How Ai Works

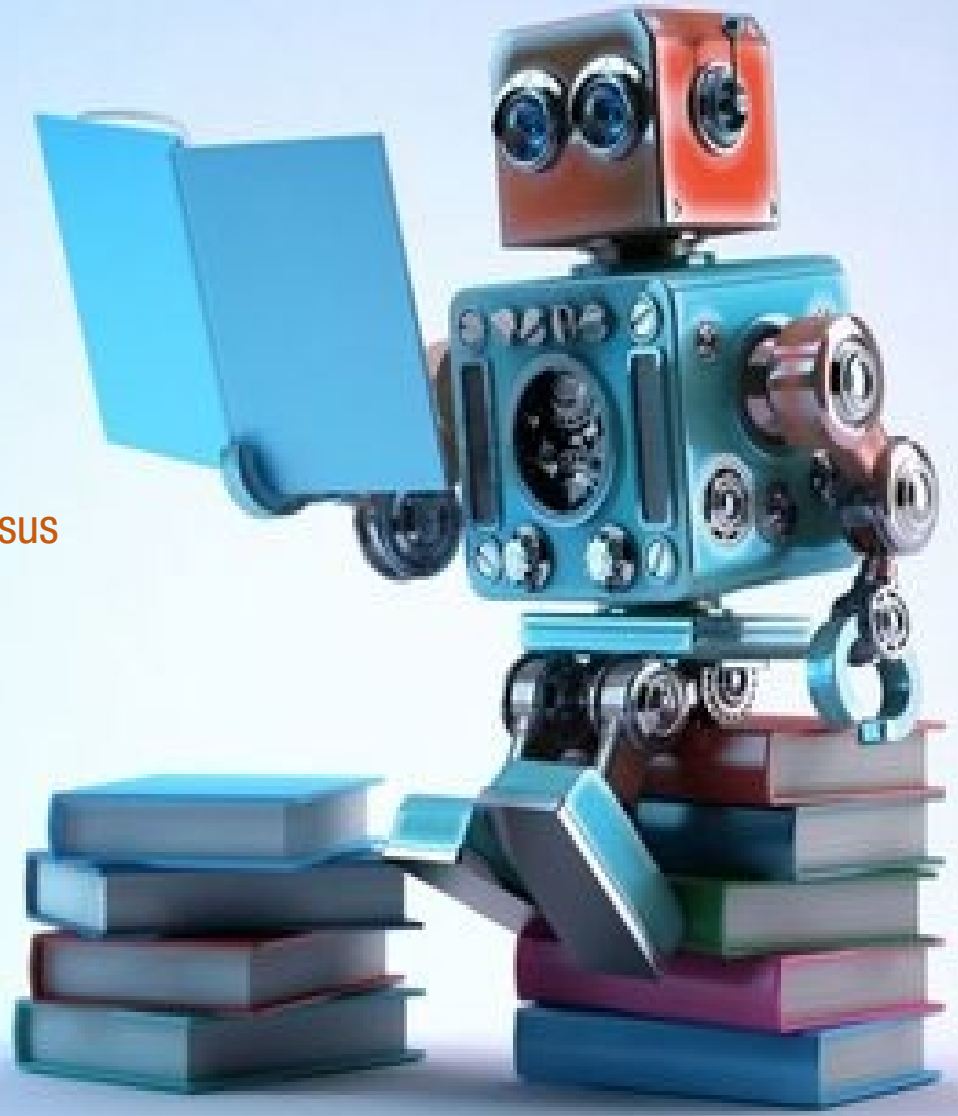
useful basics for librarians and researchers

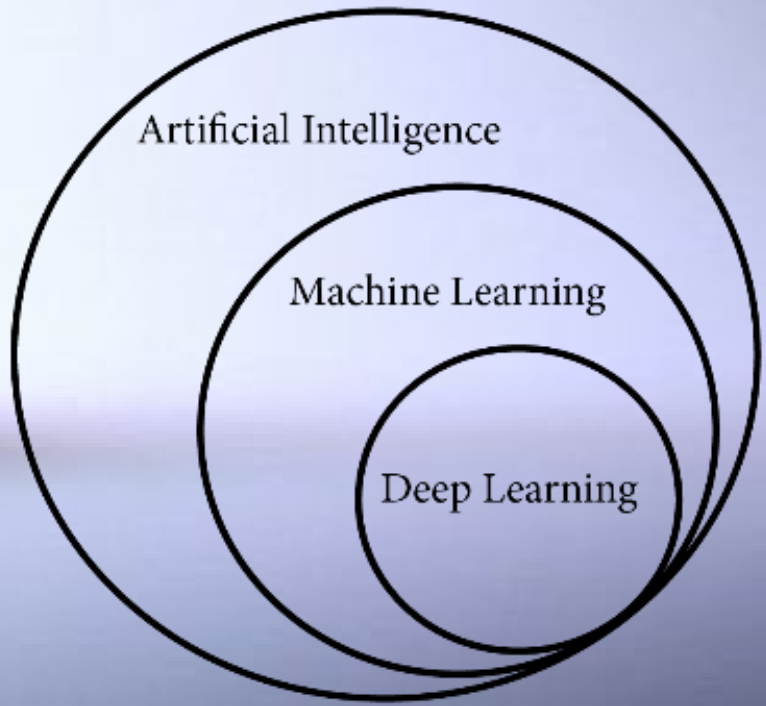
Wayne de Fremery
Dominican University of California

AI

overview

General versus Narrow





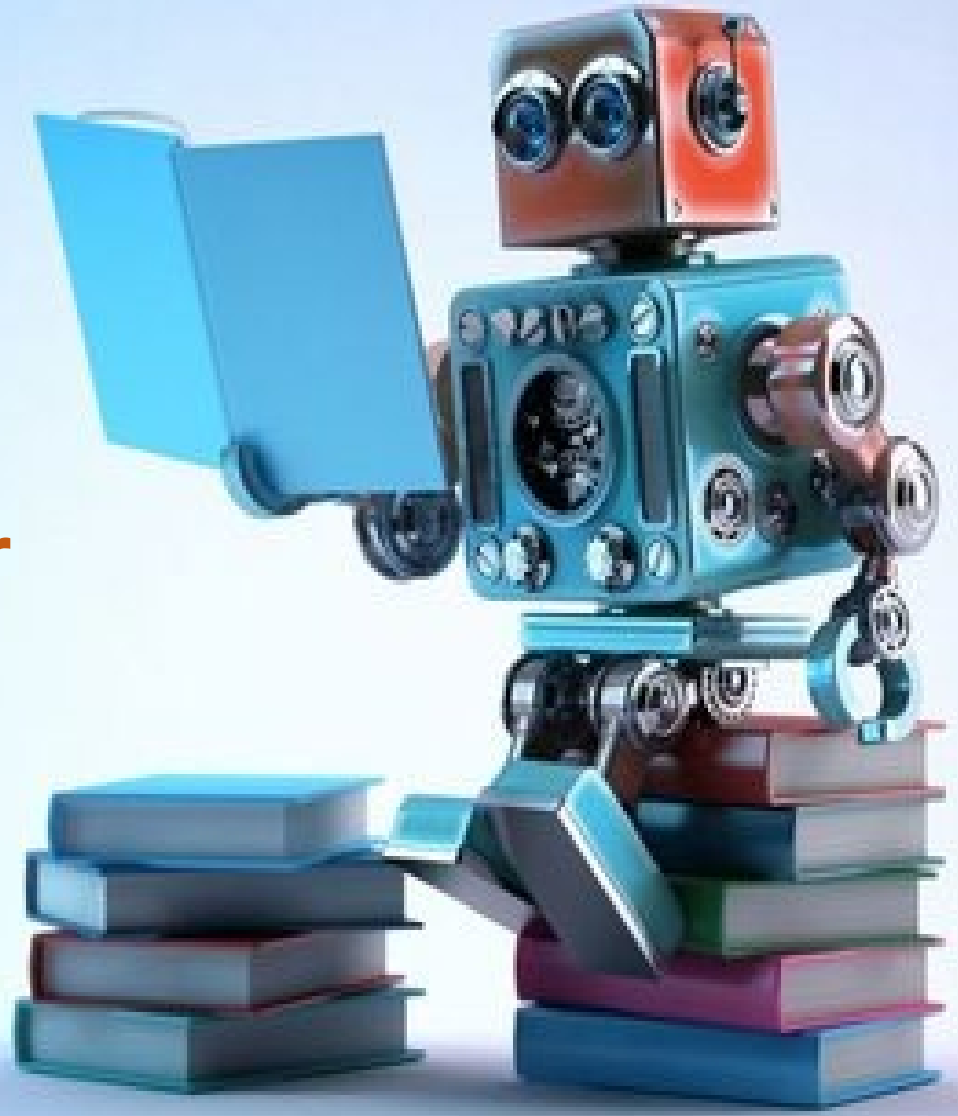
Kinds of Data & Tasks

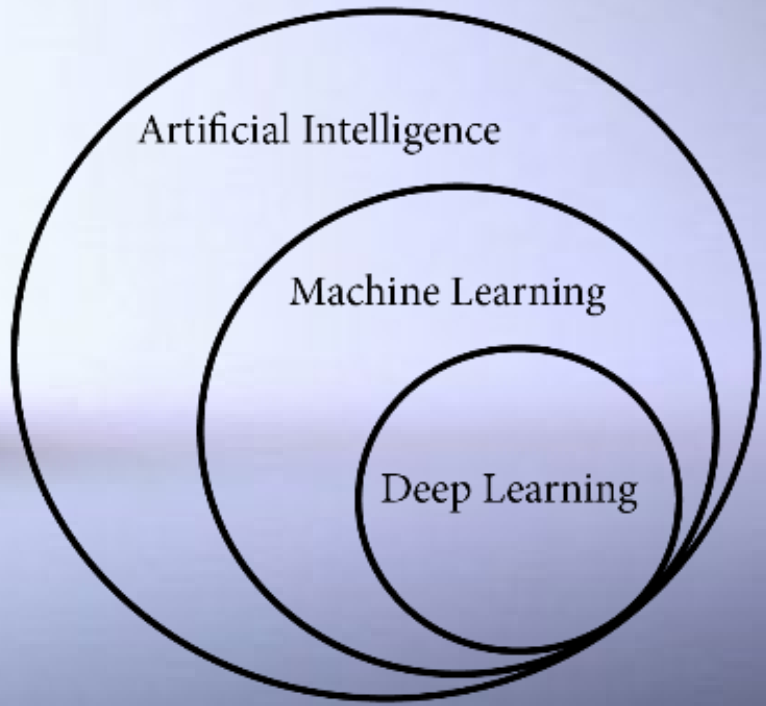
/

Kinds of Algorithms

AI & Deep Learning

Computer Vision

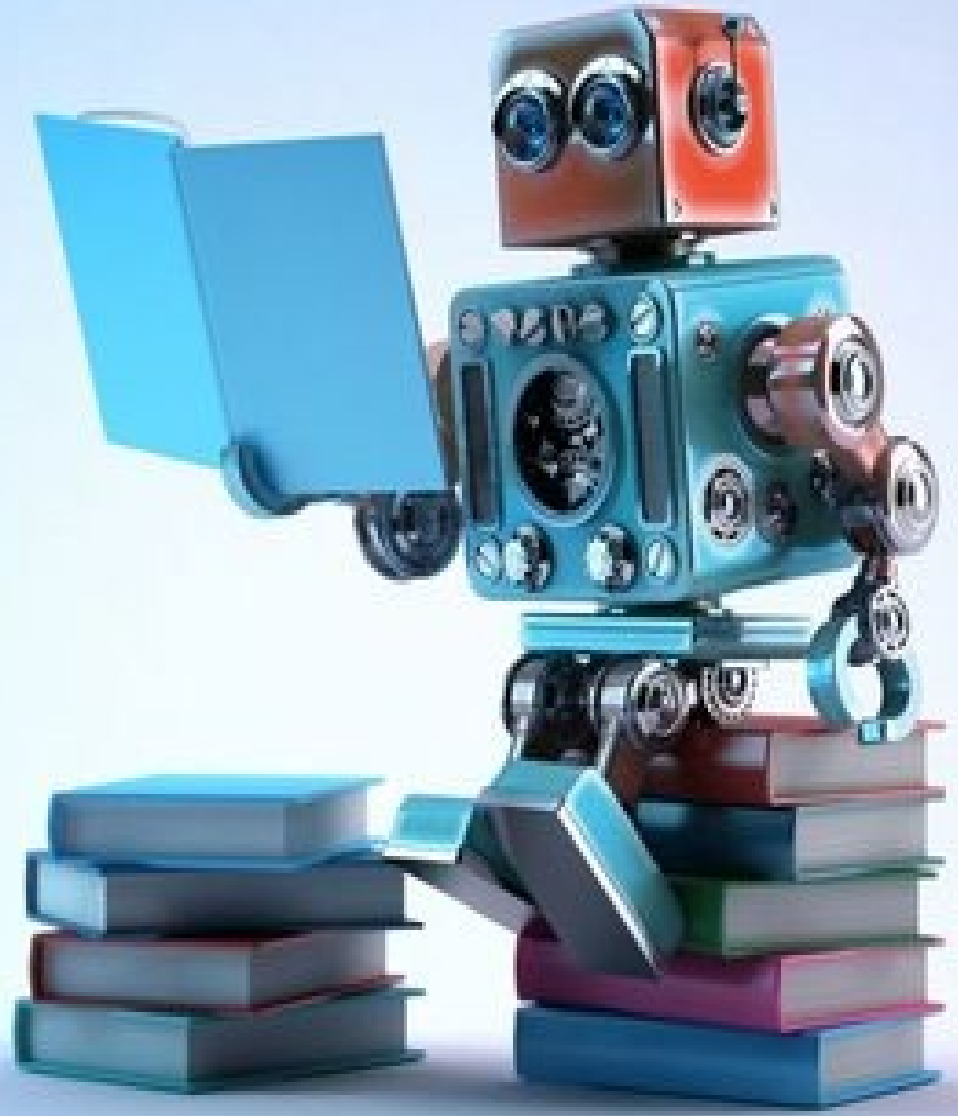


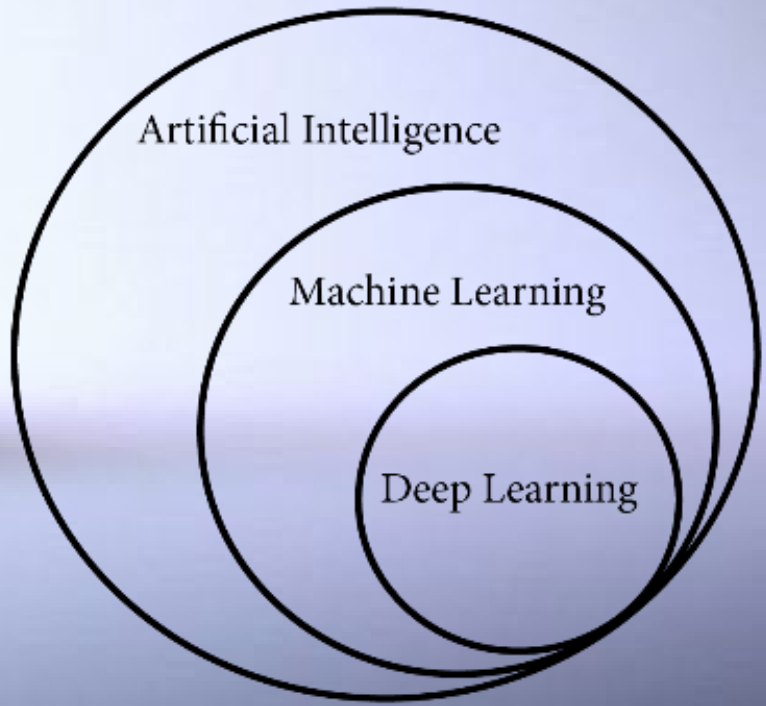


CNN (convolutional neural network)

AI & Deep Learning

(N)atural
(L)anguage
(P)rocessing





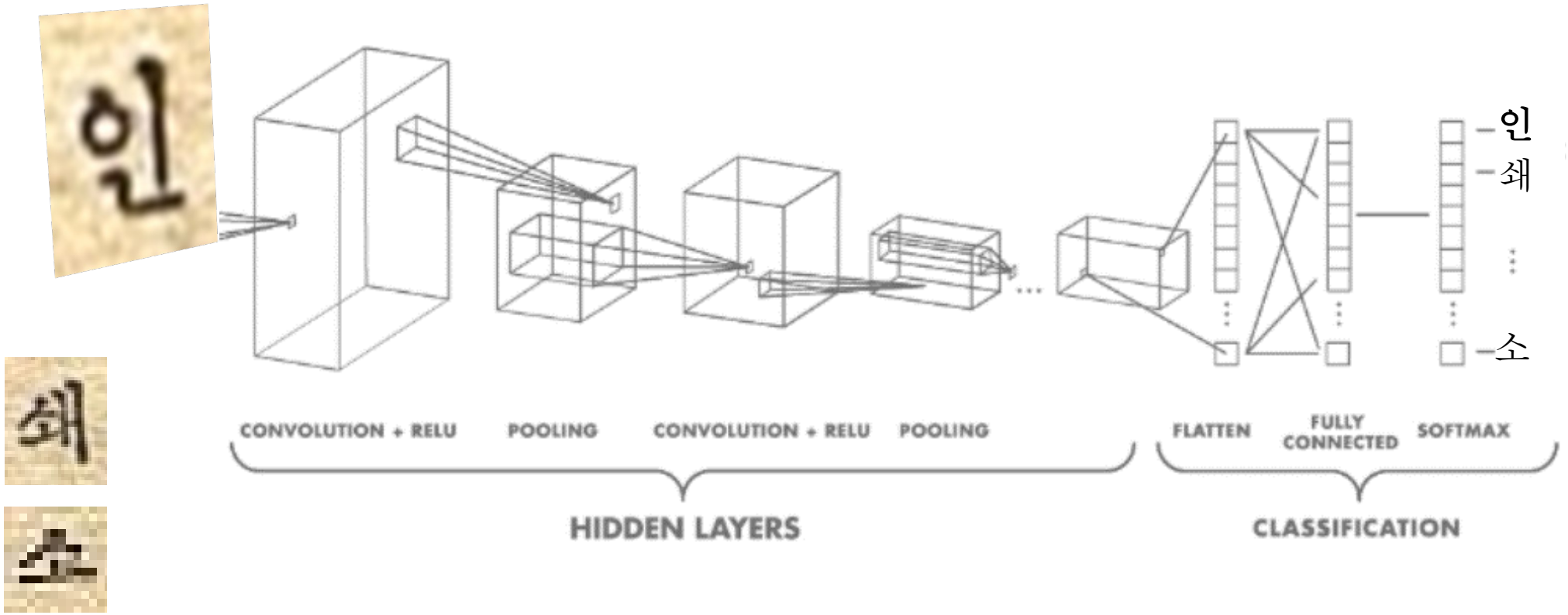
RNN (recurrent neural networks)
BERT (Bidirectional Encoder Representations from Transformers)
GPT (Generative Pretrained Transformers)

AI

Optical Character Recognition (OCR)

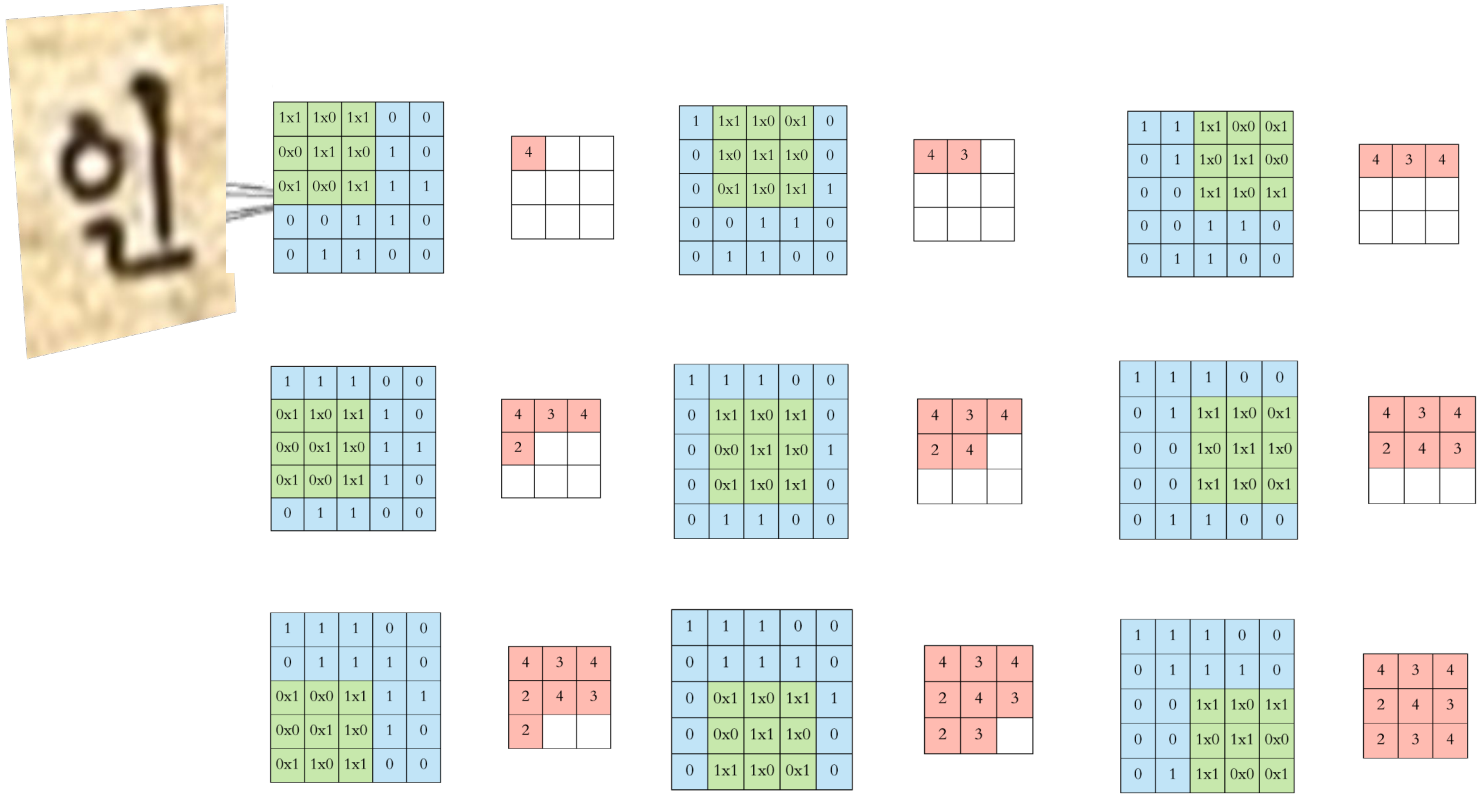
(using computer vision)

Deep Learning



convolutional neural network

Convolution



1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Input

plus

1	0	1
0	1	0
1	0	1

Filter / Kernel

Descriptive features

4	3	4
2	4	3
2	3	4

Feature (map)

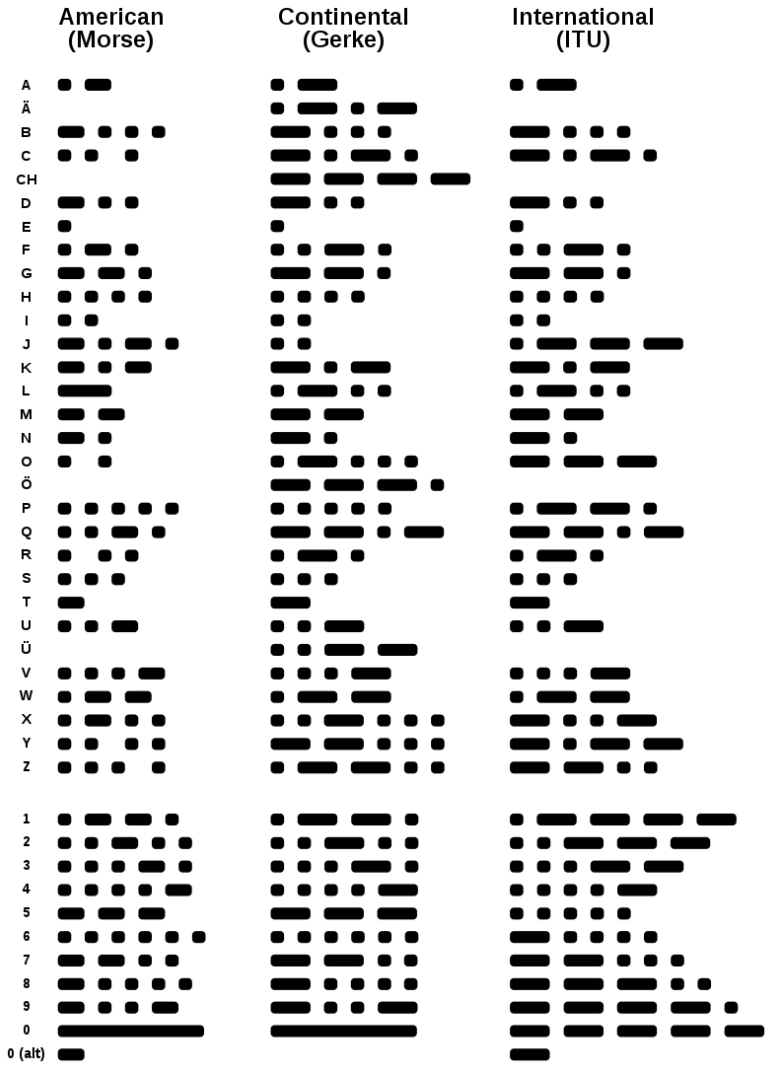
target

인

Text (Character) Encoding

Morse Code

(American, 1844; Continental, 1848; International, 1865)



Bits $b_7 b_6 b_5$ →	000	001	010	011	100	101	110	111
b_4 b_3 b_2 b_1 ↓ Row / Column →	0	1	2	3	4	5	6	7
0 0 0 0	0	NUL	DLE	SP	0	@	P	`	p
0 0 0 1	1	SOH	DC1	!	1	A	Q	a	q
0 0 1 0	2	STX	DC2	"	2	B	R	b	r
0 0 1 1	3	ETX	DC3	#	3	C	S	c	s
0 1 0 0	4	EOT	DC4	$	4	D	T	d	t
0 1 0 1	5	ENQ	NAK	%	5	E	U	e	u
0 1 1 0	6	ACK	SYN	&	6	F	V	f	v
0 1 1 1	7	BEL	ETB	'	7	G	W	g	w
1 0 0 0	8	BS	CAN	(	8	H	X	h	x
1 0 0 1	9	HT	EM	)	9	I	Y	i	y
1 0 1 0	10	LF	SUB	*	:	J	Z	j	z
1 0 1 1	11	VT	ESC	+	;	K	[	k	{
1 1 0 0	12	FF	FS	,	<	L	\	l	|
1 1 0 1	13	CR	GS	—	=	M	]	m	}
1 1 1 0	14	SO	RS	.	>	N	^	n	~
1 1 1 1	15	SI	US	/	?	O	_	o	DEL

American Standard Code for Information Interchange

(1960s)

