시삼백수》전체를 텍스트 마이닝해서 가장 많이 쓰인 시어를 살펴보았다. 《당시삼백수》에 쓰인 단어는 총 4536개인데 그 가운데 가장 많이 쓰인 단어는 ‘人’으로 총 223회 사용되었다. 다음으로 山(178회), 天(123회), 夜(122회), 日(131회), 月(130회), 風(119회), 白(119회), 長(116회), 雲(108회), 春(96회) 등으로 시어가 많이 쓰였다. 《당시삼백수》에서 50회 이상 쓰인 글자는 36개로 이를 인포그래픽으로 시각화하면 다음과 같다.<sup>6)</sup>

6) 人(223회), 山(178회), 天(123회), 夜(122회), 日(131회), 月(130회), 風(119회), 白(119회), 長(116회), 雲(108회), 春(96회), 君(92회), 生(91회), 水(87회), 見(87회), 花(85회), 秋(80회), 歸(80회), 明(74회), 落(70회), 空(70회), 年(66회), 去(64회), 馬(61회), 清(61회), 西(61회), 南(59회), 家(58회), 心(57회), 別(56회), 飛(56회), 寒(56회), 黃(52회), 知(51회), 望(50회), 流(50회).





이 도표는 《당시삼백수》의 詩題를 대상으로 추출한 자료이다. 여기서 ‘送’자가 가장 많이 쓰였다는 것은 《당시삼백수》에는 송별시가 가장 많이 수록되어 있다는 것을 보여준다. 또한 ‘寄’(19수), ‘別’(17수), ‘贈’(10수), ‘題’(10수), ‘奉’(6수)자 등이 많이 쓰인 것을 보면 전별시·교유시·화답시 등도 많이 쓰였음을 알 수 있으며, 더불어 시제를 대상으로 추출한 정보라는 것도 동시에 알 수 있다. 시어 가운데 ‘山’(27수)자가 두 번째로 빈도수가 높았는데, 이는 唐詩 작가들이 산을 소재로 쓴 시가 월등히 많았음을 말해준다. 또한 시제 안에 계절을 지칭하는 단어인 ‘春’자가 15수에, ‘秋’자가 10수에 보였고, 도표에는 나와 있지 않지만 ‘夏’자가 2수에 보였고 ‘冬’자는 시제 속에 한 수에도 보이지 않았다. 이를 통해 시인들이 즐겨 노래했던 계절은 역시 봄과 가을이었음을 알 수 있다. 《당시삼백수》의 전체 단어와 《당시삼백수》의 詩題를 텍스트 마이닝해 추출한 시어들 가운데 출현 빈도수가 매우 높은 ‘人’, ‘山’, ‘天’, ‘送’, ‘春’, ‘秋’ 등의 단어

들은 당시를 대표하는 주요 키워드라 할 수 있겠다.

## 2. 당시와 인공지능 시 작품 비교 분석

앞서 《당시삼백수》에 가장 많이 쓰인 시어는 人과 山이었고, 《당시삼백수》의 詩題로는 送이, 계절을 나타내는 단어로는 春이 주요 시어로 쓰였음을 알 수 있었다. 필자는 여기에 나타난 주요 시어들을 기반으로 시 선별을 위한 기준을 삼고자 한다. 당시 가운데 봄(春)을 주제로 한 시, 누군가(人)와 헤어지며 송별(送)한 시, 산(山)을 배경으로 한 시를 주제로 정해 당시와 인공지능의 시를 비교 분석하기로 하겠다. 《당시삼백수》가운데 이러한 주제를 가진 시가 많이 있었으나 특정 인물이 언급되거나 고유지명 또는 고유명사가 들어간 시제들은 제외하였다. 특정하거나 고유한 뜻을 가진 단어들은 인공지능이 그 뜻을 살려 시를 생성하기가 어려울 수 있기 때문이다. 필자는 이런 기준으로 《당시삼백수》에서 〈春思〉, 〈送友人〉, 〈山居秋暝〉 시를 선별하였는데, 이 시제를 키워드로 당시와 AI 시를 비교 분석하기로 한다.

먼저 봄을 주제로 한 ‘春思’를 키워드로 이백의 〈春思(봄날의 그리움)〉와 AI 度秘의 시를 비교 분석해 보기로 한다.

