



Virtual Dave Lankes

Presentation

<http://www.DavidLankes.org>

TITLE: Reference Authoring

AUTHOR(s): R. David Lankes

PUBLICATION TYPE: Presentation

DATE: 2005

VENUE: Friday Seminar SIMS, UC Berkeley, Berkeley, CA

ABSTRACT: Discussed the intersection between digital reference and information retrieval.

KEYWORDS: digital reference, virtual reference, reference authoring, information retrieval, credibility

Reference Authoring

R. David Lankes

<http://www.DavidLankes.org>

Agenda

- ~ The Reality of Knowledge Bases
 - ~ Draw Backs of a Deductive Approach
- ~ The Concept of Transformation
- ~ Search and the Road to Reference
EXTRACT
- ~ Key Concept: Credibility

Realities of KB's

- ~ Many Doubt their Usefulness
 - ~ Exceptions Like Mad Scientist, Ask Dr. Math
- ~ Yet, We Use them All the Time
 - ~ Statistics, Patron Follow-up, Evaluation, etc.

A KB Continuum



Raw



Edited



Refined

Draw Backs

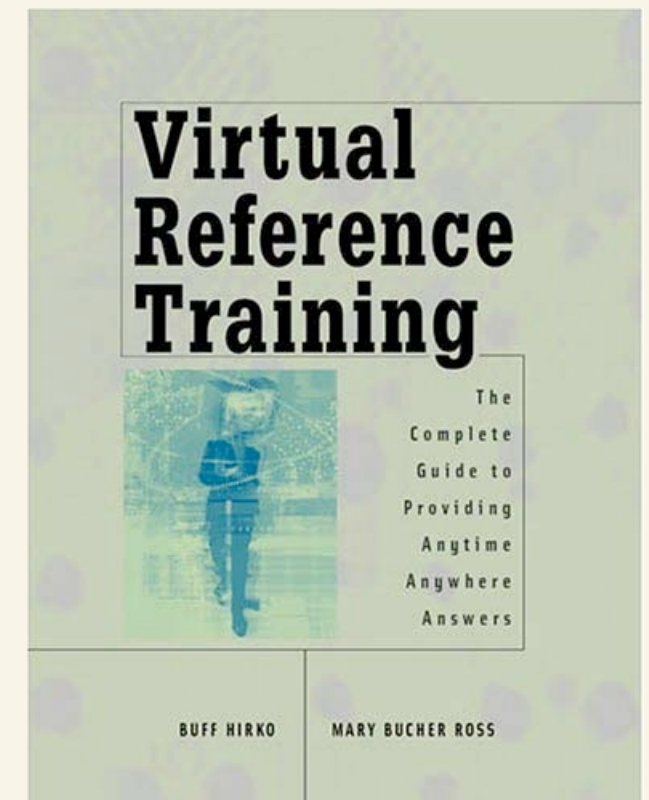
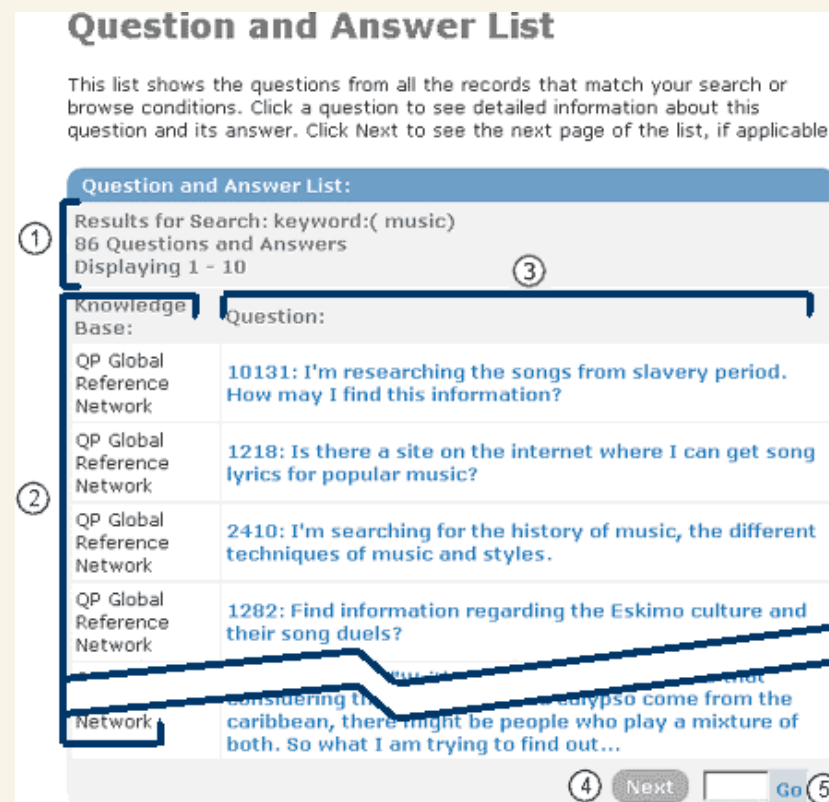
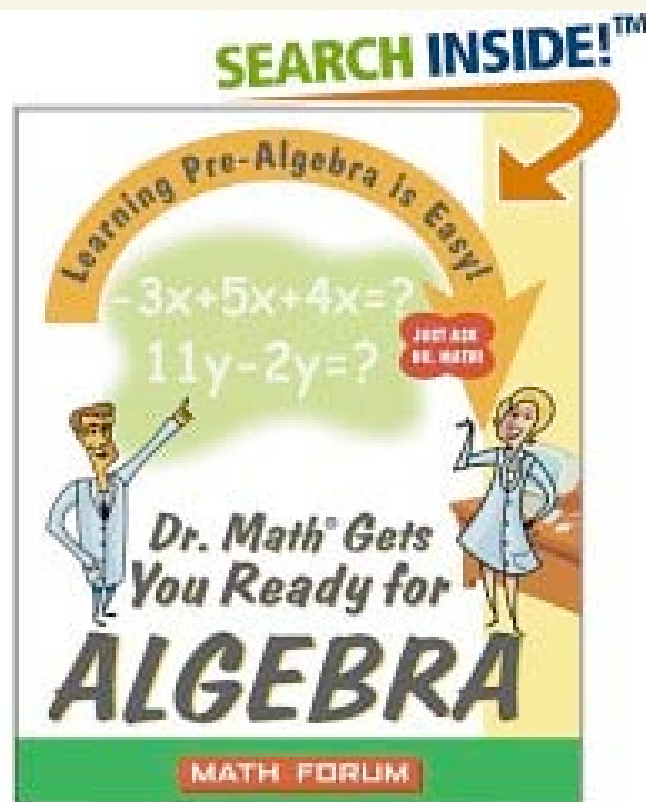
- ~ Context Dependencies
- ~ Metadata Creation
- ~ Chunking
- ~ Fact Shifting and Temporal Dependencies

