



Collaboration space, Alexandria, VA

# CRYPTOMARKET DATA FOR RESEARCH AND REGULATORY REPORTING

*Logistics and Implications for Abuse Potential Evaluation*

12 OCTOBER 2018

CCALC – BETHESDA, MD

# Cultural Context

Online discussion, debate and distribution



# Evolution of Markets

A confluence of ideas and technologies

## **TOR**

Access anonymously



## **Cryptocurrencies**

Pseudonymous transactions



## **Peer-to-Peer Encryption**

Provably private connections



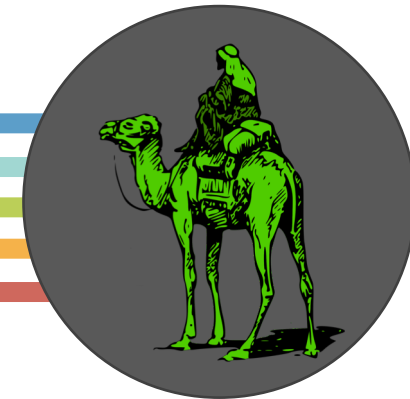
## **Machines That Forget**

Amnesiac Operating Systems



## **eCommerce**

Internet mediated sales



# Claims and Aims

Reasons for online discussion distribution

**Quality**  
Purity and potency

**Safety**  
Avoiding violence

**Freedom**  
Preserving liberty



**Access**  
Seeking sources

**Community**  
Seeking solidarity

# Market Mechanics

Logistics and Liabilities in Online Sales

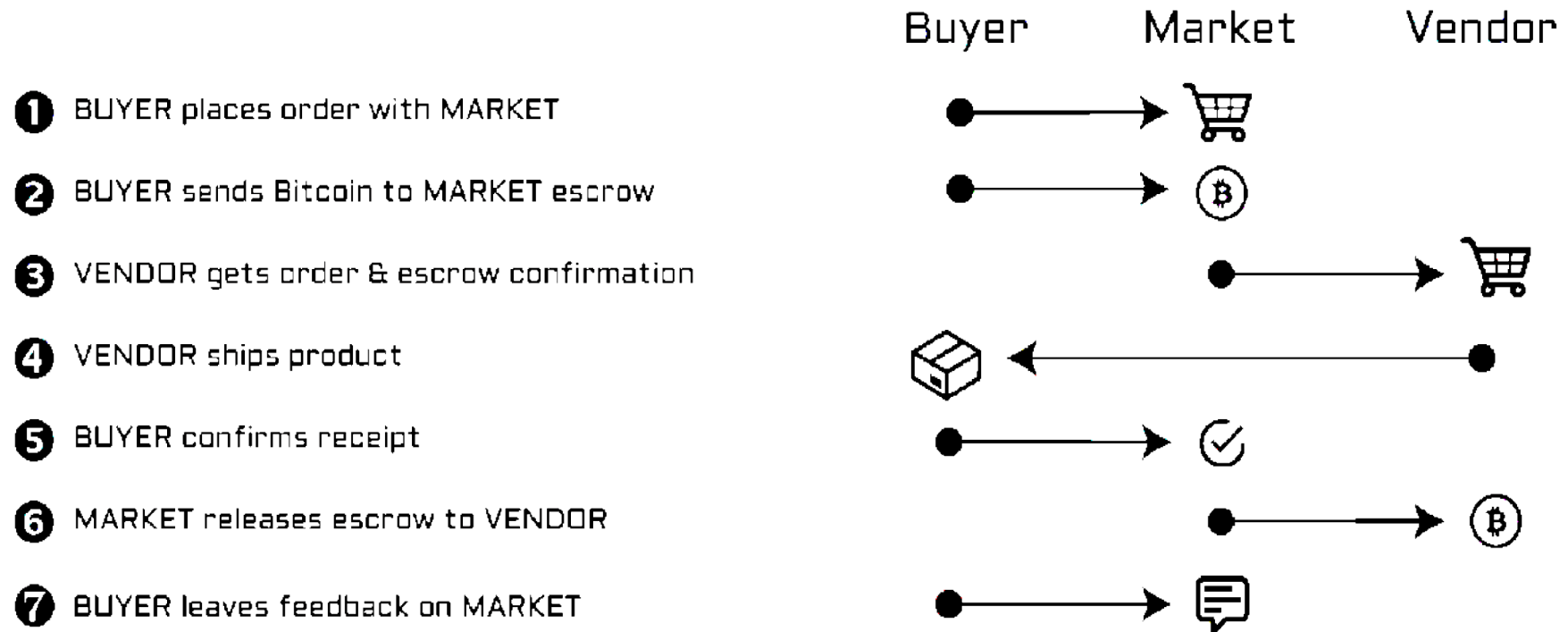


Image Credit: [http://www.ijdp.org/article/S0955-3959\(17\)30156-1/fulltext](http://www.ijdp.org/article/S0955-3959(17)30156-1/fulltext)