	이백 〈春思〉
燕草如碧絲,	연나라의 풀이 파란 실과 같을 때
秦桑低綠枝.	진나라의 뽕나무 푸른 가지 낮게 드리웠네.
當君懷歸日,	그대 돌아갈 생각하는 날
是妾斷腸時.	그때는 첩이 애끊는 때지요.
春風不相識,	봄바람은 나를 알지 못하면서
何事入羅幃.	무슨 일로 비단 휘장 안으로 불어오는지.

	AI 度秘 〈春思〉
春夢春心春色暖	봄 꿈 봄 마음 봄 색은 따뜻한데



다음 시는 ‘春思’라는 詩題로 AI가 지은 것으로 봄의 정서를 담아낸 7언고시이다. 제1-2구는 봄에 꾸는 꿈, 마음, 색깔은 모두 따뜻한데, 이런 봄을 생각하노라니 가을의 그림자와 달빛은 생각만 해도 차갑다고 말하고 있다. 제3-4구는 꽃에 미련을 두는 것은 진실한 감정이 아니며, 봄에 꼭 미인과 함께 할 필요도 없다고 말하고 있다. 제3구에서의 꽃은 봄이 되어 보게 된 꽃일 수도 있겠으나, 뒤 구와 연결해서 살펴보면 미인을 지칭하는 말로 쓰인 것이다. 따라서 3-4구는 모두 미인에 연연할 필요가 없다는 내용이다. 전체적인 내용을 살펴보면 제1-2구는 봄의 따뜻함을 말하고 있고, 제 3-4구는 미인에 연연하지 말라는 내용이라 의미에서 서로 연관성이 적다. 또한 제1-2구에서의 봄과 가을의 대조 방식도 매우 조악하다. 시구 안에 ‘春’자가 네 번, ‘影’자가 두 번이나 쓰인 것도 시의 맛을 떨어뜨린다. 暖·寒·娟자를 압운자로 쓰고 있는데, 暖은 上聲 ‘阮’운, 寒은 上平 ‘寒’운, 娟은 下平 ‘先’운이어서 압운도 어색하다. AI에 시제를 입력하면 신경망 알고리즘의 원리로 우선 압운을 정하고 주제와 비슷한 시어를 여러 개 골라서, 서로 대구를 이루는 곳에는 같은 글자를 반복하지 않는다는 원칙에 따라 시를 작성한다. 그러다 보니 이처럼 1-2구와 3-4구의 내용이 잘 이어지지 않는 작품이 만들어지기도 한다. 지금의 경우는 두 작품의 수준을 비교하기 보다는 이런 경우 AI가 왜 이런 시를 짓게 되었는지를 AI의 알고리즘을 상기하며 이해하는 편이 나올 것이다. 물론 시를 조합하는 과정에서 기발한 시도 나올 수 있으며, 다른 시제를 입력하면 전혀 다른 시가 나올 것이다.

《당시삼백수》의 시의 제목으로 ‘送’자 가장 많이 등장한다는 것은 시인들 사이에서 송별시가 많이 지어졌다는 방증이기도 하다. 송별시의 경우 특정인을 지칭하는 경우가 많은데 여기서는 친구를 송별하라는 의미의 ‘送友人’이란 키워드로 시를 살펴보자 한다. 李白의 〈送友人〉시와 AI 九歌의 시를 함께 살펴보기로 한다.

李白 〈送友人〉<sup>7)</sup>

青山橫北郭,  
白水繞東城。  
此地一爲別,  
孤蓬萬里征。  
浮雲游子意,  
落日故人情。  
揮手自茲去,  
蕭蕭班馬鳴。

푸른 산은 북쪽 성곽 위로 빗겨있고  
맑은 물은 동쪽 성을 감싸며 흘러간다.  
이제 여기서 자네가 한번 떠나면  
외로운 다복쑤처럼 만리 타향 떠돌겠지  
뜬 구름은 나그네의 마음이고  
지는 해는 친구의 심정이네.  
손을 흔들며 이제 떠나려 하니  
히잉, 떠나는 말의 울음소리.

AI 九歌 〈送友人〉

一片離愁緒,  
悠然萬里身。  
春風隨客路,  
落日照江津。  
楚岫雲連樹,  
吳門草滿塵。  
相期在此別,  
明月是誰鄰。

한 조각 이별을 앞둔 근심의 마음  
유유히 만 리로 떠나가는 신세.  
봄바람 그대 가는 길에 불고  
떨어지는 해는 강 나루를 비추네.  
초나라 골짜기의 구름에 나무 이어지고  
오나라 입구 풀 속에는 먼지가 가득하네.  
서로 여기서 헤어질 것을 기약하니  
밝은 달은 누구와 이웃할까?