Transformation



- ~ Effort: The Amount of Resources (Time, Money, Automation) Expended to Transform
- ~ Effect: The Utility of the End Result

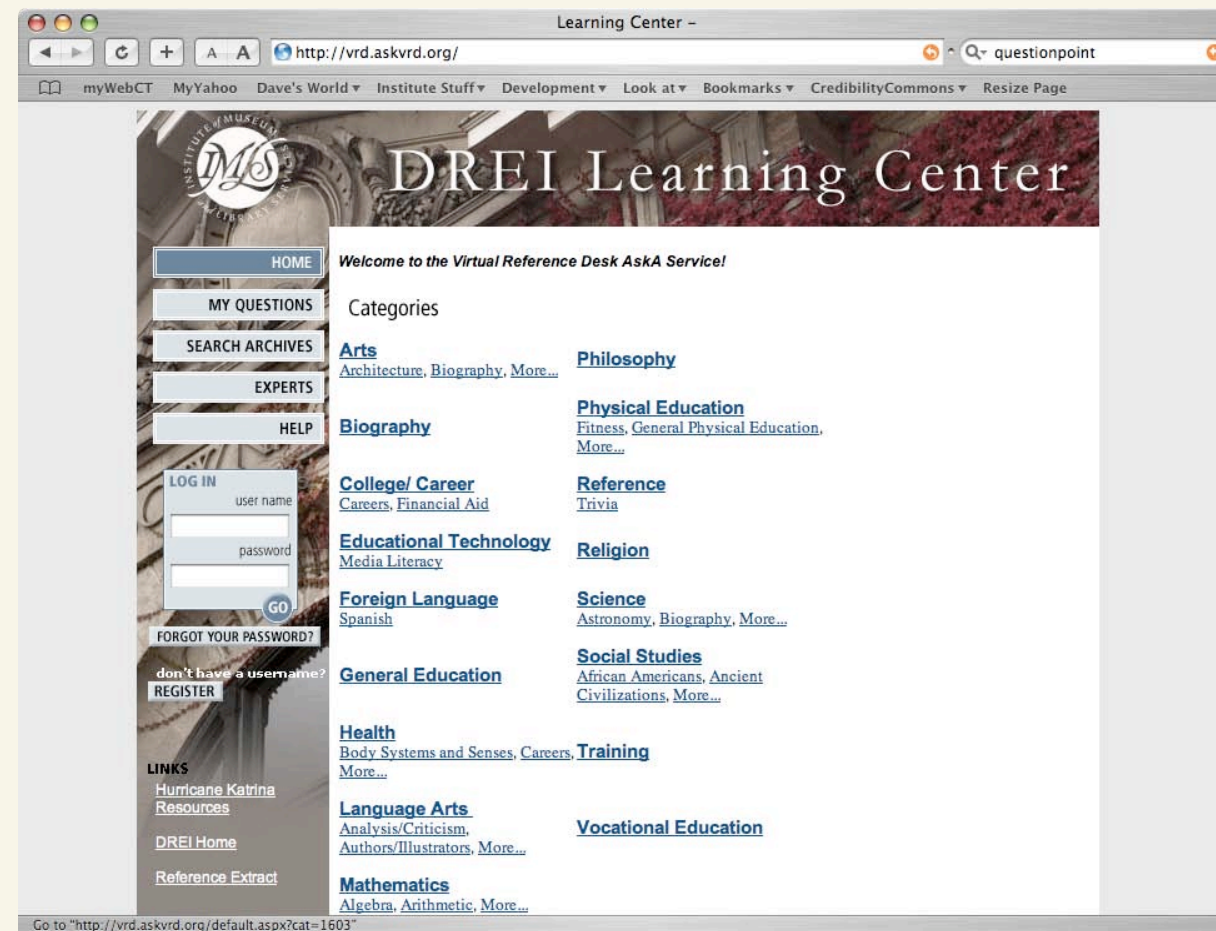
Examples: High Effect, High Effort



~ Each Requires a High Amount of Effort to Effect

Examples: Low Effect*, Low Effort

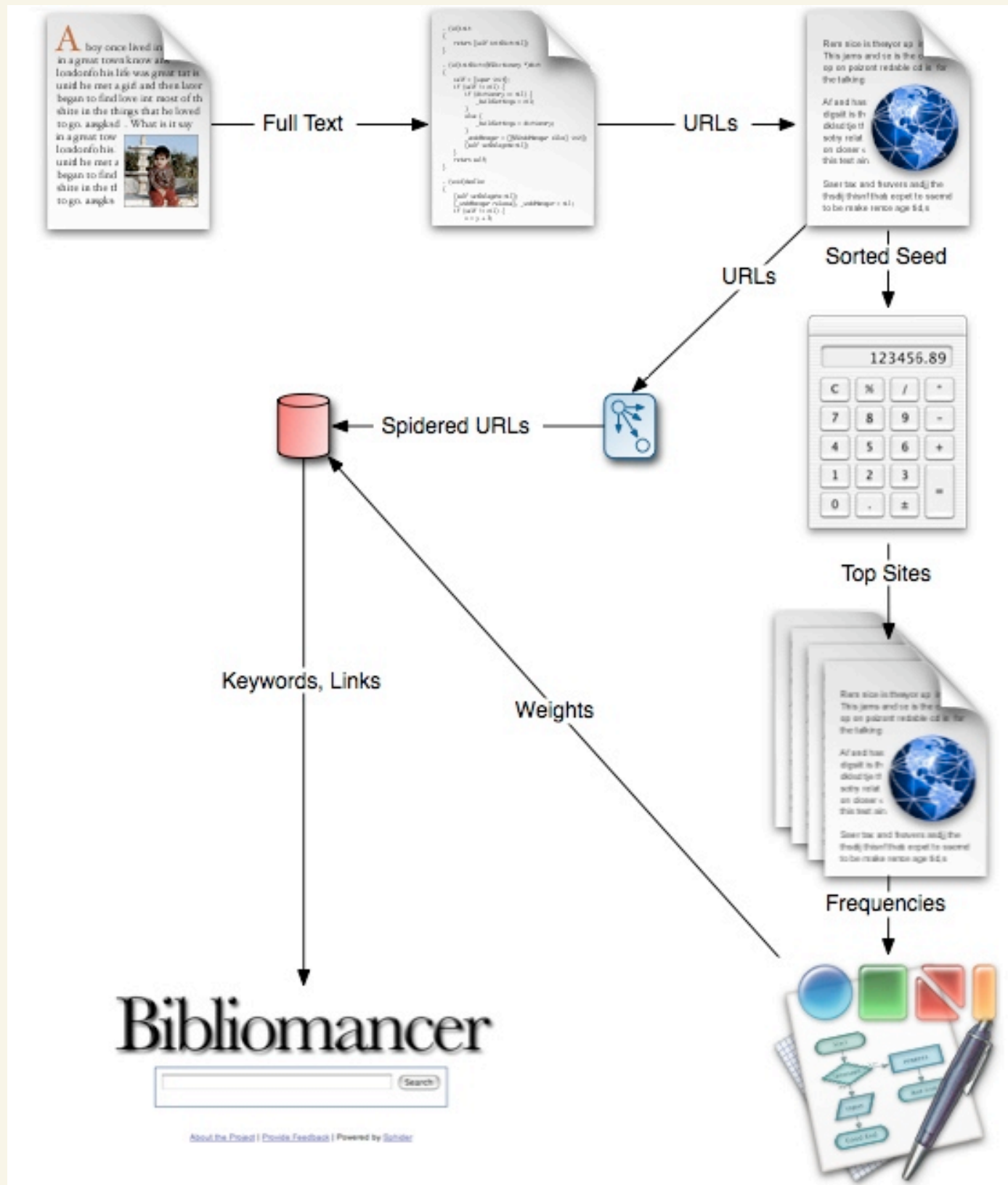
* Low Effect Simply Reflects Susceptibility to the Draw Backs of Deductive Knowledge Bases



~ Requires Low or Distributed Effort to Effect

Question: What has Low Effort and High Effect?

- ~ Search Seems Promising
 - ~ High Use in the Internet
 - ~ Matches Reference Domain
- ~ Preserves Patron Privacy
 - ~ URL Stripping