Text (Character) Encoding

American Standard Code for Information Interchange

Binary	Oct	Dec	Hex	Abbreviation			Unicode Control Pictures[b]	Caret notation[c]	C escape sequence[d]	Name (1967)
				1963	1965	1967				
000 0000	000	0	00	NULL	NUL		NUL	^@	\0	Null
000 0001	001	1	01	SOM	SOH		SOH	^A		Start of Heading
000 0010	002	2	02	EOA	STX		STX	^B		Start of Text
000 0011	003	3	03	EOM	ETX		ETX	^C		End of Text
000 0100	004	4	04	EOT			EOT	^D		End of Transmission
000 1010	012	10	0A	LF			LF	^J	\n	Line Feed
000 1011	013	11	0B	VTAB	VT		VT	^K	\v	Vertical Tab
000 1100	014	12	0C	FF			FF	^L	\f	Form Feed
000 1101	015	13	0D	CR			CR	^M	\r	Carriage Return[h]

Binary	Oct	Dec	Hex	Glyph		
				1963	1965	1967
010 0000	040	32	20	space		
011 0000	060	48	30	0		
100 0001	101	65	41	A		

10 = line feed

13 = carriage return

32 = space

48 = 0

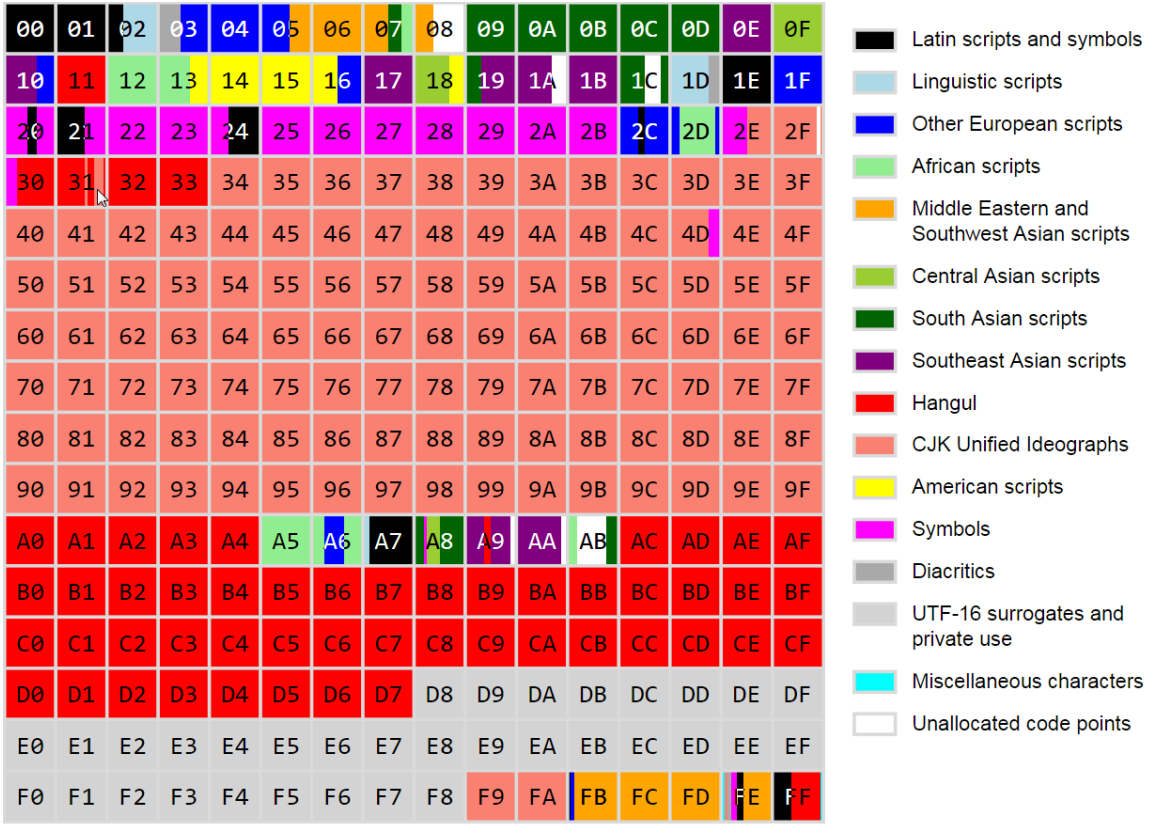
65 = A

? = 인

Unicode

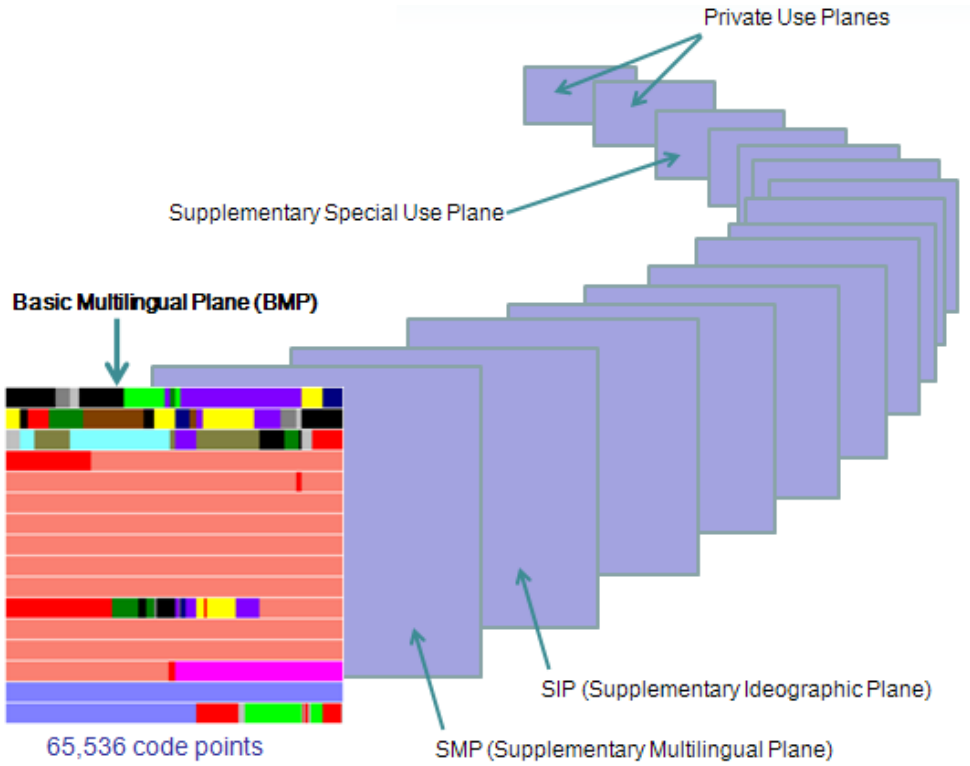
Unicode Consortium 501(C)3 Mountain View CA

A universal character encoding scheme
initiated in 1987. The Unicode Consortium was
incorporated in California 1991.



Sources:
(above) http://upload.wikimedia.org/wikipedia/commons/8/8e/Roadmap_to_Unicode_BMP.svg (modified by author);
(right) http://www.w3.org/International/articles/definitions-characters/

Unicode Code Spaces



Digital Fonts

Glyphs mapped to encoding

65 = 𝒜

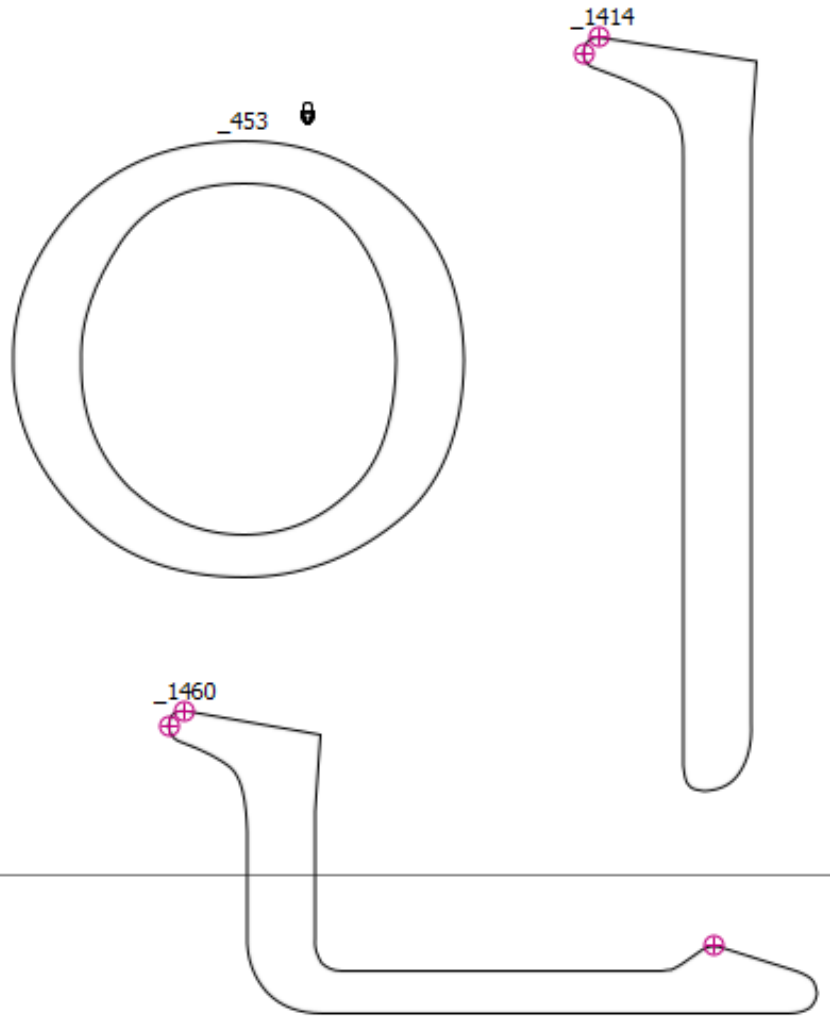
65 = 𝓐

65 = 𝔄

65 = 𝕬

65 = A

Digital Fonts



236 157 184 = 인

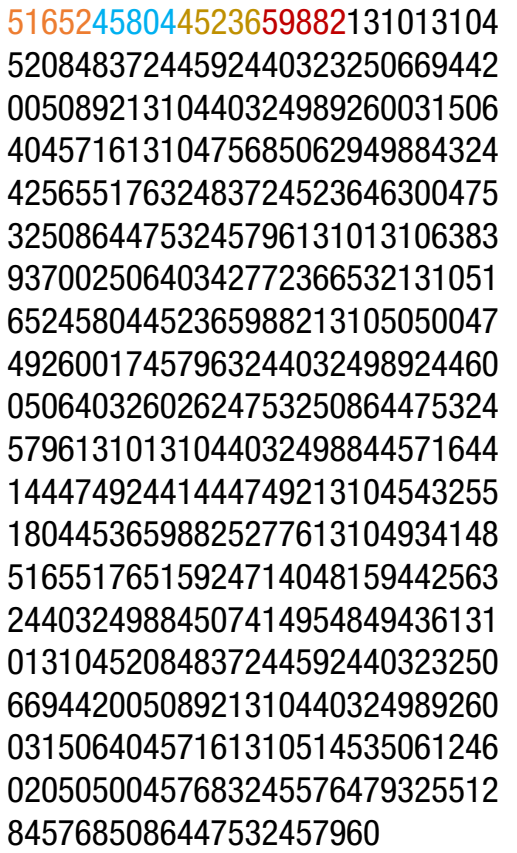
236 157 184 = 인

236 157 184 = 인

236 157 184 = 인

Glyphs mapped to encoding

Text Encoding



Haansoft Batang

진달내옷

나보기가 역겨워
가실쌔에는
말업시 고히 보내드리우리다

寧邊에藥山
진달내옷
아름싸다 가실길에 쌰리우리다

가시는거름거름
노힌그옷츨
삽분히즈려밟고 가시옵소서

나보기가 역겨워
가실쌔에는
죽어도아니 눈물흘니우리다

맞다 틀리다

진달내□

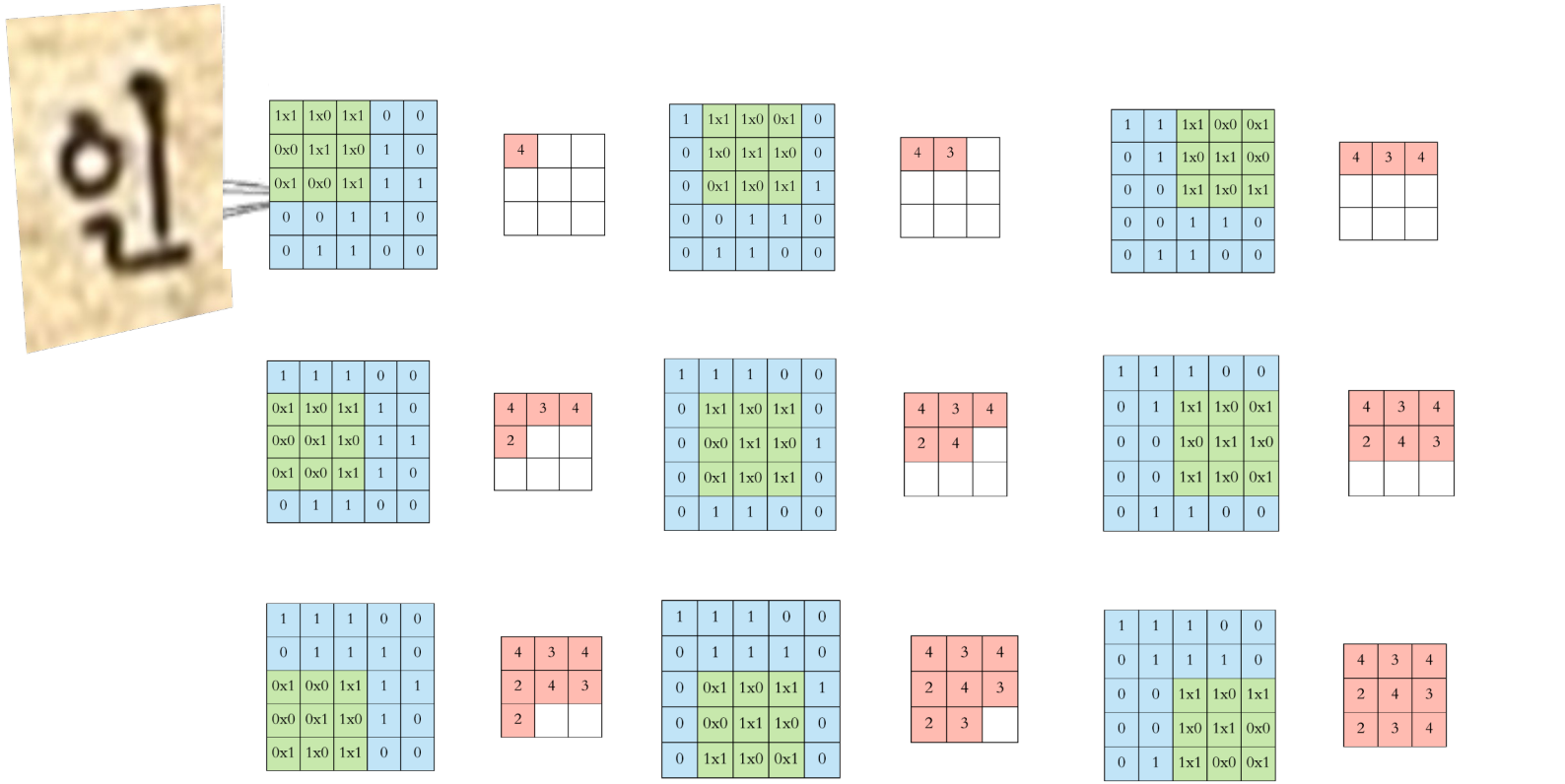
나보기가 역겨워
가실□에는
말업시 고히 보내드리우리다

寧邊에藥山
진달내□
아름□다 가실길에 □리우리다

가시는거름거름
노힌그□츨
삽분히즈려밟고 가시옵소서

나보기가 역겨워
가실□에는
죽어도아니 눈물흘니우리다

Convolution



1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Input

plus

1	0	1
0	1	0
1	0	1

Filter / Kernel

Descriptive features

4	3	4
2	4	3
2	3	4

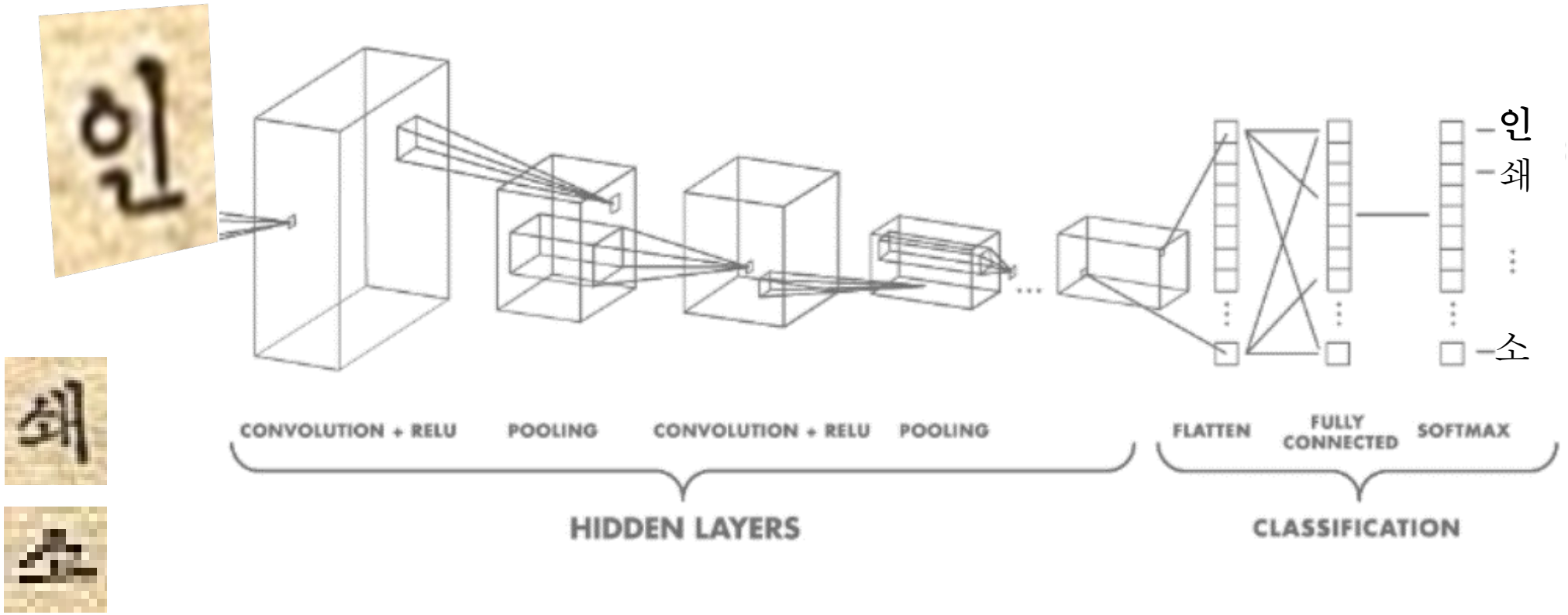
Feature (map)

target

236 157 184

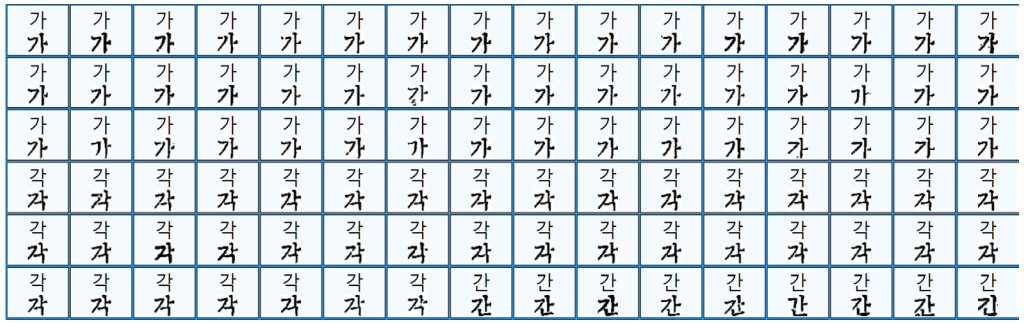
(인)

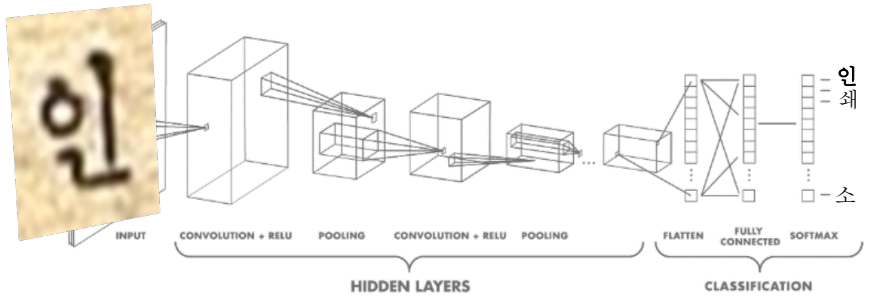
Deep Learning



convolutional neural network

CNNs & Inference





Weight File

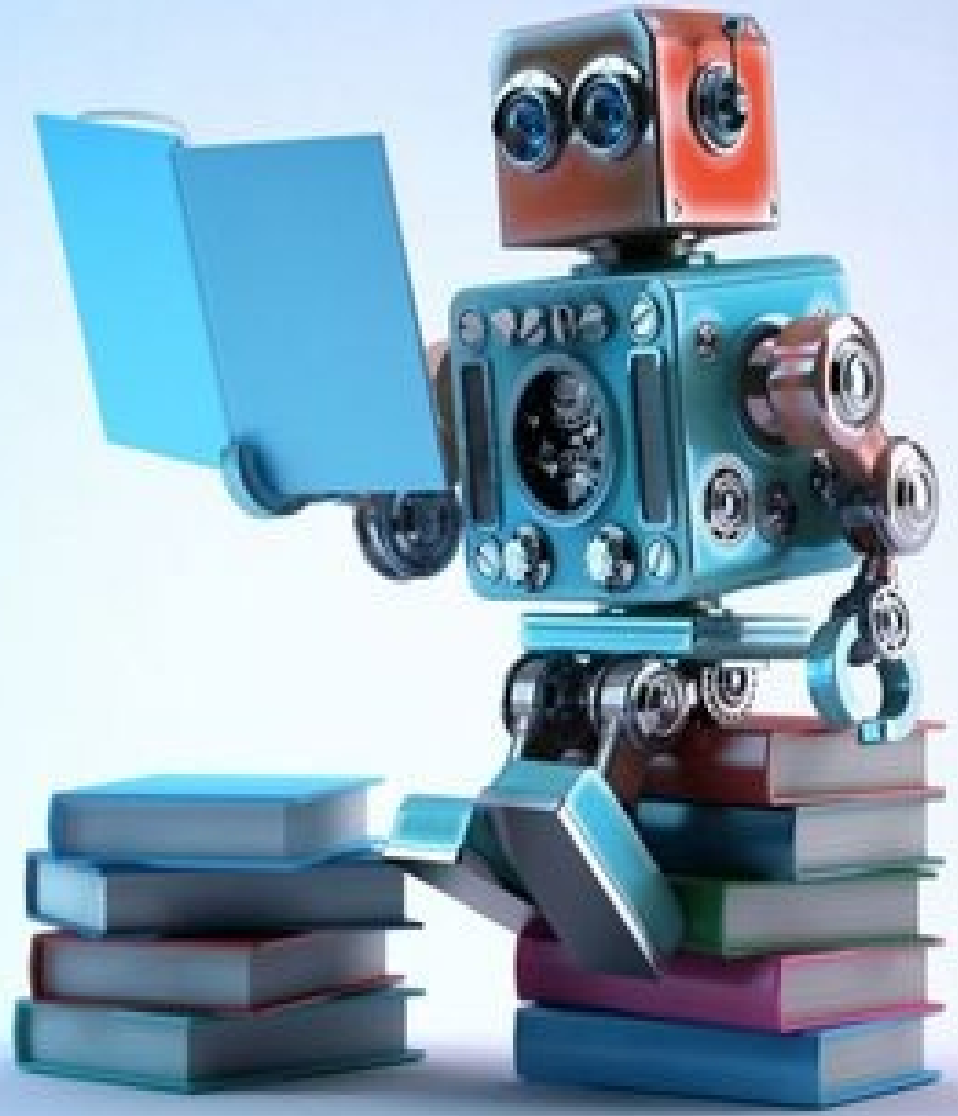
(record of which features best predict/ describe objects in training set)