7) 이 시는 747년(天寶 6년) 金陵(지금의 江蘇省 南京)에서 지은 것으로 전해지나 확실한 근거는 없다.

이백의 〈送友人〉 시는 친구가 멀리 떠나가는 것을 송별하며 지은 5언율시이다. 송별하는 친구가 누구인지는 분명하지 않다. 이 시는 크게 제1-2구에서는 경물을 묘사하였고, 제3-8구에서는 이별에 대한 서정의 마음을 표현하였다. 좀 더 자세히 살펴보면 제1-2구는 이별의 배경을 산과 강으로 나누어서 묘사했는데 여기에는 ‘靑’과 ‘白’의 색채 대비 및 ‘北’과 ‘東’의 방향 대비가 적절히 조화를 이루고 있어 묘사가 선명하면서도 짜임새가 있다. 제3-4구는 먼저 ‘此地’라는 표현을 통해 앞의 연과 매끄럽게 연결하고 있다. 다복쑥(孤蓬)은 바람이 불면 뿌리채 굴러다니는 특징을 가지고 있는데 외로이 떠도는 나그네를 상징적으로 표현하고 있다. 이별의 슬픔은 앞 연의 아름다운 경치와 대조되어 슬픈 감정을 배가시키고 있다. ‘一’과 ‘萬’의 대비 역시 선명하다. 제5-6구는 ‘浮雲’과 ‘落日’로써 길 떠나는 친구와 그를 떠나보내는 자신을 비유하였다. ‘此自’라는 표현 속에 이별의 장소와 사람이 농축되어 있다. 제8-9구는 말울음 소리를 통해 간접적으로 이별의 슬픔을 나타내고 있는데 직접적인 슬픔의 표현보다 더 긴 여운을 남긴다. 마지막 구는 말울음 소리로 작가의 슬픔을 대신한 연외의 표현이라 할 수 있다. 이 시는 오언율시로 압운자는 城, 征, 情, 鳴으로 모두 下平 ‘庚’운을 사용하였다.

다음 시는 九歌가 생성한 〈送友人〉으로 이별을 노래한 5언율시이다. 이 시는 크게 1연과 4연은 서정을 노래하고 2연과 3연은 경물을 묘사하여 짜임새 있는 구조를 갖추고 있다. 자세히 살펴보면 제1-2구는 친구를 송별하며 생기는 아쉬운 마음과 떠나가는 친구의 모습을 통해 이별의 상황을 묘사하고 있다. ‘一片’과 ‘萬里’로 표현된 일편의 마음과 만리의 길은 보내는 이와 떠나가는 친구의 마음을 선명하게 대비해 보여주고 있다. 제3-4구는 ‘春風’과 ‘落日’을 통해 이별 장소의 경물을 묘사하였다. 이별을 앞둔 해지는 강 나무에 봄바람이 불어오고 있는 풍경을 표현하였다. 제5-6구는 ‘楚岫’와 ‘吳門’의 표현으로 친구가 멀리 강남 지역으로 떠난다고 말하고 있다. 이곳은 구름마다 나무가 이어지고 풀 속에 먼지가 가득한 곳으로 깊고 우거져 있어 평탄한 길이 아님을 시사하고 있다. 제7-8구는 여기서 친

구와 헤어지게 되면 함께 보았던 밝은 달을 이제는 감상할 수 없게 된다. 이별의 슬픈 마음을 비유적으로 표현하였다. 이 시의 압운자는 身, 津, 塵, 鄰으로 상평 ‘眞’운을 사용하였다. 이 시는 전체적으로 친구를 송별하는 내용을 경물과 서정의 방식을 사용하여 내용이 무난하게 이어지며 구조 역시 짜임새가 있다.

