

# The Authentication of Textual Assertions

**Roger Clarke**

Xamax Consultancy Pty Ltd, Canberra  
Visiting Professor, ANU RSCS, UNSW Law

<http://www.rogerclarke.com/ID/ATS-ISF> { .html, .pdf }

**10th ISF Workshop – ANU  
10-11 September 2024**

# Motivation

- The origins of the IS profession and of the IS discipline are in structured data
- Other data forms have been less of a focus
- **Text has now become central to IS**
- There is a great deal of concern about the reliability of textual statements
- **IS needs a framework within which textual statements can be authenticated**
- We need to draw on relevant meta-theory in order to expand the IS discipline's horizons

# Data Forms

- Structured data: files of records of data-items, each with defined formats
- Audio
- Images, Video
- Vector Graphics, Animations
- Software
- ...
- Text

# Text in IS

- Tightly-defined data-items containing text-data since the early 1960s
- Free-form 'comments' data-items since the mid-1960s
- Text-string handling facilities since the late 1960s
- Free-text search since 1970 (Docu/Master, ICL Status, IBM STAIRS)
- Email and Bulletin Board Systems since the end of the 1970s
- Library systems, initially cataloguing metadata, by 1980 with content
- Internet text-handling capabilities in the 1980s
- Self-publication of text on the World Wide Web from the early 1990s
- Discovery of text-content from the early-to-mid 1990s
- Evidence management systems

## Text in IS

- Tightly-defined data-items containing text-data since the early 1960s
- Free-form 'comments' data-items since the mid-1960s
- Text-string handling facilities since the late 1960s
- Free-text search since 1970 (Docu/Master, ICL Status, IBM STAIRS)
- Email and Bulletin Board Systems since the end of the 1970s
- Library systems, initially cataloguing metadata, by 1980 with content
- Internet text-handling capabilities in the 1980s
- Self-publication of text on the World Wide Web from the early 1990s
- Discovery of text-content from the early-to-mid 1990s
- Evidence management systems
- Text surveillance as a sub-field of data-/cyber-security, since 2000
- Synthetic textual statements by Generative AI since mid-2010s / 2022

## The Objective

- Provide guidance to IS practice and practice-relevant IS research for **the authentication of textual statements**
- Draw on descriptive linguistics
- Apply a generic theory of authentication
- Account for relevant complexities

## Some Examples of Textual Statements

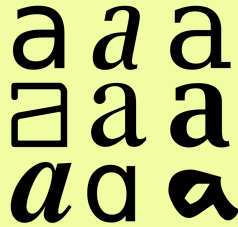
- This investment will return you 16% p.a.
- My only borrowings are a mortgage loan on my house, and a credit card
- I've made no insurance claims in the last 5 years
- We can deliver that item within 48 hours from stock held in your nearest capital city
- I have a doctorate in IT from Somerset University
- I earned \$160 in the last fortnight from casual work in my local cafe

## Structure of the Presentation

1. **Text**, its Representation, its Interpretation
2. **Assertions**, Textual Assertions
3. **Reliability**, Misinformation, Disinformation
4. **Authentication** Theory, Evidence
5. Authentication of **Textual Assertions**
6. **Reliability Criteria** for Textual Assertions
7. **Tests of Efficacy** of the Reliability Criteria

## The Nature of Text

- Natural Language, visibly expressed using characters, symbols or glyphs



## Fonts / Scripts

texting

texting

Text

text

text

نص



## The Nature of Text

- Natural Language, visibly expressed using characters, symbols or glyphs
- Logographic Scripts
- Alphabets
- Syllabaries

## Logographic Scripts

**Logogram:** A written character that represents a semantic component of a language

Many began as a picture, came to represent a word(-part)

- Cuneiform script  (1,000 Sumerian chars)
- Egyptian hieroglyphics  (several sets)
- Chinese kaishu 中文 (c. 50,000 characters)
- Japanese Kanji 漢字 (c. 50,000 characters)  
(c. 4,000 for literacy)
- Korean Hanja 한자 (c. 8,000 characters)
- Other ancient Asian scripts, and Mayan and Aztec