# Recursive Disruption

---

Responding in intrigue and interdiction

**Multisig**  
Automated Contracts

**Chainbreaking**  
Fungibility focused coins



**Decentralization**  
Distributed Networks

**Redundancy**  
Resilience through persistence

---

# Learning About Products

---

Understanding products and production



## **Products**

Identifying products as they emerge

## **Formats**

Understanding how drugs are sold and used

## **Effects**

Characterizing desirable and adverse events

## **Provenance**

Understanding product supply chains

## **Production**

Learning how products are made

---

# Learning About Populations

---

Understanding knowledge, attitudes and norms



## **Questions**

Examining  
information  
seeking

## **Transactions**

Identifying  
practices and  
expectations

## **Guidance**

Discerning  
peer-to-peer  
guidance

## **Motivations**

Building  
perspective on  
motivations

## **Governance**

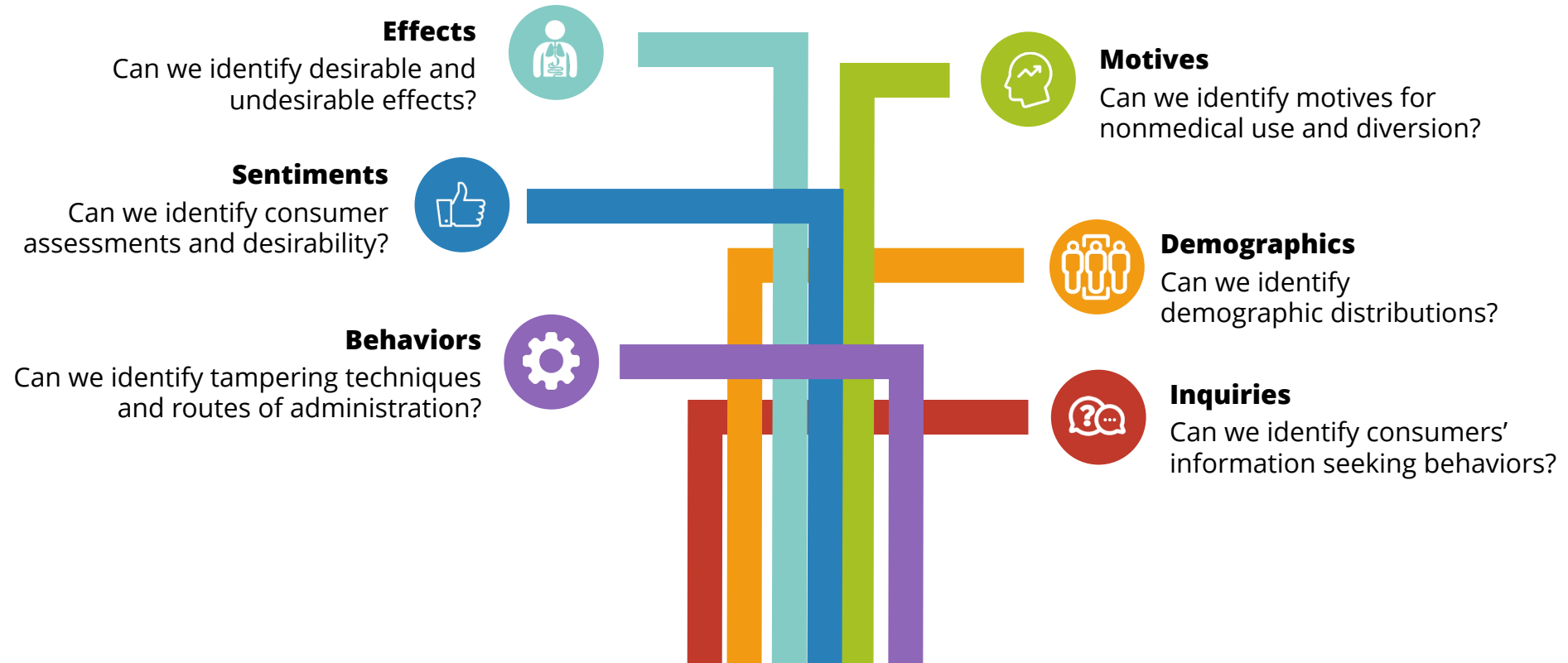
Exploring how  
communities  
self-govern

---



# Defining Aims

Identifying research questions



# Source Strategy

Selecting sources for data collection

## **Short Form Social Media**

Twitter, Instagram

## **Generalized Platforms**

Reddit, Facebook

## **Topic-specific Forums**

Bluelight, Subreddits

## **Experience Vaults**

Erowid, PsychonautWiki

## **Market Listings**

Cryptomarkets, 'RC' Vendors



# Search Strategy

Selecting terms for collection and annotation

## Prior Ontologies

MedDRA or findings from prior literature can be used to identify constructs and referent dictionaries

## Machine Derived Dictionaries

Given examples, similarity scoring tools like word2vec can identify novel and comparable terms



