

Íāðáúē íðíāēò — Áāçā ñðāáíāāāēíāíāí òðāíóóçñēíāí (BFM) — áúē íā-àò á 1989 ā. á Áúñøáē ííðíāēúííē øēíēā áóíā Marchello-Nizia). BFM — ííēííòāēñòíāy āāçā āāííúō, íñííāíāíy íā ñíāðāíāííúō ēðēòē-āñēēð ēçāāíēyō. Óāēüp íðíāēòā áúē òāēñòíāíúō āāííúō.

Éñííēúçíāāíēā á āāçā ííēííúō òāēñòíā íðíēçāāāāíēē íáúyñíyāòñy òāí, ÷òí ííā ēçíā-āēúíí áúēā íðēāíóēðíāāíā íā íāēñēíāēúíí øēðíēēē ēíðíòēēā ñðíēē ñòíðíēðíāāòú āíñòāòí-íí ēðóííúē ēíðíóñ. Íðē yòíí āēy āēēp-āíēy á āāçó íóāēðāēēñú íðāæāā āñāāí ēçāāíēy, ííāííí òāēòē-āñēēā íøēāēē ē íāííāāðēāy āñā ēñíðāāēāíēy á ñííñēāō.

Íííēíāíēā āāçú íñóúāñòāēyēíñú áēāāíāāðy ñēāíēðíāāíēp òāēñòíā ēçāāíēē, á òāēæā íāíāíāí ē āāçāíçíāçāíúí āēēāāāí ñíāōēāēēñò Íñííāíúí íāðēíāíí, íā ēíðíðúē íāðāííā-āēúíí íðēāíóēðíāāēñy íðíāēò BFM, áúē ñòāðíðòāíóóçñēēē (IX–XIII āā.), íāíāēí ā āāçó òāēñòú ñíððāíyēēñú á òíðíāòā «íðíñóíē òāēñò», á ēíāēðíāēā DOS. Ñ 2001 ā. òāēñòú āāçú ííñòāíāíí íāðāāíāyòñy á òíðíāò XML

Á íāñòíyúāā āðāíy íñíāííē ēíðíóñ òāēñòíā BFM íāñ-ēòúāāāò 74 òāēñòā íāuēí íāuāííí íēíēí 3 000 000 ñēíāíóíðòðāāēāíēē. Yēñíēó Ñ. Yēāāííí (S. Heiden). Áíñòóí ē āāçā íðāāíñòāāēyāòñy āāñíēāòíí āñāí ó-āíúí, íðāííāāāāòāyēí ē ñòóāíòāí, íðēíēíāpūēí òñēíāyēy.

Áòíðíē íðíāēò — ēíííēāēñííā íóēüòēíāāēēííā ēçāāíēā ðóēííēñāē òðāāēòēē ðííāíā Éðāòúāíā āā Óðóā «Ðúòāðú òāēāāē» (& Charrette), íñóúāñòāēāííā á 1990–2003 āā. á Íðēíñòííñēíí óíēāāðñēòāòā ííā ðóēíāíāñòāíí íðíāññíðā Ę. Pòòē (K. Uitti).

Áāí òāēüp áúēí ñíçāàòú íðēíóēíēāēúíí ííāíā ēçāāíēā ñðāāíāāāēíāíāí òāēñòā, ñíāíāúāpūāā á ñāāā ēðēòē-āñēē áúāāðāííúē òāēñòú íðòíēíāē-āñēíē ē òēēííāē-āñēíē ðāçíāòēíē ē ñíāāæāííā íúííē íñēñíāíē íāøēííē.

Íāðāííā-āēúíí òðāíñēðēíòēē ðóēííēñāē á íðíāēòā Charrette íñóúāñòāēyēēñú á òíðíāòā SGML á ñííòāāòñòāēē ñ ðāēííāíāāòēyēíē T íðēāēíāēāíē ē áíāñāíēāí āíííēíēòāēúííē ēíðíðíāòēē. Íāðāēēāēúíí íā ñíííāā ēðēòē-āñēíāí ēçāāíēy áúēā ñíçāāíā āāçā āāííúō ííyòē-ā

Áāçā āāííúō íí ííyòē-āñēēí òēāóðāí, āēíēííāòē-āñēēā òðāíñēðēíòēē ē òíðíāðāòēē ðóēííēñāē āíñòóííú íā ñāēòā íðíāēòā.

Á ðāíēāò Éííñíðòēóíā CCFM á íāñòíyúāā āðāíy āāāāòñy áúðāāíòēā āāēíúō ííðí ēíāēðíāāíēy ē íñēñāíēy ñðāāíāāāēíāíúō òðāíóóçñē āāí ó-āñòíēēíā. Ñíāíāñòíúíē òñēēēyíē ðāøāpòñy íðíāēāíú íðòíēíāē-āñēíē ðāçíāòēē ē ēāííāòēçāòēē òāēñòíā, āāāāòñy ēññēāííā

īōēīōēōāōīūā īāīōāāēāīēÿ èō ðāñøēðāīēÿ.

Summary

In this paper we will give a review of the current projects on the Medieval (9th — 15th centuries) French text databases, corpora and electronic editions. We will present: different choices of composition and sources for such databases (full text vs. sample corpora; using critical editions vs. original manuscripts); the principles of markup and the technologies used. Two projects will be considered in more detail: the Base de Français Médiéval (<http://bfm.ens-lsh.fr>) and the Charrette project (<http://lancelot.baylor.edu>)

Nīēñīē èèōāðàòóðū

Heiden S., Guillot Ñ. Capitalisation des savoirs par le web : une application de la TEI pour l’encodage et l’exploitation des textes de la Base de Français Médiéval // Ancien et moyen français sur le Web, enjeux méthodologiques et analyse du discours / Kunstmann, P. (ðää.). — Ottawa: Les éditions David, 2002. — Ñ. 77–92.

Heiden, S., Lavrentiev, A. Ressources électroniques pour l’étude des textes médiévaux : approches et outils // Revue française de linguistique appliquée. — 2004. — IX(1). — Ñ. 99–118.

Kunstmann, P. Ancien et moyen français sur le Web: textes et bases de données // Revue de linguistique romane. — 2000. — 253–254. — Ñ. 17–42.