## Alphabets

- **Alphabet:** A set of glyphs, each representing short sound(s): consonants (mouth closed), vowels (open)  
Sinai c.1700BC, via the Phoenicians c.1000-500BC (via Ancient Greek and Latin, to Western Europe)
- **English Alphabet** (a successor to Latin / Roman): 26 letters, for 24 consonantal and 20 vowel sounds
- **Other Alphabets:** >370, some with diacritics (ç ü)  
Flexible, with many used for multiple languages  
Alphabet size 21-74 characters (74 in Khmer)
- **Abjads** – consonantal characters, mostly not vowels
- **Abugidas** – indicate vowels with diacritics
- ? Chinese pinyin (21 initials, 38 finals, 5 tones)

## Syllabaries

- **Syllabary:** A set of glyphs, each representing an extended sound with consonants and vowels
- Mycenaean Greek, Linear B (87 characters)
- Japanese Hiragana (48 characters)
- Japanese Katakana (48 characters)
- Cherokee (85 characters)
- ? Chinese pinyin (initials x finals x tones = c.400)

## The Diversity of Textual Forms

- **Written Natural Languages:**
  - >3,000, c. 30 of them with >1m users
- **Alphabets, Syllabaries, Logographic Scripts:**
  - c.300, >150 of them in contemporary use
  - >1500 languages use the Latin Alphabet
- **Unicode representations**
  - c. 150,000 characters

## The Interpretation of Text

- **Data-Item Domain Definitions**  
**Controlled Vocabularies**  
Permitted Values of Colour = red, green, blue  
and/or rot, grün, blau and/or 赤 緑 青, etc.)

## The Interpretation of Text

- **Data-Item Domain Definitions**  
**Controlled Vocabularies**  
Permitted Values of Colour = red, green, blue and/or rot, grün, blau and/or 赤 緑 青, etc.)
- **Semiotics:** study of symbolic communication
  - **Syntactics:** Symbol-Symbol relations
  - **Semantics:** Symbol-Thing relations
  - **Pragmatics:** Symbol-Interpreter relations

## Syntactics versus Semantics

Lightning strikes and bush-fires  
not permitted at this installation

## Semantics versus Pragmatics

Dept of Environment Property  
Trespassers will be prosecuted

## Semantics versus Pragmatics

Dept of Environment Property  
Trespassers will be prosecuted

**CROCODILE-INFESTED SWAMP**  
Survivors will be prosecuted

## Pragmatics

I asked someone how glutenous is spelt

I was a bit bemused when they answered "very"

## Pragmatics



## Pragmatics



## Challenges for Utterers

- Choice of Words and Expressions
- Style
- Context-Relativity
- Culture-Relativity
- Time-Relativity
- Nuance, Ambiguity, Playfulness
- Transcription from Verbal Speech

## The Hermeneutic Proposition

- The Utterer's intention is not directly conveyed
- Rather, text is interpreted by the Reader
- Interpretation is based on the Reader's
  - state-of-mind (Heidegger 1962)
  - pre-understanding (Gadamer 1976)
  - background (Winograd & Flores 1986)

## xkcd on Hermeneutics



## The Hermeneutic Proposition

- The Utterer's intention is not directly conveyed
- Rather, text is interpreted by the Reader
- Interpretation is based on the Reader's
  - state-of-mind (Heidegger 1962)
  - pre-understanding (Gadamer 1976)
  - background (Winograd & Flores 1986)
- Interpolated sub-texts
- Re-reading using a different value-set
- Deconstruction (Derrida)

## Structure of the Presentation

1. **Text, its Representation, its Interpretation**
2. **Assertions, and Textual Assertions**
3. **Reliability, Misinformation, Disinformation**
4. **Authentication Theory, Evidence**
5. **Authentication of Textual Assertions**
6. **Reliability Criteria for Textual Assertions**
7. **Tests of Efficacy of the Reliability Criteria**