앞서 度秘가 지은 시에 비하여 시의 내용과 구성에 있어 많이 발전되었음을 알 수 있다. 이는 인공지능 度秘의 경우 2016년에 개발되었고 위에 작품 〈春思〉의 경우 2017년에 생성시킨 것이어서 알고리즘이 지금보다는 긴밀도가 떨어지고 덜 유기적으로 표현되었다. 반면 九歌의 〈送友人〉시는 최근에 생성시킨 것으로 그 사이에 최신 딥러닝 기술을 통해 시구 사이의 의미적 연관성을 가지도록 훈련·연습시킨 결과라고 할 수 있다. 이백의 〈送友人〉시는 아름다운 경물과 이별의 상황을 선명하게 대비하여 슬픔을 강조하였다. 또한 ‘외로운 다복쭉(孤蓬)’으로 나그네의 신세를 비유하고 말 울음소리(馬鳴)를 통해 작가의 슬픔을 언외의 방식으로 표현한 부분은 시인이 시로써 구현해 낸 창의적인 수법이라 할 수 있다. 이런 기법은 독자들에게 시에 대한 재미와 감동을 선사한다. 반면 구가의 〈送友人〉은 친구와 이별의 상황을 무난하게 묘사하고 있어 인공지능의 알고리즘이 발전하고 있음을 새롭게 느끼게 된다. 그러나 AI 시에서 기존 시와 다른 새로운 기법, 언외의 뜻, 행간의 의미 등을 찾기를 기대하기는 아직 어려운 것이 현실이라 하겠다.

다음은 《당시삼백수》에서 두 번째로 많이 쓰인 시어인 ‘山’을 키워드로 한 시를 살펴보기로 하겠다. 왕유의 〈山居秋暝(산 속의 가을 저녁)〉과 AI 九歌의 시를 살펴보기로 한다.

	왕유 〈山居秋暝〉
空山新雨後,	텅 빈 산에 막 비가 내린 후
天氣晚來秋,	날씨는 저녁이 되자 가을이 된 듯 하네.
明月松間照,	밝은 달이 소나무 사이로 비치며

清泉石上流,  
竹喧歸浣女,  
蓮動下漁舟.  
隨意春芳歇,  
王孫自可留.

맑은 샘물 돌 위로 흐르네.  
대숲이 시끄러운 것은 빨래하던 여인들이 돌아오느라  
연꽃 움직이는 것은 고기잡이 배 내려가느라.  
순리대로 봄의 향기는 시들어도  
그대는 그런대로 머무를만 하시오.

AI 九歌 〈山居秋暝〉

樹影深雲落,  
林陰夕照新.  
孤燈明靜夜,  
萬籟起流塵.  
寒氣生窗外,  
涼聲隔水人.  
空階聽露坐,  
無寐欲傷神.

나무 그림자 깊은 곳에 구름은 내려앉았고  
숲 그늘에 저녁 빛이 새롭게 비추네.  
외로운 등 고요한 밤 비추고  
만개의 소리가 떠다니는 먼지 속에 일어나네.  
차가운 기운 창 밖에 생겨나고  
청량한 소리가 물과 사람 사이에 있네.  
빈 섬돌에 앉아 이슬 맞으며  
상심에 잠 못 들어하네.



왕유의 〈山居秋暝〉은 가을 저녁 산속에서의 생활을 노래한 오언율시이다. 이 시의 제1-4구는 가을 산의 경물을 묘사하였고, 제5-8구는 산 속에서의 생활을 서사하였다. 다시 자세히 살펴보면 제1구는 ‘산에 거주(山居)’하는 상황을 제2구는 ‘가을 저녁(秋暝)’의 모습을 그려내고 있는데 이는

시의 제목을 담아낸 것이다. 제3—4구는 ‘明月’과 ‘清泉’을 통해 시선을 위아래로 분산시키고, 산과 샘이라는 평범한 소재로 자연의 의경을 표현하였다. 제5—6구는 대나무 소리를 듣고 빨래하던 아낙이 돌아가고, 연잎의 흔들림을 보고 고깃배가 돌아가는 것을 안다고 표현하고 있는데, 이는 시인의 섬세한 관찰력과 뛰어난 표현력이 잘 드러난 부분이라 하겠다. 동시에 산중 생활이 사람들과 어우러져 있어 시인이 있는 곳이 정감 있는 곳임을 느끼게 한다. 제7—8구는 다른 사람에게 산 속에서의 생활을 권유하는 형식을 통해 쓸쓸한 가을이라도 산은 은거하기에 충분하다고 말하고 있다. 소식은 왕유 시와 그림을 평하여 “시 속에 그림이 있고, 그림 속에 시가 있다(詩中有畫, 畫中有詩)”고 하였는데, 이 시 역시 한 폭의 동양화를 연상시킨다. 이 시의 압운자는 秋, 流, 舟, 留로, 下平 ‘尤’운을 사용하였고, 전형적인 5언율시의 구조를 잘 보여주고 있다.

