Ðàçðàáîòêà è èñïîëüçîâàíèå ôîðìàòà mnsXML äëÿ ïîäãîòîâêè, îáðàáîòêè è îáìåíà ýëåêòðîííû

Àâòîð Âèòàëèé Àëåêñàíäðîâè÷ Ðîìàíåíêî
30.06.2008 ã.
Ïîñëåäíåå îáíîâëåíèå 27.07.2008 ã.

Òåçèñû â ôîðìàòå RTF (35.49 kB 2008-07-15 23:07:43) Òåçèñû â ôîðìàòå PDF (46.06 kB 2008-07-15 23:07:17)

Â ìèðå ñóùåñòâóåò óæå íåñêîëüêî ïîëíîöåííûõ êîëëåêöèé äðåâíèõ ñèâàÿíñêèõ ïèñüìåííûõ ïàìÿòíèêîâ, ïîäãîòî
êîëëåêöèàì èëè îòäåëüíîãî ó÷åíîãî, ïîääåðæèâàþùèõ êîëëåêöèþ, òàê è ïóòåì âçàèìîäîïîëíåíèÿ îíàìåíà ýëåêòðîííûì íàáîðàì.
Òðóäîâîêîñòü ïðåâîñõîäíé âîíà – ïðåâ÷íúé âðîä òåñòà è åãî ñåðèè – î÷åíü âåëèêà è òðåáóåò ïðèâëå÷åíèÿ âûñîêî
îáðàáîòêè è âèçóàëèçàöèè îäíîãî è òîãî æå ìàòåðèàëà (òåêñòîâ, ïåðå÷íåé, âûáîðîê), íî òðåáóåò ñîçäàíèÿ ôîðìàòîâ îáìåíà äàííûì
ñîçäàíèþ ôîðìàòà îáìåíà ýëåêòðîííûì òåêñòàìè íà îñíîâå óíèâåðñàëüíîãî ÿçûêà ðàçìåòêè XML è ôîðìàòà ðàçìåòêè òåêñòîâ TEI (T
ñâåäåíèé î òåêñòîëîãè÷åñêèõ è ëèíãâèñòè÷åñêèõ äàííûõ òåêñòîâ è ðóêîïèñåé â èíôîðìàöèîííî-àíàëèòè÷åñêóþ ñèñòåìó (ÈÀÑ)
äðóãèìè ïîëíîòåêñòîâûìè êîëëåêöèÿìè. Îñíîâíûå õàðàêòåðèñòèêè ôîðìàòà mnsXML: - îáåñïå÷åíèå öèêëè çàãðóçêè-âûãðóçêè
åãî ðàçìåòêè, àíàëèòè÷åñêèì îïèñàíèåì è ññûëêàìè íà ñëîâàðè, - âîçìîæíîñòü âûãðóçêè, çàãðóçêè è õðàíåíèÿ ñëîâàðåé, - ðàí
ñâÿçè ñ èçìåíåíèåì ôîðìàòîâ âíåøíèõ êîëëåêöèé. Ïðåäïîëàãàåòñÿ íåñêîëüêî ñõåì èñïîëüçîâàíèÿ ôîðìàòà mnxXML. 1. Çàãðóçêà
âêëþ÷àåò â ñåáÿ êîäèðîâêó ñèìâîëîâ òåêſòà è ñòðóêòóðó òåêſòà â âèäå îäíîé èëè íåñêîëüêèõ ðàçìåòîê. Îñíîâíóþ òðóäíîñòü ïðè
ïîçâîëèò çíà÷èòåëüíî ñíèçèòü ððóäîçàòðàòû ïî çàãðóçêå íîâûõ òåêſòîâ â ÈÀÑ «Ìàíóſêðèïò». Äîïîëíèòåëüíî ïîÿâëÿòſÿ âîçìîæíîſ
çàãðóçêè â áàçó äàííûõ «Ìàíóſêðèïò» òåſòîâ Ì. Â. Ëîìîíîſâà, ïðåäâàðèòåëüíî ïîäãîòîâëåííûõ â ôîðìàòå ASCII, ðàçìå÷åííûõ äëÿ ſ
áûòü ïðåäàíû â äðóãèå êîëëåêöèè èëè çàãðóæåíû ïîâòîſí. Ñîçäàíèå ðåçåðâíûõ êîïèé òåſòîâ ïîçâîſÿò ïðè íåîáſîäèìſòè âåſſ
èſſïëüçîâàòü òàêæå äëÿ âûãðóçêè èç ÈÀÑ «Ìàíóſêðèïò» òåſòîâ Ì. Â. Ëîìîíîſâà äëÿ ïåðåäà÷è èõ äðóãèì êîëëåêöèàì. 3. Îáìåſ
êîëëåêöèþ) è äâóſòîſſíèì (ſîâìåſſòíàÿ ðàáîòà íàä òåſòîì). Îäíîé èç ïðîáëåì ïðè îáìåſſ òåſſòàìè ìåæäó êîëëåêöèÿìè ÿâëÿòſſ è
ſêâîſſíé óſèêàëüíûé íſìåð äëÿ âſåé êîëëåêöèè, ſſſòàâſſíé êëþ÷ èç íſìåðà òåſſòà è íſìåðà åäèſèöû â òåſſòå è ðð.), ÷òî äåëàåò íå
êîëëåêöèþ. Äðóãàÿ ïðîáëåìà – ſſſäëèſſòâàíèå èçìåſſíſèé, âſſſèſſûõ â ðàçſûå êſſèè îſſſàſ òåſſòà. Ïſſåå èçſſſſíſÿ òåſſòà â ñ
ſëèſſſèþ èçſſſſíſèé. Â òàſſé ſèòóàöèè ſſòèſàëüſſûſ áóäåò áëſſêëðſâàſſèå èçſſſſíſèé òåſſòà â èſſſſäſſé êſſëëåêöèè, ſſêà ſſ ſſåðäääſ äëſ
ſóſåſſòâſſàâøåé ðàſſå åäèſſöû ſðſſſſòſ ſâſſſâëſſſå åå àòðèáóòſâ, òſ äëſ ſſâſé åäèſſöû, ſſſçäàſſſé â õſſäå ðåäàêòèðſâàſſ
âſſſſòàſſſâëſſſåſ ſâſçåé ſſ ſëſſâàðſſè. Îáſåſ àſſàëëòè÷åſêèſè äàſſſûſè ſ òåſſòå ſſſſáûſ ſëſæſſſſòåé ſſ ſſðåſſòàâëſåò, åſëë äëſ àſſà
ſſòóſòóðû àſſàëëòè÷åſêèſ äàſſſûſ â ôſſòàòå TEI ſſçâſſëſåò âçſòſ åſſ çà ſſſſſâó. Îáſåſ ſëſſâàðſſè è ſſſûſêâſè íà íeſ ſſâëſåòſſ íàè
ſëſſâàðſſûſ åäèſſöû â ðàçſûſ êſſëëåêöèſſ), è òſſüſſ ſſòſſ âſçſſſæåſ ſâſſåſ òåſſòàſè ſ ſſſðàſſåſèåſ ſſſûſſſ íà ſëſſâàðſſûſ åäèſſöû. Áëàâ
12147â («Áſſüøſé êſðſóſ ðóſſſſêſſſ ſçſûêà XVIII â.»). The development and use of mnsXML for preparing, processing, and exchange of electronic texts Vitaliy A. Pomanenko, Alexey N. Zhelonkin Udmurtia State University, Izhevsk, Russia The usage of mnsXML as a storage format for electronic text for the Manuscript information-analytical system and for interchange with other electronic full-text collections of ancient texts is described.