

新疆维吾尔自治区乌鲁木齐市2022届高三三模语文试卷

现代文阅读（36分）

1. 现代文阅读

阅读下面的文字，完成问题。

人们习惯用史料来描述文学史。而在当今大数据时代，我们或可尝试用大数据来显示和验证，并透视数据背后的文学史意义，进而反思数据存在的问题和展望未来努力的方向。唐宋诗词世界华茂繁荣，大数据可带我们领略其中新的风采。

从个体诗人来看，唐诗宋词里谁的作品最多？数据显示，白居易名列唐诗作品量的榜首，诗作近三千；杜甫和李白紧随其后，都超过千首大关。宋词中辛弃疾的词作量位居第一，有六百余阕，其次是苏轼、刘辰翁的。宋诗的篇数，则由陆游称雄，凡九千多首，其次是刘克庄和杨万里的。

数据不仅标示谁第一、谁领先，还蕴含更深层的文学史意义，可以探究其背后隐藏的历史过程 and 变化动因，揭示规律性认识。比如，通过检视和分析唐诗名篇崔颢《黄鹤楼》的各种数据，可以发现第一接受者李白的评价和创作比拼发挥了重要作用。可见，除自身的艺术含量外，故事传播和意见领袖的评价也是名篇生成的重要因素。

数据还反映出诗人影响力与作品量的关系。杜甫、李白的诗作量名列前茅，辛弃疾、苏轼的词作量也是第一、第二。由此可以说，影响力大的诗人，其作品量也很大，但不能说作品量大就一定影响力大。白居易诗作量位居唐代诗人之首，综合影响力却没进入前十名。陆游的诗篇比苏轼多出两倍，综合影响力却在苏轼之后。由此可见，精品名作多才是成为影响力大的诗人的决定性因素。

虽然数据能在一定程度上描述、显示文学史的发展面貌和进程，揭示隐藏在历史背后的真相与意义，但数据也有明显的局限性。一是效度有限，只能说明文学史的部分问题，而无法有效解决所有问题。二是信度受限，无论是人工挖掘还是机器自动提取，受人的认识水平和技术水平的制约，数据的信度会有优劣高下的差异。

由于指标种类少，目前衡量名篇名家的数据还不充分完备，尤其缺少内质型数据，因此需要持续不断地挖掘和充实。挖掘数据的难度，不仅在于到哪里找数据，更在于要找哪些数据。文学史上从来没有数据意识和数据存量，什么样的数据对衡量一个时代或一个时期的文学繁荣程度、评价作家作品的优劣高低有用有效，这既是实践操作问题，也是理论认识问题。

大数据时代的文学数据，需要分类分层建立起文学史数据的指标体系，以确保数据的信度和效度。也许应该考量以下三个层面：形式层面，体裁体式的创新度、丰富度；内容层面，表现主体内在精神世界和外在社会生活世界的深广度；艺术层面，表现技巧的难度和艺术表现的完美度等。有了完整的指标体系，才有目标和途径去寻找发掘数据，才有可能借助人工智能技术，让计算机自动分析提取数据，不断扩大数据的种类和体量。在此基础上，通过计算机自动提取获得海量数据，不断满足大数据时代学术研究的需求。

（摘编自王兆鹏《大数据里的唐宋诗词世界》）

(1) 下列关于原文内容的理解和分析，不正确的一项是（ ）