다음은 九歌의 〈山居秋暝〉시를 보기로 한다. 이 시는 조용한 밤 숲속에서 잠들지 못하는 작가의 모습을 오언율시로 노래하였다. 이 시의 제1구—6구는 고요한 저녁 숲의 모습을 경물로 묘사하였고, 제7—8구는 시인의 서정을 묘사하였다. 다시 제1—2구는 저녁이 되어 나무 그림자가 길게 드리워진 숲에 석양이 비친 모습을 묘사하고 있다. 구름이 내려앉았으나 새로 저녁 빛이 든다는 표현에서 ‘落’과 ‘新’이 서로 대조를 이루고 있다. 제3—4구는 밤이 되어 고요해지니 온갖 소리가 먼지 속에서 들리는 상황을 묘사하였다. ‘孤燈’과 “萬籟”의 시어는 시각에서 청각으로 옮겨가는 공감각적인 요소를 만들어내었다. 제5구—6구는 창밖으로 차가운 기운 생기고, 청량한 물소리가 들리는 상황을 묘사하였다. 제7구—8구는 섬돌에 앉아 잠 못 이루는 시인의 마음을 담아냈다. ‘空’, ‘霜’, ‘傷’ 등의 시어는 근심의 감정을 대변하고 있다. 이 시의 형식은 오언율시로 압운자는 新, 塵, 人, 神이며, 上平 ‘真’운을 사용했다. 구성은 1.2.3연에 걸쳐 산속 저녁의 경물을 묘사하고 마지막 4연에 작가의 상심을 표현하였다. 따라서 지나치게 경물 묘사 위주로 표현되었고, 시제인 ‘山居秋暝’이라는 주제에 맞는 내용을 담아내지 못하였다. 크게 보면 산속을 배경으로 한 작가의 시름이란 정도의

내용이라 할 수 있다. 왕유의 〈山居秋暝〉이 한 폭의 그림과 같은 표현으로 작가의 자연관을 잘 드러낸 작품성이 뛰어난 작품이라면, 구가의 〈山居秋暝〉은 숲이라는 자연적 배경을 늘어놓고 마지막은 이유를 알 수 없는 작가의 시름을 말하고 있어 독자의 공감을 얻기 어려운 작품이라 하겠다.

#### Ⅳ. 중국시 연구에서 인공지능 활용방안 탐색

중국시의 전통성은 공자가 《論語·季氏》에서 “시를 공부하지 않고서는 말할 게 없다(不學詩, 無以言)”라고 한 것에서 알 수 있다. 즉 시를 배움이 말배움임을 뜻하면서 시가 말의 모든 것을 갖추고 있음을 말하고 있는 것이다. 이처럼 중국시는 인문적 전통의 중심에 자리하고 있었기 때문에 시는 오랫동안 인문교육의 대상이자 그 훈련의 방편이 되어왔다.

중국의 전통시는 《시경》 이래로 중국 문인들에 의해 창작되어왔으며 唐代에 이르러 예술적 성취는 절정에 이르렀다. 중국시는 문자를 매체로 이룩한 예술창작의 활동의 최고의 결정체라 할 수 있는데, 이런 이유로 중국문학에서 시는 고전문학의 정수로 자리매김하였다. 중국시는 현재까지도 고전문학의 보고이자 인문정신의 중심으로 그 역할을 하고 있다. 이병한 교수는 《唐詩選》 서문에서 현대를 살아가는 우리에게 시는 어떤 의미가 있는지에 대해 다음과 같이 언급하였다: “시를 읽는다는 것은 오늘날처럼 기계문명이 고도로 발달하여 기계가 사람을 부리는 상황하에 어찌 보면 어리석은 것으로 보일지도 모른다. 그러나 기계는 시를 창작하고 감상하는 주체적 능력을 지니지 못하며, 짐승이 인간의 존엄에 도전할 수 없는 것이나 마찬가지로 도구에 지나지 않으며, 시는 여전히 인간 존엄의 상징으로 남는다.”<sup>8)</sup> 이처럼 고도의 예술작품을 창작하고 이를 감상할 수 있는 능력은 동물과 기계로 대체할 수 없는 인간만이 할 수 있는 영역이고 이는 인

8) 李炳漢·李永朱 譯解, 《唐詩選》 서울: 서울대학교출판부, 1998, 3쪽.

