

数字出版技术国家重点实验室 让移动阅读 “赏心悦目”

文·本报记者 管晶晶

6月26日,中华数字书苑再次被作为国礼,送给俄罗斯。这对北大方正集团数字出版技术国家重点实验室主任、北京大学计算机技术研究所研究员汤帆来说,已经不是什么新鲜事了。2009年2月,时任国家总理温家宝首次向英国剑桥大学赠送中华数字书苑,迄今,中华数字书苑已经第15次成为国礼。

与传统的丝绸、瓷器等不同,中华数字书苑是持续出现在国礼上的第一个数字资源产品。

这也表明我国数字出版技术的先进程度。

不过,数字出版技术国家重点实验室的研究内容可不止于此。该实验室结合国家数字出版领域的发展战略和应用需求,重点研究数字出版领域中结构化版式文档技术、自动排版技术、数字出版内容的自适应重组技术等关键技术,促进数字出版产业的技术进步,并在电子书、数字报等领域进行产业化转换和推广,推动数字出版产业相关标准的制定。

CEBX技术堪称“中国的Adobe”

随着移动互联网的兴起,移动阅读逐渐成为大趋势,人们随时随地使用手机、PAD、Kindle等终端进行阅读,但是其终端呈现格式却纷繁复杂。PDF、EPUB、TXT、HTML等格式无一具备在电脑等大屏上保证美观排版的同时,又可以在手机、PAD等小屏幕上根据屏幕大小自动换行和重排的功能。

“我们的CEBX技术在版式和流式的融合技术方面是领先的,既可以实现原版原式,又可以支持流式阅读。”汤帆告诉科技日报记者。

CEBX是方正研发的新一代普适性文档技术,可以做到复杂排版情况的矢量处理,如数学公式、化学公式、物理公式。

目前,国内比较流行的阅读呈现格式主要有两种,PDF格式主要呈现书籍报刊的版式,EPUB格式主要用于手机等小屏幕的移动阅读。这两种格式都是由美国的公司主导开发,市场份额巨大。但是这两项技术并不能很好的融合起来。为满足不同终端的阅读需求,出版社需要生成多种格式的电子图书或数字报纸,读者也需要在使用不同屏幕阅读时安装不同的阅读器,甚至需要转换文件类型,非常不便。

自动排版技术助力“全民阅读”

今年6月,方正阿帕比全民阅读机系统发布;而此次与“中华数字书苑”共同作为第15次“国礼”的就是这款新产品。目前的全民阅读机具备基础版和扩展版两个版本。基础版由触摸屏、标准化的数字资源、标准化的界面、阅阅盒子等组成,适用于大型文化推广项目,大量部署于场馆、社区、街道等地,内容由项目负责人统一挑选、安装、维护便利。扩展版由触摸屏硬件和能自定义的系统及能自由选择的资源组成。拓展版全民阅读机的系统不仅具备资源展示阅读和推送功能,还可以加入资讯发布、单位介绍、楼宇导航、馆藏查询、预约等多种展示与交互的需求。拓展版全民阅读机的资源类型不仅包括基础版的电子书、数字报,还包括图片、视频等其它资源,且通过知识服务数字挖掘技术,能够实现包括专题在内的内容运营服务,同时资源数量上还能与中华数字书苑全库的海量资源对接。扩展版全民阅读机更适用于图书馆、企事业单位等有比较多个个性化展示需求的机构。

表面上看,承载这些功能的只是一个小小的

阅阅盒子和触摸屏硬件;而这背后的支撑却是实验室多年研发的自动排版技术。

自动排版技术主要内容是,将结构化或半结构化的数字内容与版面样式结合,对排版结果进行智能化的校正,最后输出适合不同发布渠道需要的排版结果。同时要保证制作结果的标准化和高质量,与专业交互式制作系统所达到的水准相一致。

各图书馆可以根据实际情况选择不同尺寸的触摸屏显示屏,读者使用的手机、PAD也是大小各异。怎么能让不同终端上显示的电子书都排版精美,且符合中文使用习惯?自动排版技术功不可没。

“全民阅读机迎合了移动互联网时代下,读者对多屏展示需求、多屏互动分享、移动阅读的需求。图书馆使用的屏幕大,分辨率也高,导致数据展示时需要生成的数据量大,自动排版技术会根据不同屏幕分辨率大小,自动计算适合在屏幕上显示的文字数量,对全文内容进行重新排版,以舒适的方式最终呈现给读者。”汤帆介绍说。