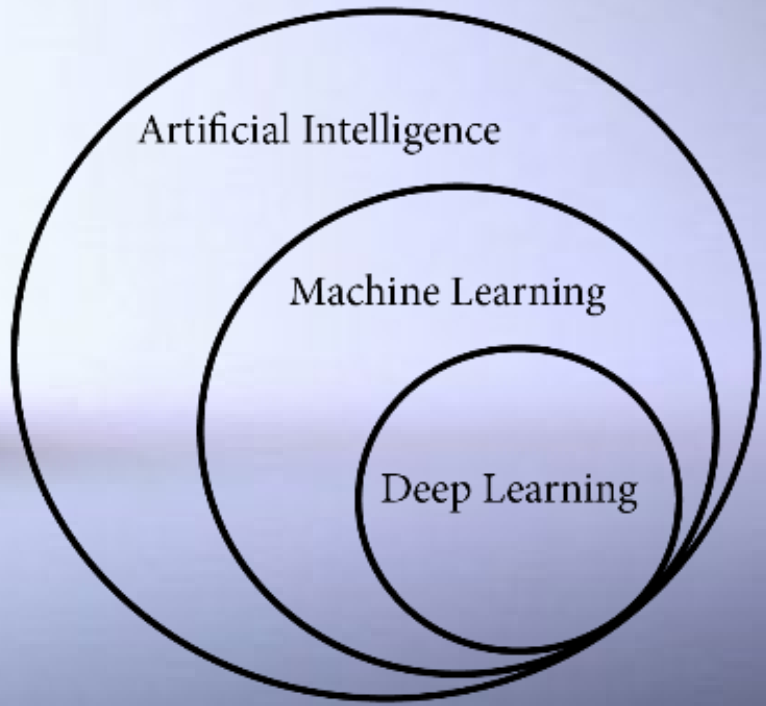
Something.weights

인

236 157 184

(N)atural
(L)anguage
(P)rocessing

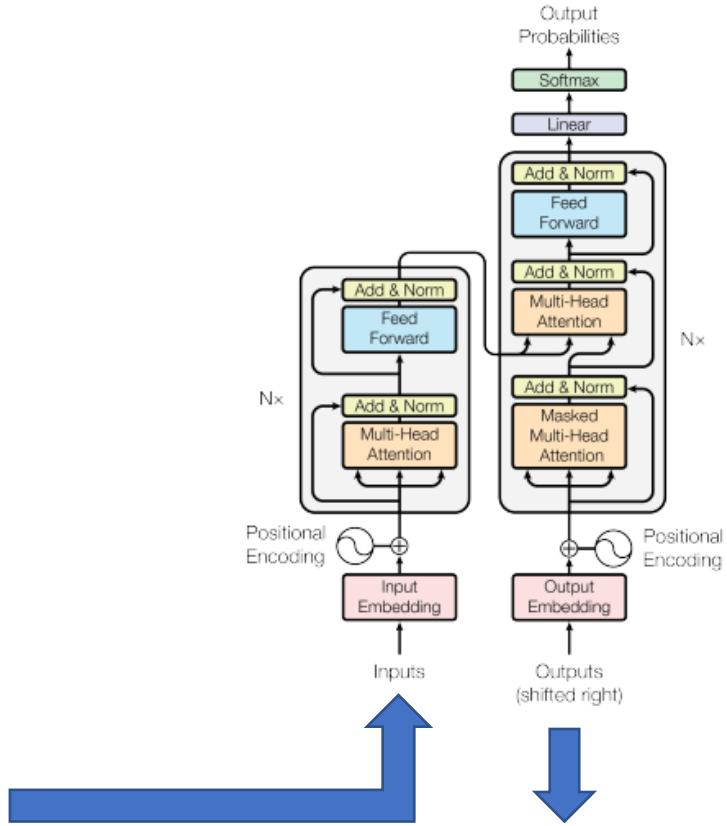




(G)enerative
(P)retrained
(T)ransformer

<?xml version="1.0" encoding="UTF-8"?>
<topics>
<topic>
<perio>로동신문</perio>
<date>1950년 6월 25일</date>
<side>3</side>
<ltitle></ltitle>
<part>보도기사</part>
<field></field>
<region></region>
<title>철도수송력의 강화 위해 기관차실동률 제고에!</title>
<stitle>신북청기관구 로동자들</stitle>
<text>
철도수송력의 강화위해 기관차실동률 제고에！ 신북청기관구 로동자들 신북청기관구 로동자들은 8.15해방 5주년을 철도수송력의강화로써 뜻깊게 맞이하고자 기관차 수리의 질적제고를 위한 투쟁을 가강히 전개하고있다. 그들은 우선 기관차의 질적수리제고를 위하여 낡은것을 용감히 퇴치하고 새로운 작업방법을 채용하는데 창발적노력을 기울이고있다. 수리직장로동자들은 40여종에 달하는 수리공구들을 품종별로재검토한후 마모되고 위본손상된공구들을 전부 자체제작하여 새것으로 보충하였으며 수리작업의 중요부속품인 각종바네등 종전에는 철 . 특히 리시모동무의 3분선반 쟉기수리개조와 신문식동무의 6척선반 쟉기 기고있다. 이곳 로동자들은 사업과정에서 일상적으로 호상 련계를 긴밀히하다 생산협의회의 역할을 높이여 성능이 좋치못한 기관차를 성능이 우수한기관 면 파시너 5호 미가 44호등은 종전에는 자체성능을 70%밖에 발휘하 였으며 변조정이 규격대로되여있지않다는 로동자들의개별적 의견들을 종합 를 정시로 견인하고있다. 특히 이곳 로동자들은 지난 5월중에는 을동휴차로 리하여 재생시킴에 성공하였으며 국가에 막대한 리익을 주었다. 그들은 푸렛치 4호 재생을 위하여 정기적으로 되는 생산협의회에서 이의 재생에대한 구체적 방법들을 충분히 토의한후 매개 로동자들의 기능수준에 엄격히 기초하여 분공을 정확히 조직하였다. 이리하여 6월에 들어서면서부터 기관차의 실동률을 부단히 높이고있는바 현재 기대당 실동률은 1.4분기에 비하여 7%나 더 제고시키였다. 이곳 로동자들은 기관차수리의 질적 제고에 더한층 창발적 노력을 기울임으로써 년간주행계획을 자기들이 맹세한 10월15일까지에 반드시 초과완수할 굳센 결의밑에 계속 힘차게 싸우고있다. 통신원 윤지월</text>
<author>통신원 윤지월</author>
</topic>
</topics>

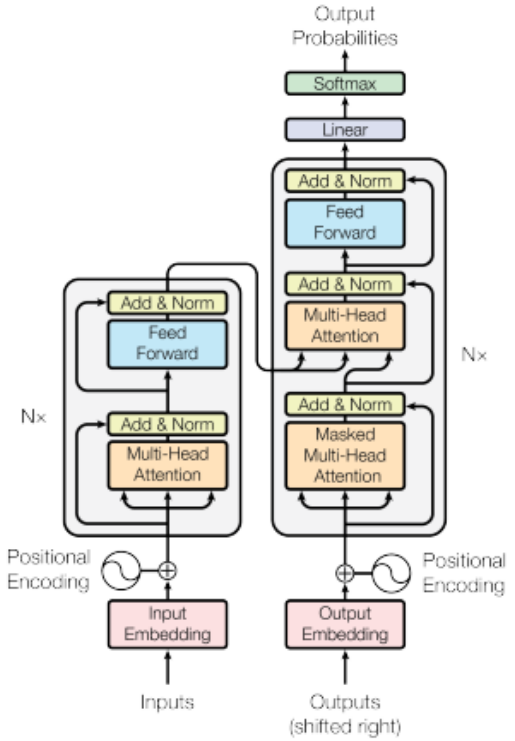
Vast amounts of historical, encoded text as input



<?xml version="1.0" encoding="UTF-8"?…

<철도수송력의 강화위해 기관차실동률 제고에！ …

<?xml version="1.0" encoding="UTF-8"?>

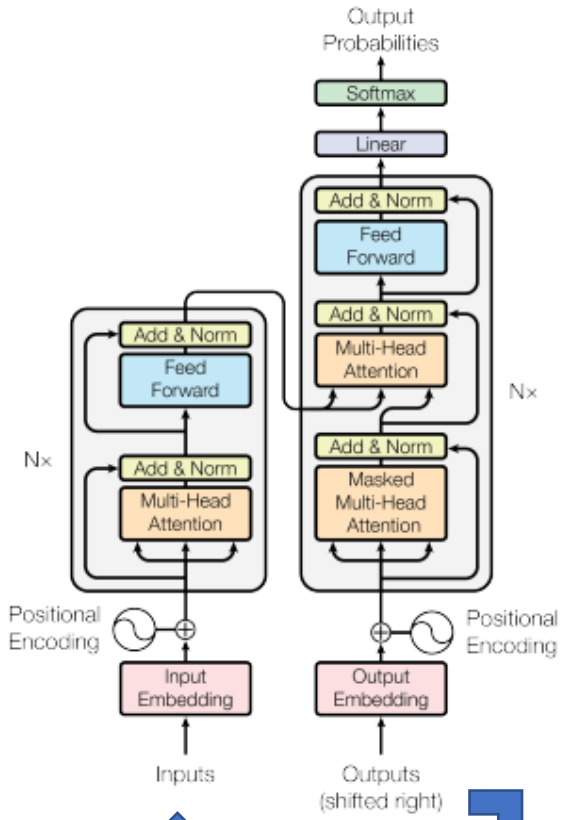


(G)enerative
(P)retrained
(T)ransformer

<?xml version

="1.0" encoding="UTF-8"?>

철도수송력의 강화위해 기관차실동률 제고에 !



(G)enerative
(P)retrained
(T)ransformer

철도수송력의

강화위해 기관차실동률 제고에 !

Optical Character Recognition (OCR)

Computer
Vision

Optical Character Recognition (OCR)

Computer Vision

(+)

Natural Language Processing

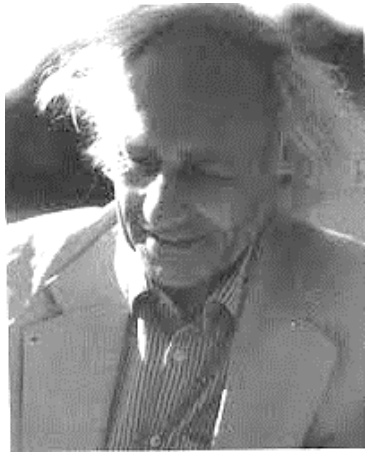
AI

a bibliographic process

copies &

accounts of difference

“Bibliography is the discipline that studies texts as recorded forms, and the processes of their transmission, including their production and reception.”



D. F. McKenzie. *Bibliography and the Sociology of Texts*

“The domain of information science is the transmission of the universe of human knowledge in recorded form, centering on manipulation (representation, organization, and retrieval) of information, rather than knowing information”

Association for Information Science and Technology website, “What is information science?” accessed February 3, 2022, https://www.asist.org/about/what-is-information-science/.

— Tefko Saracevic. (2009). Information science. In M.J. Bates (Ed.), *Encyclopedia of library and information sciences* (3rd ed.) (pp. 2570–2585). New York: Taylor and Francis.

Bibliography

Βιβλιογραφία book-writing

the writing out of books

書 글 서　　誌 기록할 지　　學 배울 학

writing-record

recording/ documenting writing / study of

Probably Japanese neologism. *Chosŏn ilbo* makes reference to a new Japanese journal called <<書誌學>> on January 25, 1933.

Bibliography

“List Makers”

Enumerative

Descriptive

Information Science

“Studiers of Texts”

Analytical

Critical

Humanities

Patrick Wilson *Two Kinds of Power: An Essay on Bibliographical Control* (1968)

Bibliography

Enumeration

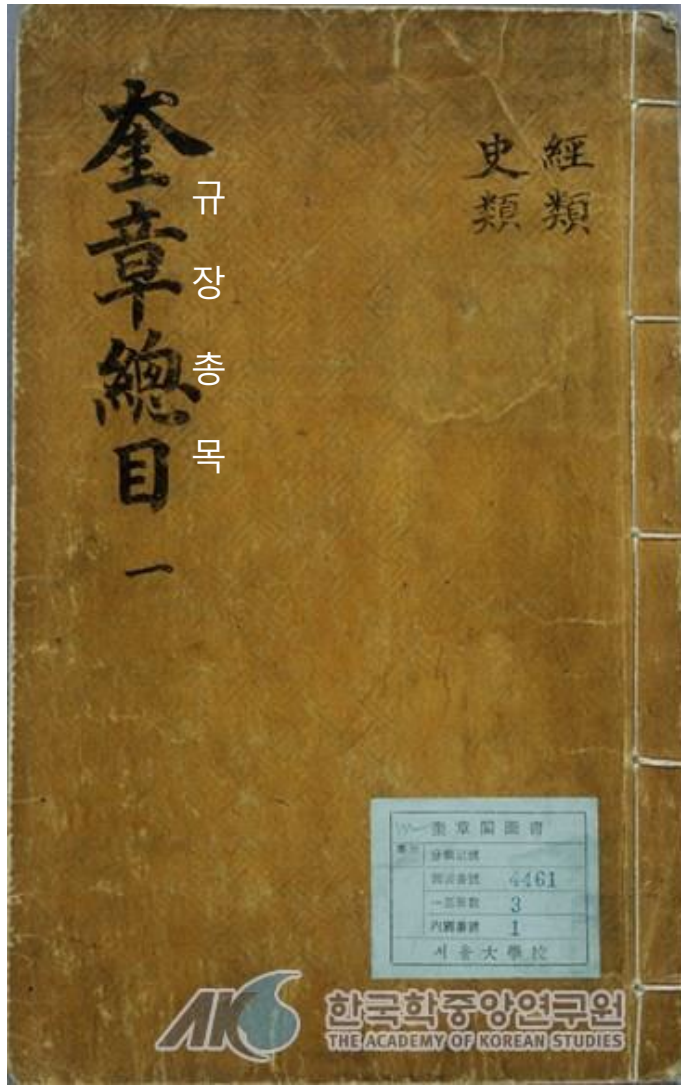
Description

Analysis

Critique

of

Copies



目錄學

Enumerative Bibliography

예위밧디잔봄
黑房秘曲
응마의鮮朝
無窮花
Gitanjali.
苦痛의 束縛
青年詩人百人集
頌然自

Descriptive
Bibliography

How Poetry Mattered in 1920s Korea

1. ***Onoe ŭi mudo***
(Dance of anguish)

Author(s):
Paul Verlaine,
Remy de Gourmont,
Albert Samain,
Charles Baudelaire,
William Butler Yeats, et al.

Translator:
Kim Ŏk

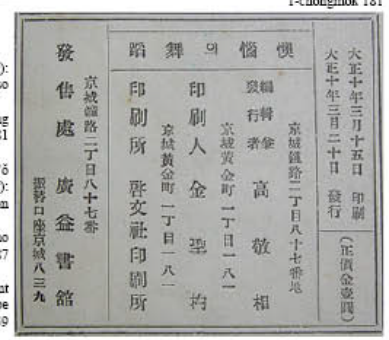
(In the Hwabong collection)

Dimensions (w x h in cm): 13.8 x 19.8; spine 7.4 (mm); 4.6-*p'anhyŏng*; proportions (width to height): 1:1.435
Cover materials: coated card (thickness: .34 mm)
Color(s): three (magenta, orange, green)
Image: flowers and musical staff

[1] Except when I discusss *Chindallaekkot* I have used the term "edition" as it is used in Korean bibliographic practice to suggest *p'anbon* 版本. Please see my discussion of "*p'anbon*," "edition," "issue," and "state" in Chapter Five.

Descriptive Bibliography

Onoe ŭi mudo title page



Onoe ŭi mudo
Colophon

Onoe ŭi mudo

inswae (printing): March 15, 1921
parhaeng (release date): March 20, 1921

price: 1 *wŏn*

p'yŏnjip kyŏm/ parhaengja (editor/ publisher): Ko Kyŏng-sang

Kyŏngsŏng Chongno 2-chŏngmok 181

inswaein (printer): Kim Sŏng-p'yo

Kyŏngsŏng Hwanggŭmjŏng 1-chŏngmok 181

inswaeso (place of printing): Kyenmusa Inswaeso

Kyŏngsŏng Hwanggŭmjŏng 1-chŏngmok 181

palsuch'ŏ (publisher/distributor): Kwangik Sŏgwan

Kyŏngsŏng Chongno 2-chŏngmok 87

chinch'e kujwa (account number to which money can be sent): Kyŏngsŏng 839

Onoe ŭi mudo
Bibliographic Notes

Number of Pages—174; extra sheets: (+1 front endsheet/ +1 tissue paper/ +1 title page/ +1 authors names/ +1 half title)

Endsheets Notes
_two different kinds.
a) .07 mm feels coated.
b) .03 mm tissue paper.

Title Page and Front Matter Notes
_title page:
single color.
3 stamps.
heavy stock (.12 mm).
relief printed, can feel.
impressions of type.

_sheet with authors names:
heavy stock (.13 mm), but does not feel coated.

_half title page:
same as body paper; feels slightly coated.

Notes on Margins
–

Paper Notes
_appears that many different kinds of paper were used: pp. 3-18 low opacity (.36 mm); pp. 19-34 more opaque (.37 mm); pp. 35-end, different paper.
_no chain lines.
_glue holds pg. 18 to pg. 19 at top and bottom.
_fibers clearly visible.

Notes on Typefaces
_pg. 3 title face *O* 懊 .5 cm x .5 cm; *sam* 삼, the first syllable in

Paper (body)
1 sheet .04 mm
8 sheet .63 mm

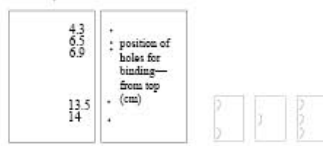
Margins (in cm)

Page No.	Top	Bottom	Outside	Gutter
pg. 3	2.2	1.9 to folio	2 to dates	
pg. 6	2.6 to text; 4.4 to title	1.5 to folio; 2.3 to text	2 to title; 3.5 to text	1.7
pg. 15	3.4 to text; 4.5 to title	1.6 to folio	2.1	1.7

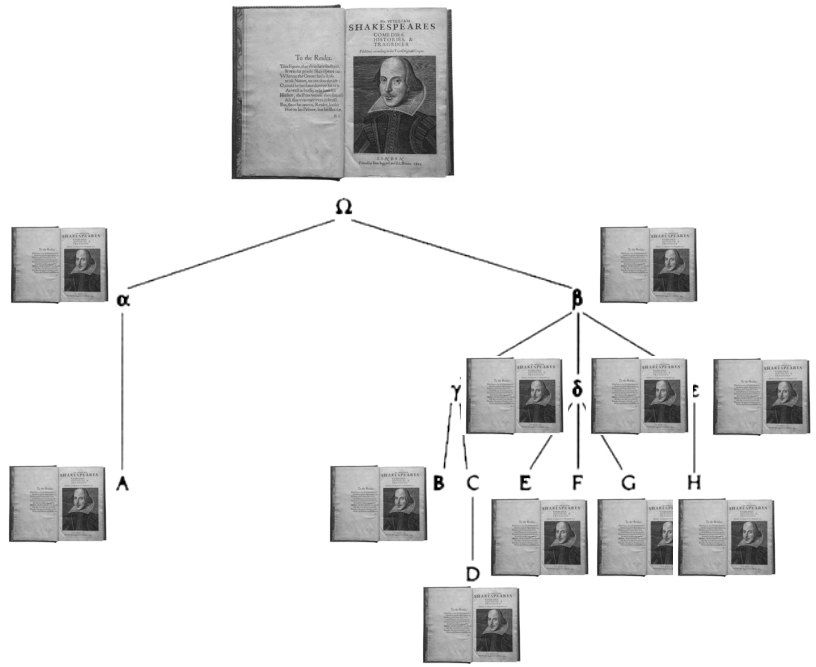
the body face is .3 cm x .3 cm.
_on pg. 15 *ka* 가 in title case is .4 cm x .4 cm; *ka* 가 in first line of body is .3 cm x .3 cm.

General Notes
_someone had fun "correcting" the text with a red pen. On pg. 19 *pul* 불 is annotated with *put* 붑.

Binding Notes
_*yangjang*.
_paper over boards; endsheets glued to cover stock.
_sewn so that string is visible between pp. 10-11, 26-27, 42-43, 58-59, 74-75, 90-91, 106-107, 122-123, 138-139, 154-155, 170-171.



New Bibliography / Analytical Bibliography

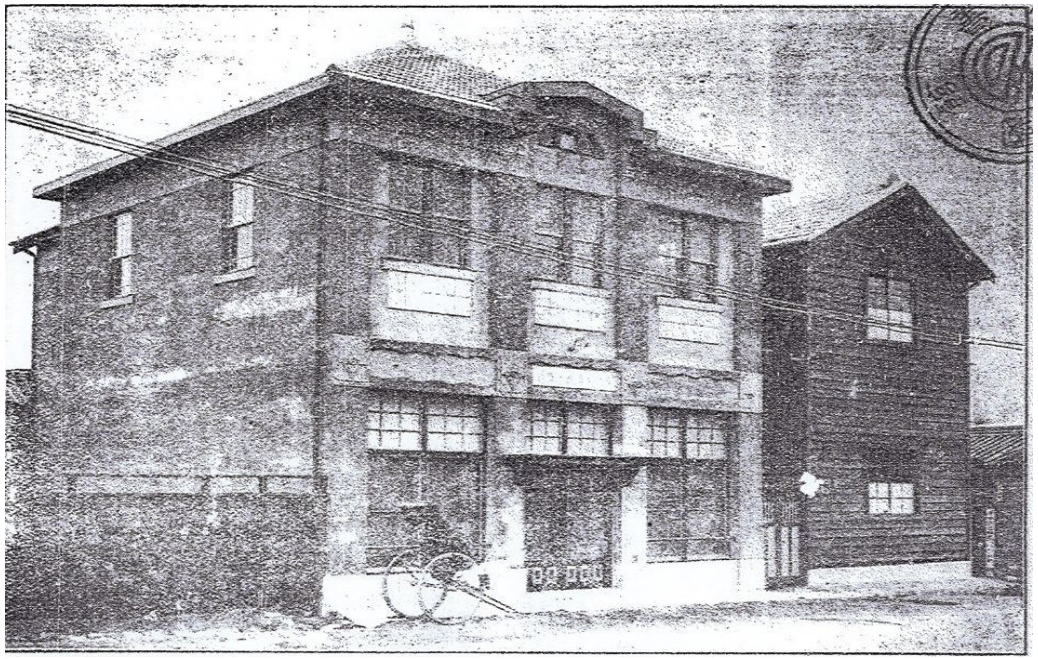


"Greg's authority [was] … rooted in classical and medievalist methodologies of textual criticism[; it] pivoted on the organicist historicism of stemmatics."

The Cambridge Companion to Textual Scholarship (Cambridge Companions to Literature) (Page 79). Cambridge University Press. Kindle Edition.

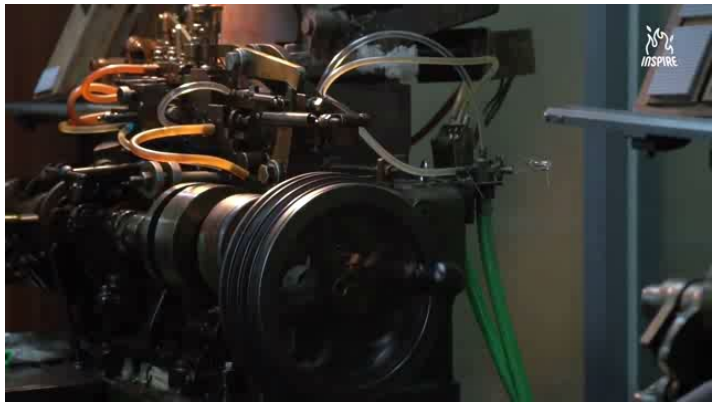
[Books] are products of a particular manufacturing or technical process, and the study of that process for printed texts is usually called *analytical bibliography*."

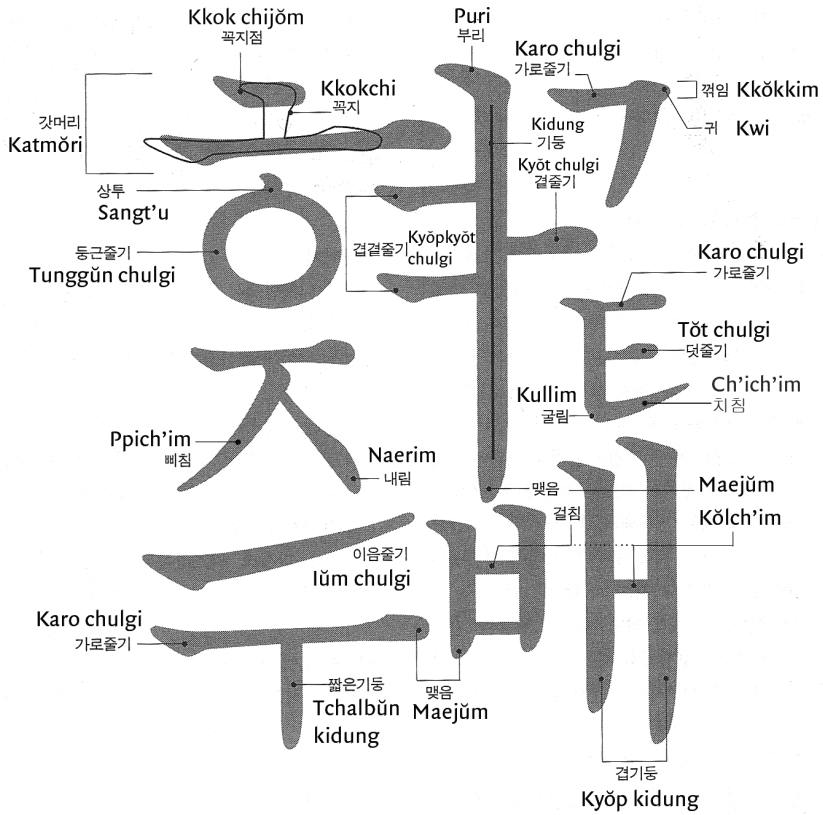
Greetham, David. *Textual Scholarship: An Introduction* (Garland Reference Library of the Humanities) (Kindle Locations 438-441). Taylor and Francis. Kindle Edition.



Newly built Hansŏng Tosŏ Chusik Hoesa building, *Haksaenggye* (December 1920), unnumbered front matter (image from microfilm at the National Library of South Korea). The work of establishing Hansŏng Tosŏ's printing facility was completed in April of 1921 and the first materials were printed there in May by No Ki-jŏng. Before May of 1921, materials published by Hansŏng Tosŏ were printed at Ch'oe Nam-sŏn's Sinmun'gwan.

Analytical Bibliography





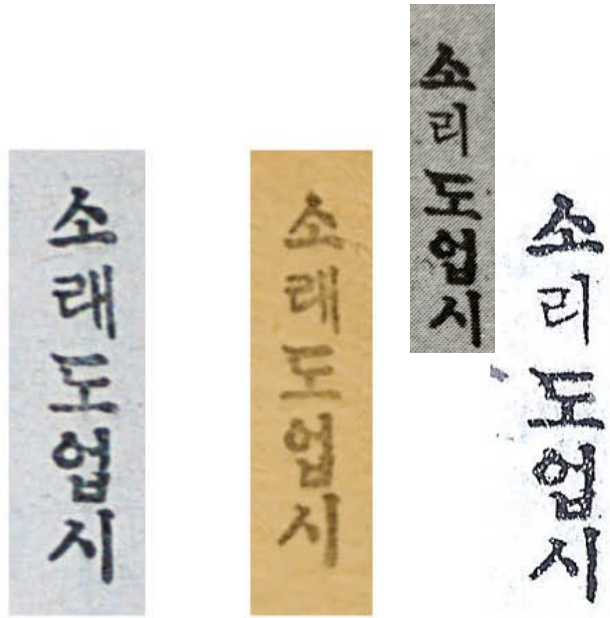
Source: Han Chae-jun 한재준, ed., *Han'gŭl kŭlkkol yongŏ sajŏn* 한글글꼴용어사전 (A dictionary of typographic terms for *han'gŭl*) (Seoul: Sejong Taewang Kinyŏm Saŏphoe, 2000), 46.

Critical Bibliography (Textual Studies/ Critical Editing)

The Covers of the Hansŏng Tosŏ issue (left) and the Chungang Sŏrim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Han'guk Hyŏndaesi Pangmulgwan), respectively.

Spine of the Hansŏng Tosŏ issue (left) and the Chungang Sŏrim issue (below).

Sources: Images of Hansŏng Tosŏ issue are from the Hwabong Mun'go (far left) and Appenzeller-Noble Memorial Museum collections (right). Image of the Chungang Sŏrim issue is from the collection of Ch'oe Ch'ŏr-hwan.



Sorae/sori 소래/소리 (sound) page 73 in the Chungang Sŏrim issue (left, in the collection of Ch'oe Ch'ŏr-hwan collection), Hansŏng Tosŏ issue (center, in Ŏm Tong-sŏp collection) and in the Munhak Sasang *yŏnginbon*.

