

数字人文+ OpenGLAM:
大模型背景下的大资源建设和共享

《*OpenGLAM*概念下利用“人工智能”
进行数字人文研究的思考》



2023年9月15日

阎焰

深圳望野博物馆

数据浪潮下，美术馆 (Galleries)、图书馆 (Libraries)、档案馆 (Archives) 及博物馆 (Museums) 多链共享后，看似人文研究所能利用的数据文本和图像信息以几何数级的规模膨胀。原本以为可以扩大的便捷，瞬间又被海量信息标的对应时间、空间的标位甄别苦恼所覆盖。

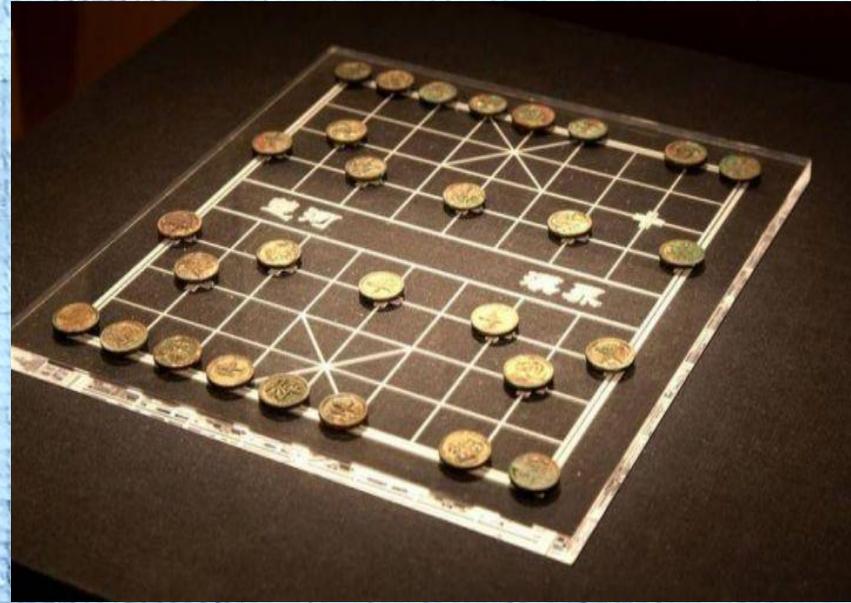
数字人文数据同物理、化学、植物、动物等等类目的标定数据有着明显的不同。

如何对“人工智能”进行数字人文研究的标的确准问题，让人文数据更准确、更有效；陡然变成一个全新的迫在眉睫的课题。

数字浪潮参与者必须深入思考。

数字人文+ OpenGLAM 是否是“人文研究的福音”呢？

- 有了前面的疑问。我们将直接面对一个关键问题。
- “人工智能”基于互联网的、可用数据来训练的、文本生成的深度学习模型——GPT，可以代替“人”读书吗？
- 或者说GPT，可以代替“人”学习吗？



阿尔法Go

谷歌人工智能机器人



VS

柯洁

世界围棋等级分第一棋手



- 通过前面的思考和了解，我们可以发现 GPT，可以代替“人”的部分阅读、思考和学习。
- 但将此思考放到更进一步地“人文”深入研究中观察时，会出现什么情况呢？

忘開襟春葉短分手夏條長獨有幽庭桂年年空自芳

奉使巡檢兩京路種果樹事畢入秦因詠歌

鄭審

聖德周天壤韶華滿帝畿九重承渙汗千里樹芳菲陝
塞餘陰薄關河舊色微發生和氣動封植衆心歸春露
條應弱秋霜果定肥影移行子蓋香撲使人衣入徑迷
馳道分行接禁闈何當扈仙蹕攀折奉恩輝