■ 创业故事

余福元:做武汉最大最暖心的“二房东”

文·谢晓丽 雷宇

清晨6点的江城武汉,天已放亮。位于武汉市徐东大街的“福星惠誉·国际城”小区开始苏醒,习惯早起的余福元,喜欢站在小区花园里向上仰望,看到租户窗口点点温馨的光晕,这位“大家长”的心才踏实。

随着小米科技CEO雷军的1000万元风投到位,余福元的“可遇”青年公寓即将突破1000间,他也成为武汉出租市场最大的“二房东”。

进入毕业季,余福元有点抑制不住内心的兴奋:武汉30万大学生毕业,成为搅动租房市场的“活鱼”。这个武汉科技大学的硕士研究生有一个朴实的梦想:在钢筋混凝土的大城市里,为青年白领安一个温暖的家。

“二房东”的互联网思维

白色家具、绿色墙纸、精致花瓶、可爱的玩偶摆件、年轻人喜爱的照片墙……走进余福元打造的“可遇”青年公寓,简约时尚的风格让人眼前一亮。“可遇”,就是可遇亦可求,遇到属于自己的家,求得更好的居住环境。”余福元解释。

据他介绍,每套公寓面积约15平方米,有卫生间、厨房、活动室,住宿、休闲、娱乐都很便利。在租客眼里,这个“二房东”很有互联网思维,比较“潮”:房屋中介、装修公司、互联网电商、

物业保洁等一条龙服务,他的公司都包了,还为年轻人打造社交圈。但最吸引年轻人的,还是他提供的贴心管家服务,定期巡房、保洁、组织房客互动交流。“租户上传保洁、报修、交房租、收快递等需求后,管家会在第一时间上门服务。”余福元有点小得意。就在上个月,他们邀请徐东片区的年轻租客们唱K、打桌球,这让大量繁忙而找不到对象的年轻人,颇为心动。“可遇”被租客们戏称为武汉版的“爱情公寓”。

一分钟路演引来两百万风投

在老师和同学眼里,余福元“很不安分”,他的创业之路从进大学就开始了。几番潮起潮落后,2012年,余福元选择来到武汉科技大学读研,缴完学费,经济上就已捉襟见肘。“租赁市场一片混乱,有些中介挂羊头卖狗肉,让身在异乡的我找一处落脚之地都颇费周折。”余福元感慨。

一次,研究生同学偶然说起服务式公寓行业的发展,让他心里一亮。联想起在北京、武汉的租房租房经历,商业细胞敏感的他,发现了商机。事后,他做市场调研、估算成本,决定试试水。

他先在武昌的南湖附近租下一套毛坯房,按照自己的设想精装修。随后,他试着把房子放到租房网,没想到一个星期就租出去了,利润丰



俄罗斯圣彼得堡国立经济大学学生体验方正阿帕比全民阅读机

知识内容重组技术打造“知识图谱”

15次成为国礼的中华数字书苑是方正阿帕比推出的专业优质华文数字内容整合服务平台,收录了250万本电子书,500多种报纸,3000万篇新闻,2000多种工具书,2000多种年鉴,35万张艺术图片,11000多集高教视频和10000张中国对外经济贸易数据。

但它不只是简单存储和提供文献及信息资源,更重要的是“利用知识服务技术将有关的知识关联起来,有机地组织起来。”

汤帆在手机上输入了“秦始皇”,立刻出来了一张有关秦始皇的“知识图谱”,除了秦始皇的生平、事迹等,还有各种关联知识,如与秦始皇相关的人物、秦国的知名人物,研究秦始皇的文献资料等等,仿佛打开了若干条“知识通道”,可以任选一条“边走边学”。

这种“知识图谱”在未来的需求也许会越来越多。

■ 行业观察

数字出版格式亟待统一

文·本报记者 管晶晶

近年来,当当、京东、中文在线等网站都发力电子阅读平台,将大量的图书资源从纸本形式转换为电子书形式销售。但是电子书加工费用居高不下,致使收支失衡。当当网副总裁王曦、多看副总裁胡晓东都表达过类似的观点:电子书的加工成本远远超过销售收入。这已成为电子书发展的一个障碍。