- ~ Raw Text File
- ~ Extracted URL's
- ~ Frequencies

Reference EXTRACT

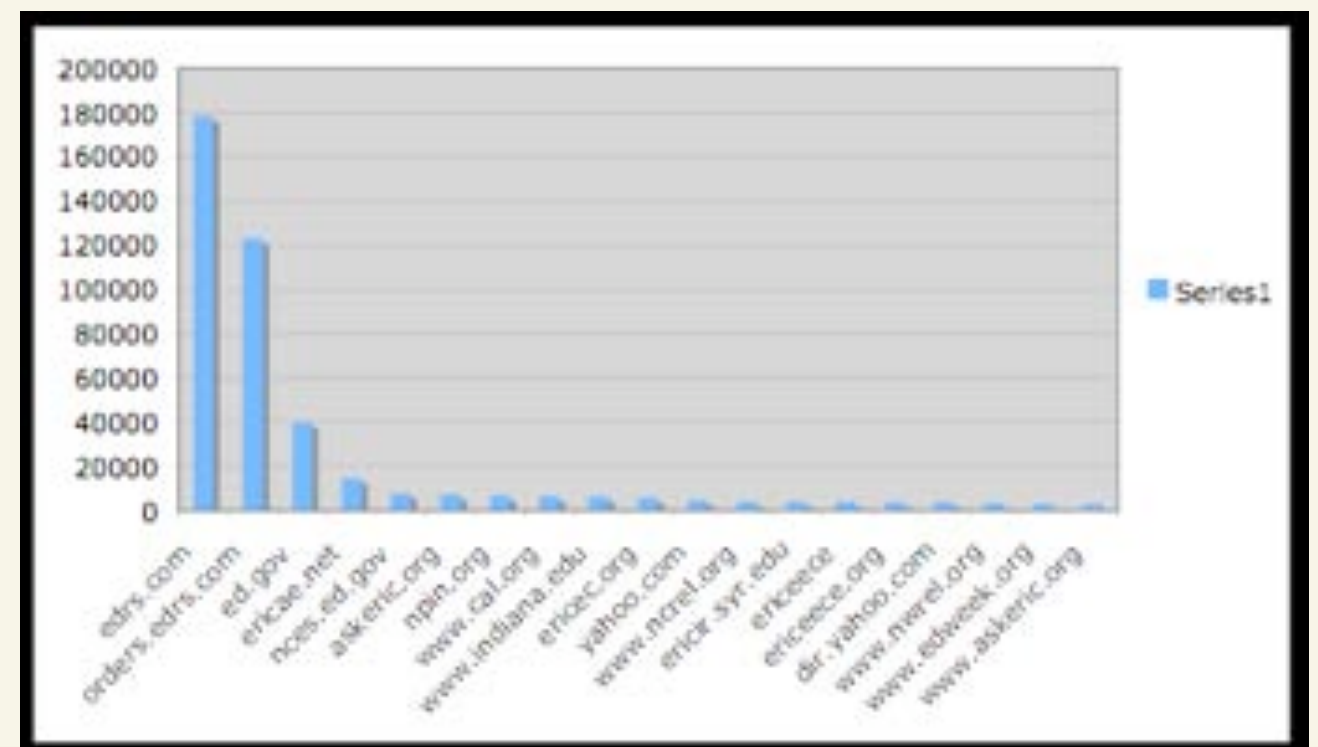
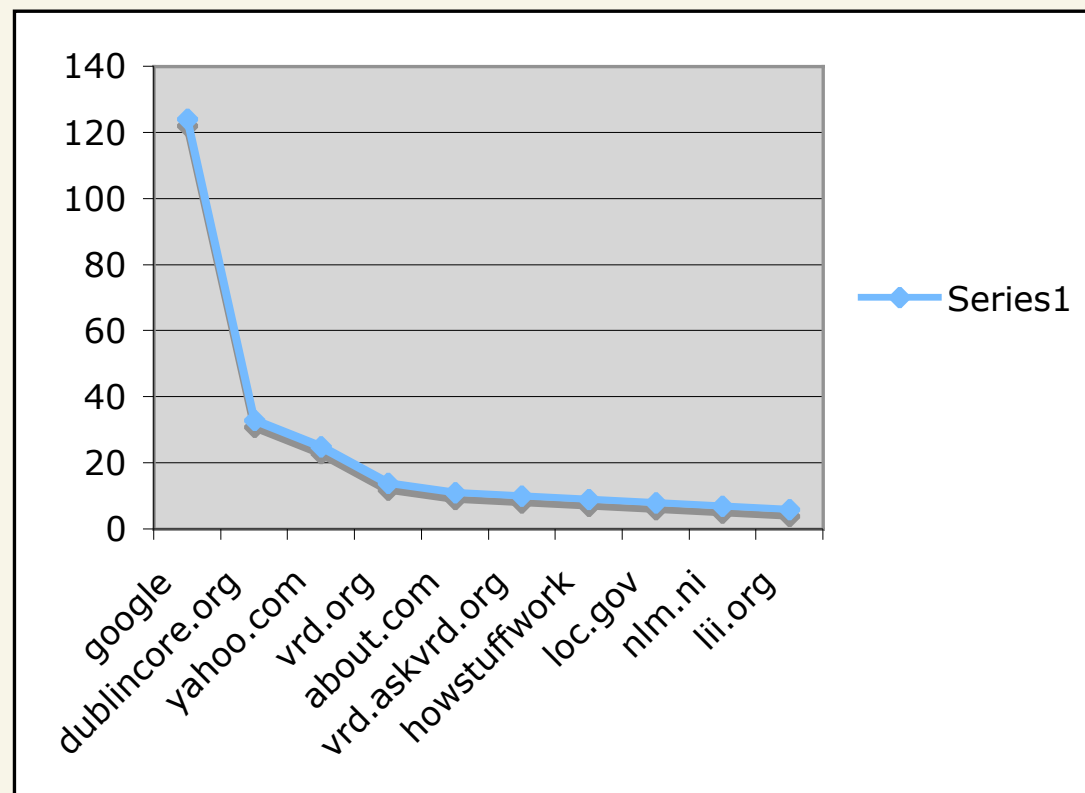
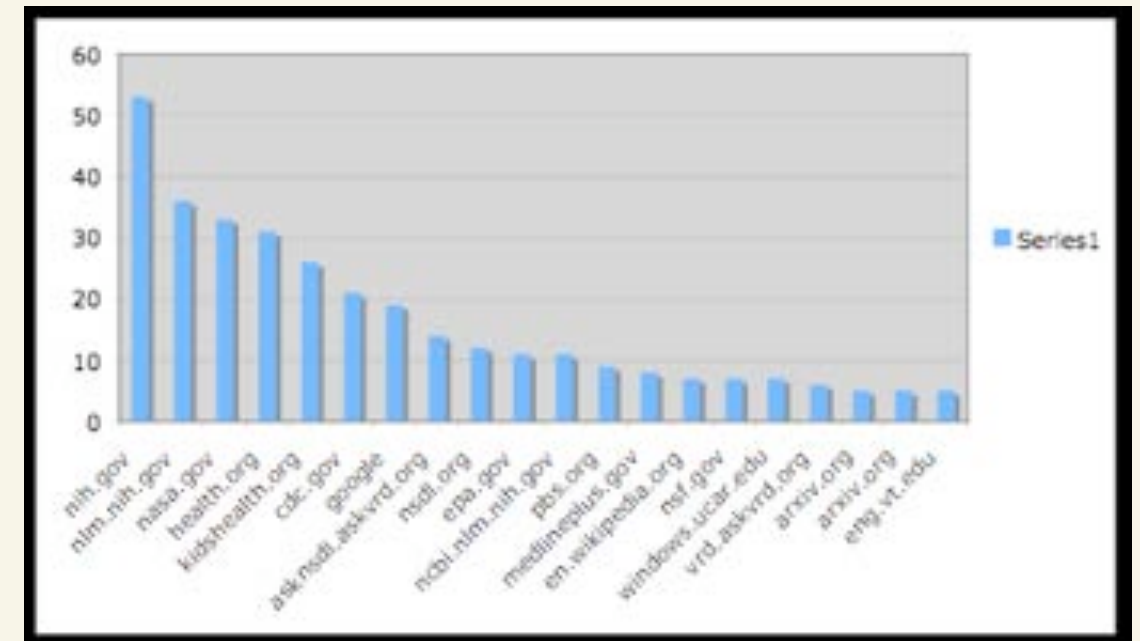