Table 5.2 Discrepancies between the Two 1925 Issues of *Chindallaekkot* and Important Collected Works of Kim So-wŏl

No.	Page No.	Hansŏng Tosŏ Issue	Chungang Sŏrim Issue	Munhak Sasang *Yonginbon*	Hanyang University *Yonginbon*	Kwŏn Yŏng-min*	Kim Chong-uk†	Kim Yong-jik‡	O Ha-gŭn§	Cho Tong-il and Yun Chu-ŭn**
1	73	In the last line of the poem, the word "sound" is printed "*sorae* 소래"	In the last line of the poem, the word "sound" is printed "*sorae* 소래"	In the last line of the poem, the word "sound" is printed "*sori* 소리"	In the last line of the poem, the word "sound" is printed "*sori* 소리"	In the last line of the poem, the word "sound" is printed "*sori* 소리" (pg. 132)	In the last line of the poem, the word "sound" is printed "*sori* 소리" (pg. 222)	In the last line of the poem, the word "sound" is printed "*sori* 소리" (pg. 73)	In the last line of the poem, the word "sound" is printed "*sori* 소리" (pg. 83)	In the last line of the poem, the word "sound" is printed "*sori* 소리" (pg. 57)
2	120	No *mo chŏm* appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven	No *mo chŏm* appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven	A smudge/*mo chŏm* appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven	A smudge/*mo chŏm* appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven	A comma appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven (pg. 198)	A comma appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven (pg. 344)	A comma appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven (pg. 113)	A comma appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven (pg. 117)	A comma appears between *pyŏgae* 벼개 and *hŭryŏnan'ga* 흐렷난가 in line seven (pg. 80)
3	126	In the first line on the page "*talpit* 달빗" appears normally	In the first line on the page "*talpit* 달빗" appears normally	In the first line on the page "*talpit* 달빗 appears as *malpit* 말빗." This appears to be an artifact of the process of photocopying	In the first line on the page "*talpit* 달빗 appears as *malpit* 말빗." This appears to be an artifact of the process of photocopying	"*Talpit* 달빗" appears as "*malpit* 말빗." In a footnote, Kwŏn writes "Kim Yong-jik suggests this is an error" (pg. 205)	"*Talpit* 달빗" appears as "*malpit* 말빗" (pg 354). Kim has a footnote that says *mal* is an error in the "original" (pg. 367)	"*Talpit* 달빗" appears as "*malpit* 말빗" (pg 117). Kim has a footnote that says *mal* is an error (pg. 177)	"*Talpit* 달빗" appears as "*malpit* 말빗" (pg 119). Kim has a footnote that says *mal* is an error (pg. 435)	"*Talpit* 달빗" appears as "*talbit* 달빗" (pg. 81)

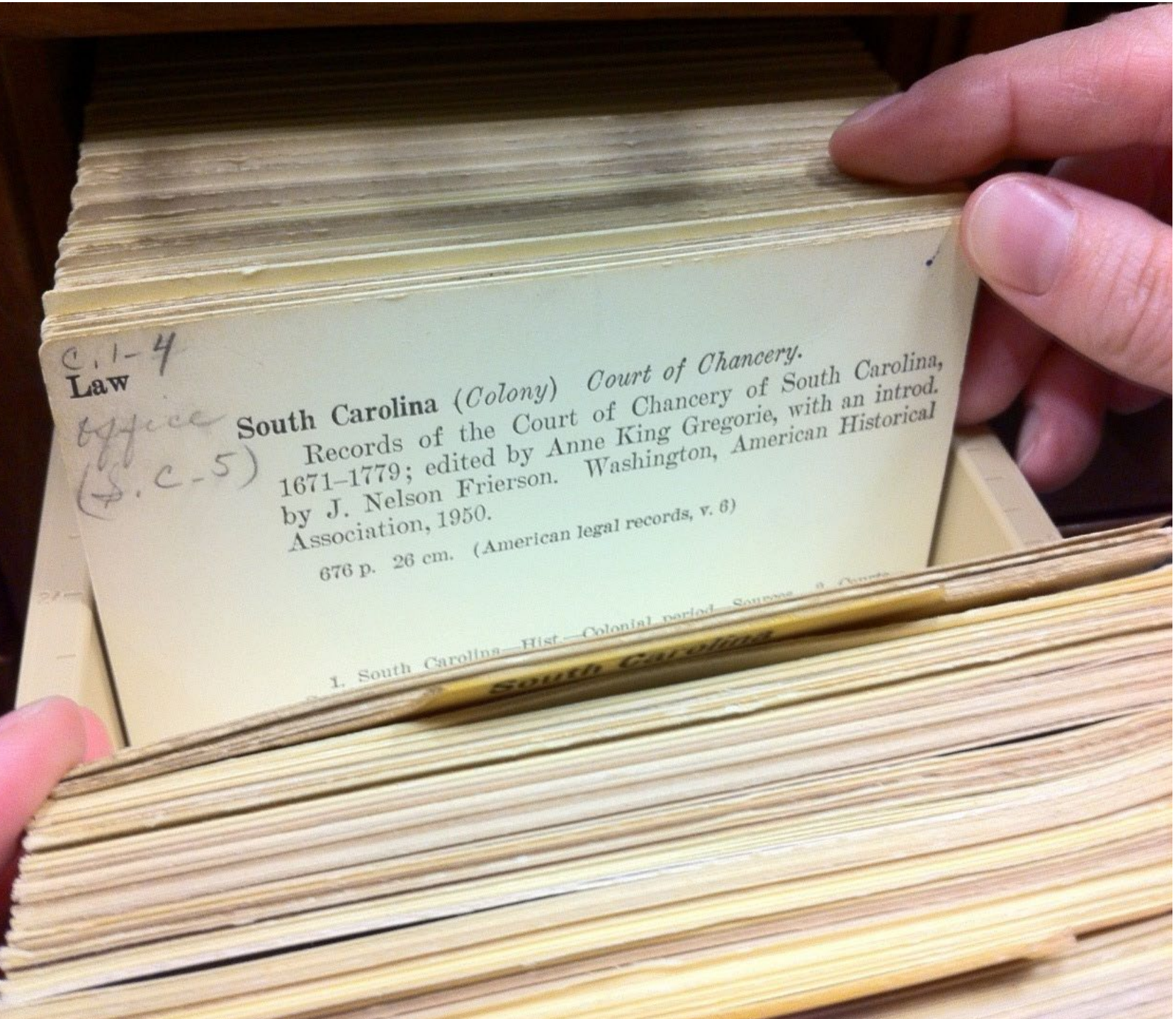
Bibliography

“List Makers”

Enumerative Descriptive

Information Science

Patrick Wilson *Two Kinds of Power: An Essay on Bibliographical Control* (1968)



Leader 01034cam a2200277Ia 4500
001 990024593990203941
005 20161109021656.0
008 760227s1950 dcu b 001 0 eng c
010 ##$aa^^52009246^
035 ##$a(OCoLC)2017627$z(OCoLC)2212604$z(OCoLC)5605193455$z(OCoLC)771364318$z(OCoLC)9764668545$z(OCoLC)1001185451
040 ##$aDetroit. Public Library$cTJC$dSLU$dHLS
043 ##$an-us-sc
050 #4$aF272$b.S67
090 ##$aF272$b.S67
110 1#$aSouth Carolina.$bCourt of Chancery$0http://id.loc.gov/authorities/names/n880718070$0http://viaf.org/viaf/sourceID/LC|n88071807$2LCNAMES$041-LIBRARY_OF_CONGRESS-n88071807$9Y
245 10$aRecords of the Court of Chancery of South Carolina, 1671-1779 /$cedited by Anne King Gregorie ; with an introduction by J. Nelson Frierson.
260 ##$aWashington, D.C. :$bAmerican Historical Association,$c1950.
300 ##$a676 p. ;$c26 cm.
490 1#$aAmerican legal records ;$vv. 6
504 ##$aIncludes bibliographical references and index.
648 #7$a1600-1775$0http://id.worldcat.org/fast/fst0135564$2FAST$041-OCLC-fst0135564$9Y
648 #7$aColonial Period (United States)$2FAST$041-OCLC-fst0135564$9N
650 #0$aCourts$zSouth Carolina$xHistory.$0http://id.loc.gov/authorities/subjects/sh85033571$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9Y
650 10$aEquity pleading and procedure$zSouth Carolina$vCases.$0http://id.loc.gov/authorities/subjects/sh85044559$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9Y
650 #0$aLaw reports, digests, etc.$zSouth Carolina$0http://id.loc.gov/authorities/subjects/sh2009129296$2LCSH$041-LIBRARY_OF_CONGRESS-sh2009129296$9Y
650 #0$aCourt records$zSouth Carolina.$0http://id.loc.gov/authorities/subjects/sh85033556$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9Y
650 #0$aCourts$vRecords and correspondence$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9N
650 #0$aRecords of court$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9N
650 #0wgaArchives$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9R
650 #0wgaEvidence (Law)$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9R
650 #0wgaPublic records$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033556$9R
650 10$aChancery$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9N
650 10$aEquity pleading and procedure$xLaw and legislation$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9N
650 10wgaCivil procedure$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9R
650 10wgaEquity$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9R
650 10wgaPleading$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9R
650 10wgaProcedure (Law)$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9R
650 10wgaTrial practice$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85044559$9R
650 #0$aCourts$xLaw and legislation$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9N
650 #0$aJudiciary$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9N
650 #0wgaDispute resolution (Law)$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0wgaJudicial districts$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0wgaLaw$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0wgaProcedure (Law)$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0$aJudicial power$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0$aJurisdiction$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
650 #0$aJustice, Administration of$2LCSH$041-LIBRARY_OF_CONGRESS-sh 85033571$9R
651 #0$aSouth Carolina$xHistory$yColonial period, ca. 1600-1775$xSources.$0http://id.loc.gov/authorities/names/n790229140http://viaf.org/viaf/sourceID/LC|n79022914$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 79022914$9Y
651 #0$aSouth Carolina (Colony)$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 79022914$9N
651 #0$aSouth Carolina (Province)$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 79022914$9N
651 #0$aIU2hnaÏa¯Karolina$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 79022914$9N
655 #7$aAbstracts$0(OCoLC)fst0142368$0http://id.worldcat.org/fast/fst0142368$2FAST$041-OCLC-fst0142368$9Y
655 #7$aHistory$0(OCoLC)fst0141162880http://id.worldcat.org/fast/fst0141162882FAST$041-OCLC-fst0141162889Y
655 #7$aSources$0(OCoLC)fst0142390050http://id.worldcat.org/fast/fst0142390052FAST$041-OCLC-fst0142390059Y
655 #7$aTrials, litigation, etc.$0(OCoLC)fst0142371250http://id.worldcat.org/fast/fst0142371252FAST$041-OCLC-fst0142371259Y
655 #7$aCompendia$2FAST$041-OCLC-fst0142368$9N
655 #7$aCompendiums$2FAST$041-OCLC-fst0142368$9N
655 #7$aCompends$2FAST$041-OCLC-fst0142368$9N
655 #7$aEpitomes$2FAST$041-OCLC-fst0142368$9N
655 #7$aPrelïds$2FAST$041-OCLC-fst0142368$9N
655 #7$aPublication summaries$2FAST$041-OCLC-fst0142368$9N
655 #7$aSummaries of publications$2FAST$041-OCLC-fst0142368$9N
655 #7$aSynopses (Abstracts)$2FAST$041-OCLC-fst0142368$9N
655 #7wgaDerivative works$0(OCoLC)fst0191991720FAST$041-OCLC-fst01423683$9R
655 #7wgaInformational works$0(OCoLC)fst0191993050FAST$041-OCLC-fst01423683$9R
700 1#$aGregorie, Anne King$0http://id.loc.gov/authorities/names/n880718060$0http://viaf.org/viaf/sourceID/LC|n88071806$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 88071806$9Y
710 2#$aAmerican Historical Association$0http://id.loc.gov/authorities/names/n79063583$0http://viaf.org/viaf/sourceID/LC|n79063583$2LCNAMES$041-LIBRARY_OF_CONGRESS-n79063583$9Y
710 2#$aAHA$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 79063583$9N
776 08$iOnline version:$aSouth Carolina. Court of Chancery.$tRecords of the Court of Chancery of South Carolina, 1671-1779.$dWashington, D.C. : American Historical Association, 1950$w(OCoLC)758496938
830 #0$aAmerican legal records$vv. 6.$0http://id.loc.gov/authorities/names/n83705700$0http://viaf.org/viaf/sourceID/LC|n83705700$2LCNAMES$041-LIBRARY_OF_CONGRESS-n 83705700$9Y
900 ##$a002459399
901 7#$aAMC2250
903 ##$aMHAMC22507HU
905 ##$aF^^0272
948 ##$aORC3170ASWID2Box03253$b19960321$c19960322
991 ##$b2022-01-11$cWorldCat record variable field(s) change: 700
991 ##$b2022-06-12$cWorldCat record encoding level change - WorldCat record variable field(s) change: 655
998 ##$aNO-OVERLAY
AVA ##$099002459399020394 1$822185178679000394 1$a01HVD_INST$bWID$cOld Widener$dUS 19596.71$eavailable$f1$g0$jGEN$kZHCL$p1$qWidener Library
AVA ##$099002459399020394 1$822185178679000394 1$a01HVD_INST$bLAW$cOffsite Storage$dKFS1849 .C5x 1850$eavailable$f1$g0$jHD$k0$p2$qHarvard Law School Library
CAT ##$aHistory_Archaeology$bRegions_Countries$cAmericas_United_States_Local_History
INST ##$a01HVD_INST
INT ##$aP

Compare BIBFRAME converted to MARC

Identifier ◉ Bib ID ○ LCCN **Serialization** ○ XML ◉ Text

20987659

Search | Clear

BIBFRAME XML

```
<rdf:RDF xmlns:bf = "http://id.loc.gov/ontologies/bibframe/" xmlns:bflc = "http://id.loc.gov/ontologies/bflc/" xmlns:madsrdf =
"http://www.loc.gov/mads/rdf/v1#" xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs = "http://www.w3.org/2000/01/rdf-
schema#" >
 <bf:Instance rdf:about="http://id.loc.gov/resources/instances/20987659" >
  <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Print" />
  <bf:issuance >
   <bf:Issuance rdf:about="http://id.loc.gov/vocabulary/issuance/mono" >
    <rdfs:label >single unit</rdfs:label>
   </bf:Issuance>
  </bf:issuance>
  <bf:provisionActivity >
   <bf:ProvisionActivity >
    <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Publication" />
    <bf:date rdf:datatype="http://id.loc.gov/datatypes/edtf" >1950</bf:date>
    <bf:place >
     <bf:Place rdf:about="http://id.loc.gov/vocabulary/countries/dcu" >
      <rdfs:label xml:lang="en" >District of Columbia</rdfs:label>
     </bf:Place>
    </bf:place>
   </bf:ProvisionActivity>
  </bf:provisionActivity>
  <rdfs:label >Records of the court of chancery of South Carolina 1671-1779 :</rdfs:label>
  <bf:responsibilityStatement >Edited by Anne King Gregorie</bf:responsibilityStatement>
  <bf:title >
   <bf:Title >
```

MARC (Text)

```
    cam a22     5c 4500
 2  $aSouth Carolina$1http://id.loc.gov/rwo/agents/n79022914$1http://id.loc.gov/rwo/agents/n79022914$4http://id.lo
c.gov/vocabulary/relators/ctb
001    20987659
003    DLC
005    20190530144430.0
008    190528s1950||||dcu           00| 0|eng
040    $aDLC$beng$erda$cDLC
084    $aLAW U.S. South Carolina 5 1950$qdlc
245 10 $aRecords of the court of chancery of South Carolina 1671-1779$cEdited by Anne King Gregorie
264  1 $aWashington, D.C.$bThe American Historical Association$c1950
300    $a6, 676 pages$c26 cm
334    $asingle unit$bmono$0http://id.loc.gov/vocabulary/issuance/mono
336    $atext
337    $aunmediated
338    $avolume
490 0  $aAmerican Legal Records ; Vol. 6
884    $aDLC bibframe2marc v2.3.0 (MarkLogic Corporation)$g20231028205518.0$qDLC$uhttps://github.com/lcnetdev/bibf
rame2marc
```

Conversion: **DLC bibframe2marc v2.3.0**

Records of the Court of Chancery of South Carolina, 1671-1779

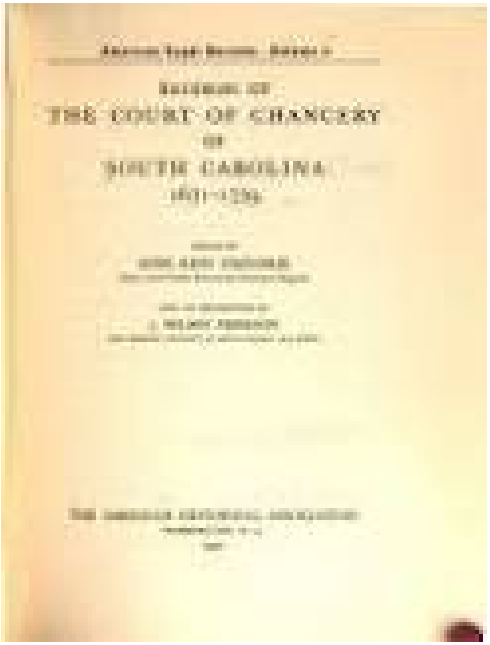
By South Carolina. Court of Chancery · 1950

Preview unavailable | Search inside | + Add to my library

Overview | Get the book | Publisher collection

Preview unavailable
Learn more

Google Books



THE UNIVERSITY OF IOWA LIBRARIES
INFOHAWK+

E-BOOK

Records of the Court of Chancery of South Carolina, 1671-1779 / edited by Anne King Gregorie ; with an introduction by J. Nelson Frierson.

South Carolina. Court of Chancery, court governed. ; Anne King Gregorie editor.; American Historical Association, issuing body.
Washington, D.C. : American Historical Association 1950

Available at Law Library Stacks (KFS2315 .A545 1671) ›

Online access may be available ›

複
寫

複 복 겹칠 (overlap, compound)
寫 사 베끼다 (copy, imitate)

베
끼
다

Copy

I. a. Abundance, plenty.
   b. Fullness, plentitude.

II. A transcript or reproduction of an original.

Copy

Copy serving information retrieval

Distance of some kind
(e.g. space)

Copy

Distance of some kind
(e.g. space)

Copy

Copy

Distance of some kind
(e.g. time)

Transform. 化.

Latin *transformāre*, < TRANS- *prefix* + *formāre* to form>

trans, ‘across, to or on the farther side of, beyond, over’

化 to become, to change

Accounting

I. To render a reckoning.

Etymology: < Anglo-Norman *acounter*, *aconpter*, *accomptier*, *accomptere*, Anglo-Norman and Old French *acunter*, Anglo-Norman and Old French, Middle French *aconter*, Middle French *acompter*, *accompter* **to count, enumerate (c1100), to relate, tell, to include (something) in a reckoning, count in, to calculate (all early 12th cent.), to consider, deem, to value, consider important (c1280), to present financial accounts (early 14th cent.), to matter, be of account** (mid 14th cent.) < *a-* A- *prefix*[5] + *cunter* , *conter* COUNT *v.* Compare Old Occitan *acomtar*, *acontar* (c1150), post-classical Latin *accomputare* (1661 in a British source).The β. forms reflect Anglo-Norman and Middle French *accompter* , *acompter* , etc.: for the development of forms of this type and their history in French see discussion at **COUNT *v.***

Copy

Copy

Distance of some kind (e.g. time and/or space)

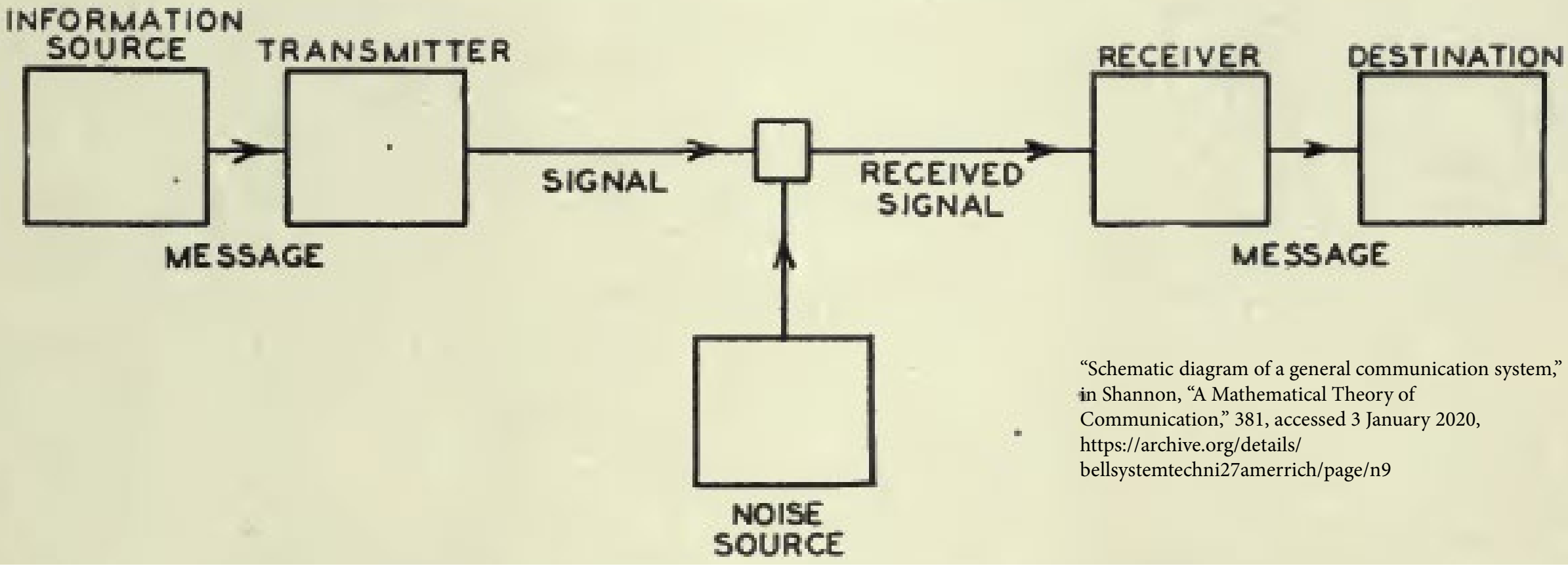
Similarity, similitude

Particularity, Individuality

Who or what determines?
For what purposes?
In which contexts?

Copies and Claude Shannon

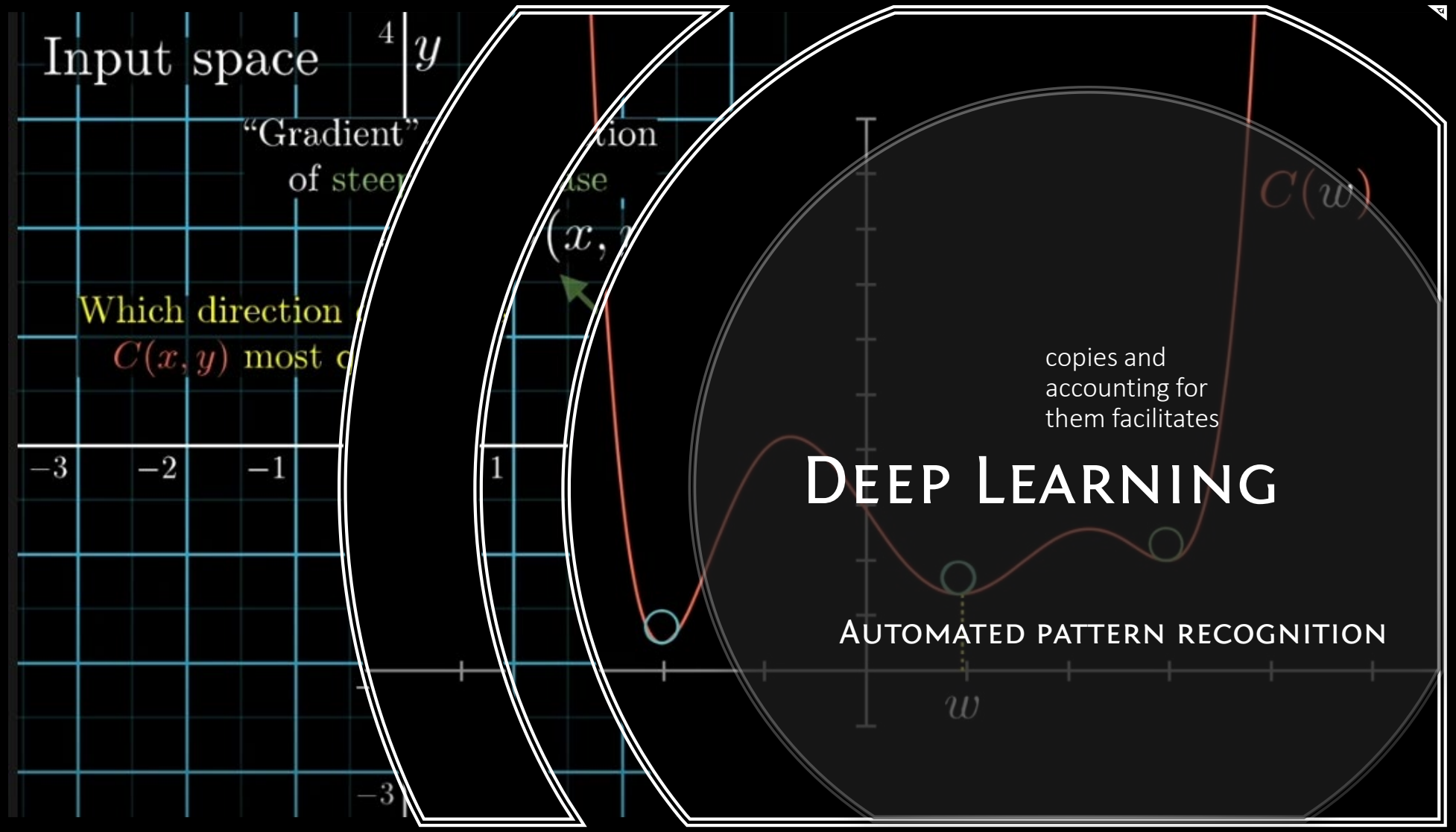
Shannon's theory of communication became known as "information theory." As a practical matter it is about making a copy at a distance. It would better be named Copy Theory . . .