## Emergent Dictionaries

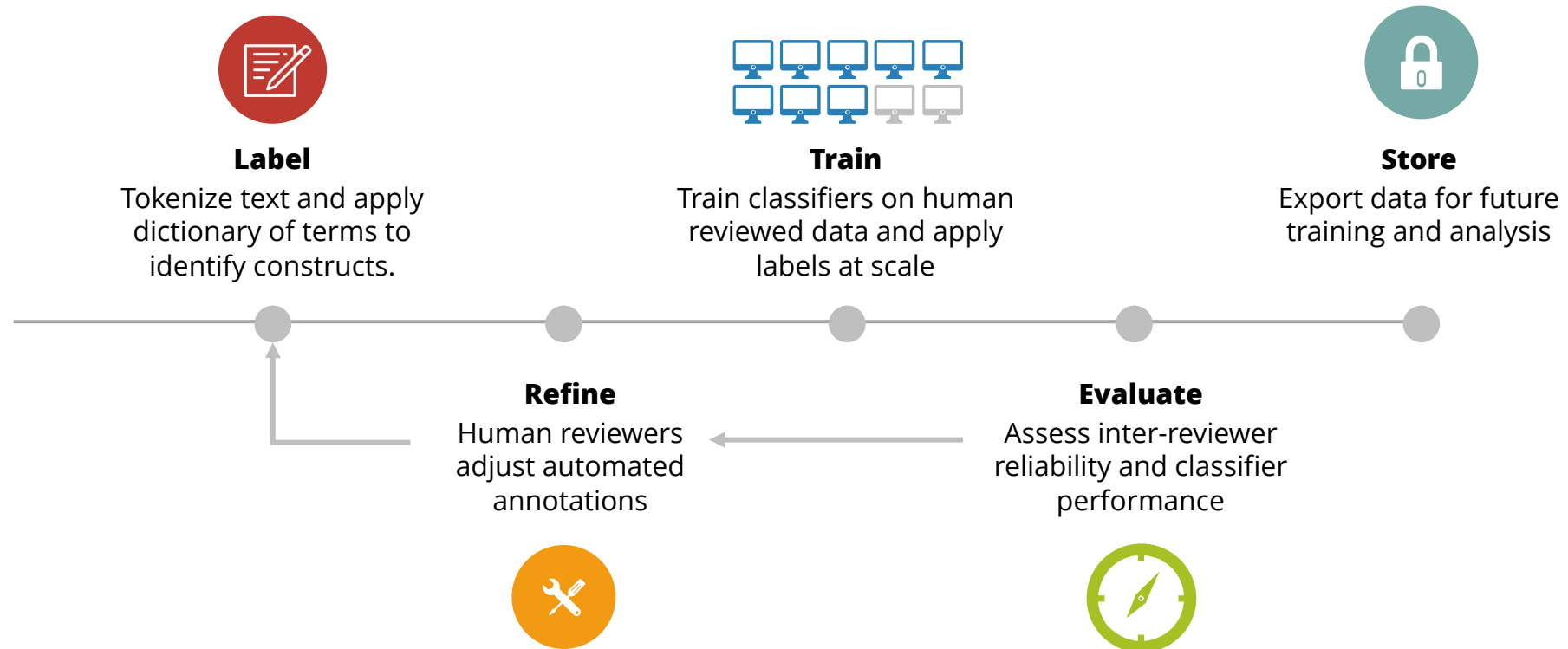
Qualitative review can be used to identify terms and phrases for implementation in a search query.

## Metadata Based Search

Metadata such as timing, network addresses, etc. can be used as alternatives or complements to textual terms

# Automating Annotation

Identifying and labeling entities

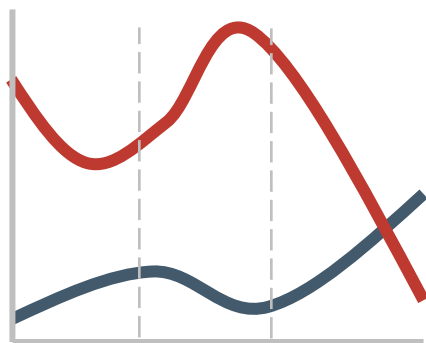


# Making Meaning

Post-processing analyses

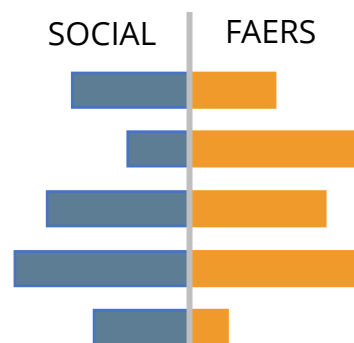
## Descriptive Comparisons

Track trends and correlations among constructs



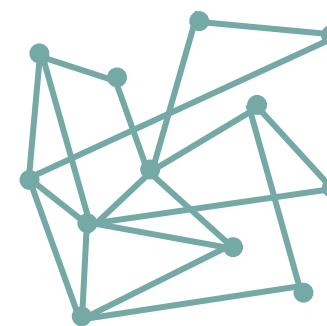
## Signal Detection

Apply PRR analyses against alternate data sources



## Network Analyses

Identify clusters within online communities



# Challenges and Limitations

Considerations for further work



## **Representativeness**

Who's in the room

## **Granularity**

Persons and personas

## **Persistence**

Enduring disruption

## **Reportability**

Disclosing exposures

# Implications for Research

---

Considerations for further work



# Comment and Contact

---

Continuing the conversation



Gilbert\_Michael@bah.com

---