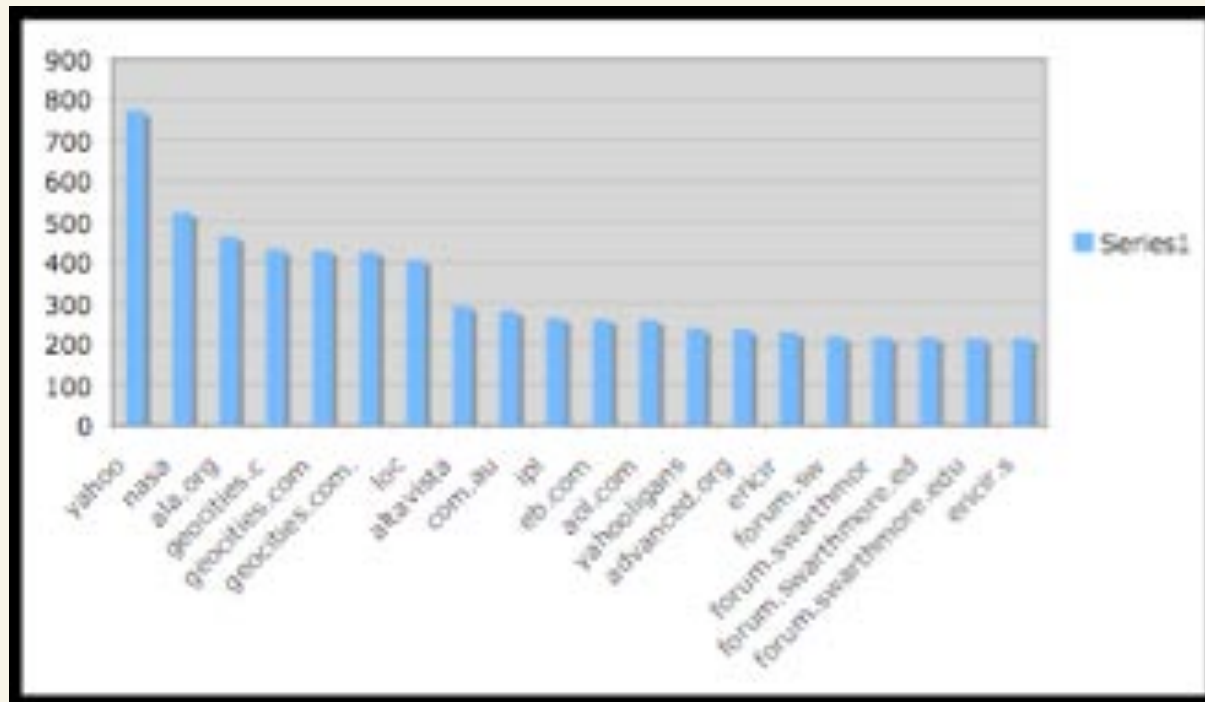
- ~ “What is the utility of a knowledge base?”
- ~ High Utility, Low Cost, High User Privacy
- ~ <http://digref.org/search/en/deathmatch.html>
- ~ Low Context