銜命使本國

胡衡

銜命將辭國非才忝侍臣天中戀明主海外憶慈親伏
奏違金闕駢驂去玉津蓬萊鄉路遠若水故園鄰西望
懷恩日東歸感義辰平生一寶劔留贈結交人

送金鄉歸新羅

陶翰

奉義朝中國殊恩及遠臣鄉心遙度海客路再經春落
日誰同望孤舟獨可親拂波銜水鳥偶宿泣珠人禮樂
夷風變衣冠漢制新青雲已千侶知汝重來賓

和中丞出使恩命過終南別業

贈李郎中

仙郎早朝退直省卧南軒院竹自成賞階庭寂不喧焚
香開後閣起草閉前門禮樂風流美光華星位尊榮兼
朱紱貴交乃布衣存是日登龍客無忘君子恩

胡衡

銜命使本國

銜命將辭國非才忝侍臣天中戀明主海外憶慈親伏
奏違金闕駢駮去玉津蓬萊鄉路遠若木故園鄰西望

懷恩日東歸感義辰平生一寶劍留贈結交人

張謂

同諸公遊雲公禪寺

共許尋難足誰能惜馬蹄長空淨雲雨斜日半虹霓簷
下千峰轉牕前萬木低看花尋徑遠聽鳥入林迷地與
喧聞隔人將物我齊不知樵客意何事武陵溪

梁鍾

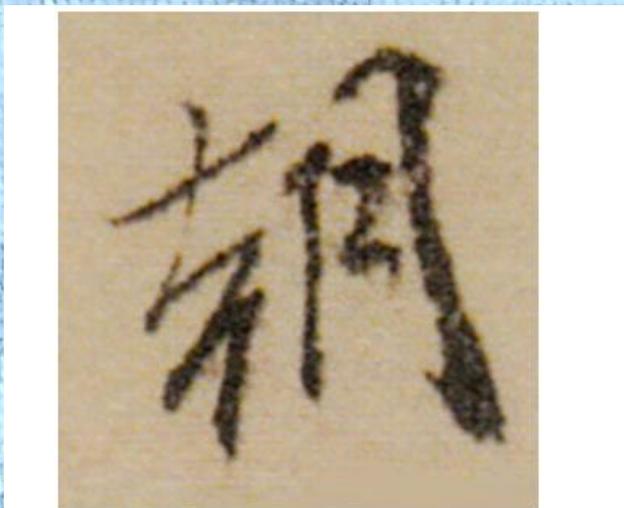
方進士恒春草



唐·颜真卿·祭伯父文稿(朝)



宋·米芾·珊瑚帖(朝)



宋·米芾·褚临黄绢本兰亭序跋赞(朝)



宋·米芾·方圆庵记(朝)