"Schematic diagram of a general communication system," in Shannon, "A Mathematical Theory of Communication," 381, accessed 3 January 2020, https://archive.org/details/bellsystemtechni27amerrich/page/n9

Deep Learning

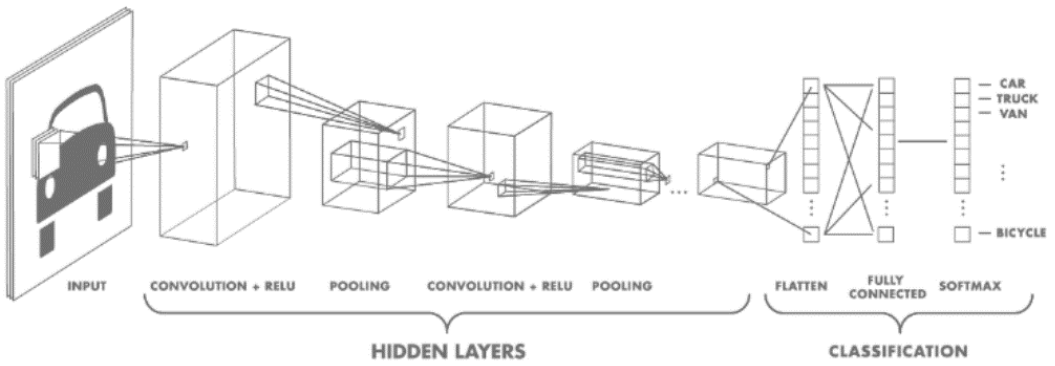
As Bibliographical Processes



copies and
accounting for
them facilitates

Deep Learning

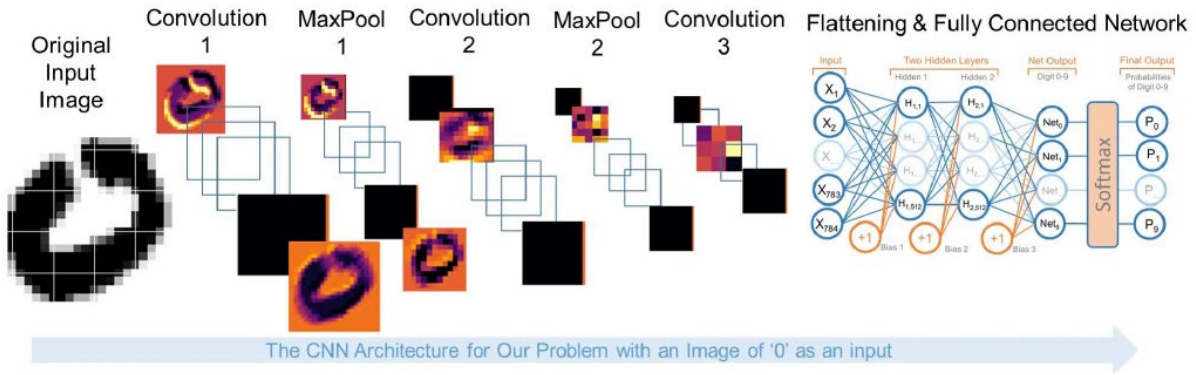
Automated pattern recognition



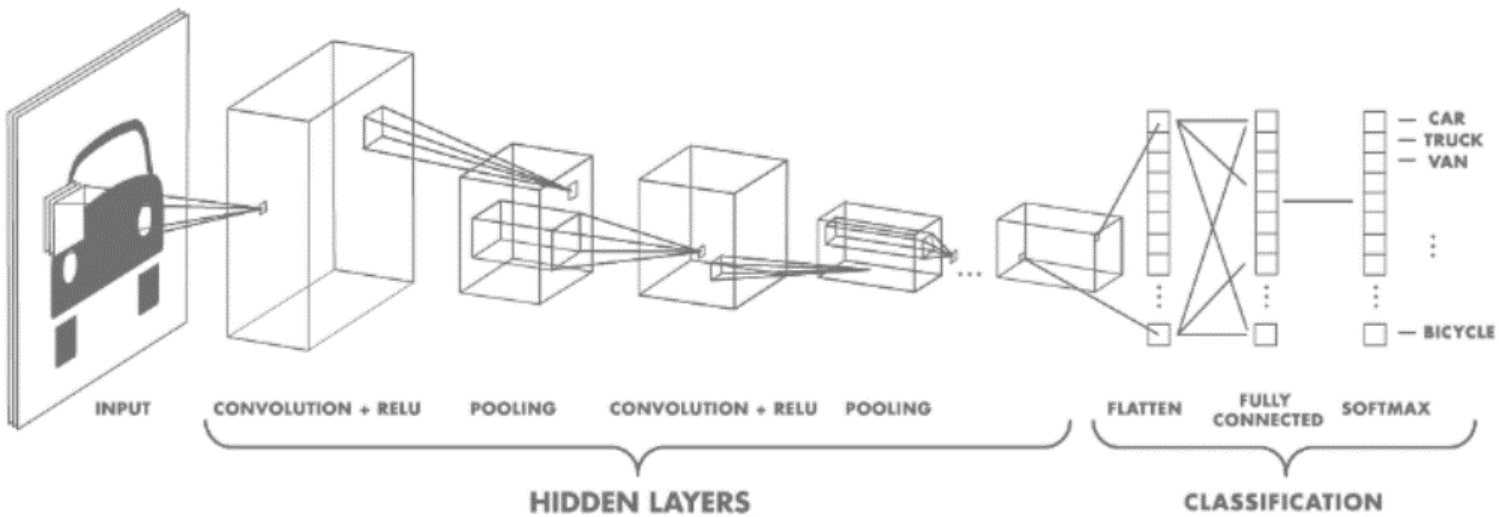
Architecture of a CNN. — Source: https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Deep Learning

Convolutional Neural Networks

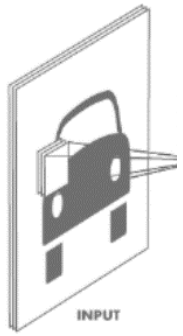


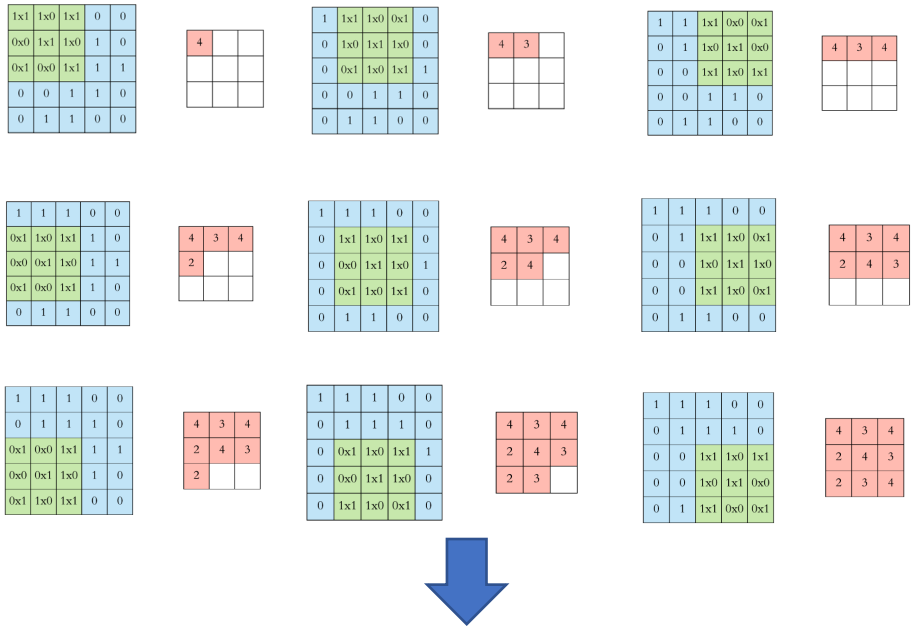
Convolutional Neural Networks



Architecture of a CNN. — Source: https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolution





1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Input

plus

1	0	1
0	1	0
1	0	1

Filter / Kernel

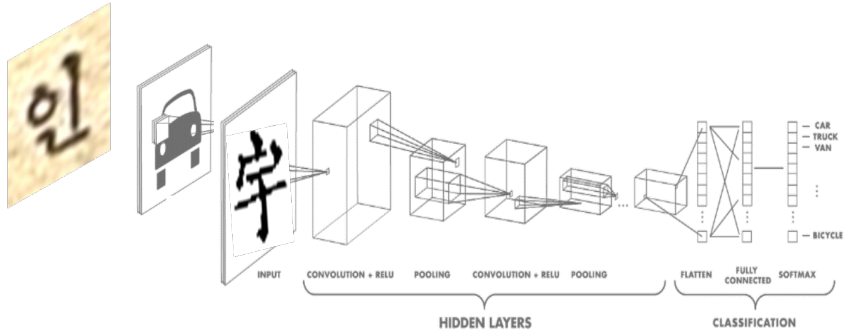
Useful similarity a copy

4	3	4
2	4	3
2	3	4

Feature (map)

Deep Learning

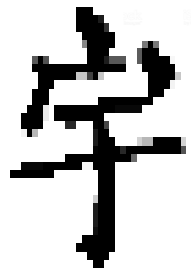
copies (like turtles) all the way down



Weight File

(record of which features best predict/ describe objects in training set)

Something.weights



229 173 151

car

99 87 114

인

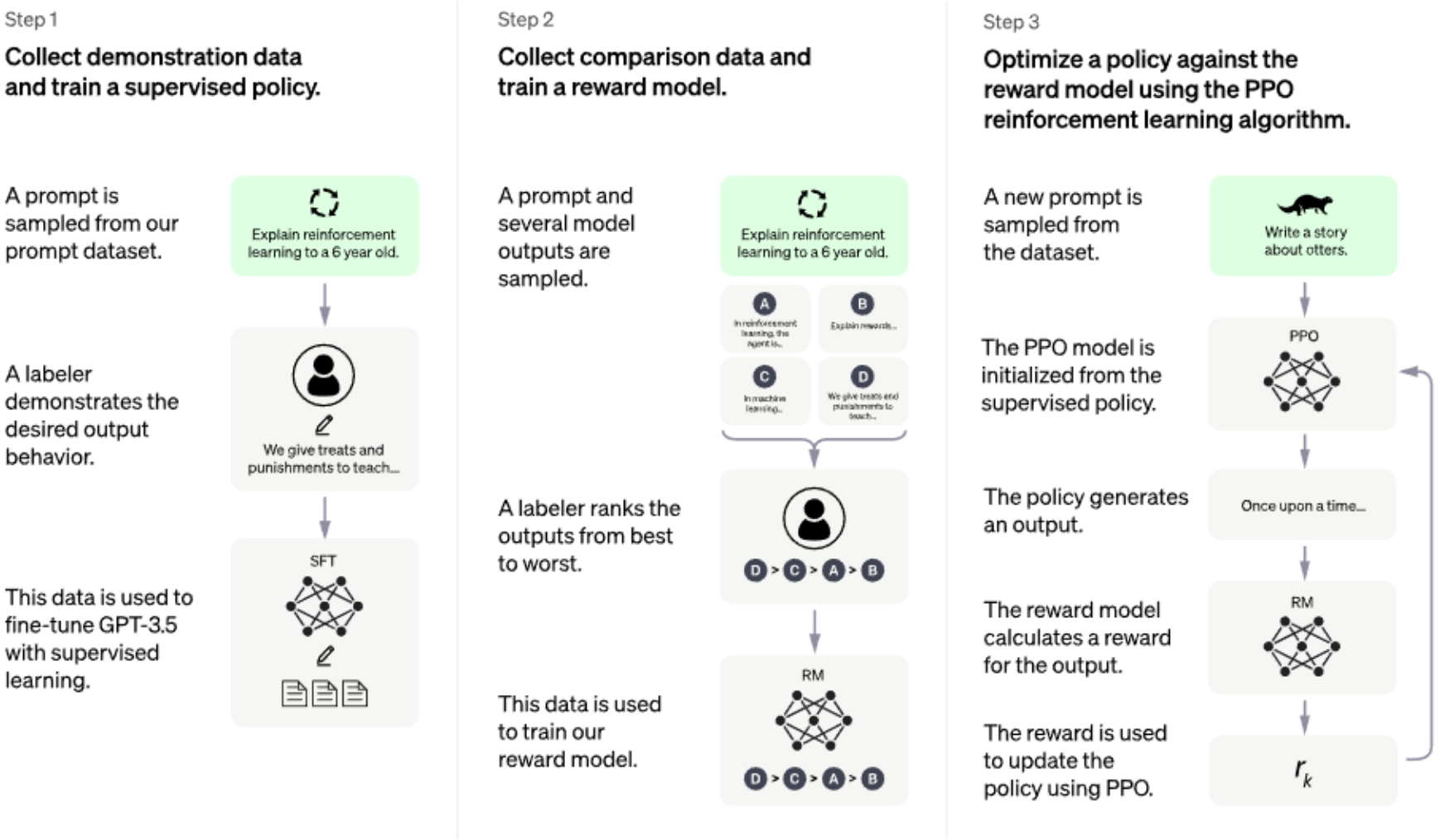
236 157 184

Enumeration

Analysis

Description

Critique



opportunities for deep learning

Early-to-Mid Twentieth-Century Korean Periodicals

談野

創刊號

등록
1963 2 15
37793
국립도서관

十二月號

學生界

第八號

納本 朝鮮總督府警務局 大正 年 月 日 第 號

京城 漢城圖書株式會社 發行

大正十年十一月十九日約束郵便物認可

中學講義錄

朝鮮通信中學館

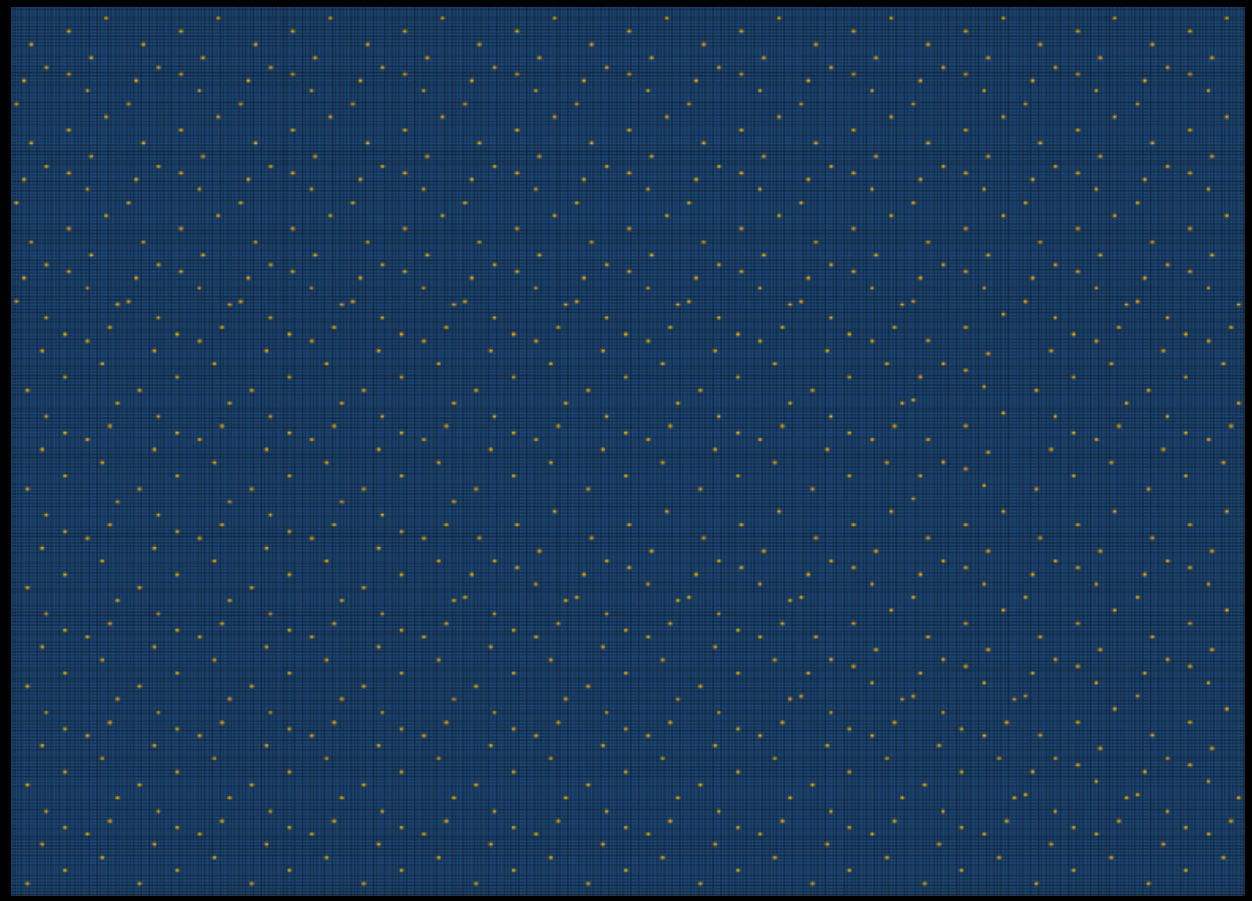
第十二號

檢閱濟

檢閱者

Deep Learning Project
at the National Library of Korea

89,981 images



1,009 rare
(*kwijungbon*)
periodical issues

932 colophons
identified & transcribed
automatically

1910

1960

一九四六年二月廿日印刷
一九四六年三月一日發行 第一號
(本號定價拾圓)
(送料 二〇錢)
編輯兼發行人 金東鎭
서울市御成町八六 凡閣章內
印刷所 朝鮮精版社
서울市長谷川町七四
定價普通號 一部 拾圓 送料二〇錢
半年分六十圓 (但、先金・送料添付)
增大號일時는超過額을別受함

駐文規定

★本誌의註文은一切先金으로하시오
(送料添付)
★送金은반드시小爲管를利用하시오
(서울市御成町八六番地에支拂)
★一切의問議는郵票同封或은往復葉書로하시오
★郵票代用은一切拒絕합니다

發行所 서울市御成町八十六番地 凡章閣
電話本局②四六五九

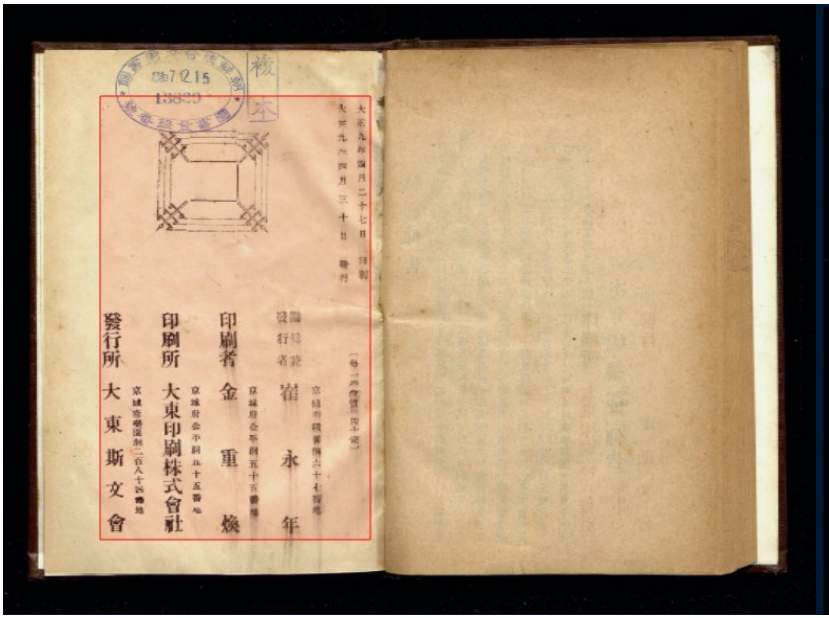
를 [illegible]입니다 (郵票添付要)

◇本誌를읽고저하시는분은地方支局에 미리注文豫約하여주십시요 部數에 制限이있을으로 늦으면 希望者全部에게 드리기困難합니다。

◇萬一 地方에支局이없어서 不便한분은 直接本社에先金으로 豫約하여주십시요。送料 二圓을 添付하여 振替 京城二〇二又는小爲替切手代用 주었어던지 좋습니다。

西紀一九四七年六月二五日 印刷
西紀一九四七年七月一日 發行
서울市乙支路二街一八〇番地
編輯兼發行人 金正修
印刷所 學生社印刷部
서울市乙支路二街一八〇
發行所 學生社
編輯局 電(本)六一六番
電話本局九一九番
振替京城二〇二番
臨時 定價五拾圓也

(購讀料先金)
一個月 十錢
六個月 五十錢
一個年一圓〇五錢
廣告料
十七字一行 三十錢
半頁 一圓八十錢
一頁 三圓
一頁以上은(每頁)一圓五十錢式加算
大正五年二月二十五日 印刷
大正五年三月五日 發行
京城府敦義洞百三十一番地
著作兼發行者 丁元燮
京城府勸農洞三十一番地
印刷者 韓養浩
京城府鍾路通一丁目三九番地
印刷所 鮮明社
京城府嘉會洞百七十七番地
發行所 侍天教報社



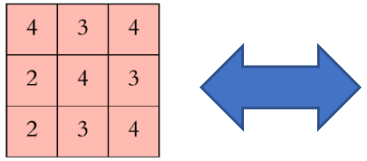
Section Model

4	3	4
2	4	3
2	3	4

Features

target

colophon



Features

target

meaningful element

Segmentation Model

東洋醫藥（第一輯）

昭和十四年七月十日　印刷
昭和十四年七月十日　發行

定價三十錢
半年先金壹圓六拾錢
壹年先金參圓

廣告料
表紙二面　八拾圓
三面　七拾圓
四面　百圓
普通面　四拾圓

京城府慶雲町九六

編輯兼發行人　田元培

京城府堅志町二一二番地

印刷人　金容圭

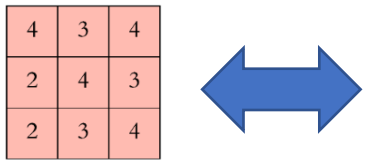
京城府堅志町二二二番地

印刷所　株式會社大同出版社

京城府慶雲町九六番地

發行所　東洋醫藥協會出版部

振替京城三〇六八九番
電話（光）二四九二番

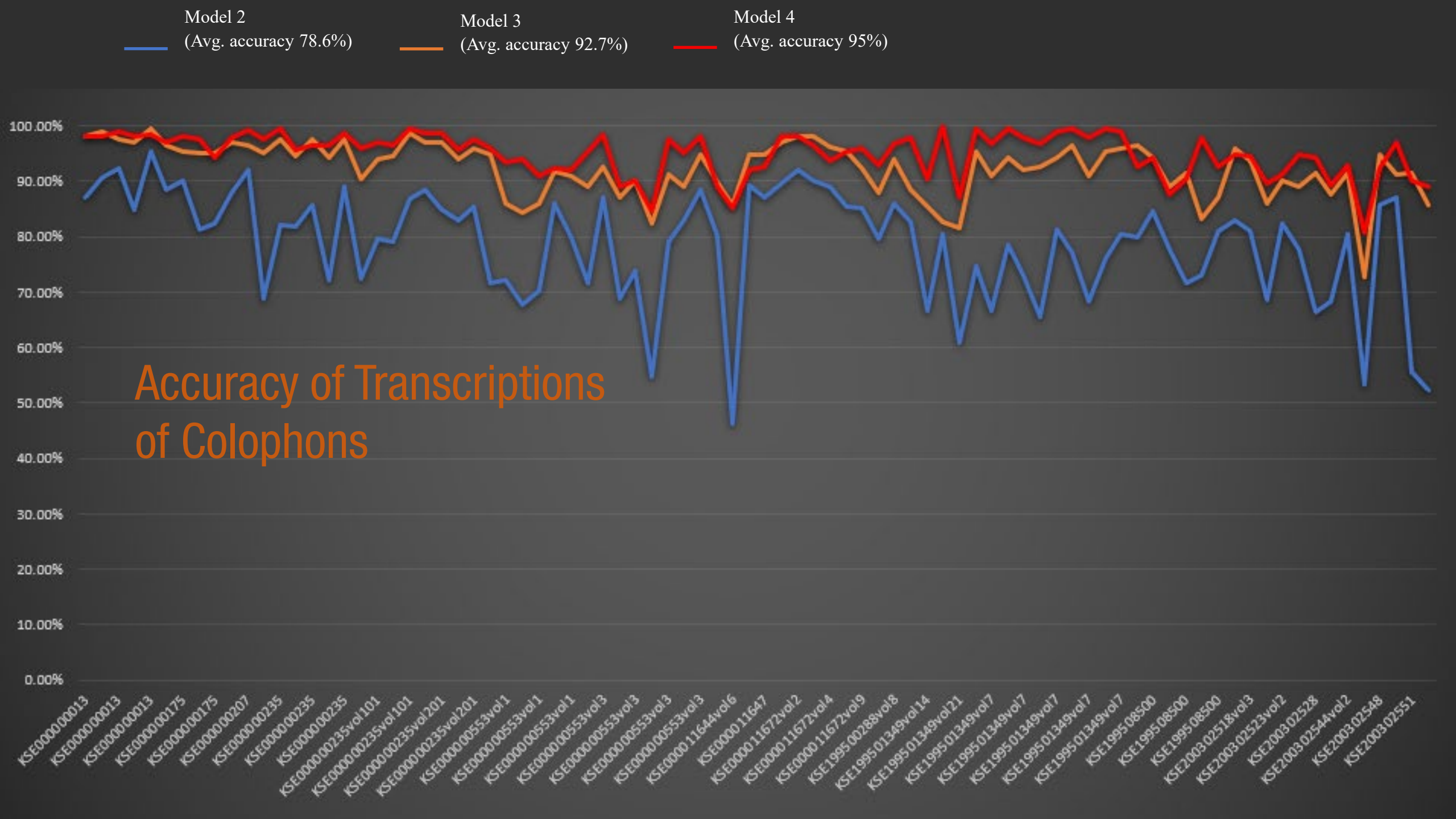


Features

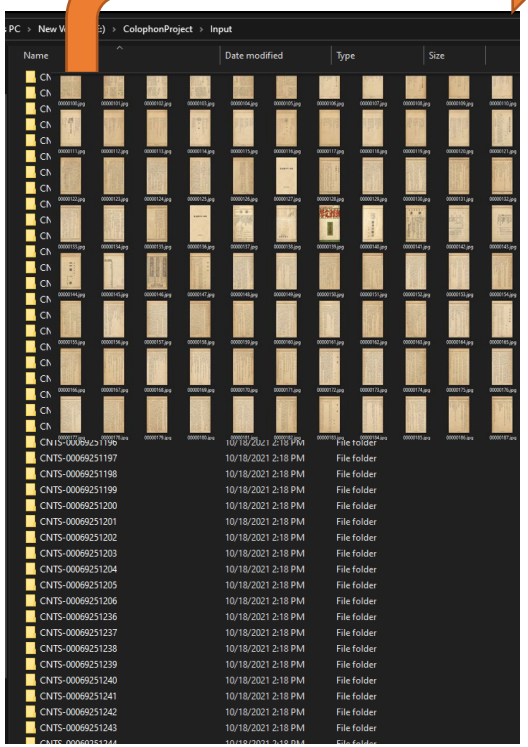
target

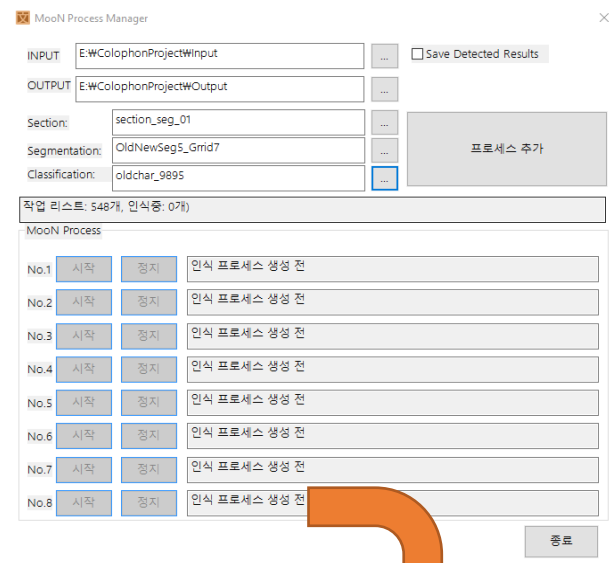
labels for meaningful elements

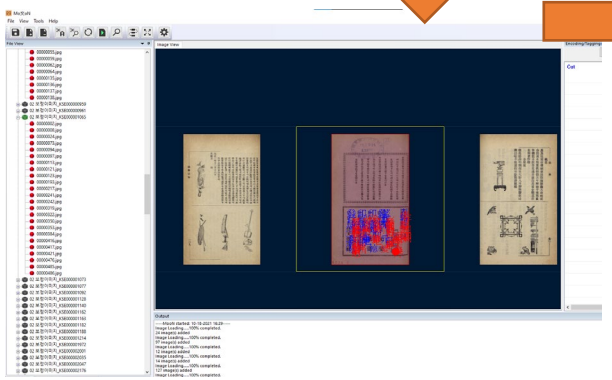
Classification Model

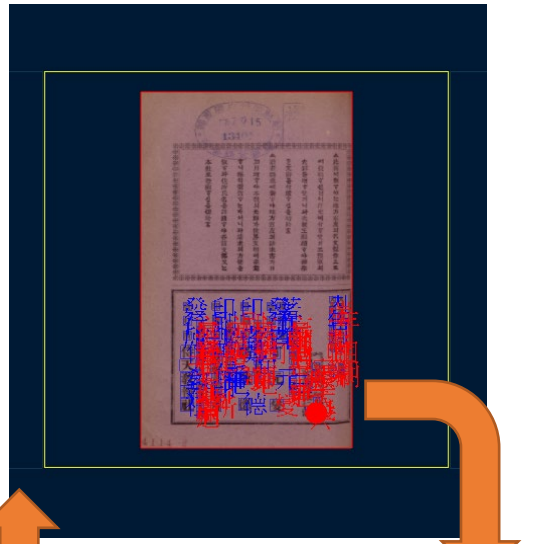

Model 2
(Avg. accuracy 78.6%)
Model 3
(Avg. accuracy 92.7%)
Model 4
(Avg. accuracy 95%)
Accuracy of Transcriptions of Colophons
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
KSE0000000013
KSE0000000013
KSE0000000013
KSE0000000175
KSE0000000175
KSE0000000207
KSE0000000235
KSE0000000235
KSE0000000235
KSE0000000235vol101
KSE0000000235vol101
KSE0000000235vol201
KSE0000000235vol201
KSE0000000553vol1
KSE0000000553vol1
KSE0000000553vol1
KSE0000000553vol3
KSE0000000553vol3
KSE0000000553vol3
KSE0000000553vol3
KSE0000011644vol6
KSE0000011647
KSE0000011672vol2
KSE0000011672vol4
KSE0000011672vol9
KSE1995002888vol8
KSE1995013491vol14
KSE1995013491vol21
KSE1995013491vol7
KSE1995013491vol7
KSE1995013491vol7
KSE1995013491vol7
KSE1995013491vol7
KSE1995085001
KSE1995085001
KSE1995085001
KSE2003025181vol3
KSE2003025231vol2
KSE2003025281
KSE2003025441vol2
KSE2003025481
KSE2003025511