So Why Isn't this LII?

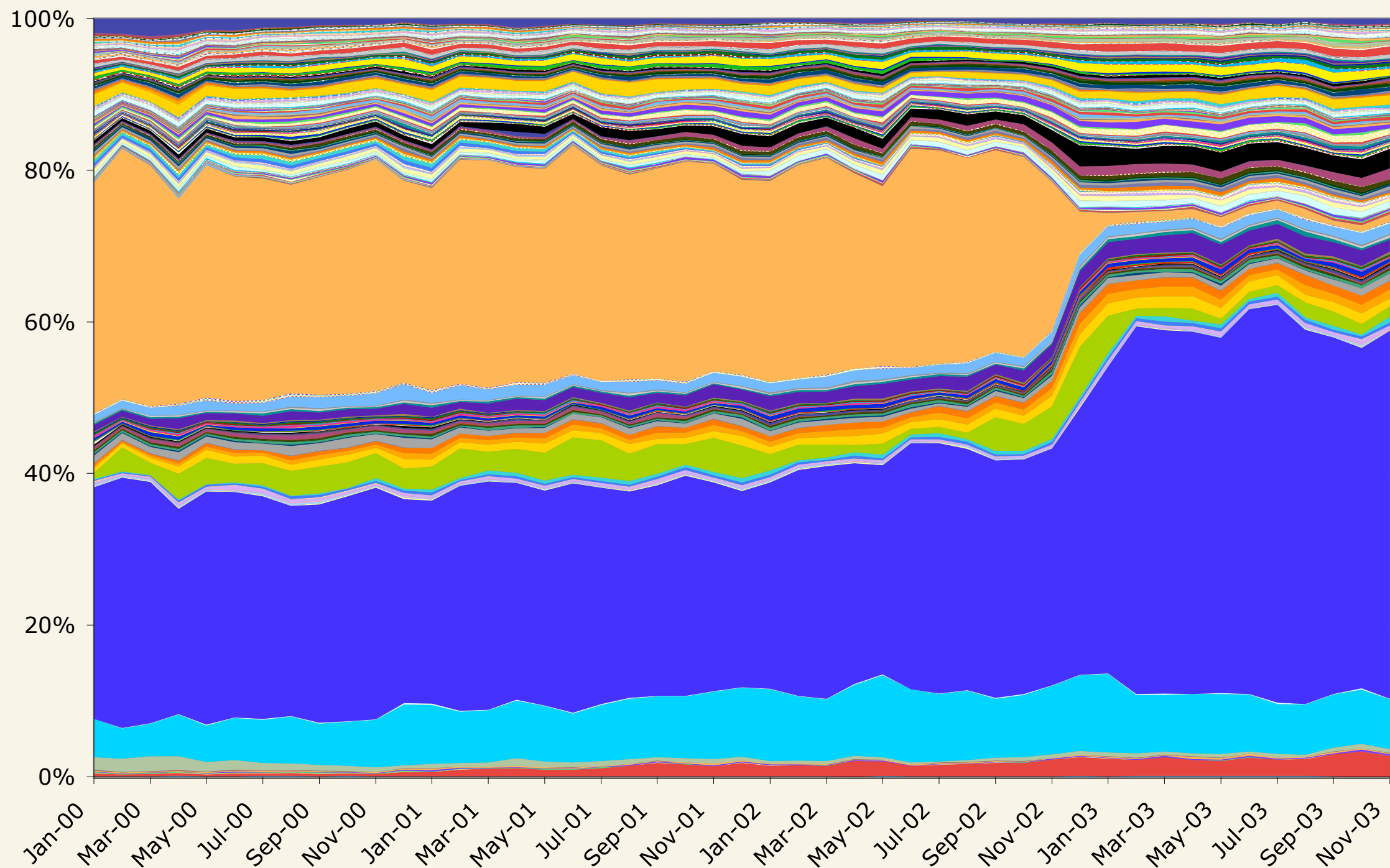
- ~ Built Inductively from Actual Use
- ~ Reference Weighting
 - ~ Using “Citedness” as a Surrogate for Relevance like Google uses Page Ranking
 - ~ “Boost” Results by Use/Citation