为什么会出现巨大的成本?目前,仅就电子书而言,市面上就有多种文本格式,这些都由单个企业自己开发。每家平台都有自己的格式,每家技术公司也有自己的格式,美其名曰“版权保护”。为了应付这些纷繁复杂的格式需求,内容商往往采用简单的办法:你要专门的格式,就自己加工。重复加工带来的是资源的浪费和成本的急剧上升,所以才有了当当网2012年“电子书卖了300万元却花了500万元的加工费”一说。这种局面已限制了各家公司在整体市场的销售收入。

为了改变这一现状,国家新闻出版广电总局早在2011年就已着手制定一系列行业和国家标准。2013年2月,我国《数字阅读终端内容呈现格式》行业标准正式发布,北大方正集团数字出版技术国家重点实验室的CEBX技术成为数字阅读呈现格式制作和阅读的行业标准。这项具有自主知识产权的技术,一直是“终结阅读行业

乱象”最被看好的成员。

“CEBX技术拥有一键解决版式和流式文件的转换困难的国际领先技术,其实这就是区别于PDF、EPUB的最大优势,也为其终结我国数字阅读终端内容呈现格式乱象提供了最坚实的技术基础,也是其能够成为行业标准最重要的原因。”汤帆告诉记者。

尽管CEBX技术被寄予了很高的期望,但目前来看,“终结阅读行业乱象”的理想并没有实现。“整个中国的电子书阅读市场还是很小,不像美国,已经有一半的阅读市场是电子书阅读。中国的电子书阅读市场更多是那些还没有出版的网络文学,而不是相对高端一些的已出版读物。”现实如此,汤帆很是无奈。

随着阅读终端的丰富,如今,出版社正面临着为不同终端要出不同版本的尴尬。“比如出一本《红楼梦》,我要出版供电脑上阅读的格式,要准备适合手机上的版本,还要准备给亚马逊Kindle的版本、给国内阅读器的版本……”汤帆说,“如果能统一标准格式,出版机构一次制作+一次发布就实现多终端应用,数字出版的成本会降低很多。希望在几年后,行业标准上升为国家标准的愿望,能够成为现实。”

只有真正解决基础性的产业瓶颈,数字出版行业才能健康有序发展。

■ 实验室动态

新疆创新实验室 落户浙大培养跨界创客

近日,全球飞行影像系统先驱——深圳市大疆创新科技有限公司(下称“大疆创新”)与浙江工商大学艺术设计学院在杭签署协议共建“大疆创新实验室”。据杭州日报报道,这是该企业在国内所建的第一个以跨界创客培养和行业产品孵化为目标的校企联合创新实验室。

“大疆创新实验室”设在浙大艺术学院。大疆创新副总裁Andy Pan表示,在未来几年内,将投入一百万建设该实验室,通过“跨、创、学、研”的全新实践教学模式,培养企业在“工业4.0时代”所需要的高素质创客类人才。同时,能为企业孵化有潜力的创新产品。

浙大艺术学院院长张建春介绍,将大疆公司的真实商业研发项目嫁接到课程中,校企协同育人,这是该院“跨界实务”教学理念的新尝试。“使学生在知识结构和理念上能与企业有效对接,尽快适应社会需求。”

此次合作缘于去年12月启动的“无梦不飞——大疆航拍无人机联合研发营”。通过首次6个月的合作,浙大艺术学院为大疆创新提供了包括产品、交互设计在内的全方位商业设计服务。

(王汝菲 方秀芬)

天津铜质线材企业重点实验室 通过天津市科委专家组现场验收

日前,由天津理工大学材料科学与工程学院与天津华北集团联合成立的天津市铜质线材企业重点实验室,经过一年时间的筹建工作,通过了天津市科委专家组现场验收。专家认定,该实验室已完全具备市级重点实验室水平。

据中国有色网报道,天津市科委专家组一行首先来到华北集团科技研发大楼,考察了实验室的硬件条件,并深入生产一线,参观了公司新引进的美国南方线材公司SCR-4500低氧铜杆连续连轧生产线。经现场查看,天津市科委领导及专家们对重点实验室建设给予了充分肯定。

天津市铜质线材企业重点实验室作为天津铜材行业唯一一家市级重点实验室,按照高技术结构、多元化组合、集散型生产、创新式管理的发展模式,以服务企业、行业为目标,以提升技术水平和产品辐射能力为基础,在“十三五”期间将使实验室的综合实力和产品研发、技术创新达到国际先进水平。进而对天津铜材企业技术进步起到有力带动作用,为加快行业发展做出积极贡献。