## RFI #1

### Are any Key Aspects of Text Missing?

- **The Representation of Text**
  - Visual Expressions of Natural Language
  - Alphabets, ..., Syllabaries, Logographic Scripts
  - The Unicode standard's c.150,000 characters
  - ...
- **The Interpretation of Text**
  - Disciplined Use (controlled vocabulary / thesaurus)
  - Grammar and spelling diversity, richness, ambiguity, 'word-plays'
  - Semiotics: Syntactics, Semantics, Pragmatics
  - Formal text cf. Transcription of spoken words
  - Hermeneutics: Interpretation; pre-understanding/state-of-mind
  - Imputed sub-texts; Alternative value-sets; Deconstruction

## Structure of the Presentation

1. **Text**, its Representation, its Interpretation
2. **Assertions**, Textual Assertions
3. **Reliability**, Misinformation, Disinformation
4. **Authentication** Theory, Evidence
5. Authentication of **Textual Assertions**
6. **Reliability Criteria** for Textual Assertions
7. **Tests of Efficacy** of the Reliability Criteria

## Relevant Occurrences of 'Assertion' in the >17,000 Corpus in AISeL

- **Specific to Assertions re Identity, Attributes**  
Security Assertion Markup Language (SAML), 'Claims'
- **Generic, encompassing Identity, Attributes**  
One directly relevant  
(+ Predecessor papers of my own at ACIS and Bled)

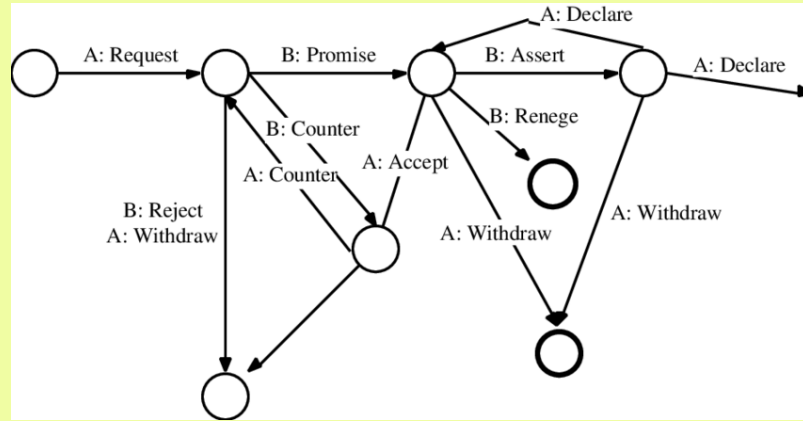
Mingers J. & Standing C. (2020) 'A Framework for Validating Information Systems Research Based on a Pluralist Account of Truth and Correctness' JAIS 21,1

## Assertions

- **An Assertion:** A statement on which some entity may place reliance, and hence may perceive value in having adequate grounds for treating the content of the utterance as being reliable
- **A Textual Assertion:** An Assertion expressed in a natural language in visual form



## Expressions that Constitute Assertions



Conversation for Action (Winograd & Flores 1986)

## Expressions that Constitute Assertions

Speech Act Theory (Searle 1975):

- 'performative' utterances constitute acts
- 'illocutionary' utterances motivate acts

## Expressions that Constitute Textual Assertions

Table 1 of 3

Mode	Mood	Description
A Fact-Based <b>Assertive</b>	Indicative or Evidential	A statement as of fact
A Belief-Based Assertive	Indicative or Evidential	A statement of belief that is intended to be understood as being determinative
An Opinion-Based Assertive	Indicative or Evidential	A statement of opinion in a context in which the statement is imbued with authority, or made by a person claiming relevant expertise

## Expressions that Constitute Textual Assertions

Table 2 of 3

Mode	Mood	Description
A Necessary <b>Implication</b>	Indicative or Evidential	A necessary implication of a statement
A Power-Based <b>Inference</b>	Indicative or Evidential	A sufficiently firm nudge by a first party with effective power over the second party
An Authoritative <b>Imputation</b>	Indicative or Evidential	An inference drawn by a party perceived to have formal authority
A Strong <b>Potential</b>	Potential	A strongly probabilistic (cf. merely possible or tentative) statement, particularly one that includes an indication of relevant factors likely to determine the outcome