간 존엄의 상징이라 할 수 있다.

그러나 4차산업 혁명 시대에 등장한 AI는 인간만이 가능하다고 여겼던 창작하고 감상하는 능력을 대체하고 있다.<sup>9)</sup> 따라서 많은 사람들이 인공지능이 어디까지 발전할 것인가에 대한 호기심과 두려움을 동시에 가지게 되었다. 중국 고전 시가의 경우 지금 시를 짓는 인공지능이 출현하였고 이를 외면할 수 없는 상황이 되었다. 필자는 이런 현재의 상황에서 인공지능 시가 어느 정도의 작품을 지을 수 있는지를 알아보기 위해 唐詩와 인공지능 시를 함께 비교 분석해 보았다. AI가 지은 시는 훌륭한 것도 있었고 반대로 아주 형편없는 시도 있었다. 이는 시를 짓도록 명령하는 알고리즘의 기술에 따라 작품 수준이 달라졌기 때문이다. 그러나 분명한 것은 딥러닝의 기술이 계속 발전하면서 시어와 시어와의 관계를 긴밀하게 만드는 패턴을 학습·훈련시키게 되면 AI 시의 수준은 더욱 정교해질 것이다.

다만 AI가 생성한 시를 분석해 본 결과 몇 가지 한계점이 있음을 알 수 있었다. 우선 인공지능은 30만 수 이상의 시를 딥러닝 학습을 거쳐 설계한 것이어서, 압운, 평측, 시가 장르 등의 시의 형식을 갖추고 내용 역시 주제에 맞게 무난한 작품은 생성할 수 있다. 그러나 형식적 요소 이외에 예술작품이 갖추어야 할 창의적 수법이나 시라는 장르에서 만들어 낼 수 있는 시 밖의 의미 즉 言外의 뜻 같은 다층적 의미를 구현해 내기에는 아직 부족한 부분이 많다. 여전히 모방과 학습의 결과물으로써 표층적인 시로 머물 수 있겠다는 생각이 든다. 다음은 인공지능 시에는 인명이나 지명과 같은 고유명사를 주제로 한 시가 나오기 어렵다는 것이다. 《당시삼백수》의 시제로 송별시가 가장 많이 등장한 것은 그 시사하는 바가 크다. 시의 기능 가운데 상대방과 서로 주고 받으며 그 사람과의 특별한 감정을 노래하는 것이 중요한 요소 중 하나이다. 그때 교류하는 사람, 머무르는 장소,

9) 2022년에 나온 챗GPT는 말 그대로 채팅을 하기 위한 AI이며 GPT라는 모델에 기초하여 작동된다. 여기에서 말하는 GPT는 ‘생성형 사전 학습 트랜스포머(Generative Pre-trained Transformer)’를 줄인 것이다. 언어에 특화된 대화형 인공지능 챗GPT는 예술작품을 창작하고 감상하기도 한다.

떠나가는 곳 등 인명과 지명 같은 고유명사가 자연스럽게 등장하게 된다. 그러나 인공지능 시는 특별한 대상을 상대로 시를 지을 수 없으니 교유로서의 기능을 전혀 할 수 없는 것이다.

이러한 한계에도 불구하고 필자는 인공지능이 중국시를 학습·훈련하여 작시하게 되는 원리를 실제 한시를 지을 때 활용한다면 많은 도움이 될 것 같다는 생각이 들었다. 우리는 시를 지을 때 먼저 주제를 생각하고 압운을 정한 다음, 시 속에 들어갈 키워드 즉 시어를 생각한다. 그리고 쓰고자 하는 내용을 압운·대구·평측 등에 잘 맞추어서 시를 작성한다. 이는 AI의 알고리즘과 매우 유사함을 알 수 있다.

다음에서 AI가 지은 〈春雪〉<sup>10)</sup>시를 통해 시의 작시 원리를 살펴보기로 한다.

飛花輕灑雪欺紅,	꽃잎 가벼이 날려 떨어지는 데 눈이 붉은 꽃을 기만 하였고
雨後春風細柳工.	비온 후 봄바람은 가는 버들 더 푸르게 하였네.
一夜東君無限恨,	하룻밤의 봄 사이에 끝없는 한 생겨나니
不知何處覓青松.	푸른 소나무 어디서 찾아야 할지 알지 못하네.