胡衡

銜命使本國

銜命將辭國非才忝侍臣天中戀明主海外憶慈親伏
奏違金闕駢駮去玉津蓬萊鄉路遠若木故園鄰西望
懷恩日東歸感義辰平生一寶劍留贈結交人

銜命使本國

胡衡

銜命將辭國非才忝侍臣天中戀明主海外憶慈親伏
奏違金闕駢駮去玉津蓬萊鄉路遠若水故園鄰西望
懷恩日東歸感義辰平生一寶劍留贈結交人

思鄉作

瀟北行人絕雲南信未還庭前花不掃門外柳誰攀坐
久銷銀燭愁多減玉顏懸心秋夜月萬里照關山

朝衡

品粟作
胡衡

朝衡字巨卿日本人開元初日本王聖武遣其臣粟田
副仲滿來朝請從諸儒授經仲滿慕華不肖去易姓名
曰朝衡歷左補闕久之歸國上元中擢散騎常侍詩一
首

欽定四庫全書

御定全唐詩

卷七百三十二

銜命還國作

銜命將辭國非才忝侍臣天中戀明主海外憶慈親伏
奏違金闕駢駢去玉津蓬萊鄉路遠若木故園林西望
懷恩日東歸感義辰平生一寶劍留贈結交人

長屋

長屋日本相國也詩一首

繡袈裟衣

緣明皇時長屋嘗造千袈裟繡偈於
衣緣來施中華真公因泛海至彼

國傳
法馬

- 銜命使本國

- [唐代] 朝衡

- 銜命將辭國，非才忝侍臣。天中戀明主，海外憶慈親。

- 伏奏違金闕，駢驂去玉津。蓬萊鄉路遠，若木故園林。

- 西望懷恩日，東歸感義辰。平生一寶劍，留贈結交人。

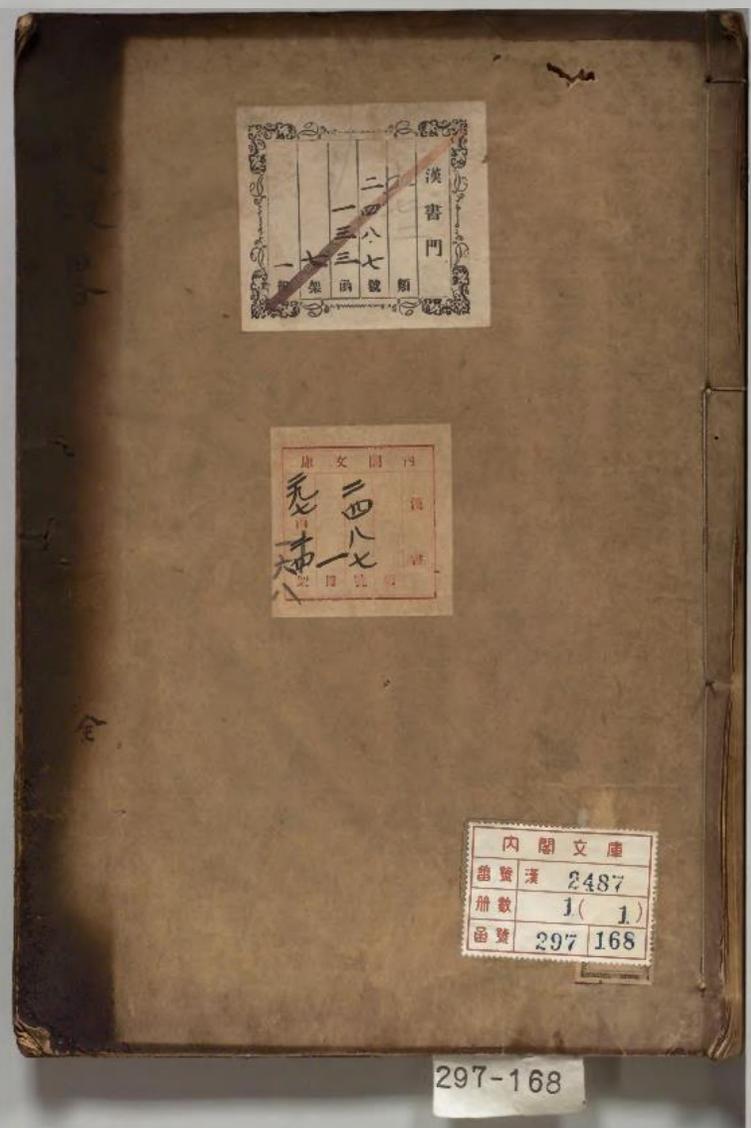


林罗山，幕府儒官，江户林家的始祖。
名信胜，法号道春，又称罗浮子。
京都人，曾入建仁寺为僧。

林罗山，第三子林恕（春斋）编撰《罗山先生文集》在卷三十七文末附录了，其四弟林靖（春德）关于“**胡衡**”是“**朝衡**”讹误的观点。

歷代叙略題辭
 淺草文庫

自有宇宙以來上下千萬年其亶載于簡冊汗
 牛充棟何可勝紀有能舉其要于一編使讀之
 者不待覽全史而瞬息之間盡非得失已瞭然
 乎心目亦可謂奇矣君臨江梁寅氏歷代叙略
 是也梁氏以史學自負其於十七史并宋元史
 皆有節略此編則節之又節者蓋自上古至有
 元國統之離合時運之盛衰年世之久速君臣
 之賢否貫之以史辭申之以議論為篇者僅十



內閣文庫
 總號 漢 2487
 冊數 1 (1)
 番號 297 168

297-168

國立公文書館
 National Archives of Japan

內閣文庫 林氏藏書 明代 (江戶初) 林羅山寫本 (國立公文書館) 來自数据库

- 前述研究使用目前已知的计算机和大型工作站或者说GPT，它无论如何也得不出这个看似极为简单的人文研究答案。
 - 但无疑从“胡衡”到“朝衡”的研究思考，可以发现数据库对“人文研究”意义巨大。
 - 数据时代的浪潮，是每个人必须面对的，这浪潮必然改变所有人的工作、学习、生活环境。
 - 但就目前所见之现实，在相对漫长的时间中，“人工智能”无法在一些独特领域取代灵长目进化的天花板——“人”的“智能”。
-
- “学习”是“人”驾驭“机械（电脑）”最强有力的武器。也是人类历史性生存的根本。

谢谢各位垂听！

部分图片来自网络，版权为原始权益人所有，谨此致谢！