## Expressions that Constitute Textual Assertions

Table 3 of 3

Mode	Mood	Description
A Directive	Imperative	A command
A Directive	Deontic	A statement of obligation
A Commissive		A commitment to an action
An Expressive		An indicator of attitude
A Declaration		An authoritative pronouncement
A Conditional	Conditional	A statement that includes an express indication, or a sufficiently clear implication, of contexts in which the statement does and/or does not apply

## Expressions that Constitute Textual Assertions Combining Tense and Aspect

- **Assertions about Now:**
  - Present tense does
  - Continuing or progressive present is doing
- **Assertions about Back Then:**
  - Past or preterit tense did
  - Imperfect tense used to do
  - Perfect tense has done
  - Pluperfect tense had done
- **Assertions about Later:**
  - Future tense will do
- **Assertions Not Time-Bound (generic, universal or existential):**
  - 'Ys are generally superior to other races'
  - 'All Xs are white'
  - 'There is at least one Z'

## Expressions that are NOT Textual Assertions

Mode	Mood	Description
Questioning	Interrogative	Asking a question, e.g. 'Is this customer entitled to a loyalty discount?'
Optative	Hope	Expression of a wish, e.g. 'I hope this driver is licensed'
Postulating	Presumptive / Hypothetical	Expression of a possibility for consideration, e.g. 'Just suppose that this job applicant had already finished the degree that they're enrolled in'
Tentative	Potential ('may', 'might')	Conveying uncertainty about an outcome and about factors that will affect the outcome, e.g. 'Customers might not be happy with delivery delays'
Conditional	Subjunctive ('would', 'could', 'should', 'ought')	Conveying conditionality, e.g. 'This job applicant would be attractive if they had a driver's licence'
Weak Assertive	Opinion or Belief	Assertives not intended to be seen as determinative and/or uttered without special expertise, e.g. 'I think it's important that applicants be able to drive', 'This candidate doesn't seem appointable'

## Structure of the Presentation

1. Text, its Representation, its Interpretation
2. Assertions, Textual Assertions
3. Reliability, Misinformation, Disinformation
4. Authentication Theory, Evidence
5. Authentication of Textual Assertions
6. Reliability Criteria for Textual Assertions
7. Tests of Efficacy of the Reliability Criteria

## The Reliability of an Assertion

- Reliability has to do with **trustworthiness**
- It may be associated with performance against spec. (in engineering), or performance consistency (stats)

In the present context, however:

- Reliability of an Assertion means the extent to which an entity has the perception that the statement is trustworthy, based on its correspondence with reality

## Related Notions To Be Investigated

- Warranted Assertibility
- Rational Assertability
- Rational Acceptability

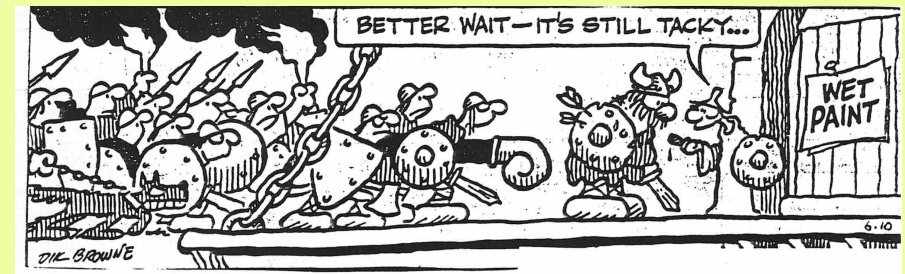
Dewey J. (1938) 'Logic: the Theory of Inquiry' Holt, 1938  
Habermas J. (2003) 'Truth and Justification' Polity, 2003  
quoted in

Mingers J. & Standing C. (2020) 'A Framework for Validating Information Systems Research Based on a Pluralist Account of Truth and Correctness' JAIS, 21,1

## Mis- & Dis-Information Contexts

- The Law
- 'Natural Philosophy'
- The Sciences
- Politics
- Public Relations
- Advertising
- Media
- Social Media