〈春雪〉이라는 시제가 입력되자 AI는 먼저 紅·工·松자를 압운자로 골랐을 것이다. 그리고 주제인 ‘春’과 ‘雪’ 또는 ‘春雪’과 관련된 시어로 ‘飛花’, ‘雪’, ‘細柳’, ‘東君’, ‘青松’, ‘雨’를 뽑아, 평측과 대구를 맞추어 배열했을 것이다. 여기서는 제1-2구의 ‘飛花’와 ‘雨後’, ‘紅’과 ‘工’을 대구로 삼아 평기식으로 시를 구성하였다. 압운자인 ‘紅’과 ‘工’은 上平 ‘東’운, ‘松’은 上平 ‘冬’운으로 通韻하였다. 이런 원리로 7언절구 시를 구성했을 것이다. 이처럼 AI가 어떤 방식으로 시를 지었는지 거슬러 고찰해 보면, 시를 구성하는 요소들을 좀 더 체계적으로 생각해보게 된다. 이런 과정은 학생들에

10) 이 시는 ‘塑料花’라고 불리는 AI가 지은 것이다. AI 시가 나오기 시작하자 ‘사람이 지은 시’와 ‘AI가 지은 시’를 고르게 하는 대회가 열렸다고 한다. 이때 제출 문제로 나온 시다.(澎湃新聞 2017년, 2.11)

게 중국시를 교육을 할 때 매우 유용할 것이다. 시의 내용을 감상하는 단계에서 더 나아가 중국시를 구성하는 주요 요소인 압운, 평측, 대구, 시어, 詩形 등에 대하여도 비교적 수월하게 이해할 수 있도록 안내할 수 있을 것이다.

〈九歌 인공지능 시 쓰기 앱 공식 웹사이트〉에 九歌 앱의 기능에 대하여 “사용자가 이해하고 배울 수 있는 풍부하고 포괄적인 시 콘텐츠를 제공하여 사용자가 시 창작 능력을 지속적으로 향상시킬 수 있도록 돕는다”<sup>11)</sup>라고 언급하고 있는데, 필자 역시 AI가 중국시 교육에 도움을 줄 수 있을 것이 생각한다. 인공지능 시가 지은 작품이 아직은 예술성이 뛰어난 당시의 수준을 뛰어넘을 수는 없겠지만, 지금 우리가 AI 시의 작시 원리를 역추적하는 방식으로 새롭게 접근한다면 중국시를 이해하는데 도움을 줄 수 있을 것이라 기대한다.

#### IV. 결론

2016년 1월 다보스포럼에서 ‘4차 산업혁명’이란 용어가 처음 등장한 이후로 그 핵심기술인은 인공지능은 정말 빠른 속도로 우리에게 다가왔다. 중국 고전 문학의 정수라 할 수 있는 중국시도 이제는 한시를 짓는 AI가 등장하여 詩題만 입력하면 바로 시를 지을 수 있는 시대가 되었다.

이런 시대적 전환기에 AI에 관심을 가지고 중국시의 연구방향을 고찰해 보는 것은 시의적이며 꼭 필요한 작업이라 하겠다. 본고는 《당시삼백수》와 AI 시를 비교 분석하고 중국시 연구에 AI를 활용할 수 있는 방안을 고찰해 보았다. 먼저 기초 작업으로 텍스트로서의《당시삼백수》의 의미를 고찰하고, 《당시삼백수》전체 내용과 시제를 텍스트 마이닝 하여 가장 많이 쓰인 시어를 추출하였다. 여기서 추출된 주요 시어는 ‘人’, ‘山’, ‘天’,

11) 주5) 참고.

