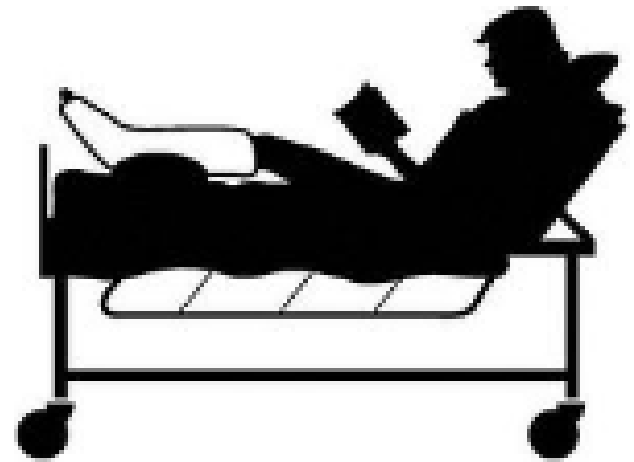


Phenotype-related information extraction from routine data for biomarker research

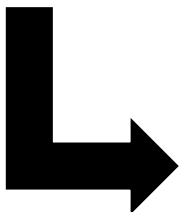
Catalina Martínez-Costa, Stefan Schulz, CBmed Project 1.2



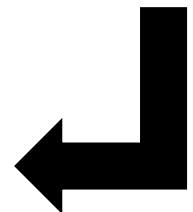
- Lab data
 - large volume
 - created by machines
 - highly structured

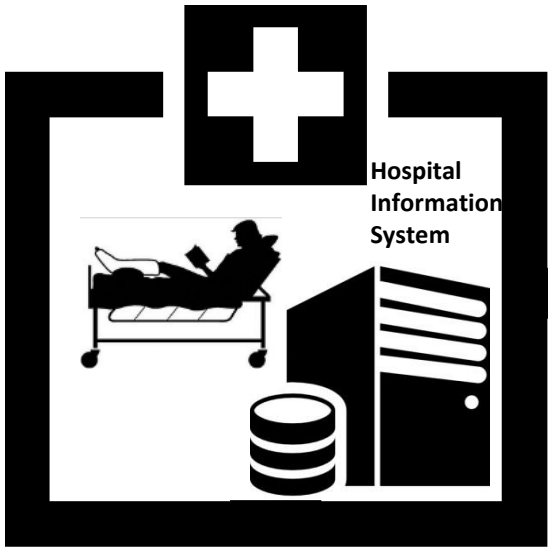


- Clinical data
 - small volume
 - created by humans
 - poorly structured
(text in local language)

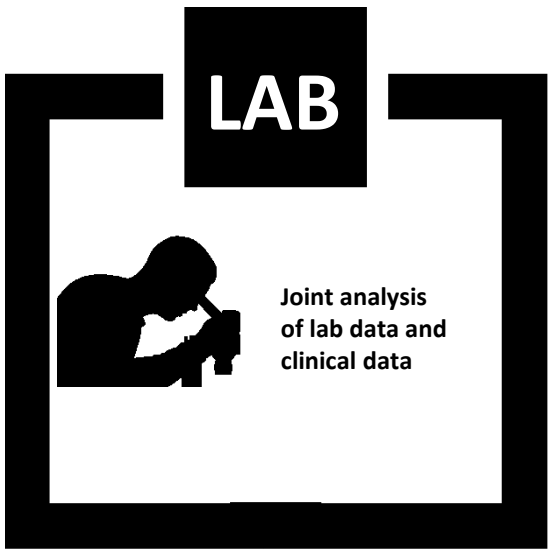
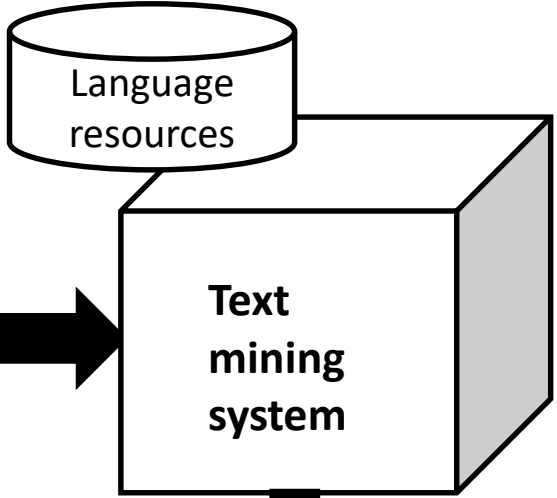


How to bring both kinds of data together for biomarker research ?





"St.p. superfiziell spreitendes malignes Melanom in Zusammenhang mit, oder ausgehend von einem Nävuszellnävus (NZN) anm Unterschenkel links 3/07(Level III, TD 1,3mm, AJCC T2a)St.p. Nachexzision mit axillärer Lymphadenektomie 3/07 - 7 von 16Lymphknoten (LK) positiv St.p. adjuvanter Therapie mit Interferon-Hochdosis nach Kirkwood vom7.5.2007 bis 11.3.2008



- Clinical data to be extracted:**
- past diseases
 - findings
 - procedures
 - lifestyle data
 - drugs
 - family history

Code (e.g. SNOMED CT)	Value	Context
254730000 Superficial spreading malignant melanoma of skin		392521001 History of
301889008 Excision of malignant skin tumor		392521001 History of
47224004 Skin of posterior surface of lower leg7771000 Left		
81827009 Diameter 258673006 millimeter	2.41	
258403002 Lymph node level IV		
94339008 Secondary malignant neoplasm of inguinal lymph nodes		15240007 Current 2667000 Absent