## The Spurious Authority of the Written (or Painted) Word



## The Mythology of Truth

- Yellow journalism  
The Tabloid Press  
Fake News  
Trumpian Post-Truth
- Air-Brushing  
Photoshop-ing  
Deep Fakes  
Generative AI
- A Metatheoretic Commitment:  
**Assuming Human-Accessible Truth is Untenable**
- The True/False Dichotomy is impractical,  
whether proposed by Floridi or by the EU  
(‘Veridicality’, ‘Alethic modalities’; ‘False Information’)

## Mis- & Dis-information

- **Disinformation** means one or more unreliable Assertions that were demonstrably intended by the utterer to mislead the reader or hearer
- **Misinformation** means one or more unreliable Assertions that were **not demonstrably intended** by the utterer to mislead the reader or hearer

## RFI #2

### Textual Assertions and Their Reliability

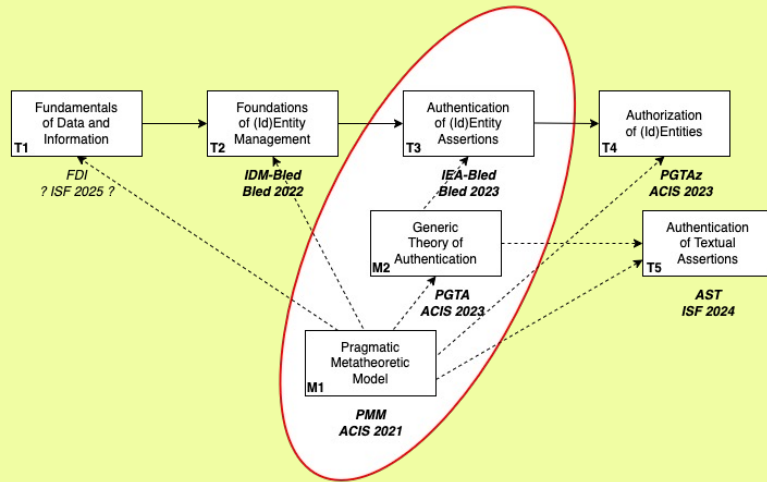
How do I find relevant prior works  
in IS and closely-associated literatures:

- Authors?
- Keywords?
- Conferences? Special Issues?
- Other?

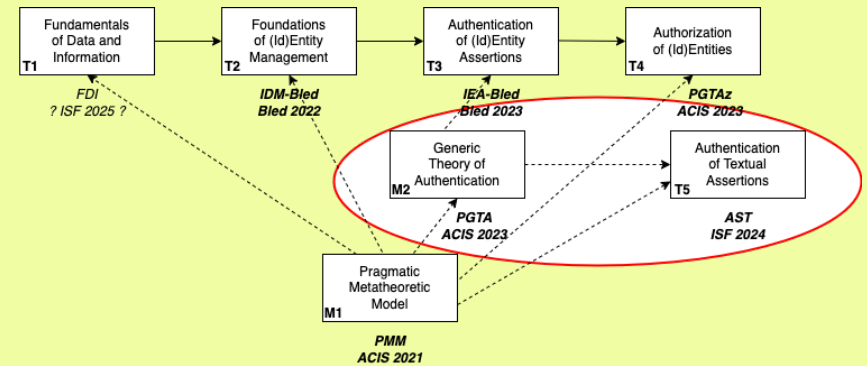
## Structure of the Presentation

1. Text, its Representation, its Interpretation
2. Assertions, Textual Assertions
3. Reliability, Misinformation, Disinformation
4. **Authentication** Theory, Evidence
5. Authentication of **Textual Assertions**
6. **Reliability Criteria** for Textual Assertions
7. **Tests of Efficacy** of the Reliability Criteria