‘送’, ‘春’, ‘秋’였는데 이를 토대로 〈春思〉, 〈送友人〉, 〈山居秋暝〉 시를 선별하였다. 다음으로 인공지능의 알고리즘의 원리를 살펴보았다. 한시 짓는 AI는 2013년 바이두(Baidu)에서 SMT(통계기반 기계번역) 방식의 1.0버전으로 처음 등장하였고, 2016년에 2.0버전인 NMT(인공신경망 기반 기계번역) 방식의 度秘가 개발되어 본격적인 AI로 발전하였다. 이후 칭화대학에서 최신 딥 러닝 기술을 채택하고, 시 생성을 위해 특별히 설계된 여러 모델을 결합한 인공지능 九歌를 개발하였다. 본고에서는 앞서 선별된 시제를 키워드로 《당시삼백수》와 AI ‘度秘’와 ‘九歌’의 시를 비교 분석하였다. 중국시를 딥러닝해서 만들어진 AI는 시의 형식과 내용에서 무난한 시를 지을 수 있지만 고도의 예술기법을 갖춘 唐詩와는 아직 많은 격차가 있음을 알 수 있었다. 또한 AI로 특정인을 대상으로 시를 쓰기에 도 아직 한계가 있었다. 그러나 AI가 작시하는 원리를 역추적해 시를 분해해 보면 시 창작과 교학에 크게 도움이 될 수 있겠다는 가능성도 발견하였다.

중국 고전시는 오랫동안 축적되어 온 고도의 문학 창작물이라고 여기는 우리에게 AI 시는 많은 생각을 하게 만든다. AI가 창작한 시를 예술작품으로 볼 수 있을까? 또는 AI 시에서 인간의 감정을 기대할 수 있을까? 등 다양한 의문이 생길 것이다. 이런 상황에서 우리는 AI의 실체를 바로 알아 장단점을 파악하고 이를 활용할 수 있는 다양한 방안을 고찰하는 등 적극적인 자세가 필요한 것 같다. 이제 중국문학 분야도 4차 산업혁명이 도래하였음을 실감한다.

### < 참고문헌 >

- 류종목·주기평·이지운 옮김, 《唐詩三百首》, 서울: 소명문화사, 2010.  
 劉若愚 著, 李章佑 譯, 《中國詩學》, 서울:明文堂, 1994.  
 유종호, 《문학이란 무엇인가?》, 서울: 민음사, 1995.

- 유종호, 《시란 무엇인가?》, 서울: 민음사, 1995.
- 李炳漢·李永朱 共編, 《中國古典文學理論批評史》, 서울: 한국방송통신대학교출판부, 1988.
- 李炳漢·李永朱 譯解, 《唐詩選》, 서울: 서울대학교출판부, 1998.
- 임동석譯註, 《唐詩三百首》, 서울: 동서문화사, 2010.
- 陳伯海 著, 李鍾振 譯, 《당시학의 이해》, 서울: 사람과책, 2001.
- 강병규, 〈AI의 중국 古典詩歌 창작-시어의 학습과 생성〉, 《중국문학》 제 100집, 2019.
- 陳伯海, 〈文學傳統與唐詩的創新〉, 《中國古代·近代文學研究》, 1988.
- 〈詩歌只是一種算法? 人工智能寫詩出詩集〉, 華夏經緯網, 2017.
- Stephen Owen, “The end of the Chinese ‘Middle Ages’ – Essays in Mid-Tang Literary Culture”, Stanford University Press, 1996.
- Zhe Wang, Wei He, Hua Wu, Haiyang Wu, Wei Li, Haifeng Wang, Enhong Chen. 〈Chinese Poetry Generation with Planning based Neural Network〉, 《University of Science and Technology of China. Hefei》, 2016.
- 전당시 DB자료.

### <Abstract>

This paper was written to research Chinese poetry using artificial intelligence. Recently, as artificial intelligence algorithms based on big data have been widely used, various attempts have been made to utilize them in the creation of literary works. One thing that cannot be overlooked in this trend is the emergence of artificial intelligence that composes Chinese poetry. This paper was written to explore Chinese

poetry using artificial intelligence. For this purpose, the author conducted a study by comparing and analyzing 《Three Hundred Tang Poems》 and artificial intelligence poetry. 《Three Hundred Tang Poems》 is the most representative book among the historical collections of Tang poetry and was compiled as a learning material for Tang poetry for children. This paper includes 1. Introduction, 2. 《The Three Hundred Poems of the Times》 and artificial intelligence as a text, 3. Comparative analysis of 《The Three Hundred Poems of the Times》 and AI poetry, 4. Exploration of ways to utilize artificial intelligence in the study of Chinese poetry, 5. Conclusion.

Key Words : 인공지능(Artificial Intelligence), 구가(Jiuge), 뚜미(Dumi), 중국시 (Chinese Poetry), 당시삼백수(Three Hundred Tang Poems), 중국시 연구(Chinese Poetry Research), 텍스트 마이닝(Text mining), 인포그래픽(Infographics).