Section Model
Segmentation Model
Classification Model







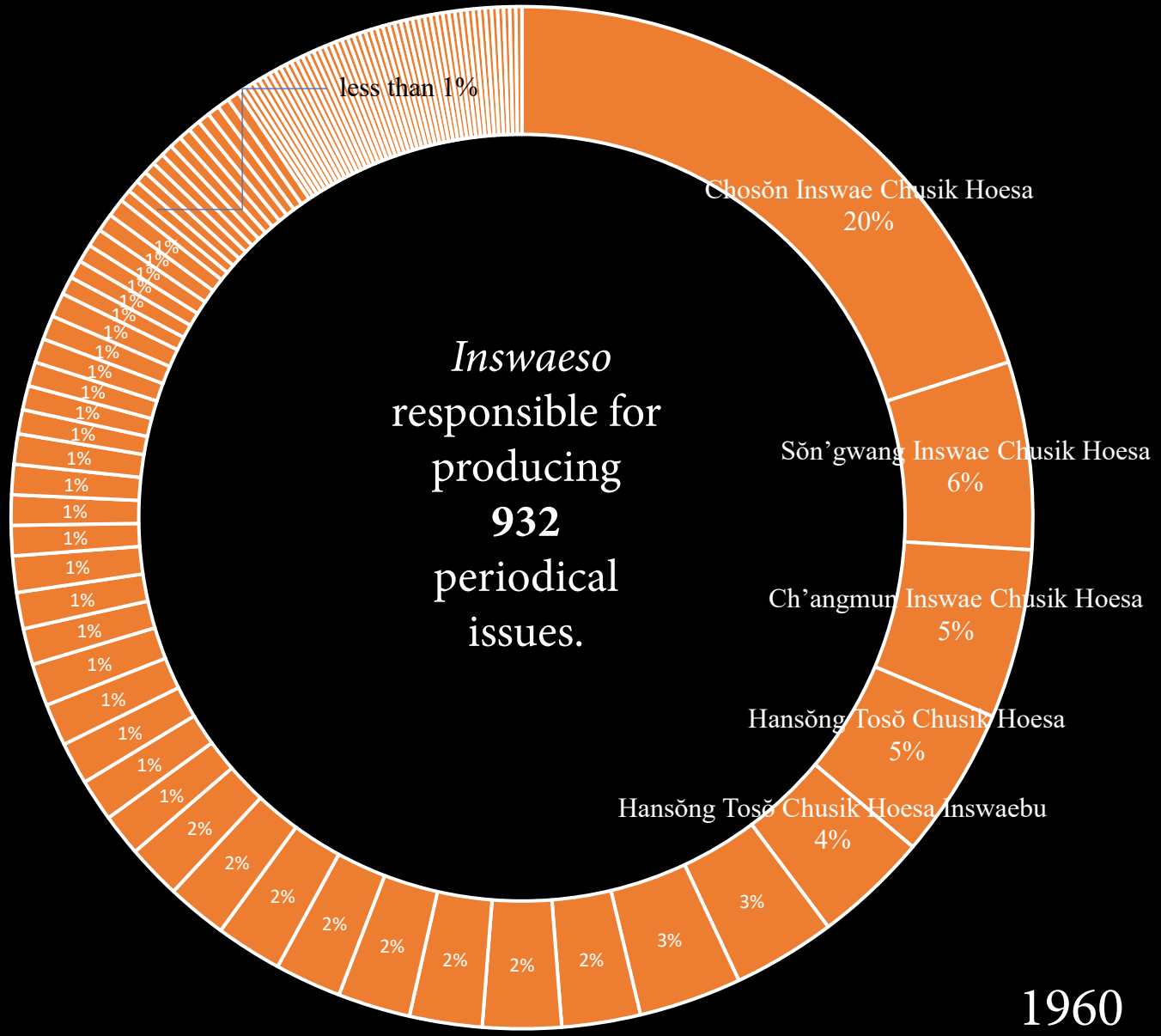


00000242.Jpg

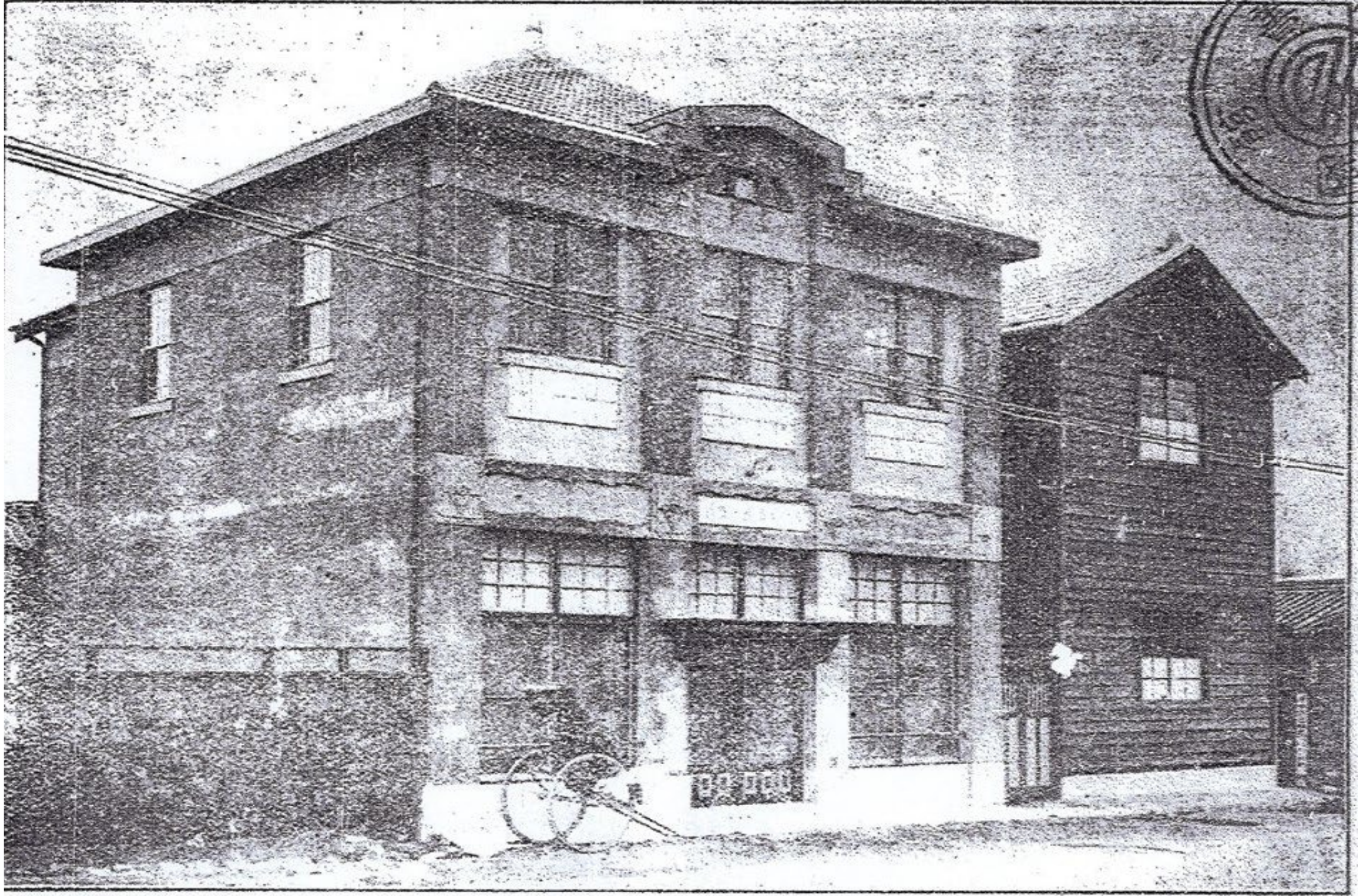
---

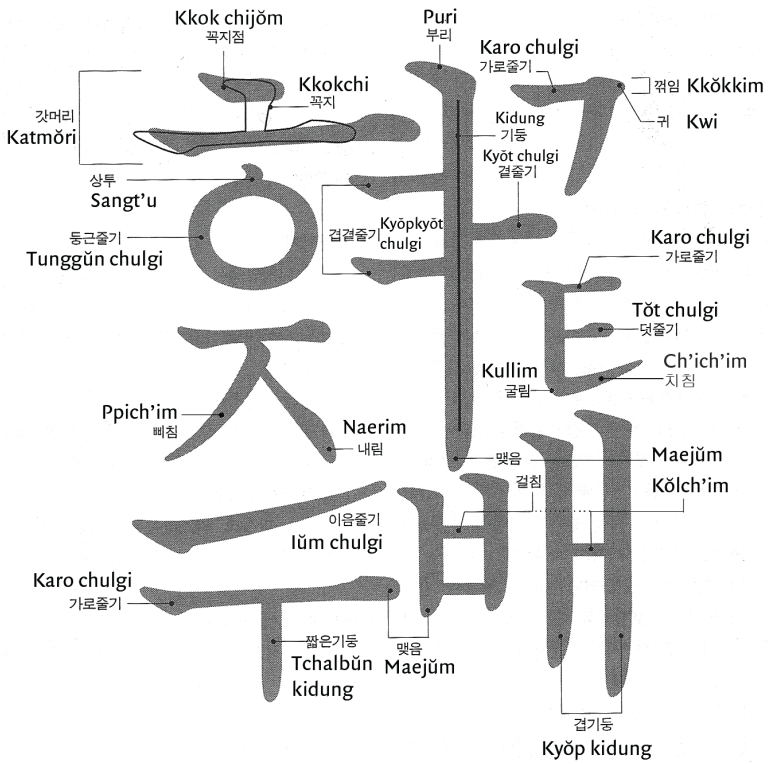
1 年 :+相 印刷
2 치 遄年二月劃卿 發가
3 定價傘 五錢宍
4 獎費 ●
5 京兼啣鄰會迵百平三番地
6 著 作 兼
7 發 行 者 丁 元 燮
8 京坡府鄉迵三面 爾地
9 印 刷 者 鄭 敬 德
10 京所府苑痾三瞑番地
11 印 刷 所 朝鮮福痦印한所
12 京撼府嘉覽載帝g七五地奮廉迵
13 發 行 所 侍天敎報社

Deep Learning Project
at the National Library of Korea


less than 1%
Chosŏn Inswae Chusik Hoesa
20%
Sŏn'gwang Inswae Chusik Hoesa
6%
Ch'angmun Inswae Chusik Hoesa
5%
Hansŏng Tosŏ Chusik Hoesa
5%
Hansŏng Tosŏ Chusik Hoesa Inswaebu
4%
3%
3%
2%
2%
2%
2%
2%
2%
2%
2%
2%
1%
1%
1

Newly built Hansŏng Tosŏ Chusik Hoesa building, *Haksaenggye* (December 1920), unnumbered front matter (image from microfilm at the NLK).

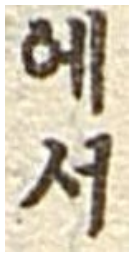


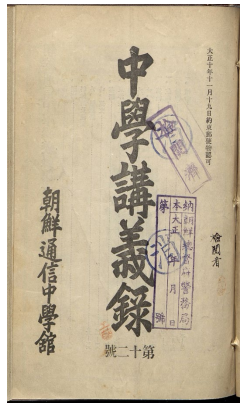


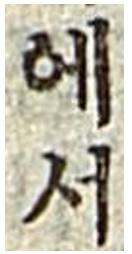
Source: Han Chae-jun, ed., *Han'gŭl kŭlkkol yongŏ sajŏn* (A dictionary of typographic terms for *han'gŭl*) (Seoul: Sejong Taewang Kinyŏm Saŏphoe, 2000), 46.

Type samples

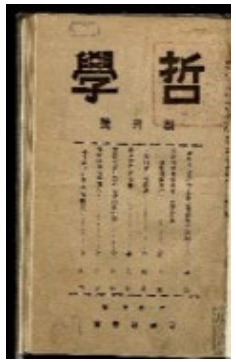
from materials printed by Hansŏng Tosŏ Chusik Hoesa

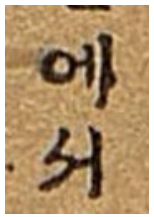






1920s

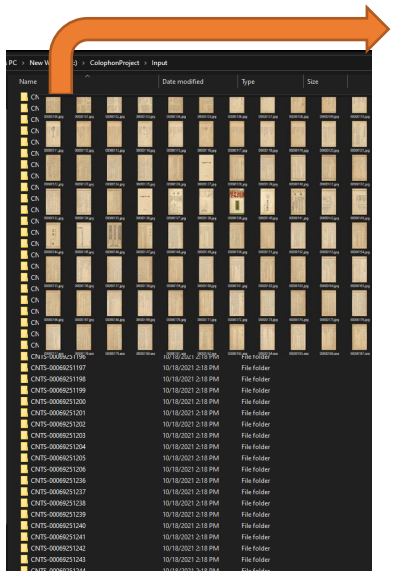


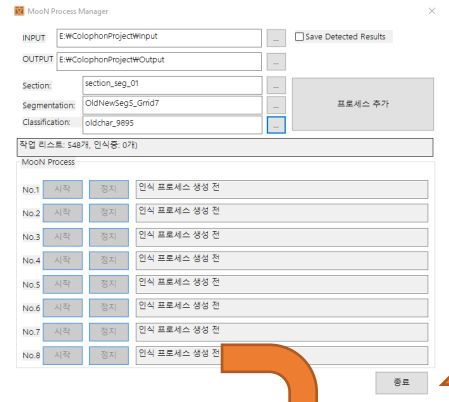


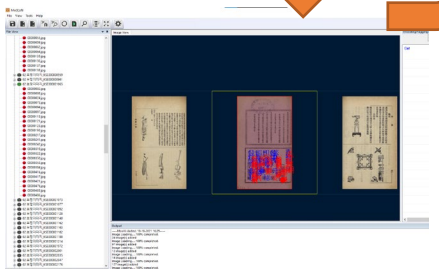
1930s

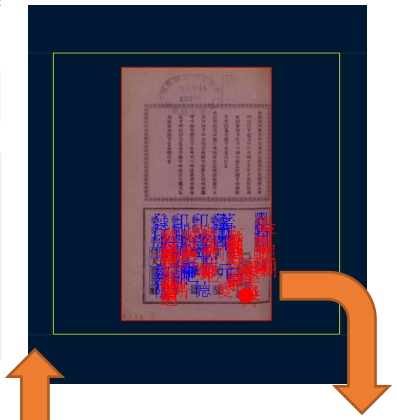
Type samples from page 101 of the May 1921 issue of *Haksaenggye* (top left), page 78 of the August 1922 issue of *Chunghak Kangŭirok* (top right), page 72 of the June 1935 issue of *Ch'ŏrhak* (bottom left), and page 157 of the November 1937 issue of *Yadam* (bottom right).

Full text transcriptions of the Periodicals held by the National Library of Korea









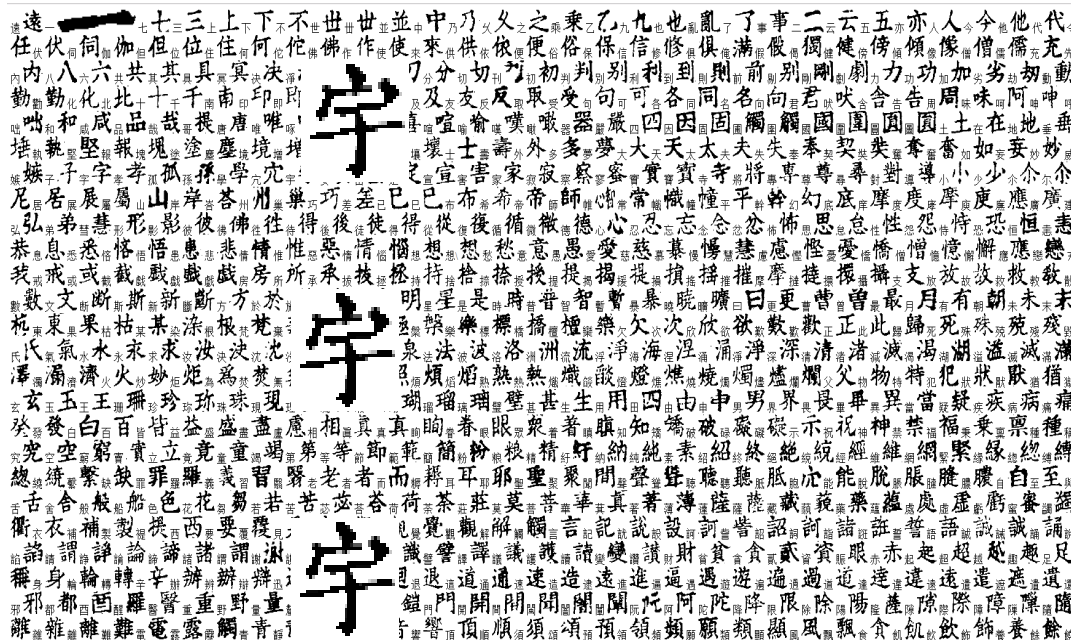
00000242.Jpg

세계를 한 집으로 보고 사해를 한 형제로 본다 하면 말이 갓고 풍속이 갓고 역사가 갓고 이해가 가튼 한 민족은 더욱이나 한 집안 한 형제로 보게 될 것입니다. 설사 하는 일은 세계 전체를 표준한다 할지라도 세계가 하로 아츰에 한 집안 한 형제가 되야 가지고 세계 전체의 일을 할 수가 업고 민족과 민족이 어울려저 내 일도 전체의 일이오 네 일도 전체의 일이라 하는데서 其實 사해 형제의 본의가 실현될 줄로 생각합니다. 그것은 오랫동안 관계를 가저온 한 민족이 스스로 어울려지기 쉽고 관계맷기 쉬울 경우에 잇는 까닭이 안이겟습니까. 이 점에서 한 가지로 우리 朝鮮 동포라 하면 그 하는 일의 표준을 세계적이거나 민족적이거나를 물론하고 스스로 동포와 동포가 손목을 잡고 일터에 나아가게 되는 것입니다. 더욱이 그 민족의 일을 그 민족이 자진하야 하지 안이하면 즉 자기의 맛튼 일을…

more opportunities for deep learning

Qisha Canon
磧砂藏

Qisha Canon Encoding Project (Fo Guang Temple, Taiwan)



大唐三藏聖教序　　天一

太宗文皇帝製

蓋聞二儀有像顯覆載以含生四時無形潛寒暑以化物是以窺天鑑地庸愚皆識其端明陰洞陽賢哲罕窮其數然而天地苞乎陰陽而易識者以其有像也陰陽處乎天地而

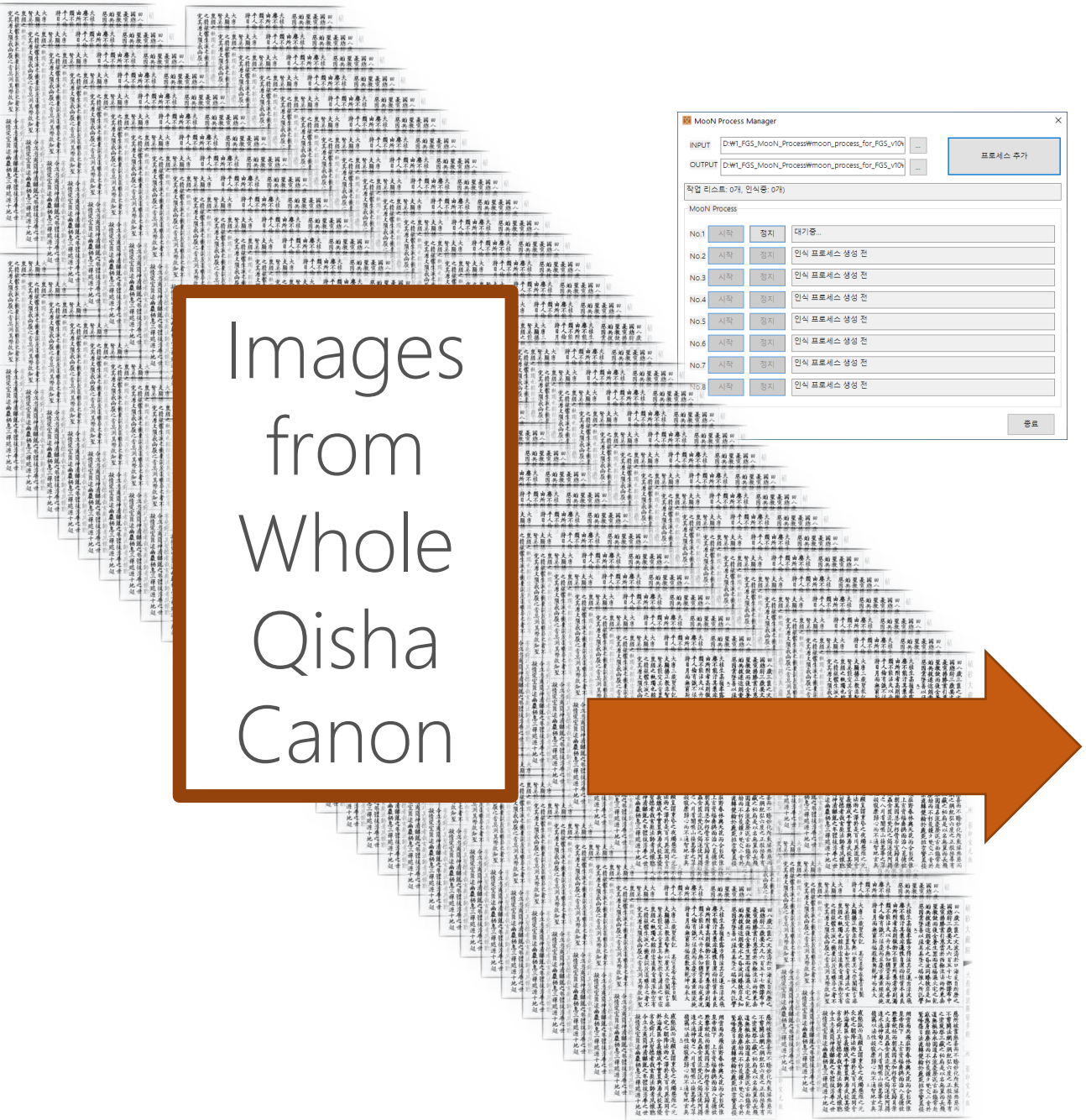
一　孫仁判

難窮者以其無形也故知像顯可徵雖愚不惑形潛莫覩在智猶迷況乎佛道崇虛乘幽控寂弘濟萬品典御十方舉威靈而無上抑神力而無下大之則彌於宇宙細之則攝於毫釐無滅無生歷千劫而不古若隱若顯運百福而長今妙道凝玄遵之莫知其際法流湛寂挹之莫測其源故知蠢蠢凡愚區區庸鄙投其旨趣能無疑惑者哉然則大教之興基乎西土騰漢庭而皎夢照東域而流慈昔者分形分跡之時言未馳而成化當常現常之世民仰德而知遵及乎晦影歸真遷儀越世金容掩色不鏡三千之光麗象開圖空端

四八之相於是微言廣被拯含類於三塗遺訓遐宣導群生於十地然而真教難仰莫能一其旨歸曲學易遵邪正於焉紛糺所以空有之論或習俗而是非大小之乘乍沿時而隆替有玄奘法師者法門之領袖也幼懷貞敏早悟三空之心長契神情先苞四忍之行

松風水月未足比其清華仙露明珠詎能方其朗潤故以智通無累神測未形超六塵而迥出隻千古而無對凝心內境悲正法之陵遲栖慮玄門慨深文之訛謬思欲分條析理廣彼前聞截偽續真開茲後學是以翹心淨土往遊西域乘危遠邁杖策孤征積雪晨飛

途間失地驚砂夕起空外迷天萬里山川接煙霞而進影百重寒暑躡霜雨而前蹤誠重勞輕求深願達周遊西宇十有七年窮歷道邦詢求正教雙林八水味道餐風鹿苑鷲峰瞻奇仰異承至言於先聖受真教於上賢探賾妙門精窮奧業一乘五律之道馳驟於心



Images from Whole Qisha Canon

001_Page_112.jpg

```
1        大唐三藏聖教序 天 一
2        大宗支皇帝製
3        蓋聞三儀有像顯覆載以含生四時無形潛
4        寒是以化物是以窺天觸地庸愚皆識其端
5        明陰洞陽賢哲罕窮其數然而天地苞乎陰
6        陽而易識者以其有像也陰陽處乎天地而
7        乘仁刊
8        難窮者以其無形也故知像顯可徵雖愚不
9        惑形潰莫覩在智猶迷況乎佛道崇虛乘斷
10       控寂弘濟萬品典御十方擧威靈而無上抑
11       神力而無下大之則彌於宇宙細之則攝於
12       毫聲無滅無生歷千劫而不古若隱若顯連
13       可福而長今妙道凝玄邊之莫知其際法流
14       湛寂挹之莫測其源故知蠢蠢凡愚迴區庸
15       鄙投其旨趣能無疑惑者哉然則大教之興
16       基乎四十騰漢庭而皎夢
17       者分形分跡之時言未馳
18       之世民何德而知遵及乎
19       世金容掩色不鏡三千之
20
21       磧 砂 大 藏 經 .大般若
```