## A Generic Theory of Authentication



## A Generic Theory of Authentication



## A Generic Theory of Authentication

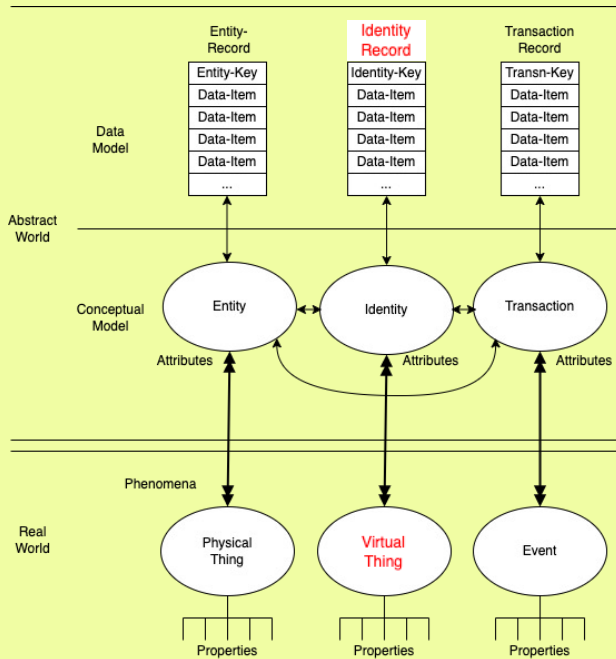
- **Authentication** is a process that establishes a degree of confidence in the Reliability of an Assertion
- **Syntactic Level**  
Compliance with definitions, content-specifications and/or the rules of logic and grammar, to ensure that there is no invalid expression or inference
- **Semantic and Pragmatic Levels**  
Consistency with reality and perception

## A Generic Theory of Authentication

- **Evidence**, in relation to the Authentication of an Assertion, is data that tends to support or deny that Assertion
- Evidence may be existing data already held, or acquired from, or accessed at, some other source that is adequately reliable and is sufficiently independent of any party that stands to gain from masquerade or misinformation
- Evidence may be newly acquired from a person or active artefact, e.g. by means of challenge-response

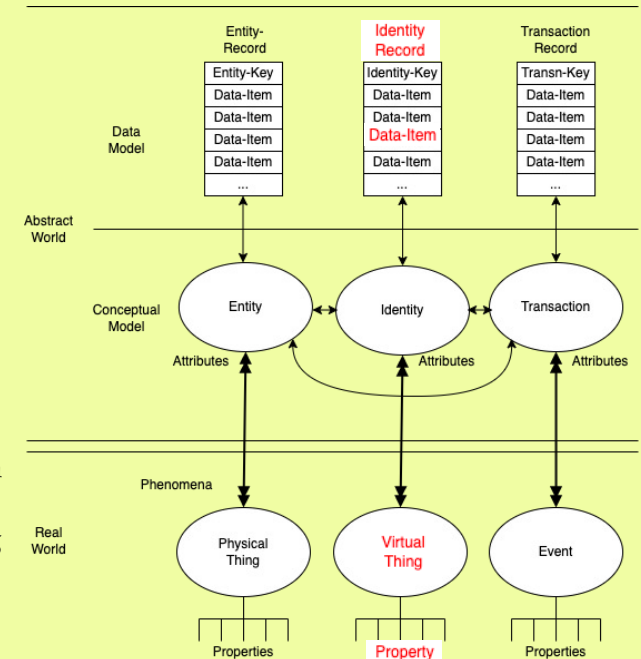
### (1) (Id)Entity Assertion

A particular **Thing** is appropriately associated with one or more (Id)Entity-Records



### (3) Property Assertion

A particular **Data-Item-Value** in a particular (Id)Entity Record is appropriately associated with, and reliably represents, a particular **Property** of a particular Thing



## Categories of Assertion Important in (e)Commerce

- (1) (Id)Entity
- (2) (Id)Entity Match
- (3) **Property (Simple)**
- (4) Property (Complex)
- (5) Principal-Agent
- (6) Monetary **Asset-Value**
- (7) Non-Monetary Asset-Value
- (8) **Content Integrity**

## Structure of the Presentation

1. **Text**, its Representation, its Interpretation
2. **Assertions**, Textual Assertions
3. **Reliability**, Misinformation, Disinformation
4. **Authentication Theory**, Evidence
5. Authentication of **Textual Assertions**
6. **Reliability Criteria** for Textual Assertions
7. **Tests of Efficacy** of the Reliability Criteria

