Supplementary material

We include detailed about the refactoring types and the coverage of Weaponized and Word Hurtlex datasets in the Harm Category.

Table 1: Refactoring that can be misused to inject harmful content

Category	Refactoring
Extract	Extract Class
	Extract Method
	Extract Superclass
	Extract Variable
Rename	Rename Class
	Rename Method
	Rename Field
	Rename Variable
Replace	Replace Command with Function
	Replace Conditional with Polymorphism
	Replace Constructor with Factory Function
	Replace Error Code with Exception
	Replace Function with Command
	Replace Magic Literal
	Replace Primitive with Object
	Replace Subclass with Delegate
	Replace Superclass with Delegate
	Replace Temp with Query
	Replace Type Code with Subclasses
Encapsulate	Encapsulate Collection
	Encapsulate Record
	Encapsulate Variable
Introduce	Introduce Parameter Object
	Introduce Parameter
	Introduce Special Case
Others	Change Method Signature
	Hide Delegate
	Parameterize Function
	Remove Flag Argument
	Separate Query from Modifier
	Split Phase
	Split Variable

Table 2: The coverage of Weaponized and Word Hurtlex datasets in the Harm Category. Y represents coverage, and N represents no coverage.

Categories	Weaponized Word	Hurtlex
Doxing	N	Y
Eating Disorder Promotion	N	Y
Extremism, Terrorism Organized Crime	Y	Y
Identity Attack	Y	N
Identity Misrepresentation	Y	Y
Insult	Y	Y
Threat of Violence	Y	N
Self-Harm	N	Y
Misinformation	N	Y
Sexual Aggression	Y	Y
Adult Sexual Services	N	Y
Child Sexual Abuse Materials	Y	Y
Scams	N	Y
Total Y	7	11

Table 3: Coverage of refactoring types and corresponding repositories in the large-scale benign-program dataset (Y = included; N = not included).

Refactoring	Include	Repository
Extract Class	Y	JUnit
Extract Method	Y	RStudio
Extract Superclass	Y	Netflix Eureka
Extract Variable	Y	JetBrains Intellij-community
Rename Class	Y	Apache Cassandra
Rename Method	Y	Spring Framework
Rename Field	N	Buck
Rename Variable	Y	OsmAnd
Replace Command with Function	N	Dropwizard Metrics
Replace Conditional with Polymorphism	N	JLine2
Replace Constructor with Factory Function	N	Graylog2-server
Replace Error Code with Exception	N	Hazelcast
Replace Function with Command	N	Google Auto
Replace Magic Literal	N	Spring Framework
Replace Primitive with Object	N	Deeplearning4j
Replace Subclass with Delegate	N	MPDroid
Replace Superclass with Delegate	N	Spring Data JPA
Replace Temp with Query	N	Fabric8
Replace Type Code with Subclasses	N	phishman3579/ java-algorithms-implementation
Encapsulate Collection	N	openHAB
Encapsulate Record	N	openHAB
Encapsulate Variable	N	Jackson Databind
Introduce Parameter Object	N	Clojure
Introduce Parameter	N	Apache Cassandra
Introduce Special Case	N	Clojure
Change Method Signature	N	Restlet
Hide Delegate	N	HdrHistogram
Parameterize Function	N	Apache Cassandra
Remove Flag Argument	N	Jenkins Pipeline-plugin
Separate Query from Modifier	N	Buck
Split Phase	N	Thunderbird for Android
Split Variable	N	Lealone