Reference Weighting



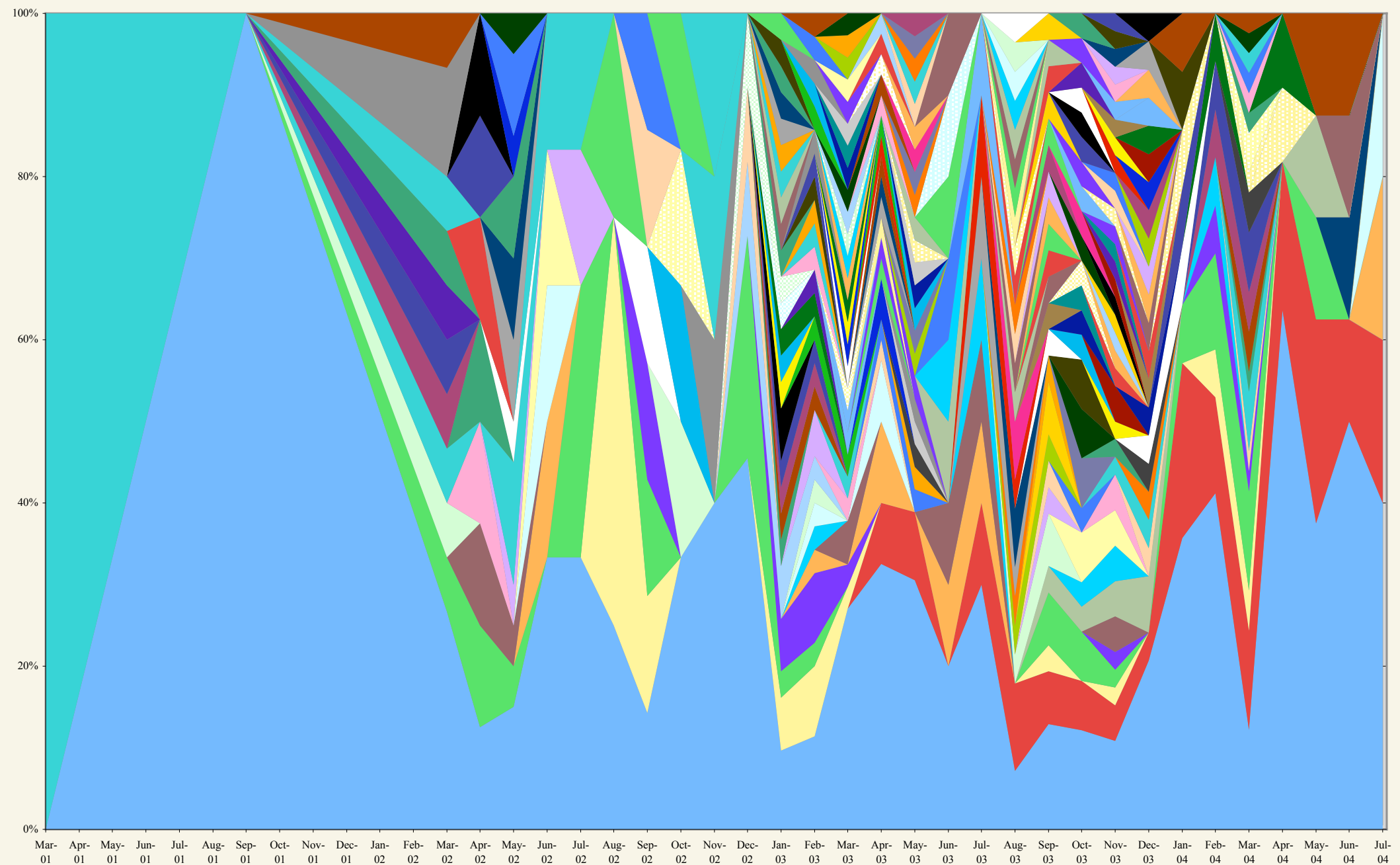
AskERIC Histogram

AskERIC Top Cited URL's 2000-2003



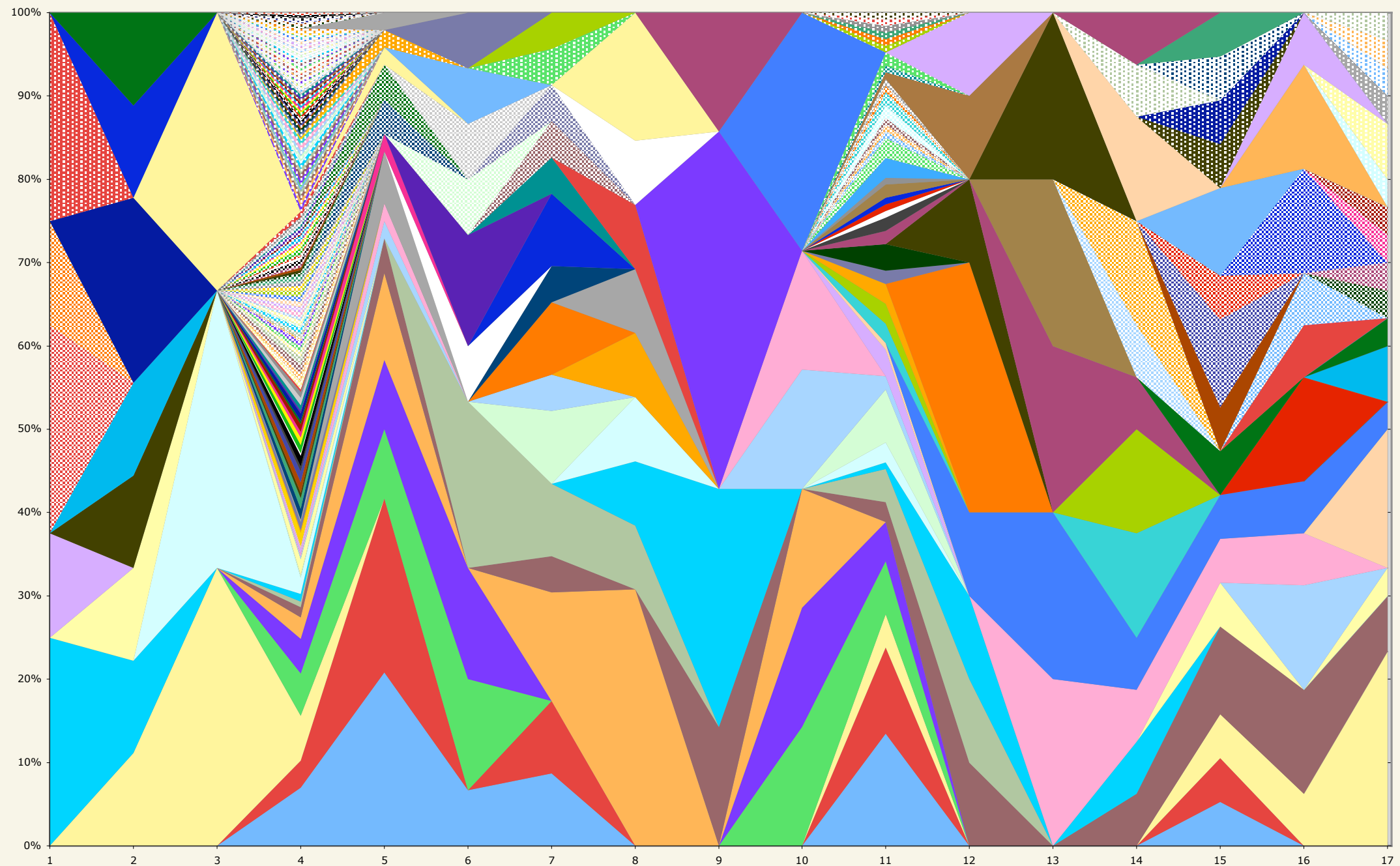
VRD Histogram

VRD Top Cited URLs 2001-2004

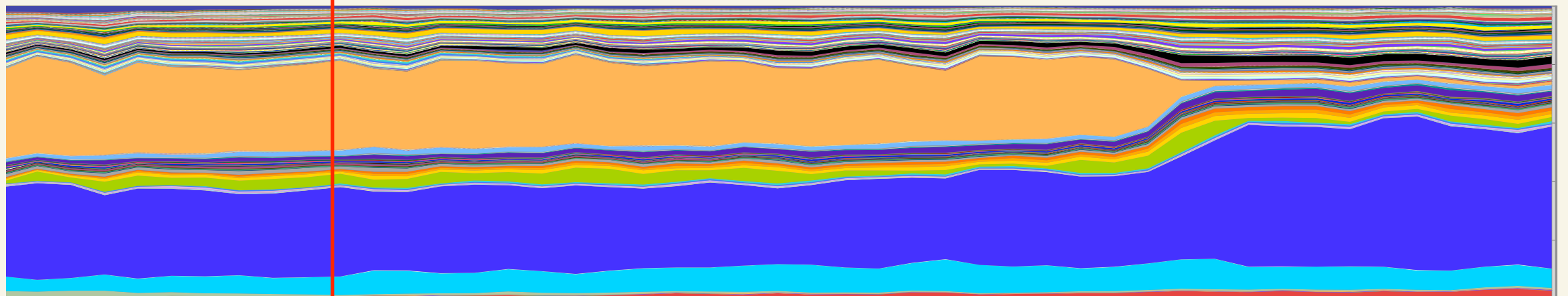
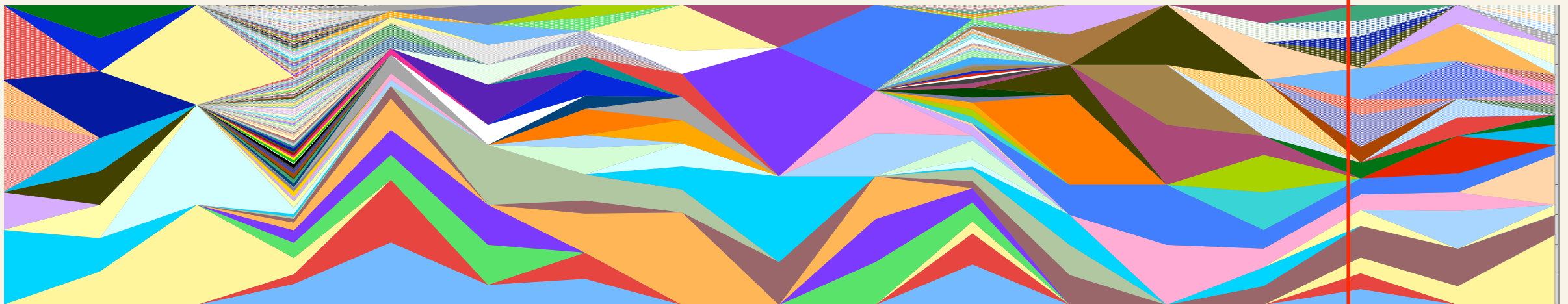


AskNSDL Histogram

AskNSDL Cited URLs 2002-2005



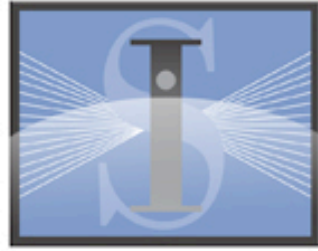
“Weighting Window”



Story Starters

- ~ High Context
- ~ Builds on Blogs
- ~ Reference as Community





The Information Institute of Syracuse's New **Weapons of Mass Instruction**



R e f e r e n c e E X T R A C T & S t o r y S t a r t e r s