## Assertions as of Fact

### A Textual Assertion that involves at least one Real-World phenomenon

- An Assertion as of Fact may or may not be:
  - Supported by Evidence
  - A reasonable inference from Evidence
  - The subject of Counter-Evidence
- Many Assertions as to Fact are merely unsupported statements that the utterer claims that it is so
- Reliance on an Assertion as to Fact is advisable only where Evidence exists, and supports it more strongly than the Counter-Evidence

## Relevant Assertions as of Fact

- **Out-of-Scope**
  - Ineffable mysteries (of a spiritual nature)
  - Evocative art-forms (poetry, ballads)
  - Graphical elements dominant over text elements
- **Within-Scope**
  - Statements about real world phenomena
  - Statements about associations between real-world properties and abstract-world data-items
  - Polemic, hyperbole and satire (because they affect attitude and behaviour)
  - Structured text (bullets, numbered segments, tables)

## Quality Factors in Authentication

- Authentication Approaches depend on context, e.g.
  - In Courts
  - In Law Enforcement Investigations
  - In Commercial Transactions
- Authentication Quality depends on Data Quality
- Data Quality reflects the data source's motivations
- False Positives and False Negatives arise
- Cost Trade-Offs and Risk Management are applied

## Structure of the Presentation

1. **Text**, its Representation, its Interpretation
2. **Assertions**, Textual Assertions
3. **Reliability**, Misinformation, Disinformation
4. **Authentication** Theory, Evidence
5. Authentication of **Textual Assertions**
6. **Reliability Criteria** for Textual Assertions
7. **Tests of Efficacy** of the Reliability Criteria

## Assertions as of Fact Authentication and Evidence

- Easy: Assertions that involve structured data (the number of people at a meeting, the date of an event)
- The Laws of (Court) Evidence
- The Norms of Criminal Investigation
- ‘Critical Thinking’
- Fallacy Theory
- ...
- ‘How to spot fake news’
- Automated Fact Checking (emergent techniques)

## Fallacy Theory

Ad hominem	Attack on some aspect of the person making an argument rather than on the substance of the argument
Ad misericordiam	Appeal to pity, a 'sob story'
Ad populum	Invocation of the majority view rather than a logical analysis
Ad verecundiam	Invocation of authority, rather than logic

Complex [or trick] question	Confusion induced by complex expression ('Do you still beat your wife?')
-----------------------------	--------------------------------------------------------------------------

Secundum quid	Assumption that a heuristic or rule of thumb (which often holds) is a universal rule (which always holds)
Gambler's fallacy	Faulty assumption that prior events influence the next event (as in successive coin tosses)
Ignoratio elenchi	Faulty assumption that a point made is relevant to the issue

## Digital Misinformation Information Reliability Criteria

- **The Author**  
incl. publishing record, background, affiliations
- **The Venue**  
incl. its nature, standing, ownership and goals
- **The Content's Origin and Nature**  
incl. sources cited, style (e.g. narrative, analysis or opinion), expression quality, coherence and cohesion of the argument
- **The Context of the Work**  
incl. declared and apparent motives, declared and apparent perspective(s) on the subject
- **Replication of the Content**  
incl. republication, review, quotation, citation

## RFI #3 What other Sources / Approaches to Reliability Assessment?

- Easy: Assertions that involve structured data (the number of people at a meeting, the date of an event)
- The Laws of (Court) Evidence
- The Norms of Criminal Investigation
- ‘Critical Thinking’
- Fallacy Theory
- ‘How to spot fake news’
- Automated Fact Checking (emergent techniques)
- ...