001_Page_113.jpg

```
1        磧 砂 大 藏 經 大般若波
2        田八藏三篋之文波濤於口海爰自所歷之
3        國紵將三藏要支凡六百五十七部譯布中
4        夏宣 勝業引慈雲於西極法法兩於東無
5        聖教缺而復全養生罪而還福濕火宅之莫
6        焰共拔迷途朗愛水之界波同臻彼岸是知
7        惡因業墜善以緣是是墜之崇惟人所託譬
8        天 三
9        夫桂生高嶺雪露方得泫其花蓮出澡波飛
10       摩不能法其棄非蓮性自潔而桂質本貞良
11       由所附者高則微物不能界所無者淨則濁
12       類不能法夫以卉木無知猶貪善而成善況
13       乎人倫有識不緣慶而求 慶方異慈經流施
14       將日月而無窮節福遐數與歎坤而永大
15       大唐三藏聖教訶 高宗皇帝在春宮見制
         正教非智無以廣其支崇闡微言非
         定其旨蓋真如聖教者諸法之玄宗
         軌躅也 捨 遠與旨遐深極空有
         體生滅之機要訶茂道曠尋之者不
         支羅義斷願之者莫測其際故知聖
```

```
1        四八之相於是微言廣被拯含類於三塗遣
2        訓遐宣導群生於十地然而眞教離仰莫能
3        一 其旨歸曲學易邊邪正於焉紛亂所以空
4        有之論或習俗而是非大小之乘乍泣時而
5        薩昔有玄奘法師者法門之領神也初懷貞
6        敏早悟三空之心長契神情先苞四忍之行
7        松風水月未足此其情華仙露明珠記能方
8        其朗澄故以智通無界神測未形超六塵而
9        逈出隻千舌而無對礙心內境悲正法之陵
10       遲栖慮玄門慨深支之訛謬思欲分修所理
11       廣彼前聞載爲讀眞開慈後學是以超心淨
12       土往遊四域乘危遠邁杖策孤征積雪晨飛
13       途間失地驚砂多起空外迷天萬里出川撥
14       煙霞而進影百重寒暑蹣霜兩而前蹤誠重
15       勞輕求深願達周遊四宇十有七年窮歷道
16       邦訶求正教雙林八水味道餐風鹿如鷲烽
17       瞻奇何畢承至言於先聖受眞教於上賢探
18       妙門精窮與業一 乘五律之道 驟於心
19       羅蜜多 經 卍 第 一 冊
```

```
         經 卍 影印宋 元 版 三
2        慈所彼業無善而不臻妙化所處緣無惡而
3        不翦開法網之網絕弘六度之 正教拯 有
4        之塗炭聲三藏之祕扃是以名無翼而長飛
5        道無根而永固道名流慶摩說舌而鎖常赴
6        感應身經摩動而不朽是鐘多梵支三竟於
7        鷲擊慧日法流轉雙輪於鹿苑排空寶蓋接
8        翔雲而共飛莊界春林與天花而令變伏惟
9        皇帝性下 上玄貪福善拱而治八流德被
10       黎斂衽而朝萬國恩如朽骨石室歸見棄
11       之支澤及足蟲金圓流梵說之偈遂使阿耨
12       達水通神甸之 八川耆闍崛山接當華之辜
13       嶺竊以法性凝寂摩歸心而不通智地玄具
14       天 三
15       感懇識而遂顯豈謂重昏之夜燭慧想之光
16       火宅之朝降法兩之澤於是百川異流同會
17       於海萬區分義慈成乎實菩與湯武校其優
18       劣堯舜此其聖德者哉玄奘法師者夙懷瞲
19       令立志夷簡神清齬離之年體拔淨華之世
20       凝情定室匿迹幽巖栖息三神巡遊十地超
```

Accuracy 98~99%

AI authoring &
the power to create
AI solutions

Mo文oN

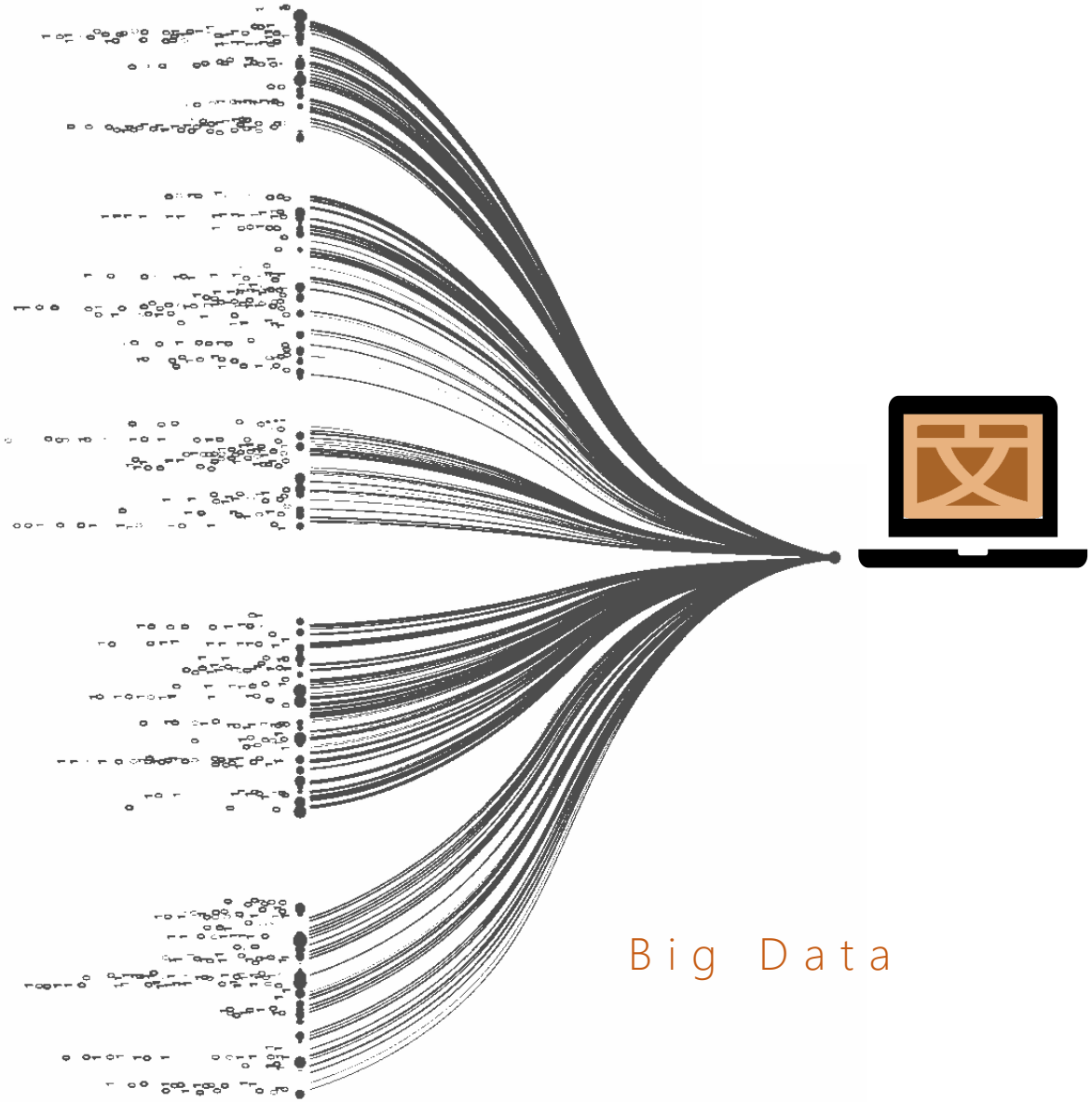
an introduction

What is AI Authoring?

Microsoft Word is an authoring tool for documents and PowerPoint is an authoring tool for presentations

MoXoN is an authoring tool for AI solutions

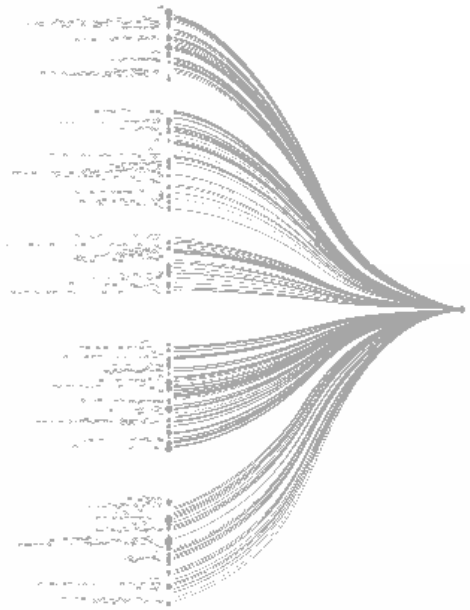
Mo文oN enables experts to quickly curate big data



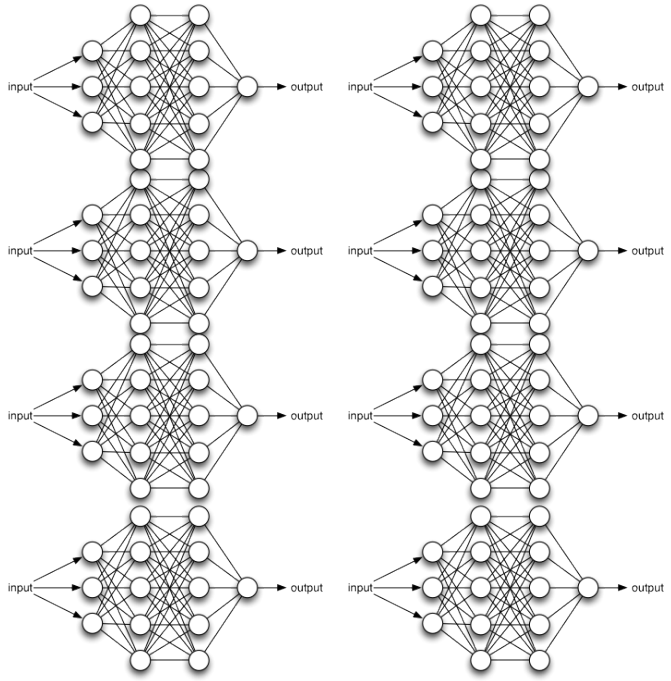
Big Data

Experts

and use various machine learning algorithms



Experts



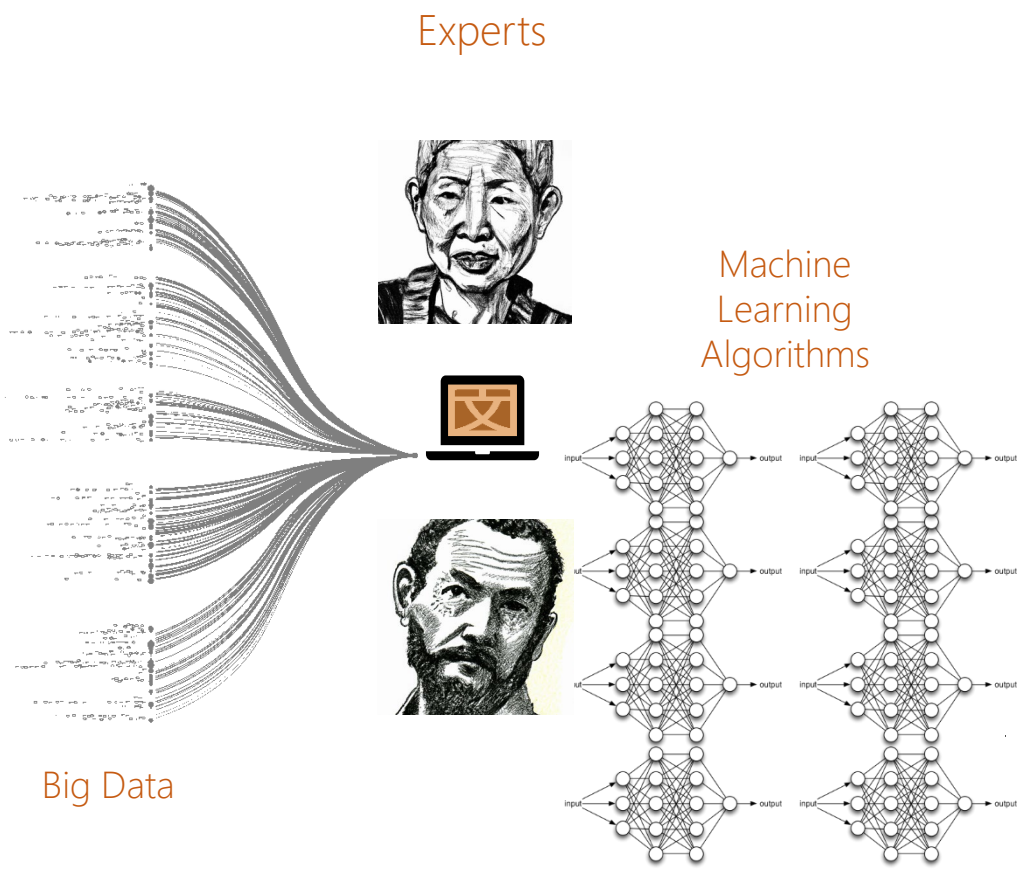
Machine Learning Algorithms

to author

custom

AI

solutions



A Few Examples,

Books of Korean Poetry

AI solutions for text encoding, an example

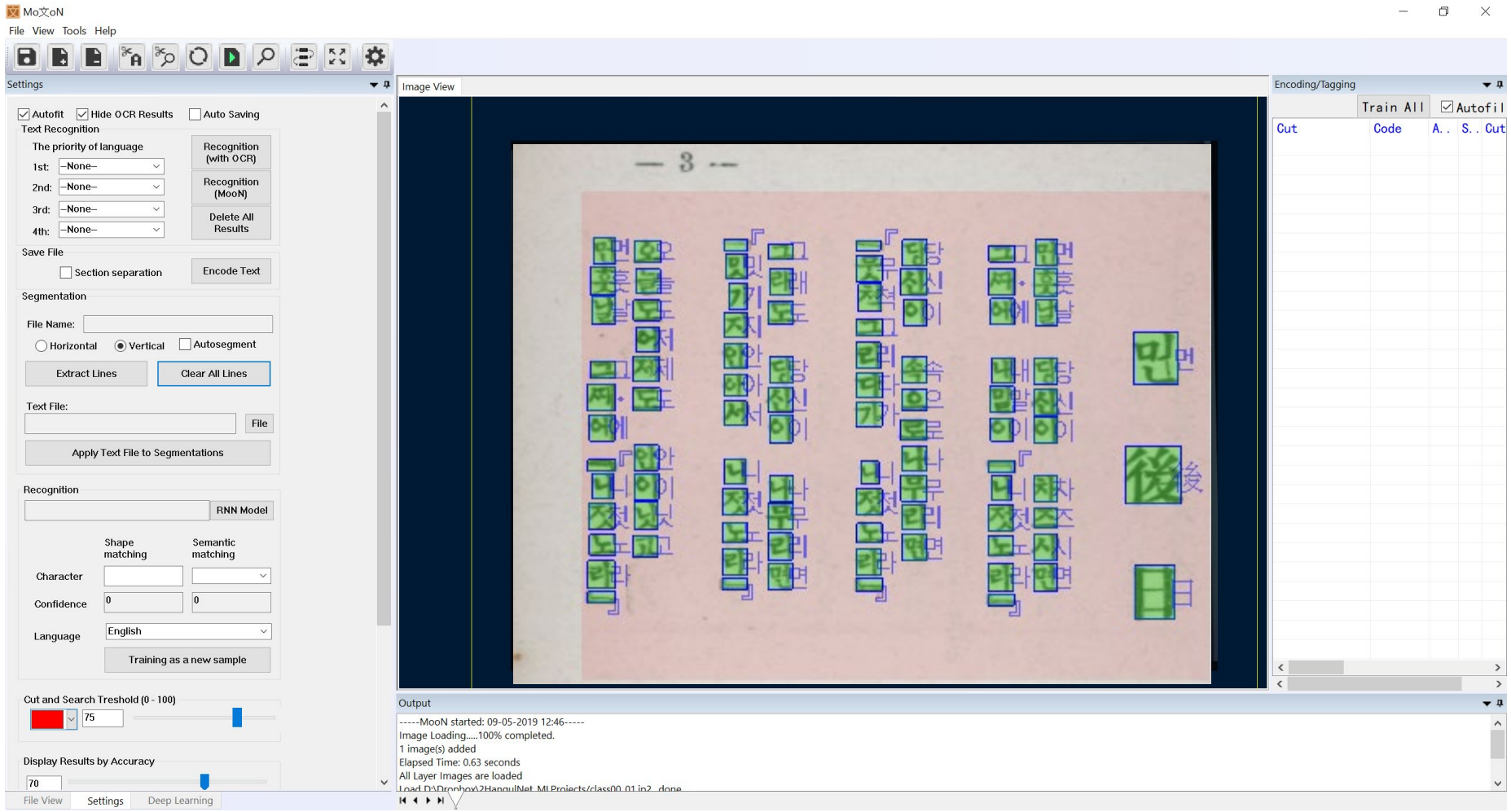
The Covers of the Hansŏng Tosŏ issue (left) and the Chungang Sŏrim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Han'guk Hyŏndaesi Pangmulgwan), respectively.

The Covers of the Hanaŭng Tosŏ issue (left) and the Changang Sŏrim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Hanʼgŭl Hyŏndaesi Pangmulgwan), respectively.

Drag and Drop images of a rare book into interface

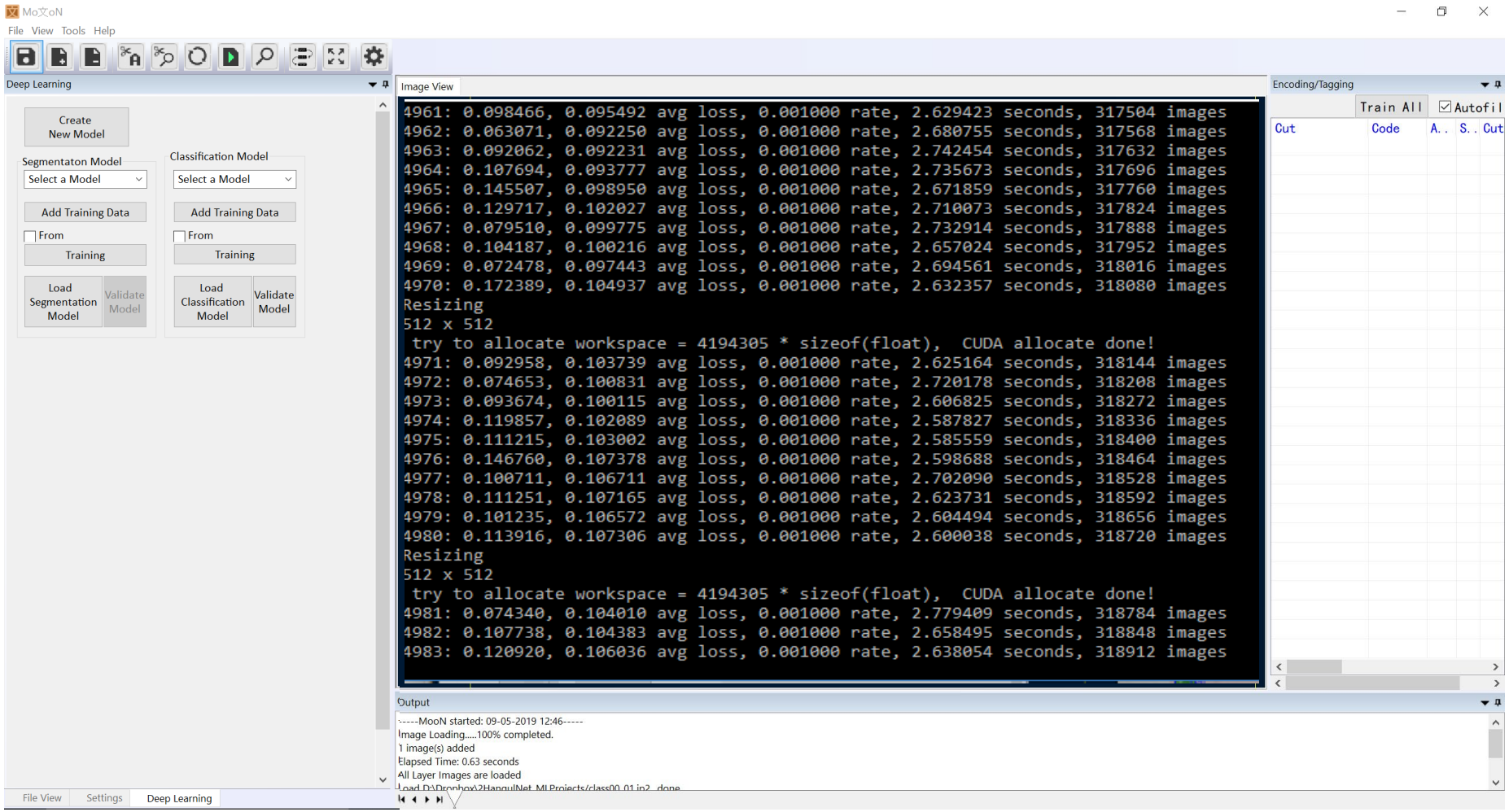
Quickly segment and label text

The Covers of the Hanaŭng Tosŏ issue (left) and the Changang Sŏrim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Hanʼgŭk Hyŏndaesi Pangmulgwan), respectively.



Use labeled images to build a custom AI solution for encoding similar rare texts, with just a few clicks

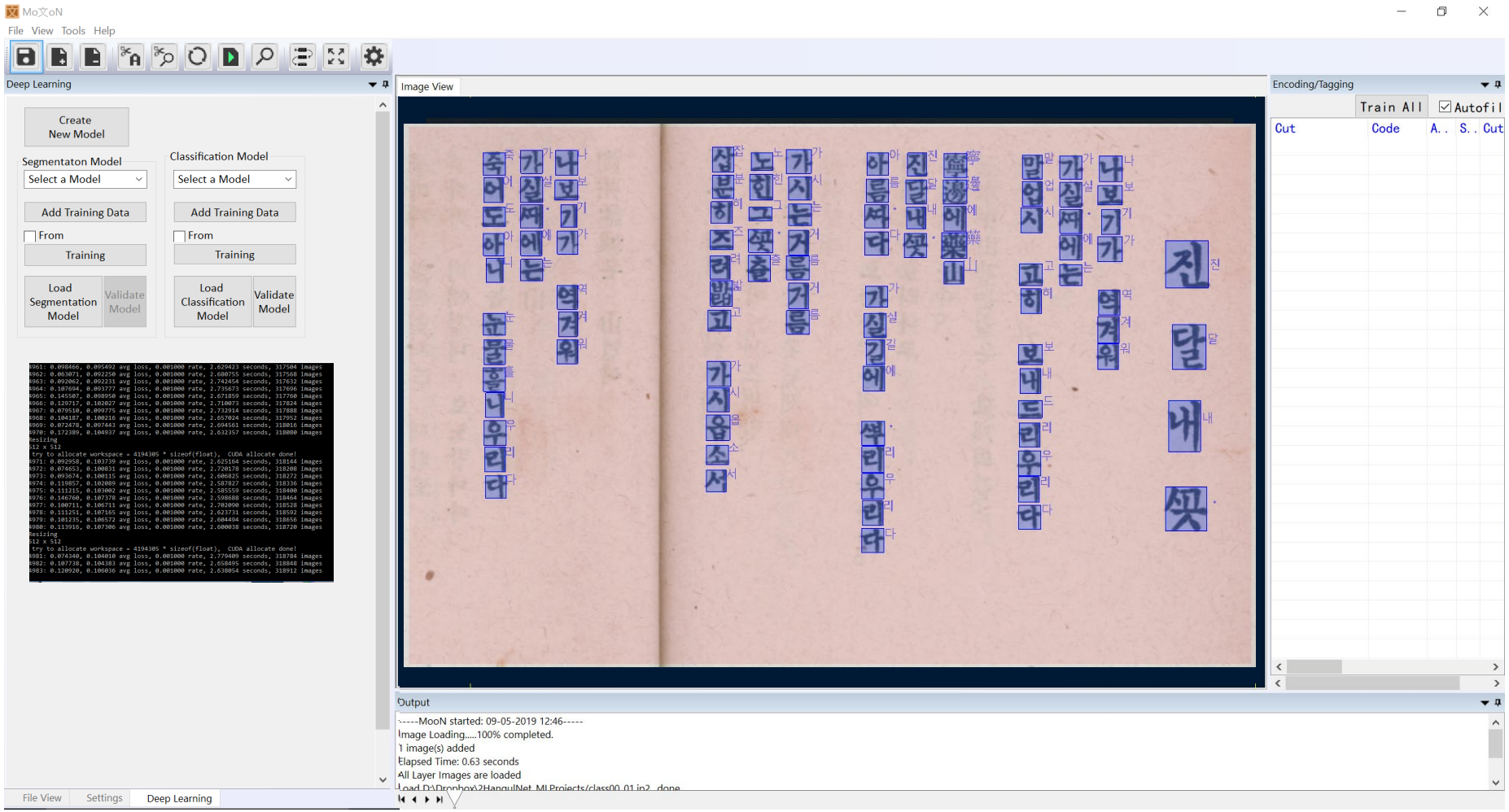
The Covers of the Hansong Yeoh issue (left) and the Changang Shim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Hanguk Hyondaesi Pangmulgwan), respectively.



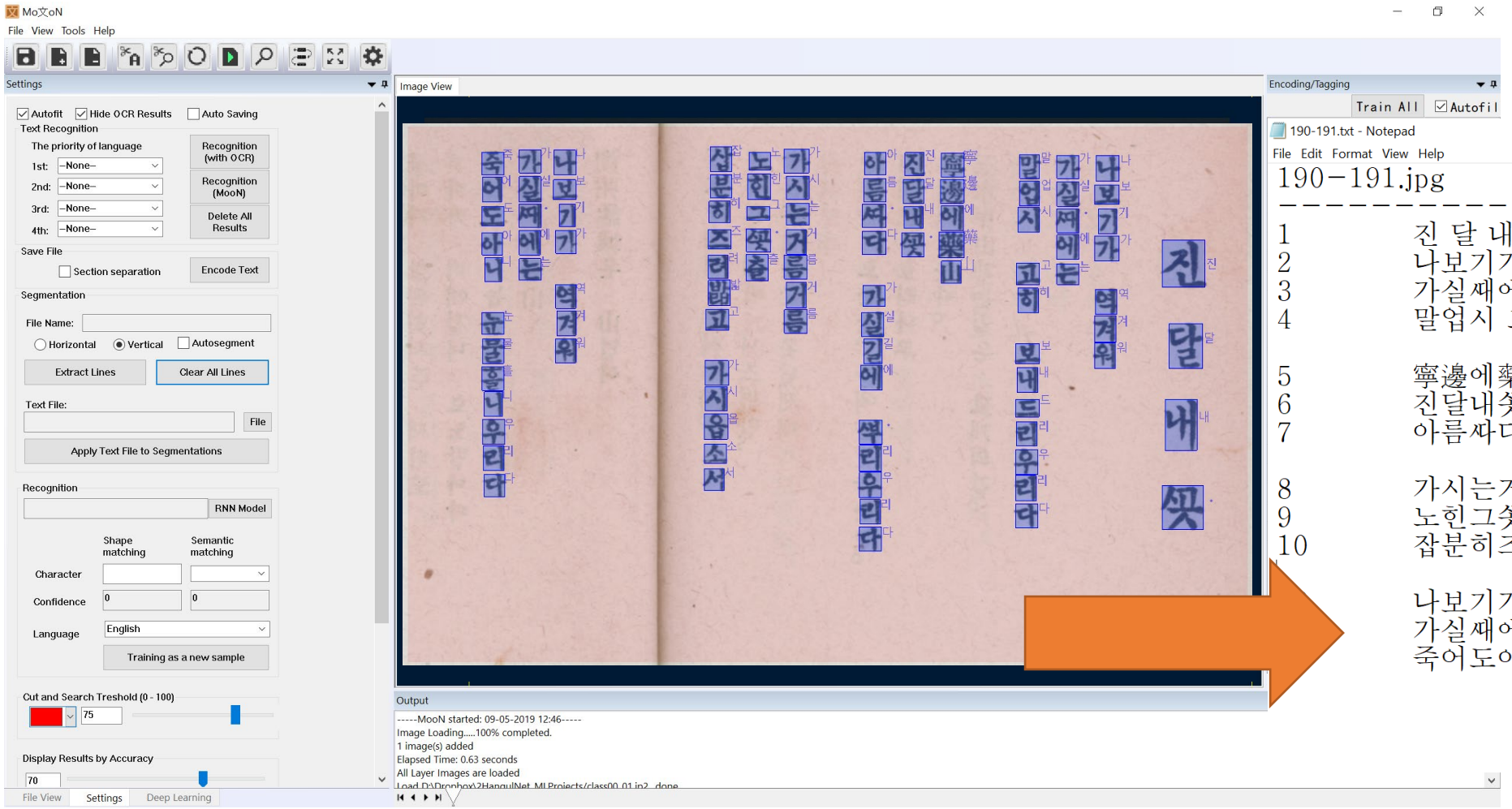
The Covers of the Hanshing Seoh issue (left) and the Changsang Shim issue (right) in the Appenzeller-Noble Memorial Museum and the Museum of Contemporary Korean Poetry (Hanʼgukʼ Hyŏndaesi Pangmulgwan), respectively.