(佚名)

360天眼实验室 用大数据保障数据驱动安全

2014年,79790家公司被黑泄露数据,2122家公司公开确认信息被窃取,500强企业超过半数……如此触目惊心的数字,是奇虎360企业安全集团天眼实验室负责人韩永刚在29日举行的2015产业互联网大会上所展现的。据中新网报道,其中,60%的案例中,攻击者仅需要几分钟就可以得手,70%—90%的样本都是有针对性的,75%的攻击会在24小时内从一个受害者快速扩散到其他受害者。

“我们已经进入了大数据时代,未知威胁的发现也将从研究‘一片叶子’,过渡到观察‘整片森林’。”韩永刚认为,依靠安全专家的个人经验已经无法满足快速发现的需求,只有从数据、技术、人员等多个方面拥抱大数据技术,才能真正有效的、快速的发现未知威胁。

中国互联网网民数已经达到了全球第一。2014年,中国互联网人口达到了6.5亿,普及率达到47.9%。人们在享互联网红利的同时,不禁会担心,如此快捷、流通的网络之下,数据驱动安全能否得到保障?

天眼实验室负责人韩永刚指出,传统安全设备从本地网络或终端数据中发现未知威胁,对于互联网而言,仿佛研究森林中的一片叶子。为了改变这种情况,2014年,360天眼实验室正式成立,专门利用大数据技术研究未知威胁。天眼系统(SkyEyeSystem)也是全球首个基于大数据的未知威胁感知系统。

“现在面临的高端攻击,不管用什么样的方式都是用免杀木马,这个很难防御。如果他盯上了你,不管你有什么样的防御手段,被人入侵只是一个时间问题。我们所希望的是能提前发现攻击。”韩永刚告诉记者,360天眼系统基于海量数据挖掘、引入机器学习算法,能够有效准确识别未知恶意软件,是人工智能技术在恶意程序自动分析领域的首次商业应用。

但在这个能力的背后,首先需要海量的数据样本。据了解,目前360拥有全球独有的样本库,有80亿的总样本,且每天新增500万样本。

有了这些海量的数据,通过一个分析的平台,就对数据进行分析处理和预处理,再利用360的搜索技术,可以将几百个TB在云端数据产生的威胁情报全部汇总,然后得出结论。

“某个IP在某一个时间,下载了什么样的软件,有什么危害,关联到哪些用户……这些都要‘可见’,我们主要是结合了数据能力,安全技术以及核心的数据搜索方面的技术能力,包括设备能力以及在整个安全数据的运营和情报运营方面的能力,最终可以发现APT的攻击。”韩永刚说。

“我们希望用大数据见你所未见,知你所未知。”最后,他道出了360天眼实验室的目标,也为以大数据保障驱动安全作出了寄语。

据了解,本次大会是世界产业互联网第一阵营中美德日的高峰对话,大会还发起在浙江杭州成立中美产业互联网联合实验室,旨在加强中美在产业互联网领域交流与合作。

(李苑露 吴雨辰)

白浪费,这一足以容纳上百家上市公司的市场,还没有得到足够重视。”余福元第二次在光谷创业咖啡举办的“青桐汇”路演中说。

“这小子有眼光。”光谷创业咖啡创始人之一李儒雄对余福元很是赞赏。今年3月,经李儒雄引荐,余福元被小米CEO雷军邀请到北京面谈,“我们只聊了10多分钟,就把投资的事搞定了,1000万。”余福元说。

余福元对发展宏图胸有成竹:通过电脑、手机APP、微信等全渠道实现租房、服务、社区的O2O闭环,重构租房市场格局,为青年人提供高性价比、有尊严的品质租住生活,成为城市精品公寓租住体验的领跑者。

“只要有流量,就有钱赚”,余福元用这句互联网金句类比,“只要有租客,就有盈利模式”。赚房租、收服务费只是他计划的盈利模式中最直接、简单的一种,巨大租户形成的盈利空间,才是他真正看中的。

“雷军提醒我,前期不要想着赚钱,先把品质做好。”余福元说,这正是他做青年公寓的初衷。

今年毕业季,余福元的“可遇”青年公寓将扩大到1000多间。“这只是第一轮投资,我们的目标是上市。”余福元信心满满。

(据中国青年报)