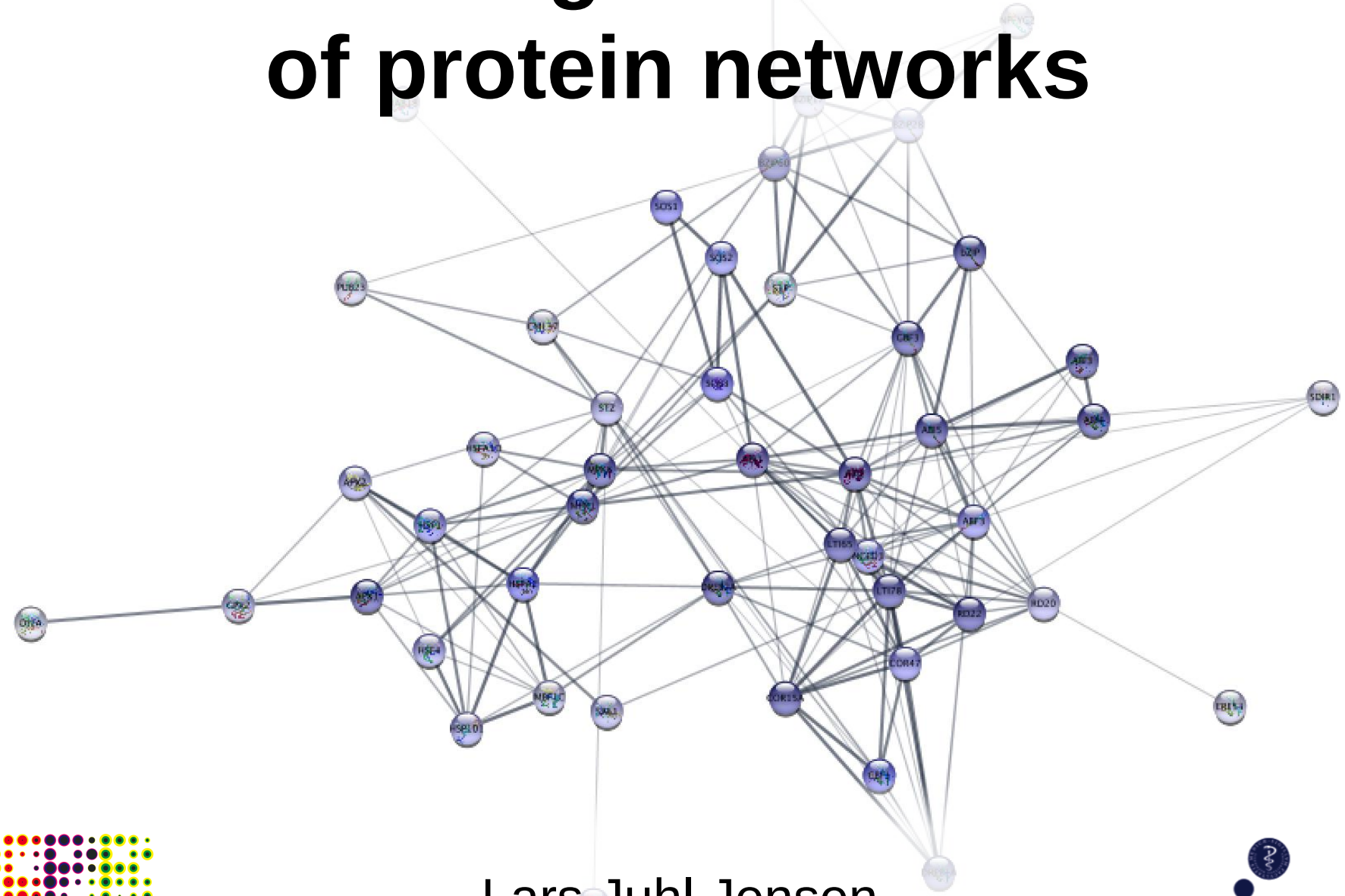
Text-mining-based retrieval of protein networks