Use custom AI solution to label and encode similar text

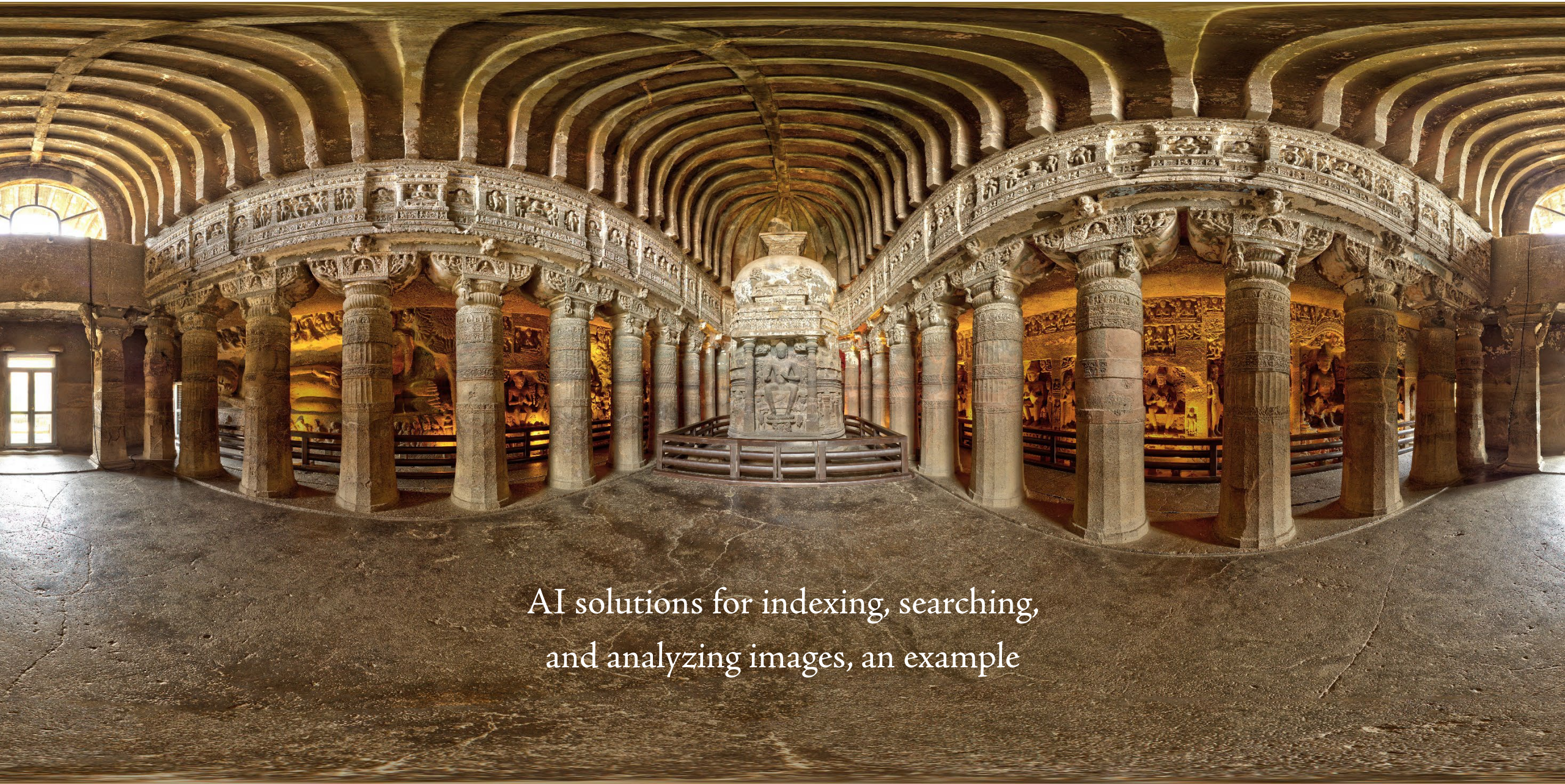
Use custom AI solution to label and encode similar text



Export encoded text and related metadata to enable additional search and analysis

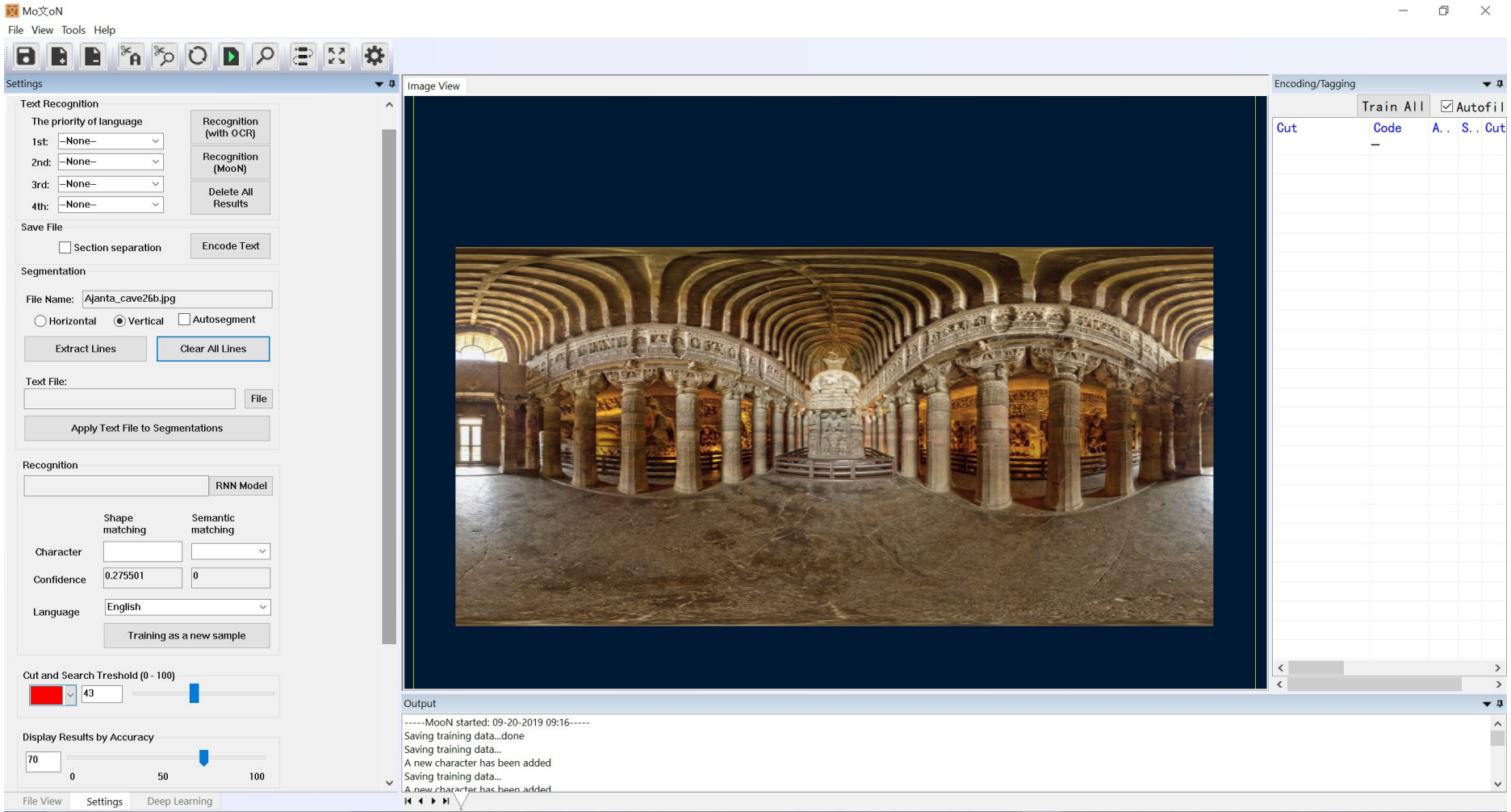


Some additional experiments,

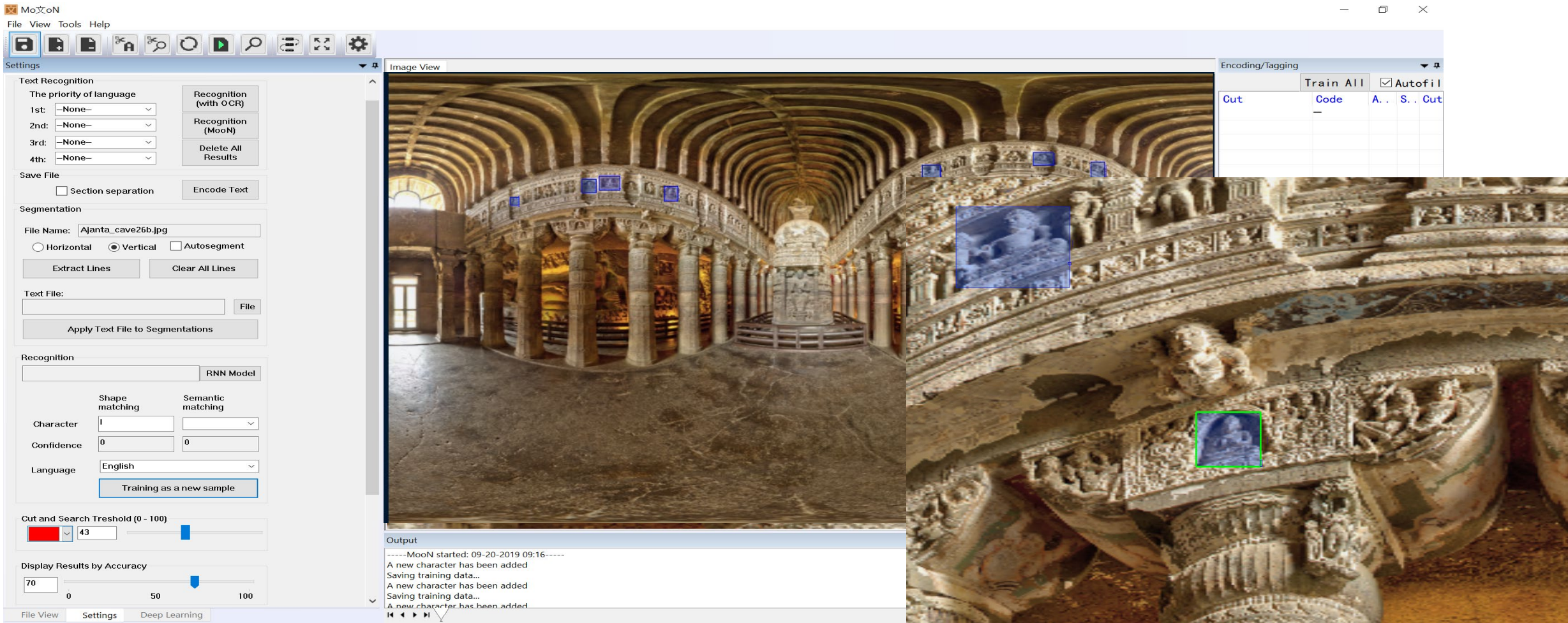


AI solutions for indexing, searching,
and analyzing images, an example

Drag and Drop image of Ajanta Cave into interface



Quickly segment and label portions of the image, the Buddha sitting in lotus position for example



Use labeled segments to build a custom AI solution for identifying similar visual elements in an image, with just a few clicks

MoonN

File View Tools Help

Deep Learning

Create New Model

Segmentaton Model

Select a Model

Add Training Data

From

Training

Load Segmentation Model

Validate Model

Classification Model

Select a Model

Add Training Data

From

Training

Load Classification Model

Validate Model

File View | Settings | Deep Learning

Image View

```
4961: 0.098466, 0.095492 avg loss, 0.001000 rate, 2.629423 seconds, 317504 images
4962: 0.063071, 0.092250 avg loss, 0.001000 rate, 2.680755 seconds, 317568 images
4963: 0.092062, 0.092231 avg loss, 0.001000 rate, 2.742454 seconds, 317632 images
4964: 0.107694, 0.093777 avg loss, 0.001000 rate, 2.735673 seconds, 317696 images
4965: 0.145507, 0.098950 avg loss, 0.001000 rate, 2.671859 seconds, 317760 images
4966: 0.129717, 0.102027 avg loss, 0.001000 rate, 2.710073 seconds, 317824 images
4967: 0.079510, 0.099775 avg loss, 0.001000 rate, 2.732914 seconds, 317888 images
4968: 0.104187, 0.100216 avg loss, 0.001000 rate, 2.657024 seconds, 317952 images
4969: 0.072478, 0.097443 avg loss, 0.001000 rate, 2.694561 seconds, 318016 images
4970: 0.172389, 0.104937 avg loss, 0.001000 rate, 2.632357 seconds, 318080 images
Resizing
512 x 512
 try to allocate workspace = 4194305 * sizeof(float),  CUDA allocate done!
4971: 0.092958, 0.103739 avg loss, 0.001000 rate, 2.625164 seconds, 318144 images
4972: 0.074653, 0.100831 avg loss, 0.001000 rate, 2.720178 seconds, 318208 images
4973: 0.093674, 0.100115 avg loss, 0.001000 rate, 2.606825 seconds, 318272 images
4974: 0.119857, 0.102089 avg loss, 0.001000 rate, 2.587827 seconds, 318336 images
4975: 0.111215, 0.103002 avg loss, 0.001000 rate, 2.585559 seconds, 318400 images
4976: 0.146760, 0.107378 avg loss, 0.001000 rate, 2.598688 seconds, 318464 images
4977: 0.100711, 0.106711 avg loss, 0.001000 rate, 2.702090 seconds, 318528 images
4978: 0.111251, 0.107165 avg loss, 0.001000 rate, 2.623731 seconds, 318592 images
4979: 0.101235, 0.106572 avg loss, 0.001000 rate, 2.604494 seconds, 318656 images
4980: 0.113916, 0.107306 avg loss, 0.001000 rate, 2.600038 seconds, 318720 images
Resizing
512 x 512
 try to allocate workspace = 4194305 * sizeof(float),  CUDA allocate done!
4981: 0.074340, 0.104010 avg loss, 0.001000 rate, 2.779409 seconds, 318784 images
4982: 0.107738, 0.104383 avg loss, 0.001000 rate, 2.658495 seconds, 318848 images
4983: 0.120920, 0.106036 avg loss, 0.001000 rate, 2.638054 seconds, 318912 images
```

Encoding/Tagging

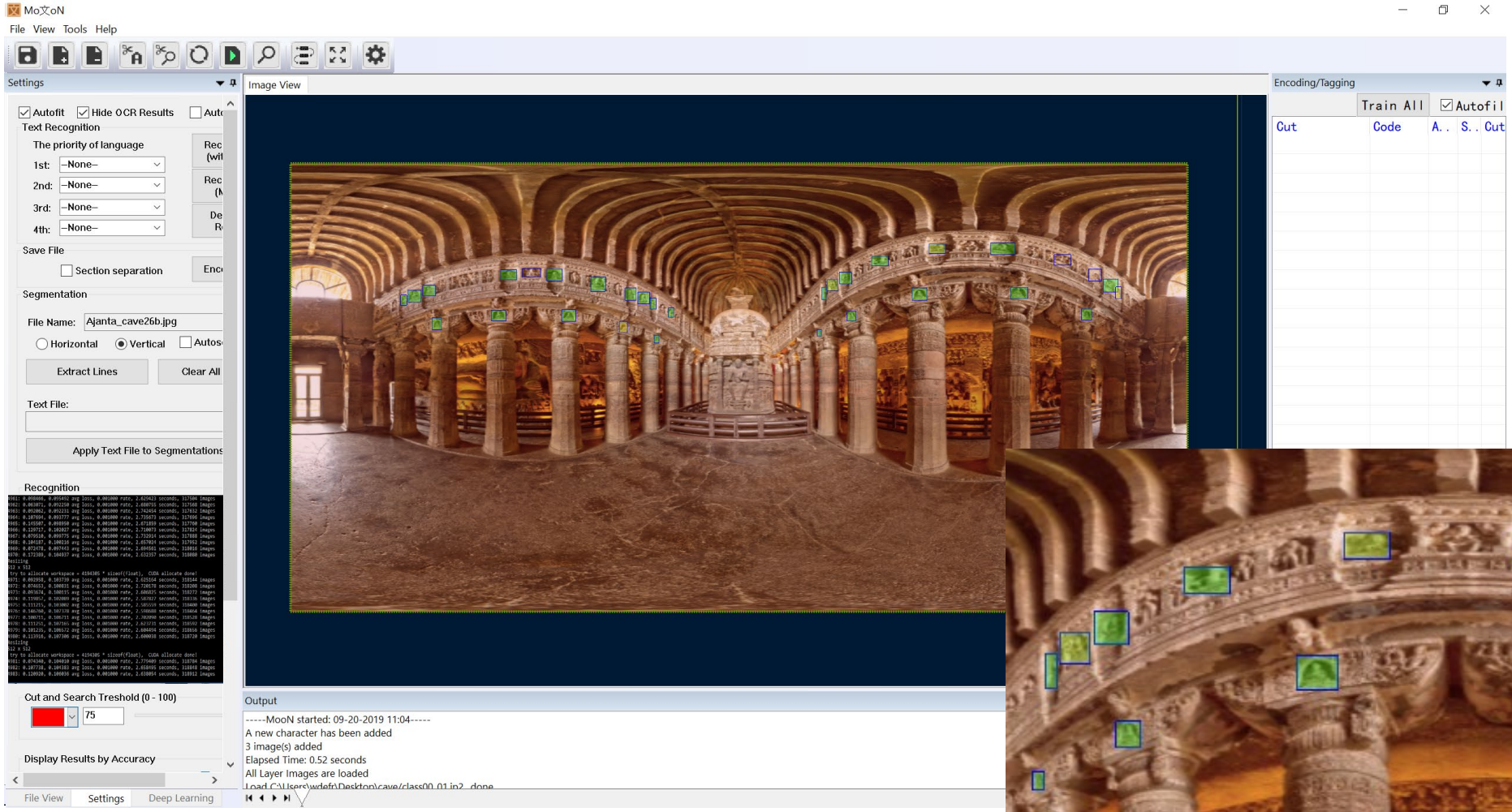
Train All | Autofil

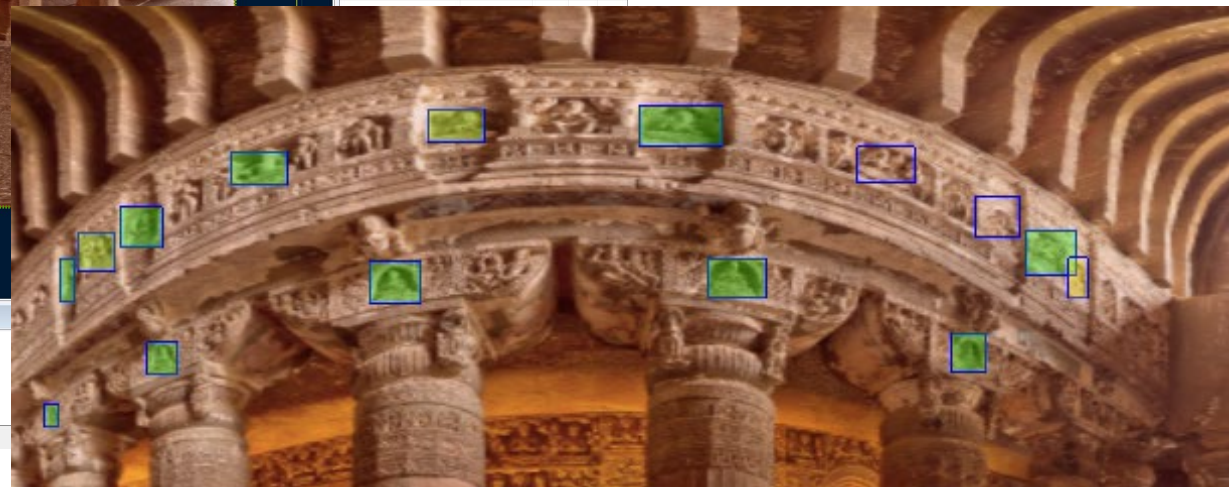
Cut | Code | A.. | S.. | Cut

Output

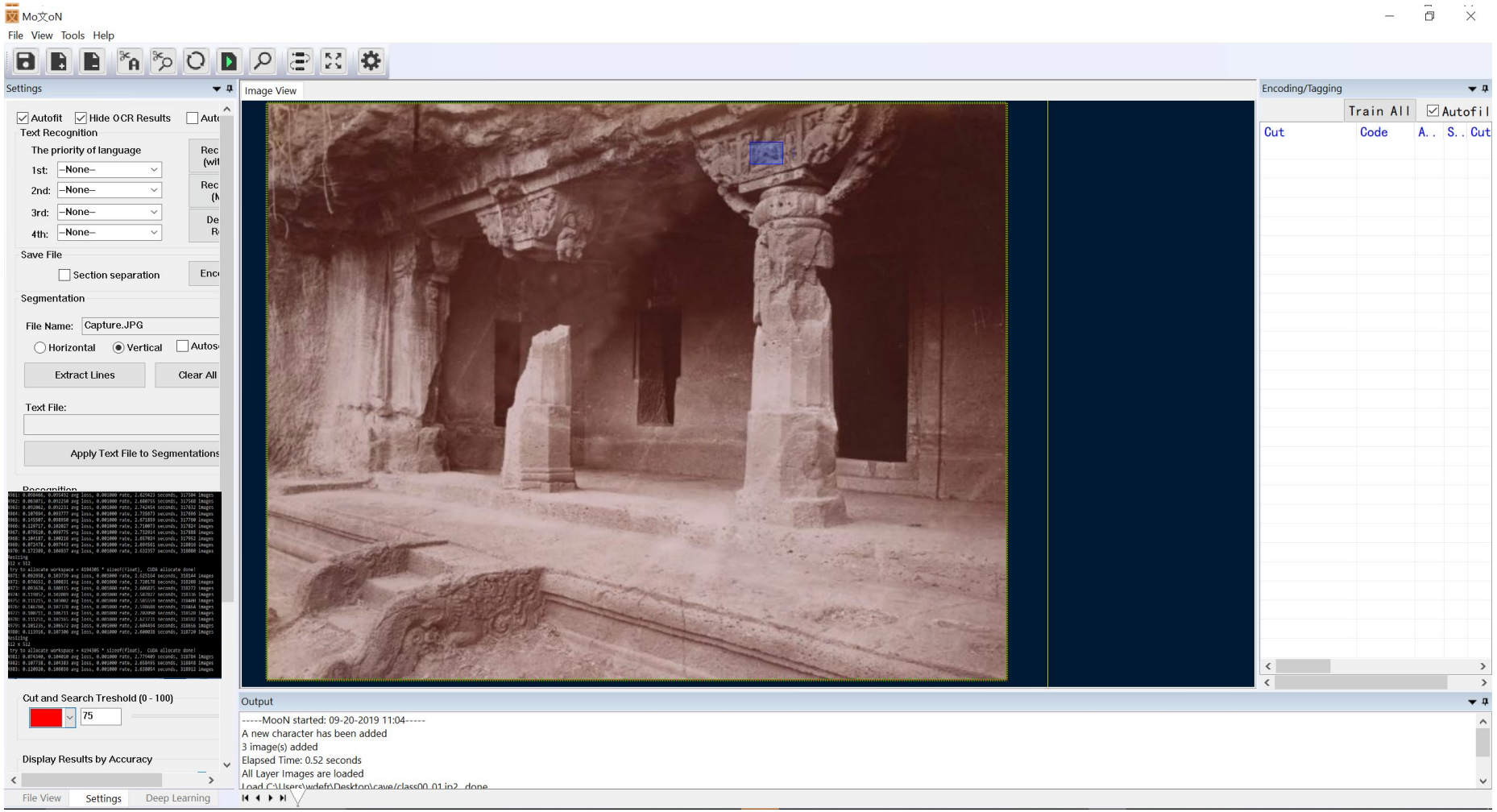
```
-----MooN started: 09-05-2019 12:46-----
Image Loading.....100% completed.
1 image(s) added
Elapsed Time: 0.63 seconds
All Layer Images are loaded
Load D:\Dropbox\2HangulNet_ML Projects\class00_01.in2  done
```

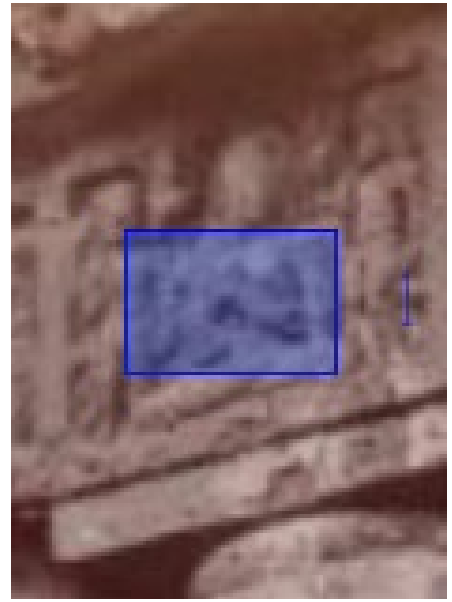
Use custom AI solution to segment and tag similar elements in the image of Ajanta Cave





Use the same custom AI solution to search, segment, and label related images, here a picture of the façade of Ajanta Cave XX from 1865





Export labeled images, segments, and related metadata from growing database of images and labels to empower additional forms of search and analysis

How Ai Works

useful basics for librarians and researchers

Wayne de Fremery
Dominican University of California