A.nalysis of S.ocially I.ntelligent S.tructures:

Titular Dottedness and ASIS.D.O.T.S.

Research has shown that users prefer systems with easy to remember acronyms. Users also prefer systems with acronyms that make them feel good about themselves.

Researchers surveyed 100 persons living in a mobile home park in Texas and determined that 80% of them believed the term "American" to be International in scope, especially now that David Hasselhof is a big hit in Germany. Interestingly enough, 100% surveyed believed that "Texan" had more International appeal than "American" and thought that the system ought to be named:

TexSIS (Texis Structure for Information Science)

This supports our hypothesis that users prefer to use systems whose names are familiar, easy to remember, and that make them feel good about themselves.

Such feel-good properties are also found in an **A.nalysis of S.ocially I.ntelligent S.tructures** in web pages.

Llewellyn Puppybreath, following the inspired study on the relationship of titular colonicity (that is, whether a title of a paper contains a colon) and the scholarship of an article, set out to establish some corresponding theses about web publications. After a detailed analysis of web sites, he came to the brilliant insight that the number of dots in a URL correlates to the intellectual substance of the web site. The title of his paper recently submitted for peer review poses the following thesis:

THESIS ONE

"Titular dottedness: the master key to web filtering?"

After painstaking examination of thousands of porn sites, he discovered that they had very few dots in their URLs whereas several universities and art museums had at least three or more. If his thesis holds true, it will fulfill the Holy Grail of information retrieval research: a pure and simple quantitative measure for relevance judgements. The mind boggles with the prospect of a flood of NSF grants!

1 dot = asis.org

2 dots = www.kent.edu

3 dots = new.york.times.com

THESIS TWO

This thesis is not the only one whose attractiveness has engaged Llewellyn Puppybreath. In fact, it lies in contention with another seductive thesis:

The information density of a web site quality is inextricably linked to the melange of mixed cases for directories, subdirectories and homepages.

In short, the greater the number of mixed cases in the URL, the greater the significance of information content of the web site.

Llewellyn sees this as a simple extension of Shannon's law:

$$I = -\sum_{n=1}^{n} p_i \log_2 p_i$$

***I* = information density of a website**

***i* = case change count, initialized to 1**

***n* = total number of case changes**

Whereas the information content of a message for Shannon is the average number of binary digits necessary to isolate a message among a set of possible messages, for Llewellyn, the information content of a web site is a function of the probability of case changes vis-a-vis the total number of case changes.

THESIS TWO: ILLUSTRATION

Example: MuseumOfNaturalHistory.org/EntranceHall

There are six case changes:

Change	***Probability***	***Log_2***	***$p*Log_2$***
1	.10	-3.32	-.3320
2	.25	-2.00	-.5000
3	.40	-3.32	-.5280
4	.20	-1.32	-.4640
5	.03	-5.06	-.1518
6	.02	-5.64	-.1128

***I* = -(-2.0886)**

***I* = 2.0886**

The higher the number the higher the density, because the higher the number of changes of cases, the higher the information quantity of the site. Taking a site like asis.org where there are no case changes, the information content = 0. In fact, so are many university sites....

I = 0

Obviously we need to change this sad state of affairs.

THESIS THREE and FOUR

His excitement with these results has led him to postulate some other theses:

> **link unpopularity and title length are indicators of web site quality**

In the first case, he argues that if few sites link to a site, the better the quality of that site. Unfortunately, this thesis is constrained by time. If, after a web page is mounted, there are few links to it for several months, this may be an indicator of web quality. If a site becomes popular after its quality is discovered, this thesis declines in value.

In the second case, length of URL is strongly correlated with information density of the site.

OCCAMS RAZOR for THESIS ONE

After reviewing these theses, Llewellyn Puppybreath has decided to opt for Occam's razor: namely the simplest thesis is the best. And so his paper centers on establishing that titular dottedness is the simplest and best measure of web site quality.

This will be known as Puppybreath's Law, a new weapon in the arsenal of bibliometrics. What makes this discovery exciting is that it is computationally convenient: web search engines will find it simple to count titular dots.

Should his paper be rejected by the dottiness of his peers, so convinced is Llewellyn Puppybreath of the truth of his thesis, he will not be dissuaded from publishing it on the web. In fact, this paper can be found at:

http://the.Truth.is.Known.and.We.have.Found.it/LlEwELyN.PuPpYbReaTh/ScHoLaRShiP.Of.diSCoVery/REcent.PaPErs/A.nalysisofS.ociallyI.ntelliGentS.tructures/TiTuLaR.Dottedness.html

PREPOSTEROUS PRECISION

If we do not factor Puppybreath's law, then we have the current state of affairs in search engine usage: preposterous precision. Raucous recall is the number of vaguely relevant documents over all documents retrieved.

	relevant	not relevant	pathetically irrelevant
retrieved	hit (A)	false drop (B)	falsetto drop (C)
not retrieved	miss (D)	acceptably non-retrieved (E)	why not? Everything else was! (F)

$$\textbf{Preposterous Precision} = \frac{A}{D + E + F}$$

With Puppybreath's Law, C and F can be considerably curtailed. And it would also have major impact on:

$$\textbf{Raucous Recall} = \frac{A}{A + B + C}$$

ASIS.D.O.T.S

Driven by his study of the role of dottedness in webpage URLs, Llewellyn has been inspired to postulate a new name for ASIS. While membership in societies may be enhanced with digital correctness, acronym memorability, and alphabetical inclusivity, one could hardly doubt that the real character of inclusivity and scholarship is the DOT. Thus, the only real viable option for a name change for ASIS should be

ASIS.D.O.T.S.

ASIS.D.O.T.S. – the dot without the domain would suggest that ASIS is a society for all domains not just .org or .com or .edu. What could be more inclusive than being domainless? (And Llewellyn cautions against detractors who would argue that this amounts to being lost in cyberspace).

ASIS.D.O.T.S. would set a precedence for high scholarly standards because of the his discovery of the significance of the scholarship associated with DOTS. Llewellyn is not unmindful of the fact that there will be attempts to make signification of the .D.O.T.S., and he lends his support so such interpretations as: **Digital Obliqueness and Technological Superiority**. Who would possibly refuse to join the American Society for Information Science, Digital Obliqueness and Technological Superiority? Who could not doubt that members of such a society would not feel good about themselves? In fact, who would not doubt that all information professionals would lust after such a membership? Among the digerati, the **dotterati,** particularly the **ASIS dotterati**, will be the creme de la creme.

In order to advance his cause for the adoption of this new moniker for the society, Llewellyn Puppybreath has promised the development of a new unique search engine, suited for the membership of ASIS.D.O.T.S.. Such an engine would not only identify quality sites, but its relevance ranking algorithm, while not ignoring query term frequency and location, would tweak the results based on titular dottedness, link unpopularity, digital malformity, titular length and titular case complexity.

Let’s all herald the arrival of the most progressive association for information professionals:

ASIS.D.O.T.S.