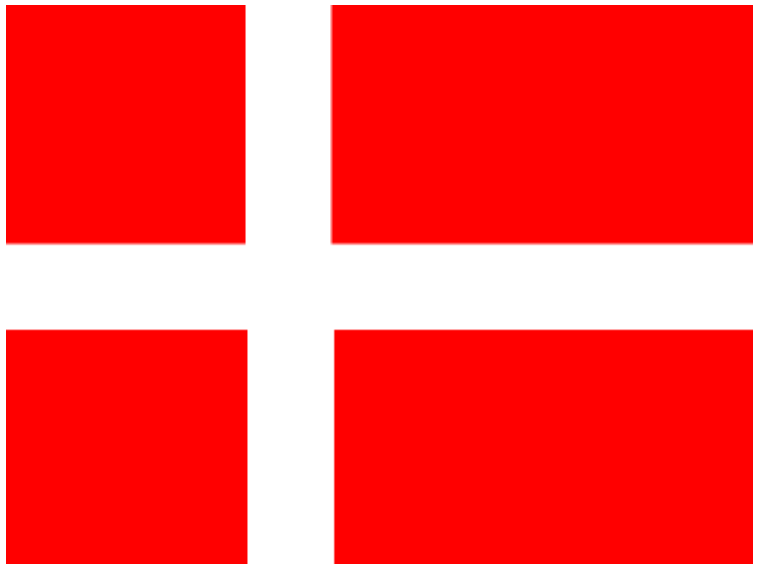


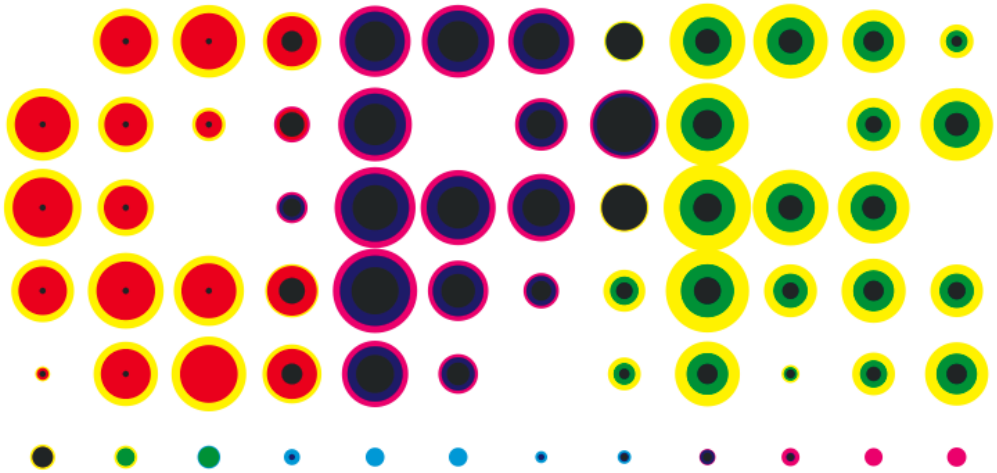
Lars Juhl Jensen

The Novo Nordisk Foundation
Center for Protein Research

FACULTY OF HEALTH SCIENCES
UNIVERSITY OF COPENHAGEN

me



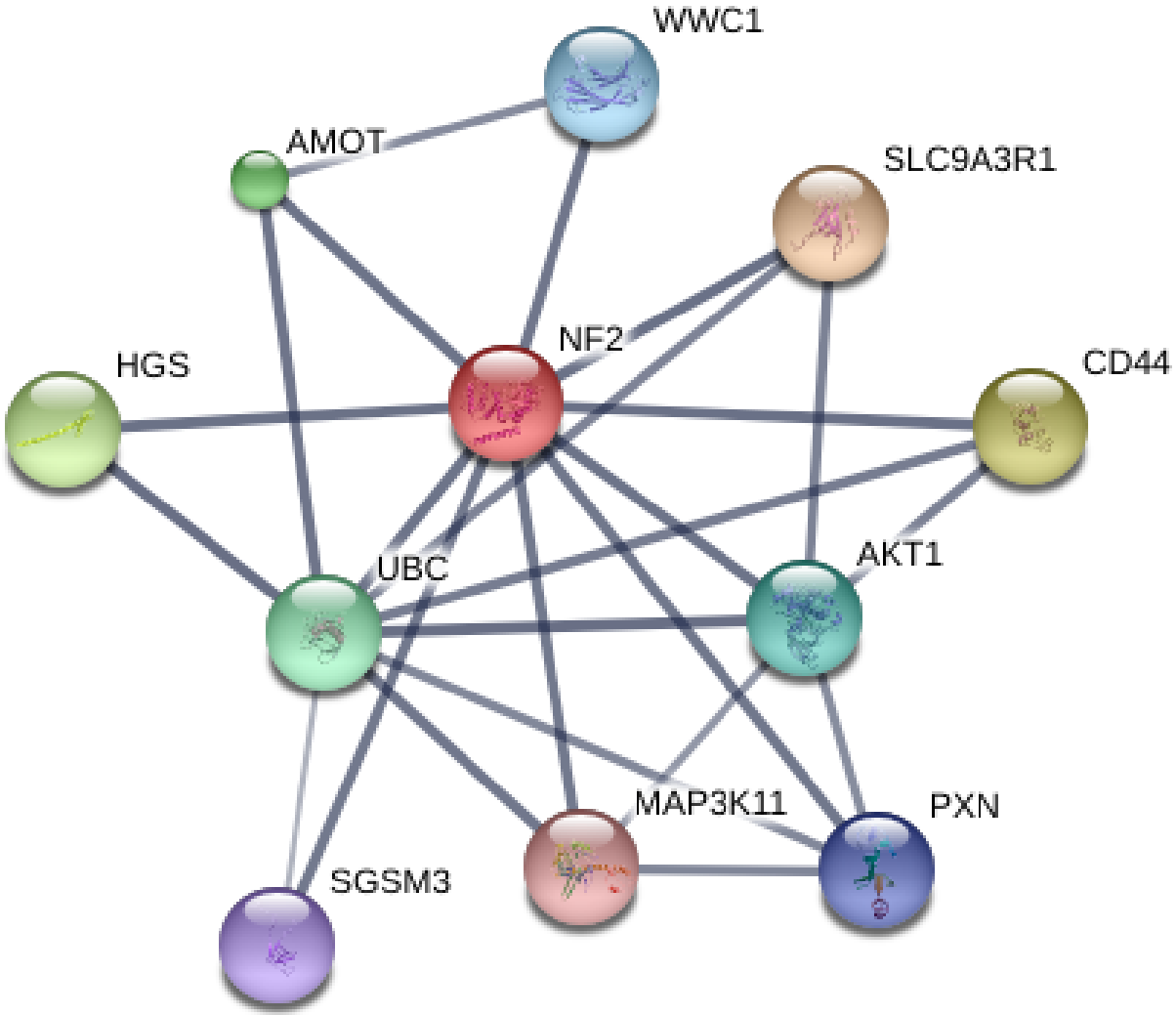


The Novo Nordisk Foundation

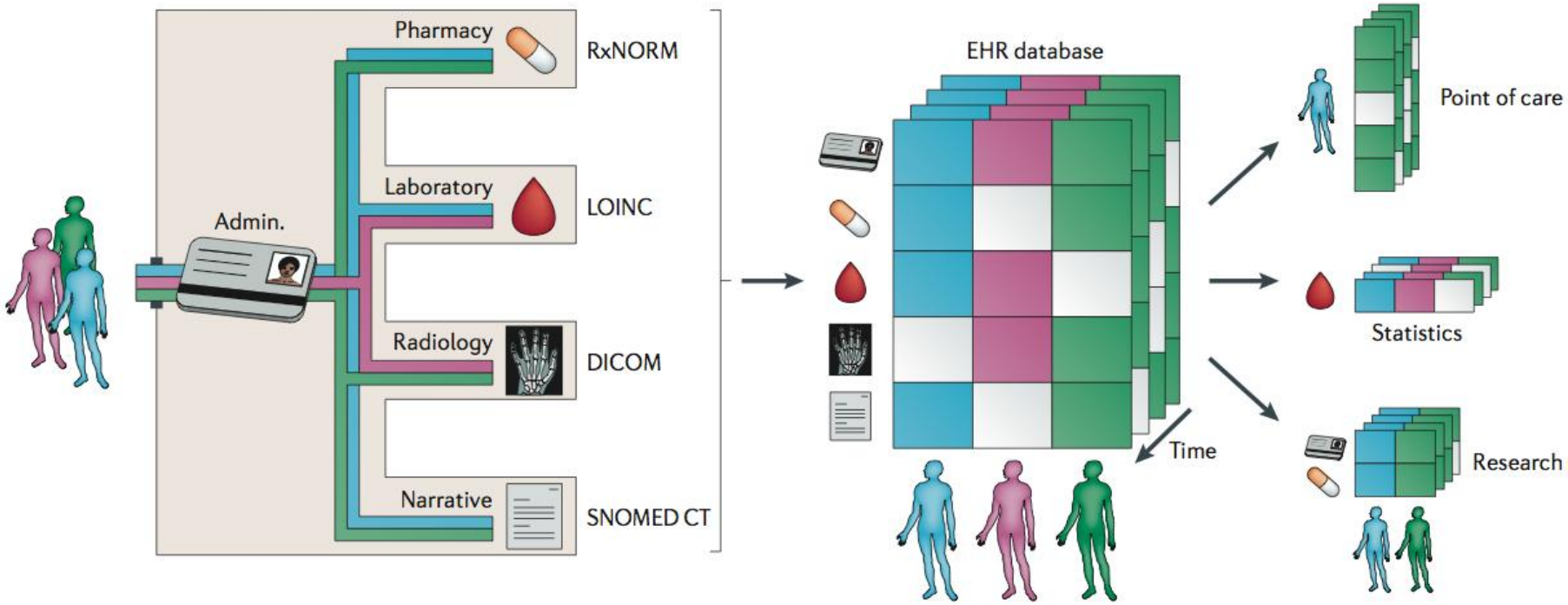
Center for Protein Research

intomics
from data to biology

protein networks


WWC1
AMOT
SLC9A3R1
NF2
HGS
CD44
UBC
AKT1
MAP3K11
PXN
SGSM3

medical data



Jensen et al., *Nature Reviews Genetics*, 2012

text mining

named entity retrieval

topic of interest

PubMed query

set of abstracts

named entity recognition

foreground count

entire PubMed

background count

enrichment score

set of entities

network construction

text mining

data integration

STRING

Cytoscape App

Acknowledgments

John “Scooter” Morris

Nadezhda Doncheva

Sune Pletscher-Frankild

Christian von Mering

Peer Bork