## Reliability Criteria for Assertions as to Fact

Criterion	Guidance
<b>The Work</b>	
• The Author	Check their identity, affiliations, prior works
• The Venue	Check its nature, ownership, declared mission, reputation
• The Style	Check for expression, rigour, polemic and satire
• The Values	Identify the values embedded in the content
• The Author's Purpose	Identify the declared and/or implied or inferred motivations for creating the Work
<b>The Author's Assertions</b>	
• Their Clarity	Isolate the key assertions, and examine for ambiguities, emotional language
• Their Internal Consistency	Identify related passages within the Work, and check for inconsistencies among them
• The Sources	Identify declared, implied and inferred sources, including by Web-searches on key expressions and passages Evaluate any key sources, using this Guidance
<b>The Author's Argument</b>	
• Its Clarity	Extract and critique the flow of the author's argument
• Its Premises and Assumptions	Identify the starting-points of the argument
• The Logic	Test for fallacious forms of argument. See #T5Table 5
<b>Evidence</b>	
• Other Sources	Perform searches independently of the Work, for sources on key aspects of the Premises and Assumptions
• Their Content	Assess the consistency of the Work with those Sources
<b>Counter-Evidence</b>	
• Other Sources	Perform searches independently of the Work, for sources on key aspects of the Premises and Assumptions
• Their Content	Identify inconsistencies between the Work and those Sources

## Structure of the Presentation

1. Text, its Representation, its Interpretation
2. Assertions, Textual Assertions
3. Reliability, Misinformation, Disinformation
4. Authentication Theory, Evidence
5. Authentication of Textual Assertions
6. Reliability Criteria for Textual Assertions
7. Tests of Efficacy of the Reliability Criteria

## Test-Cases – 1

### The Iraqi Weapons of Mass Destruction

During 2003, the US Administration of George W. Bush enlists the UK Prime Minister Tony Blair to create public belief that **Sadam Hussein's Iraq possessed 'weapons of mass destruction' (WMDs)** and is prepared to use them, and hence invasion is justified.

### Trump and the Storming of the Capitol

In January 2021, Donald Trump addresses supporters at a rally 2.5 kilometers from the Capitol, where a joint session of Congress is conducting proceedings to certify the results of the election and thereby end Trump's Presidency. The Capitol is stormed. **Trump incited or at least encouraged the storming of Congress; or conspired to obstruct Congress.**

## Test-Cases – 2

### Textual Assertions underlying Robodebt

The Australian social welfare payment each fortnight depends on the amount of income the individual has declared for the previous fortnight. Most recipients are in, at best, sporadic employment. So some overpayments occur.

**The social welfare agency intends to acquire the data declared annually to the taxation agency, divide it by 26 to impute fortnightly earnings, and demand repayment of amounts that appear from that data to have been wrongly paid.**

The Assertions to be authenticated are that:

- **The 'income averaging' method is reliable**
- **The process is lawful**

## Test-Cases – 3

### Textual Assertions that:

- A Payment has been made
- An Entity has a particular Property
- An Online User is a person not an active artefact

### Textual Assertions underlying:

- Asset Valuations
- Claims about New Sources of Scarce Raw Materials
- IPOs
- Technological Inventions
- Business Cases

### Textual Assertions in Job Applications

## The Argument

- ‘Authentication’ is used in IS to refer only to a very narrow category of Assertions about Identity
- Many other categories of Assertions are important
- Most data has been digital for several decades now, whether structured, graphical, audio or textual
- **Quite suddenly, technology has emerged that automatically generates Textual Assertions**
- **Organisations expect IS to encompass Assertion Authentication, nomatter what form the data is in**
- **IS practitioners need practice-relevant research to underpin the Authentication of Textual Assertions**

## The Contribution

- Outlines of the nature of text, of representations of text, and of the interpretation of text, which lay the **foundation for understanding Textual Assertions**
- The use of concepts from descriptive linguistics to enable **identification of the categories of Textual Assertion on which human activities rely**
- **Definitions** of Textual Statement Reliability, Misinformation and Disinformation
- The **application** to Textual Assertions of a **Generic Theory** of Authentication
- **Reliability Criteria** for Textual Assertions as to Fact

## The Authentication of Textual Assertions

### Roger Clarke

Xamax Consultancy Pty Ltd, Canberra  
Visiting Professor, ANU Computing, UNSW Law

<http://www.rogerclarke.com/ID/ATS-ISF> { .html, .pdf }

**10th ISF Workshop – ANU**  
**10-11 September 2024**