

大日本史-Dublin core 親和性の考察 新たな基本語彙セットの提示

後藤 怜

現在は、電子書籍化が盛んに進められている時代といえる。しかし、新たに発行される書籍と相反して、古書の多くは電子書籍化が進んでいない。「なぜ、それらの電子化が進まないのか」という疑問点を抱いたことが本研究における始まりといえる。では、「どのような問題があり古書の電子書籍化が進んでいないのか。」という点について考えたとき、「電子化するためのメタデータ記述法」がはっきりと確立していないため、一定の基準に沿った電子書籍の作成が行えないということが明らかになった。この一定の基準を考察し、明らかにすることが本研究の目的である。

具体的な方法としては、古書としては日本を代表する総合的な歴史書である大日本史を用い、現在において最も国際的に普及が進んでいるメタデータの編集基準である Dublin core Metadata Element Set(以下、Dublin core)との親和性を考察する。内容としては、Dublin core の項目一つ一つに対応する大日本史の要素をあてはめ、一致するものは「完全親和」、部分的に一致するものは「不完全親和」、まったく一致しないものは「排他」と三分類する。そののちに、全体や要素ごとにおいて三分類の比率がどのようになるかをデータ化し、考察を行った。

研究結果より明らかになったのは、全体における不完全親和性の部分が、その他二つに比べ約60%を占めている点である。これは、書籍全般において国際的に用いられる規格としては適応性に欠けるといわざるを得ないだろう。または部分的に所見の述べると、Dublin core が作成されてから発行された書籍については対応ができていますが、それ以前の書籍への対応ができていないともいえる。つまり、大日本史と Dublin core の親和性はかなり低いものといえる。完全親和率は20%程度であることから、いくら Dublin core が部分的に用いる規格だとしても何らかの措置をとるべきだろうと考える。

結論としては、現在の項目に加えて、拡張子を追加することや現在の項目に解説を付加するなど、要素の入れ替えではなく要素の追加において対応すべきと考える。拡張子を追加することで、収録物のより詳細な内容を登録することができ、検索する側も登録する側においてもより正確なデータ検索が行えるからである。

今後の課題としては、どのような拡張子を追加すべきか検討していきたい。

(指導教員 水嶋英治)